

Xbiom – A Workbench for Biomarker & Translational Research

February 2019, Version 1.0

*This white paper from PointCross Life Sciences Inc. (Foster City, CA) discusses the importance of **unifying clinical trial data, molecular and other biomarkers from patient bio-samples for search and a unified vision to gain insights**. We present the benefits of such a unified environment as a workbench for biomarker and translational researchers for discovering, developing, and applying biomarker for biologics based therapeutic drugs. The paper presents query masks for stratifying cohorts and search analytics on clinical and biomarker data as **available on the Xbiom solution** (www.xbiom.com). It is not intended to be an original presentation of a scientific development. This paper will be periodically republished to reflect new information and user’s insights. We welcome comments from our industry clients – please send them to insights@pointcross.com.*

PointCross Life Sciences Inc.
1291 E. Hillsdale Blvd., Suite 304
Foster City, CA 94404
Tel: +1 844-382-7257

www.xbiom.com
www.pointcross.com

Need for a Workbench for Biomarker & Translational Research.....	2
Convergence of Bio and Info Technologies	2
Traditional clinical trial data	2
Translational research and biomarkers	3
Biomarkers Purposes	4
<i>Growth of biomarker industry</i>	5
<i>Future biomarkers</i>	5
Unite patient’s clinical and biomarker data for search analytics	5
Integrate the research workbench to bio-sample tracking	7
Ingestion and loading of clinical and biomarker data	7
Xbiom™ Workbench for Translational and Biomarker Research	8
<i>Brief description of the Xbiom platform and solutions available</i>	8
Xbiom Clinical and Biomarker Insights Module – Workbench for Translational and Biomarker Research.....	10
Omics Data currently supported in Xbiom’s ontologies:.....	11
<i>Search Query Masks and Stratified Cohort Selection</i>	12
<i>Data visualization and IGO</i>	16
Getting Xbiom biomarker research workbench setup quickly and inexpensively	19
<i>Software as a Service</i>	21
<i>Contact PointCross Life Sciences for more information and demo</i>	21
References	22

Need for a Workbench for Biomarker & Translational Research

Convergence of Bio and Info Technologies

Two separate revolutionary advances in BioTech and InfoTech are now converging to impact our human experience in many ways¹. InfoTech has been following Moore's Law by doubling the number of components and capacity of electronic chips every 18 months or so since 1965. Software and network technology has taken this advance to evolve machines that learn, adapt and even mimic human intelligence. Since the completion of the first human genome sequencing in the early 2000s, the BioTech industry has entered its race with advanced high throughput assays and Next-Generation sequencers that decode the chains of four nucleotides that are part of all DNA, RNAs and the structure of amino-acid chains of proteins. This technology is moving ahead at the square of Moore's Law and the convergence is driving rapid change in the BioTech and BioPharma industry. Products of biological analysis today are as often in the form of digital data as it might be received in a vial. Recent developments in AI are equally applicable in both InfoTech and BioTech and the lines are getting blurred.

Some of these advances are bringing new targeted therapies and diagnostics for diseases such as cancer or rare and orphaned diseases. This is happening through our ability to peer into the very genomic code and molecular mechanisms that define and direct our bodies, and then using insights from that knowledge to design biologically derived molecules that can treat or mitigate the risk of diseases.

There are three important disruptions that are bringing rapid changes to the pharmaceutical drug development industry – the move to represent clinical study data in standardized digital machine readable form (CDISC); the rapid growth in the molecular, protein and metabolomic biomarkers and small molecules; and the coming tsunami of remotely, non-invasively, monitored sensor data that will generate new types of biomarkers in the near future. Many of these will prompt new advances in therapeutics.

Traditional clinical trial data

Clinical trials are traditionally run sequentially through Phases 1, 2a, 2b, 3 and a 4th phase after market. First studies in humans in Phase 1 and 2a assess safety, dose-limiting toxicity and the maximum allowable dose estimations, which are

checked with dose escalations studies. Phase 2b studies the activity and the overall response rates. Phase 3 is focused on efficacy as measured by end-points such as reversal of disease, overall survival or progression free survival in the cases of diseases such as cancer. With traditional clinical trials the bulk of the data collected were from observations in the clinic, lab tests, and various diagnostic tests such as Xrays, CTScans and such. The collected data are reported in pre-designed eCRFs (Case Report Forms), which are processed into study files (SDTM) that are processed by scripts to extract analysis data sets (ADaM).

A very positive disruption is the mandate from the FDA (and PMDA) since Dec 2016 to digitally represent the clinical study data in a standard exchange format in columnar tabulations that are designed for computer readability as opposed to human readable study reports with tabulations and figures. Clinical study data is now unlocked from proprietary formats in a clinical data management system. It is in an open CDISC SDTM format and the analysis data is available in the ADaM analysis data sets along with other descriptive files. Terminology for adverse events, lab, observational data, and pathology data are becoming standardized. This allows software to be easily analyze one study, multiple studies or to compare data across studies. It also means that virtual cohorts with patients from multiple studies can be tracked and studied. Within the constraints of informed consent and compliance to regulations, the clinical study data can be re-purposed for permitted research for re-targeting, diagnostics or other therapeutical areas.

Translational research and biomarkers

NIH² defines biomarkers as *“a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention”*

Collected lab tests (on blood, urine, serum, tissue) and observational data (gait, response, physical examination...) in clinical trials are biomarkers, by this definition.

Translational research bridges the “bench”, or lab, to the bedside (of the patient). Biomarkers, quantified at the bench or sequencing laboratories, are rapidly becoming important in diagnostics, prognosis, predictive markers are biologically derived molecules with somatic or hereditary mutations. These may be involved in assessing patient’s clinical outcome or developing targeted therapeutic drug molecules. They include genes, RNAs, Proteins and metabolomes. This disruption that has already changed the industry from that one that sought blockbuster small molecules to biologics that are very effective for precisely targeted patient

populations. These biomarkers and biologic molecules are built with an understanding of:

- Genomes that have insertions, deletions, translocations, SNPs and CNVs; or that carry epigenetic information, and allelic variations.
- Transcriptomes that include various mRNAs, miRNA, signaling RNAs, ribosomal RNAs, and tRNAs
- Proteomes that carry post translational modifications, and contribute to changes in regulatory mechanisms through activation or degradation.
- Metabolome that play a part in regulatory functions through enzyme (amino acids, lipids ...) kinetics, transport or accumulation

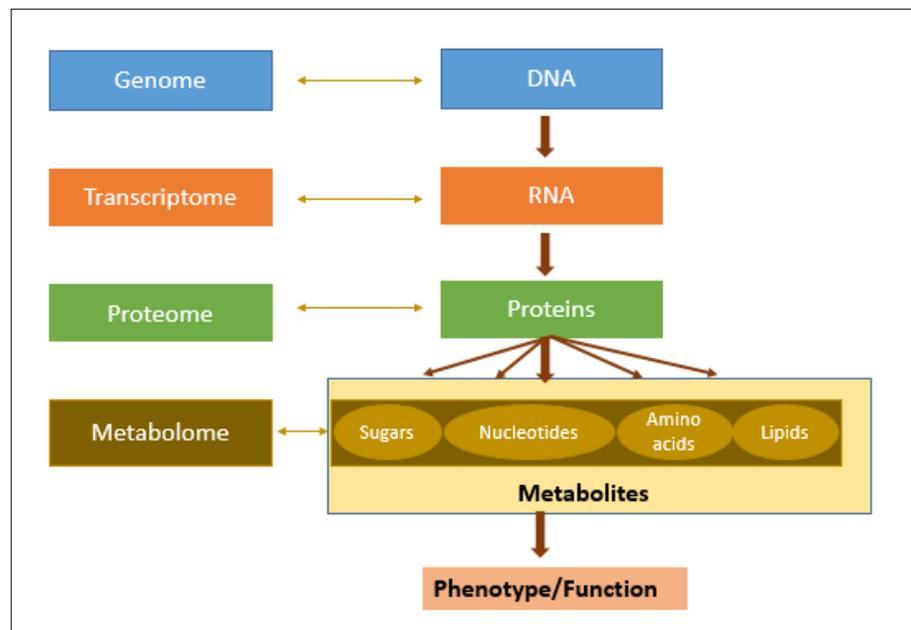


Figure 1 Omics Schema - Metabolomics (Image from Wikipedia, 2009⁶, https://en.wikipedia.org/wiki/File:Metabolomics_schema.png)

Biomarkers Purposes

A fast growing segment of industry is in identifying markers in patients through assays, sequencing and transcriptomic profiling with high-throughput laboratory technology and diagnostic services. Companion diagnostics is also a fast growing industry because increasingly precision medicines may be applied only on patients that meet very specific conditions and show specific markers that must be diagnosed.

Biomarkers play an important role in the monitoring and tracking of the disease progress, or reversal with treatment. The difference between biomarkers and therapeutic molecules can be blurred. Developing tumor specific therapeutic molecules such as in immuno-therapy and CAR-T connects diagnostic biomarkers for specific cohorts and their therapies.

Growth of biomarker industry

Translational research for Precision Medicine and demand for highly targeted therapeutics is driving rapid growth in the BioTech industry. Industry analysts^{3,4} estimate the global market for biomarkers at \$80 Billion in 2018 and growing at about 11%-12% annually. This growth is driven by the BioTech and the Big Pharma businesses focusing on biologics and precision-targeted medicines.

Future biomarkers

Another disruption to the drug development and healthcare industry is also a result of the convergence of the bio and info tech worlds. It is the availability and pervasive low cost non-invasive testing and monitoring systems. This bridges the patient to the clinic and it has major implications on how many aspects of clinical trials may be conducted in the future. Patients once visited clinics for their EKG, EEG, gait, response (to stimuli), mobility and other tests to monitor CV, CNS and other conditions. Wearables with sensors and radio technologies with smart software will soon monitor patients all the time especially during medical conditions - such as AFib heart fibrillations. The iWatch EKG is a recent example.

This data is raw and often streaming in real-time. Big data tools run algorithms in real-time on such streaming data. Deep learning applications will pick out markers of interest that will be streamed through the networks to be matched up with patient CRFs in on going clinical trials. This is emerging technology for the BioPharma and BioTech industry but these types of telemetry and extraction of markers have been in use in various industries including oil and gas exploration, autonomous vehicles, aerospace and defense. The BioTech industry is likely to benefit by leap frogging from this work. These types of biomarkers will cause disruptions to the way eCRFs are designed, the logistics of big data, and how these biomarkers will be incorporated into the standard of care of patients.

Unite patient's clinical and biomarker data for search analytics

Uniting patient data from clinical data (eCRFs, health history and baselines) and genotyping data along with the molecular and other biomarkers creates a very rich research environment. Researchers can search for, and find elusive cases where there is a correlation between the clinical end points in patients, their hereditary markers, and their markers from response to disease progress or therapeutic dosing. This is about uniting disparate data to provide the researcher

with a single, united vision for gaining insights. The integration occurs in the way the disparate data pools can be searched from a single vantage point.

Unified workbench of clinical and biomarker data offer many use-cases and benefits:

- Discover new or potential biomarkers for companion diagnostics, predicting response , or therapeutic monitoring
- Improve benefits/risk assessment for finely stratified cohorts retrospectively within and across trials
- Design Adaptive trials
- Looking for longitudinal changes in clinical and biomarker end points in patients or cohorts
- Analyzing other data collected from patients in a cohort
- Testing hypothesis of companion diagnostics retrospectively

The promise of biomarkers is very high, although success eluded many early efforts such as genome-wide association studies (GWAS) that yielded few clear markers such as a gene with a specific inherited mutation. For example, a pattern of repeated CAG in the Huntingtin gene was associated with Huntington's disease, but continuing research is revealing additional markers⁵ that are involved in the onset, monitoring and progression of this disease. Other known examples are CEA for colon cancer and PSA as a detector for prostate cancer. Markers related to a disease or its cure can often involve proteins connected to multiple or fused combinations of more than one gene with specific types of mutations. The markers related to heterogeneous tumors are also often heterogeneous. Breast cancer is an example where the inherited BRCA1 and BRCA2 genes are possible predictors. But today the heterogeneity of breast cancer tumors is better understood and the treatment of a specific tumor, or in the case of heterogeneous tumors within a tumor, can mean treatment strategies that require assays to identify multiple proteins such as ER, PR and HER2 receptors with one approved assay⁶ looking at as many as 21 factors⁶. Understanding somatic mutations that play a role in cancer was elusive until promising research of implicated oncogenes showed the complexity of the interplay of mutations in genes and the various RNA. A complex network of multiple genes and multiple mutations are often implicated.

Immuno-therapies, such as modified CARs derived from patient's T cells, are promising breakthroughs⁷ in oncology and other diseases. Dealing with divergent tumor genetics, tumor microenvironments, and immune activation is challenging for immune Oncology therapies. This will raise the need to include related

biomarkers where differentially expressed immune signatures and density of diverse immune cell population within the microenvironments of tumor cells are measured and indexed for search. These offer promising therapeutic mechanisms to improve the clinical outcome for treatment of cancer patients.

Finding these biomarkers and their relation to the patient's clinical findings, pathology, or the somatic variation of their markers require search tools that can answer specific and targeted questions of the researcher from vast amounts of data in a large number of disparate data sources unified within the workbench.

Integrate the research workbench to bio-sample tracking

Bio-samples obtained from patients in legacy studies, with their informed consent for the research purpose, can be further analyzed, possibly with newer technologies. Imagine finding a highly stratified cohort with 10 patients found in one or more studies. But the search also points out that there are additional 25 patients who might have joined this highly stratified cohort but for the fact their previous screening did not look for the RNAs or proteins that the 10 patients were screened. A workbench is a convenient place from where a researcher can conduct a search, and know what deficiencies can be corrected within the constraints of compliance and ICFs. It is the virtual, digital laboratory for the researcher.

Ingestion and loading of clinical and biomarker data

Clinical trial data is collected in multiple sites, managed and analyzed by multiple CROs and the sponsor (the BioPharma conducting the trial). Data managers face many challenges in maintaining a single point of truth for this collected, analyzed and standardized (CDISC, SDTM, ADaM, TFL) data. Clinical trials have trial specific protocols and study designs. The data structures used to be trial specific and this has meant that most of the completed legacy studies are stored in data models that must be transformed. Local terminologies need to be standardized before they can be used for translational or even longitudinal research and analysis. This requires that the data models must be transformed to a consistent data model so it is searchable. Curation and ingestion for indexing implies smart transformation and mapping algorithms with provenance and support of machine learning.

Genomic biomarkers are reported with varying terminology and conventions to denote the location and nature of mutations, and whether genes are identified as wild type. These are basic definitions that need to be harmonized. The disparate platform and technologies used may provide different levels of expressions but those are important for researchers to evaluate the data. Similar to ingesting clinical data, biomarker data must be normalized for data models and

harmonized to a common format so that they are searchable within a unified workbench.

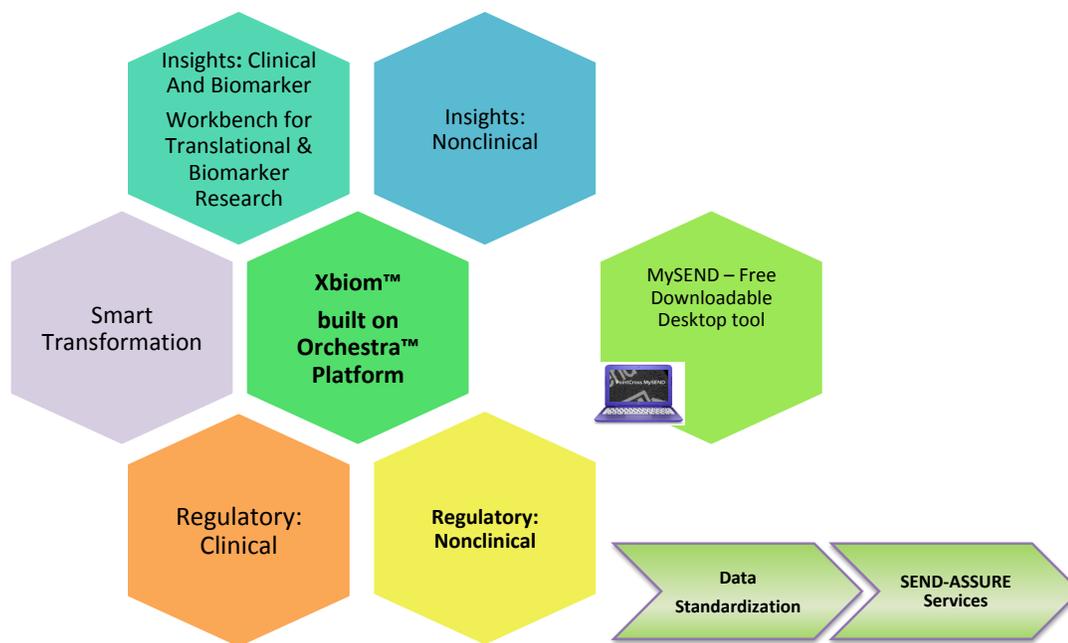
Ingestion of the data needs to be supported by a dashboard for assessing the quality of the data and driving the workflows.

Xbiom™ Workbench for Translational and Biomarker Research

Brief description of the Xbiom platform and solutions available

Xbiom™ is a commercially available stack of solutions for BioTechs and BioPharmas. It is currently in production for both. It is available for immediate use through a public cloud instance in AWS (Amazon), or as a dedicated instance in the cloud or private cloud. It is also in production as an on-premise installation at some big Pharma companies.

Xbiom[©] Suite for Research & Regulatory in Clinical & Nonclinical



Xbiom™ is built on top of the PointCross Ontology platform, Orchestra™. Xbiom™ uses a micro-services strategy to keep an evergreen architecture that is informed by biomarker and study data considerations, search, analytics, and immersive data visualization with a backdrop of security, workflow automation and artificial intelligence engines for machine learning. Xbiom has workflows and tools to

support regulatory processes in Clinical and Nonclinical; and it has workbenches for research in clinical translational and biomarkers and nonclinical. A brief description of modules supporting regulatory and nonclinical are listed below. A more detailed description of **Xbiom™ Insights – Clinical & Biomarkers**, the clinical and biomarker workbench for translational research is provided in the next section.

Xbiom™ Regulatory NonClinical automates the conversion and preparation of nonclinical studies to the required SEND IG version and the selected Controlled Terminology (CDISC). SEND and non-GLP studies can be ingested into the Xbiom Insights Nonclinical workbench for cross-study toxicology and drug safety research.

Xbiom™ Regulatory Clinical streamlines study workflows between CROs, Sponsors and the statistical teams preparing clinical trial data for submission to regulatory agencies. Clinical study data that needs to be available for translational research can be read by the Xbiom™ Insights – Clinical Biomarker workbench. Additional services and a freely downloadable tool for validation of CDISC standardized clinical and nonclinical data is available.

Xbiom™ Insights – Nonclinical is a workbench for discovery of toxicology signals and patterns across studies and species within Nonclinical studies. This module permits easy cross-study search through the use of a Nonclinical data repository, which contains normalized data through semantic integration.

Xbiom™ Smart Transformation module allows data managers to curate and transform clinical, nonclinical and biomarker data from data-lakes to a target model. The solution allows users to rapidly handle disparate data from multiple sources, harmonize and normalize into a single Development and Research, FAIR repository. This process is supported and controlled by Xbiom Metadata Governance to ensure consistency of the sponsor proprietary ontologies.

Xbiom Clinical and Biomarker Insights Module – Workbench for Translational and Biomarker Research

Xbiom™ Clinical Insights is the module for the translational and biomarker research workbench. It is a semantically unified and searchable resource of available clinical trials, molecular biomarkers, bio-sample availability and bio-sample tracking data and bioinformatics registries. The platform provides secure access to users and other applications to search for highly stratified cohorts, retrospectively or prospectively, using simple query masks and easy Boolean logic across all datasets collected during the drug development lifecycle.

Xbiom's ontology is curated and carries all the genetic variants as and when they are encountered. As biomarker data is ingested a smart transformation engine with a machine-learning algorithm maps and ingest the data before it is indexed. Clinical data is normalized and harmonized to the "global" unique identity so that the indexed search engine will respond with the right result even when a synonym is used in the search. A comprehensive dashboard to manage the ingestion is provided and client data managers can be trained on its use.

Xbiom is built on top of diversified data types from high-throughput technologies including NGS, Mass Spectrometry, Array or Seq based Expression, Flow Cytometry, and FISH technologies. This provides researchers with easier ways to assimilate and incorporate finer classification of biological sub-strata to look for precise use of a targeted therapies.

Industry continues to discover, or re-evaluate, complex genome involved phenomena such as epigenetics and cell signaling, and vast stretches of noncoding sections of the genome, once thought of as "junk DNA", in regulatory mechanisms. Rapidly reducing cost and increasing technology of Next-Generation sequencing of DNA and RNA with WGS and whole exome sequencing along with advances in immunohistochemistry and image processing is making the consumable biomarker metadata easier to obtain and use in search analytics across all the data regimes. The data is still huge and considerable technology goes into making it easy for the bench scientists to interact with it without a bevy of data wrangling programmers and statisticians. Biomarker or translational scientists need a single digital window provided by Xbiom's workbench into the bench and the bedside of clinical patients.

Biomarker data supported currently in Xbiom

Omics Data currently supported in Xbiom's ontologies:

OMICS	Measurement	Measurement Types	Assays, Technologies, Methods	Data Vendors	Reference Dictionaries	Current Biomarker Counts
Genomics	Genetic Variation	Short Variant	DNA Targeted Sequencing, Digital PCR	Foundation Medicines; Sysmex Inostics; Applied Biosystems	NCBI Gene Info, COSMIC	
		Copy Number Alteration				
		Rearrangements				
Transcriptomics	Sequencing	Targeted RNA Sequencing	Targeted RNA Sequencing	Illumina	NCBI Gene Info	59670 ^{**} (Genes)
	Expression	Gene Expression; Targeted mRNA expression	Microarray, Nanostring	AltheaDx; NanoString Technologies	NCBI Gene Info	
		miRNA expression	qRT-PCR	Covance	miRBase	2656 (miRNA)
		*Targeted mRNA detection	In situ Hybridization			
Proteomics	Expression	Expression of Proteins	Immunohistochemistry	Mosaic Laboratories	UniProt (SwissProt)	20214 (Proteins)
	Concentration	Concentration of Proteins	Immunoassay, ELISA, Multiplex Assay	Myriad RBM		
	Immune cells Population	*Cell Sorting (Cluster of Differentiation) Cell surface proteins	Flow cytometry			
Metabolomics	Concentration	Concentration of Metabolites	Liquid Chromatography-Mass Spectrometry	BIOCRATES Life Sciences AG	Human Metabolome Database (HMDB)	114100 (Metabolites)

Note: * indicates 'Work In Progress'

Note: ** Gene count of 59,670 includes:

Types of Gene (Current Total 59,670, NCBI)	Counts
Protein coding	20684
Non coding RNA	18079

Types of Gene (Current Total 59,670, NCBI)	Counts
Pseudo genes	16392
Unknown	2467
Other	866
Transfer RNA (tRNA)	642
Small nucleolar RNA (snoRNA)	431
Small Nuclear (snRNA)	63
Ribosomal RNA (rRNA)	43
Small Conditional RNA (scRNA)	3

Search Query Masks and Stratified Cohort Selection

A key feature of the Xbiom workbench is the query screens where the researcher is provided a comprehensive list of parameter to search based on a compilation of all the data available from all patients and biomarker data available. The query masks are separated into convenient segments within which the range of available data for any selected parameter is provided to the user as a prompt. The selections in the query masks can be combined with Boolean logic. The search results can be saved and the search criteria can be saved so that it can be shared with colleagues or re-applied as new data becomes available to monitor ongoing studies.

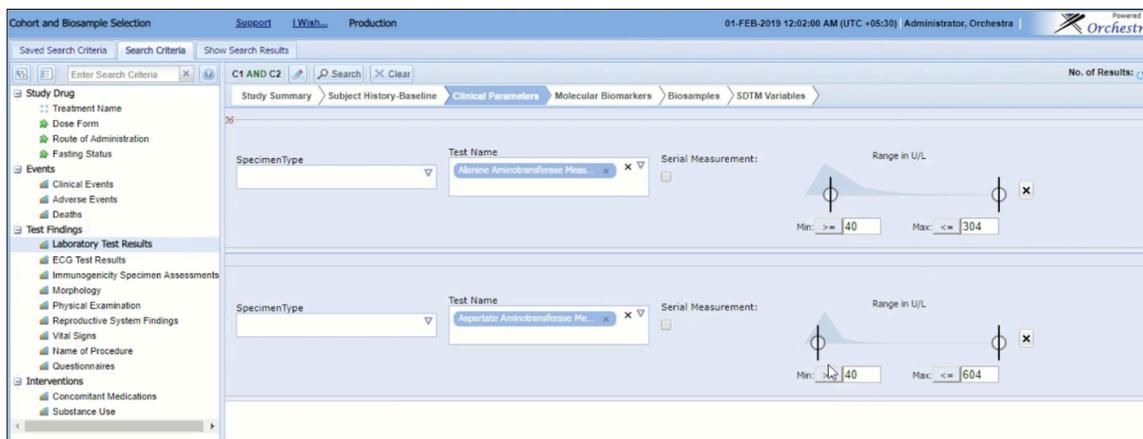


Figure 2

Study Summary

This is a simple query mask to limit the search to specific studies, or candidate drug molecule, or study type among many attributes associated with the Trial Summary.

Subject History Baseline

This query mask allows the user to set various parameters that define the cohort of interest including their age (as limited by the anonymization rules), sex, height race etc. and their genotype as well as specific conditions in their medical history. All available baseline lab work prior to the clinical trial is also cataloged here and may be set in the search criteria.

Clinical Parameters

This query set up the specific ranges of clinical findings or end-points that the researcher is looking for – e.g. ALT between two set values, or a combination of ALT, Bilirubin and other lab biomarkers that define a certain type of response on the patient. Percent changes from patient’s baseline will also be available in the next release.

Molecular Biomarkers

The user specifies DNA, RNA or Proteins and the list of available mutations in the database are made available for selection. When I-O (immuno-oncology) biomarkers are added these will include the quantitative magnitude or amount of expression. A number of screens showing how these biomarkers are selected for search are shown below.

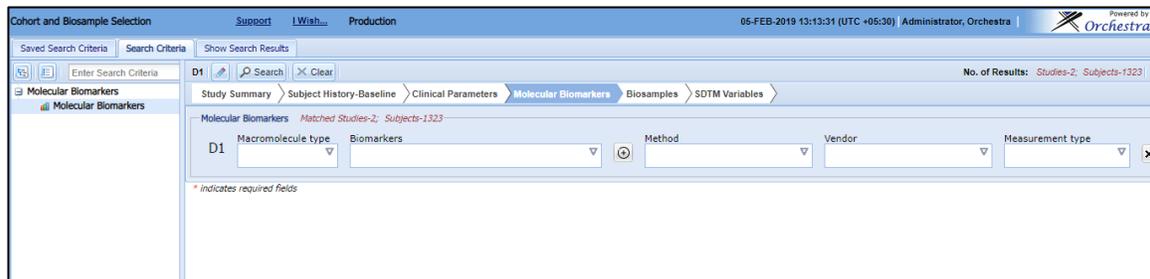


Figure 3

The screenshot shows a detailed view of 'Molecular Biomarkers'. On the left, there are tabs for 'Short Variants', 'Copy Numbers', and 'Rearrangements'. The main area is titled 'Short Variants' and includes a 'Mutational Status' dropdown menu. Below this is a table with the following columns: Gene, Gene ID, Coding Sequence alteration, Protein alteration position, Protein alteration, Genomic position, and Technique. The table contains three rows of data for the ABL1 gene.

Gene	Gene ID	Coding Sequence alteration	Protein alteration position	Protein alteration	Genomic position	Technique
<input type="checkbox"/> ABL1	25	3319A>C	1107	N1107H	chr9:133760939	DNA targeted sequencing
<input type="checkbox"/> ABL1	25	3343A>G	1115	S1115G	chr9:133761020	DNA targeted sequencing
<input type="checkbox"/> ABL1	25	3344G>C	1115	S1115T	chr9:133761021	DNA targeted sequencing

Figure 4

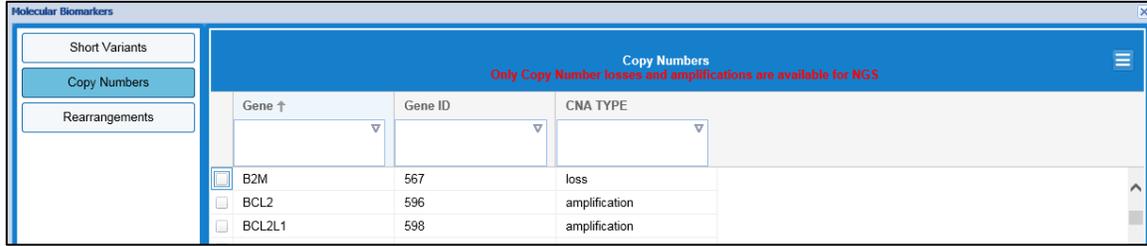


Figure 5



Figure 6

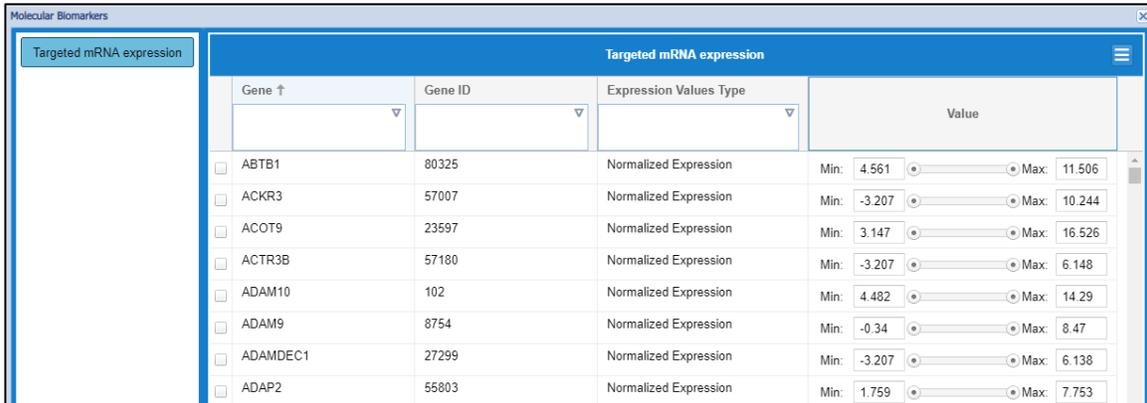


Figure 7

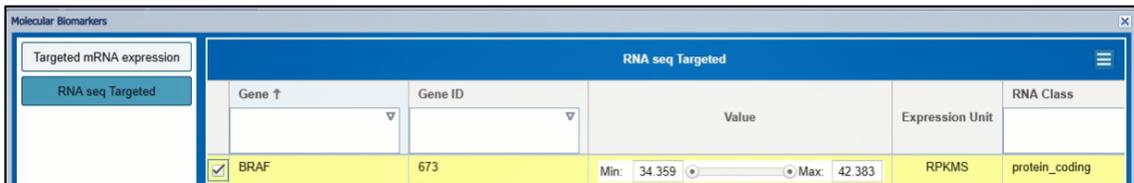


Figure 8

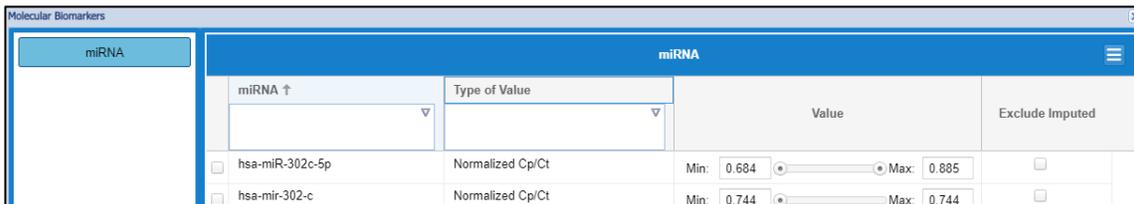


Figure 9

Protein Biomarker	Protein ID ↑	Include BLLOQ	Include Not Available	Value
<input type="checkbox"/> C-C motif chemokine 24	O00175	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Min: 111
<input type="checkbox"/> Agouti-related protein	O00253	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Min: 159
<input type="checkbox"/> Tumor necrosis factor receptor superfamily member 11B	O00300	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Min: 1.2
<input type="checkbox"/> C-C motif chemokine 21	O00585	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Min: 165
<input type="checkbox"/> C-C motif chemokine 22	O00626	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Min: 53

Figure 10

Protein Biomarker	Protein ID ↓	Examination Detail	Value	Unit	Method
<input type="checkbox"/> Ki-67	P46013	3+	Min: 35.41 Max: 35.41	%	immunohistochemistry
<input type="checkbox"/> Ki_67	P46013	1+	Min: 7.335 Max: 7.335	%	immunohistochemistry
<input type="checkbox"/> Ki_67	P46013	Percent_Positive+	Min: 0.445 Max: 41.65	%	immunohistochemistry

Figure 11

Molecule Name ↑	Include BLLOQ	Include Not Available	Concentration	Unit
<input type="checkbox"/> Chenodeoxycholic acid	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Min: 0.011 Max: 3.4	
<input type="checkbox"/> Chenodeoxycholic acid glycine conjugate	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Min: 0.03 Max: 20.9	
<input type="checkbox"/> Cholic acid	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Min: 0.013 Max: 2.4	
<input type="checkbox"/> Deoxycholic acid	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Min: 0.01 Max: 1.3	
<input type="checkbox"/> Deoxycholic acid glycine conjugate	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Min: 0.002 Max: 11.9	

Figure 12

Bio-samples

Most BioTechs or big Pharma engaged in translational research or development of biologics will likely have a strategy for managing the clinical trial patient’s bio-samples through a bio-bank partner. These samples can often be re-analyzed to obtain data that was not collected earlier. If the cohort search shows that some potential patients meet all the criteria except that their samples were not screened for a specific proteomic, transcriptomic or metabolomic assay then the availability of viable bio-samples is important in discovering the potential availability of additional data. This is where retrospective translational analysis of clinical trial data can be powerful.

SDTM Variables

Although Xbiom is a workbench that can be used directly by clinical and biomarker researchers most big Pharma companies have a legacy team of data managers, programmers, and bio-statisticians analyze their clinical data to the

statistical analysis plan. Recently that data is becoming coded digitally into CDISC SDTM standards for submission to regulatory agencies. This means that Xbiom users could be the same team that normally works with the clinical operations teams. Xbiom provides a query mask where SDTM variables can be used so that these users can make the same queries in terminology with which they are familiar.

Data visualization and IGO

Xbiom has two powerful ways for biomarker researchers to access, view and analyze the results of their search and findings.

1. Xbiom includes a built-in IGO (Interactive Graphics Object) that allows the user to interact with the data as in most modern data visualization and analytics platforms such as SpotFire, JMP, Qlik, Tableau, and others. These are point and click applications with considerable flexibility.
2. Xbiom provides an OData open interface that allows external 3rd Party Applications (SAS, SpotFire, R, Jreview, and many others) to interact and let the user to drive queries from scripts as if they were logged in to Xbiom and still get access to the search and analytics results. Users will interact from applications such as R, SAS, Python using interactive interpretive tools like Jupyter.
3. Gene Ontology (GO) Enrichment

The user specifies genes of their interest obtained from high throughput “x”RNA expression experiments (ex: Microarray, Nanostring, RNA-Seq). By Internally mapping gene to GO terms in Xbiom, the functional annotation and gene enrichment analysis is performed by Fischer’s Exact Test, to determine collective biological function which are statistically over-or under-represented within the list of interesting genes.

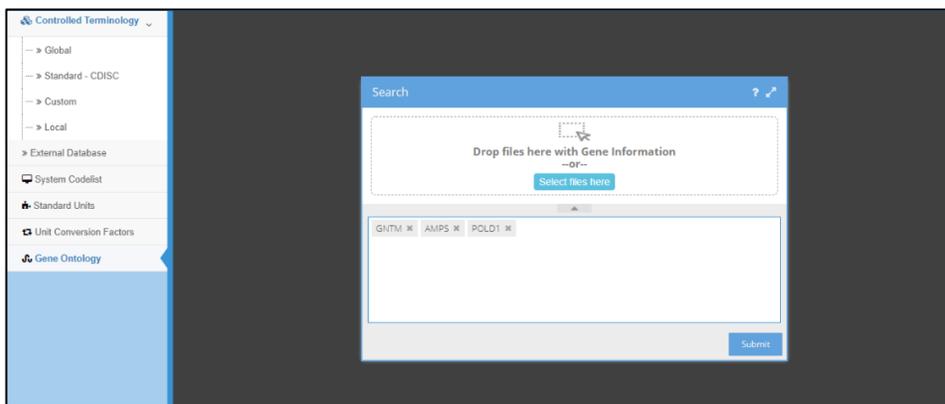


Figure 13

GO ID	GO Term	GO Category	Observed Genes	# Observed	# Expected	Diff: Sample & Observed	Diff: Human & Expected	Observed / Expected	P-Value
GO:0050892	intestinal absorption	Process	0	0	58	2	20611	0	1
GO:0006189	de novo IMP biosynthe...	Process	0	0	94	2	20575	0	1
GO:0006164	purine nucleotide biosy...	Process	0	0	186	2	20533	0	1
GO:0009168	purine ribonucleoside ...	Process	0	0	39	2	20630	0	1
GO:0006493	protein O-linked glycosy...	Process	0	0	146	2	20523	0	1
GO:0051262	protein tetramerization	Process	0	0	174	2	20495	0	1
GO:0016266	O-glycan processing	Process	0	0	74	2	20595	0	1
GO:0005975	carbohydrate metabolic...	Process	0	0	1676	2	18993	0	1
GO:0005829	cytosol	Component	0	0	14696	2	5973	0	0.514257
GO:0000139	Golgi membrane	Component	0	0	1978	2	18691	0	1
GO:0016020	membrane	Component	0	0	31512	2	-10843	0	0.156910
GO:0070062	extracellular exosome	Component	0	0	9070	2	11599	0	1
GO:0016021	integral component of ...	Component	0	0	38960	2	-18291	0	0.120161

Figure 14

Some examples of the IGO are included here.

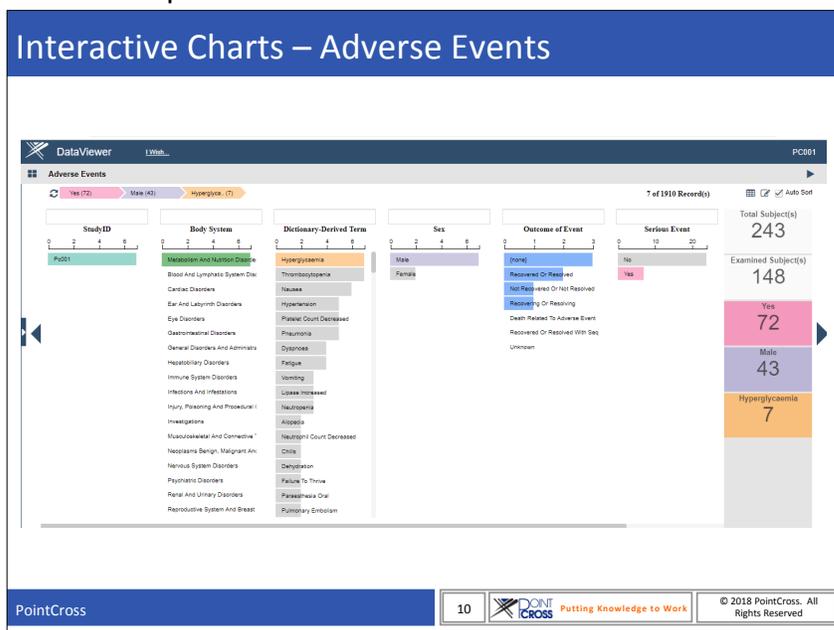


Figure 15

The following chart shows as an example:

- How many DNA copies are amplified for each gene- within or across studies?
- Compare Copy Number Alterations across genes and across cohorts

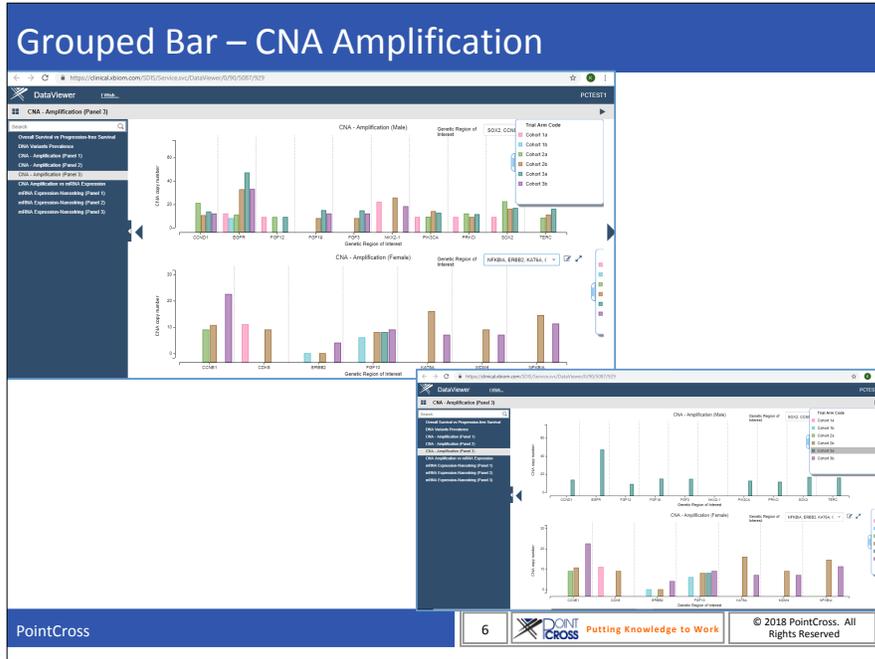


Figure 16

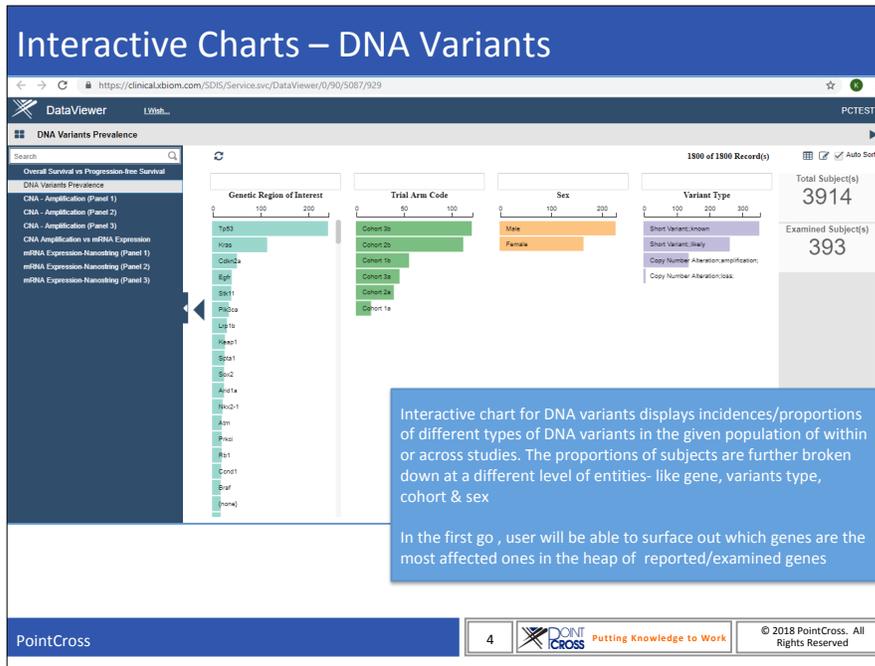


Figure 17

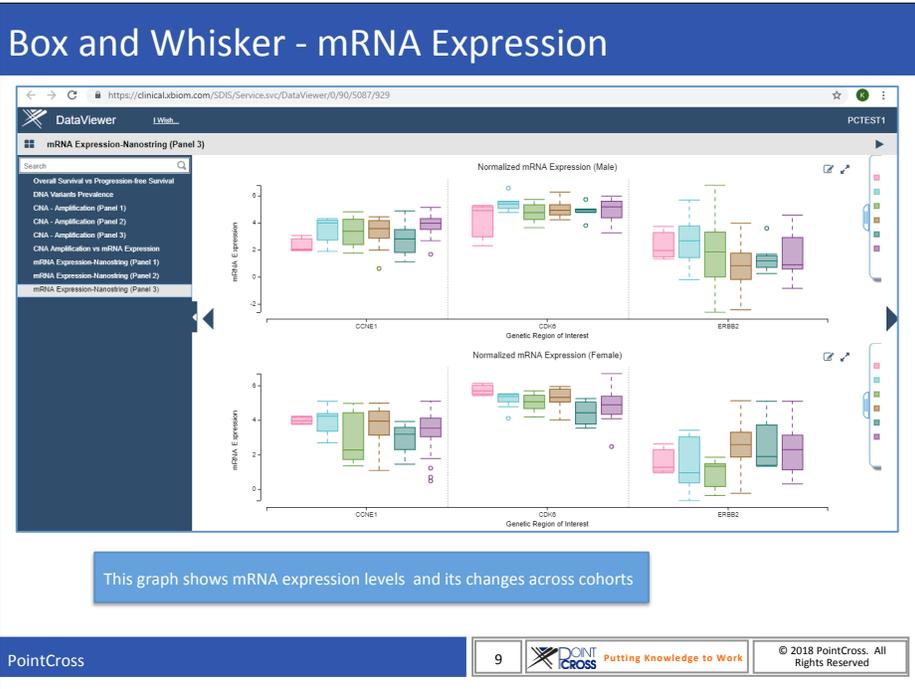


Figure 18

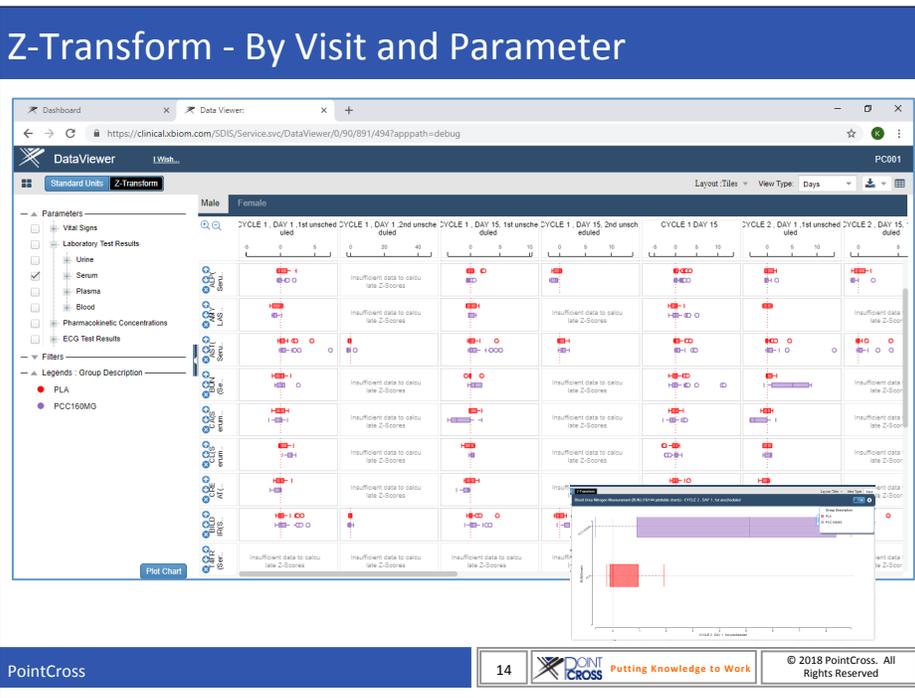


Figure 19

Getting Xbiom biomarker research workbench setup quickly and inexpensively

Xbiom is a comprehensive and powerful workbench but it can be set up and ready to be used with secured data of the BioTech's. PointCross sets up an Xbiom instance on the AWS cloud under a subscription agreement (AWS, Amazon is

partnered with PointCross) in a matter of days. PointCross also provides data services to curate and load clinical study and biomarker data for a fee based on the number of studies and data volume.

Some of the steps to getting the Xbiom workbench set up:

1. Request an online demo or a demo video from insights@pointcross.com
2. If there is interest, enter into an NDA or confidentiality agreement. You will also be provided a standard price sheet with the subscription prices and the set up fees and services for loading your data.
3. Get an inventory of all the studies and biomarker data and get a firm fixed price for the data services.
4. Begin the subscription service for Xbiom.
5. Use the Virtual Data Room to share the data. PointCross will load the data and contact the assigned point of contact to get any clarifications needed on the data.
6. User training or orientation will be provided to the key users prior to handover.
7. PointCross will periodically reach out or invite key users to its roadmap special interest group so that your technical needs can be included in the product roadmap.

A dedicated private cloud hosting or on-premise options are available on request.

Why should BioTechs and Biologics focused companies consider Xbiom?

1. The last BioTech that signed up with PointCross went live within a day after they subscribed to the Xbiom SaaS service
2. Their clinical data was curated and ingested and loaded in 2 weeks
3. It is a single workbench for biomarker research – one-stop to data, analysis and insights
4. It respects confidential clinical data
5. It integrates the knowledge available in the global registries
6. Discover new or potential biomarkers for companion diagnostics, predictors, or therapies
7. Improve benefits/risk assessment for finely stratified cohorts retrospectively within and across trials
8. Adaptive trials

9. Looking for longitudinal changes in clinical and biomarker end points in patients or cohorts
10. Analyzing other data collected from patients in a cohort
11. Testing hypothesis of companion diagnostics retrospectively

Software as a Service

Our Xbiom software as a service ensures that new capabilities that help researchers are constantly made available. We are also constantly updating the various ontologies for the biomarker lists by referencing authoritative registries. All of these are available allowing the researchers to access the latest.

Contact PointCross Life Sciences for more information and demo

A variety of resources stand ready to help get more information on Xbiom, and how to get started.

Visit our Xbiom website at www.Xbiom.com

Visit the company at www.pointcross.com

Email us at insights@pointcross.com to set up an online demo or a on-site meeting

PointCross Life Sciences Inc.
1291 E. Hillsdale Blvd., Suite 304
Foster City, CA 94404
Tel: +1 844-382-7257

Chris Miller for US based calls +1 310-528-4712
Hubert Germain for EU based calls +1 33 6 07 69 40 61

References

1. 21 Lessons for the 21st Century, Yuval Noah Harari, 2018, Spiegel & Grau, ISBN 9780525512172
2. What are Biomarkers?, Kyle Strimbu and Jorge A. Tavel, M.D., PMID: PMC3078627, NIHMSID: NIHMS259967, PMID: 20978388, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3078627/>
3. Biomarkers: Technologies and Global Markets, September 2018, BIO061E, BCC Research Staff
4. Biomarkers Market by Product, Type, Disease Indication, Application - Global Forecast to 2021, February 2017, BT2120, <https://www.marketsandmarkets.com/Market-Reports/biomarkers-advanced-technologies-and-global-market-43.html>
5. Making Meaningful Clinical Use of Biomarkers, Matthew J Selleck,¹ Maheswari Senthil,¹ and Nathan R Wall², [Biomark Insights](#). 2017; 12: 1177271917715236. <https://www.ncbi.nlm.nih.gov/pubmed/28659713>
6. Tumor Heterogeneity in Breast Cancer, Gulisa Turashvili and EdiBrogi, Published online 2017 Dec 8. doi: 10.3389/fmed.2017.00227, PMID: PMC5727049, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5727049>
7. The Breakthrough: Immunotherapy and the Race to Cure Cancer, Charles Graeber, November 2018, Hatchette Book Group, ISBNs: 978-1-4555-6850