GENOMICALLY DRIVEN CANCER TRIALS: AN EVOLUTION OF PERSONALIZED MEDICINE

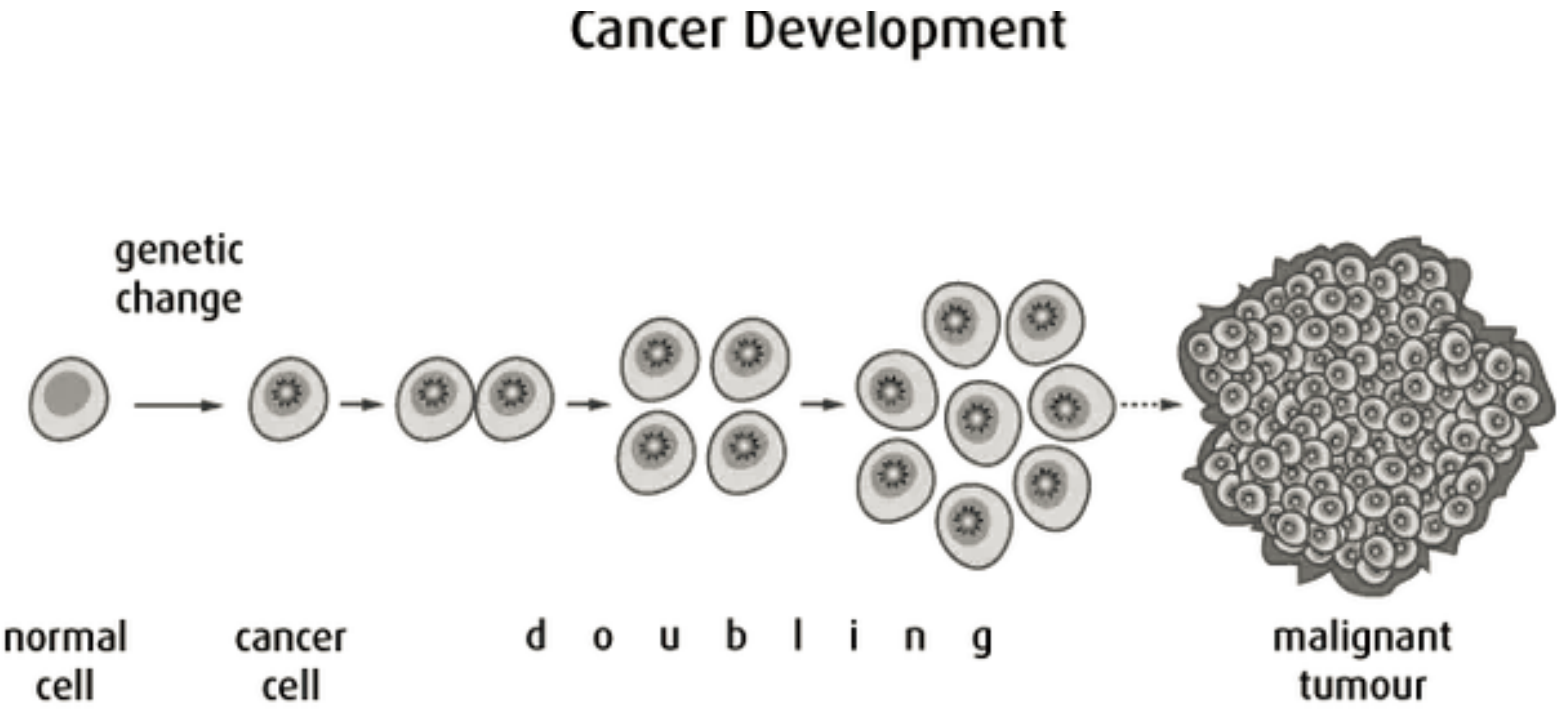
Kameshwari Peri, Senior Statistical Programmer, Syneos Health, Gurgaon

17th March 2018

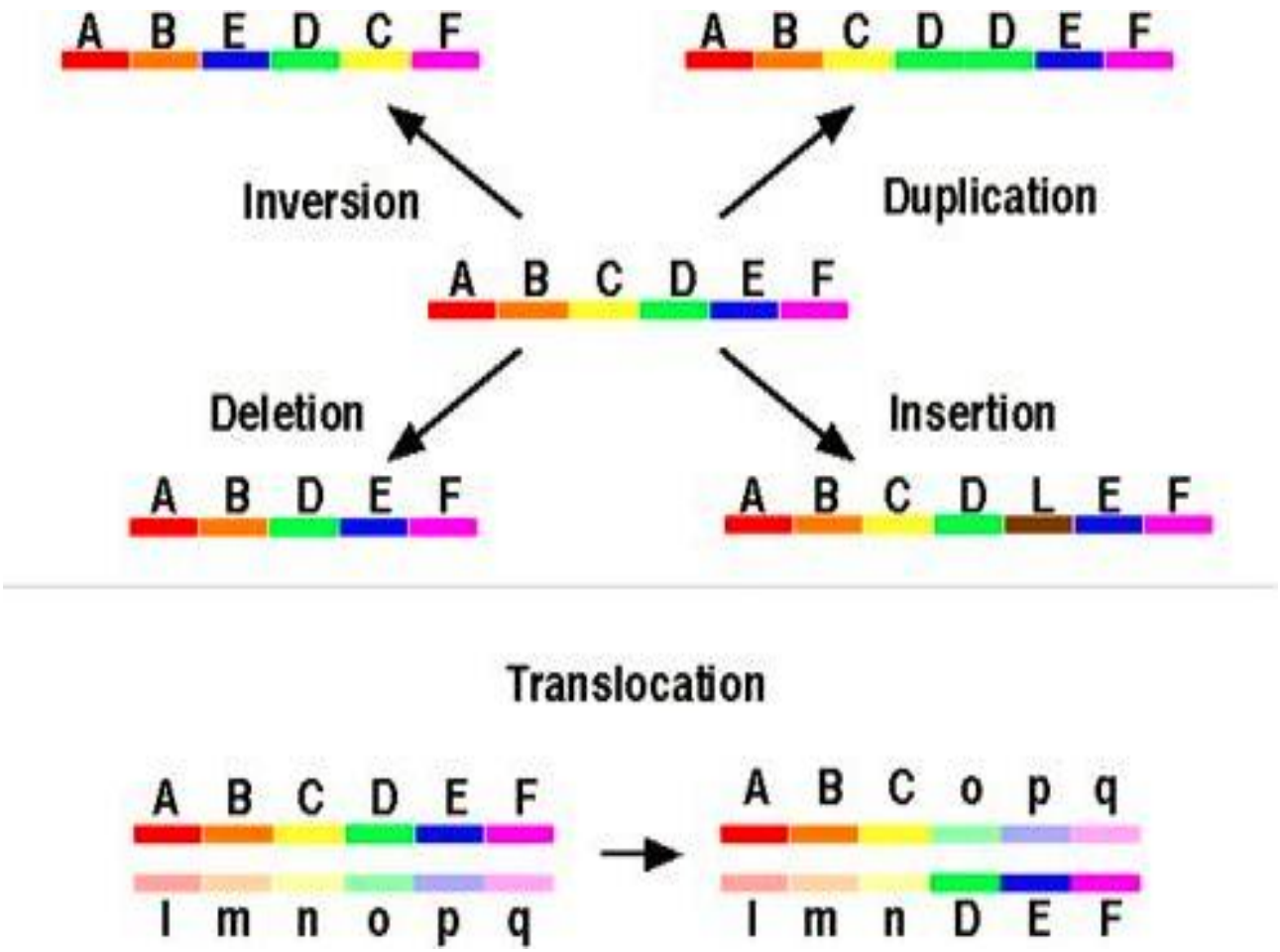
Contents

Cancer – The basics

- Uncontrolled proliferation of cells

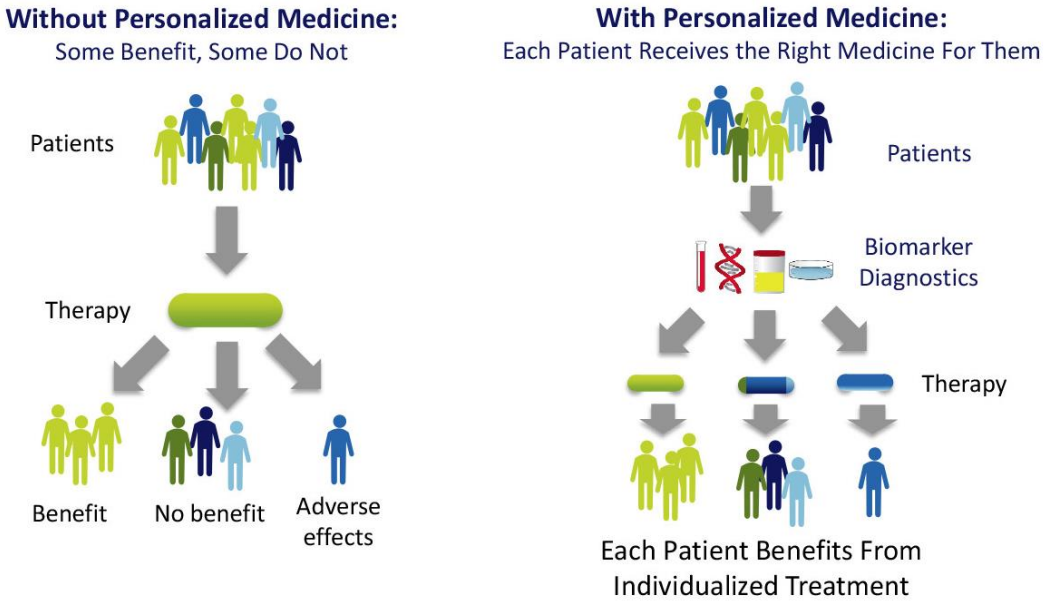


Cancer – The basics



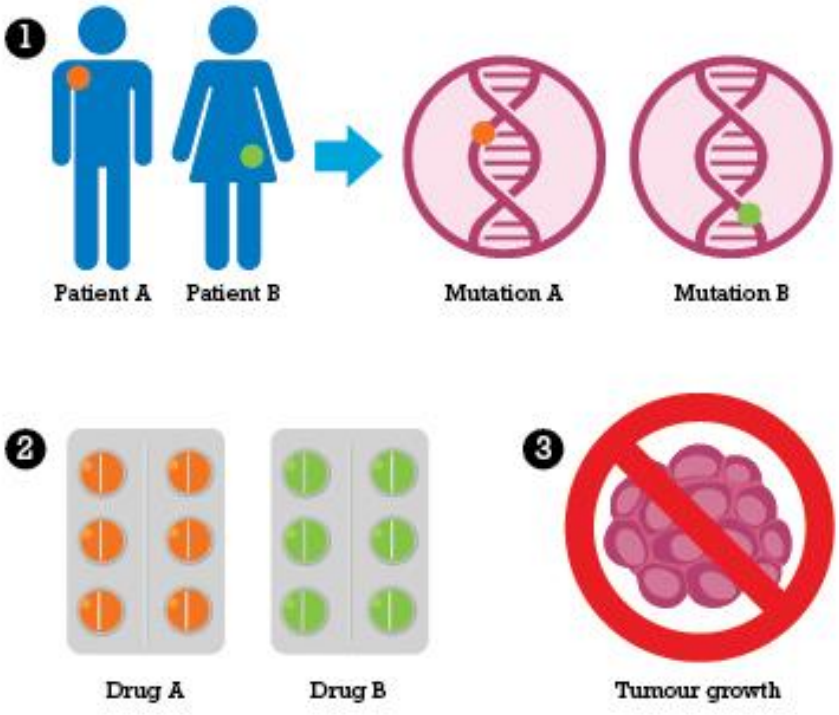
Personalized Medicine

- Helps us to
  - Understand person's genetic makeup
  - Identify effective prevention, screening and treatment strategies
  - Find treatments that cause fewer side effects than the standard options



Personalizing cancer treatment

- Every person in this world is unique and different, so are people's cancers.
- Targeted Therapy : choosing treatment that targets the specific features of that cancer



Personalizing cancer treatment

- Comparison of DNA between tumor and normal cell
- Tracking genetic profiles of patients' tumors
- Pharmacogenomics
- Precision medicine approaches consider
  - proper medication doses to maximize efficacy and
  - minimize side effects,
  - interactions with other drugs patients may be taking, as well as
  - factors from the environment, such as diet and exposure to toxins.

Precision medicine is not the future but present

- University of California, San Diego Health Sciences published in the June 6, 2016 issue of *JAMA Oncology*.

After reviewing 346 phase I clinical trials involving 13,203 patients, the authors found that in precision medicine treatment arms more than 30 percent of patients responded to treatment compared to 4.9 percent of patients enrolled in non-personalized arms. Patients treated with precision medicine also benefited from longer progression-free survival with a median of 5.7 months before their disease worsened compared to 2.95 months for the others

Table 1: Examples of successful DNA based biomarkers and targeted therapies in oncology

Target	Disease	Drug	Success
Her2Neu	Breast, stomach	Herceptin, Lapatinib	HR: +/- 0.50 for progression free survival for Her2+ breast cancer patients after surgery[3;4]
c-KIT	GIST	Imatinib	Majority of patients show impressive responses[5]
BCR-Abl	CML	Imatinib	>50% response in BCR-ABL positive CML[6;7]
ALK	NSCLC	Specific ALK inhibitor	Promising phase I data, to be confirmed.
PARP	BRCA 1 / 2 associated ovarian carcinoma, triple negative breast cancer	Multiple PARP inhibitors	Promising phase I/II data[8;9]
BRAF	BRAF mutant melanoma	Specific BRAF inhibitor	Promising phase I data[10]

Pharmacogenomics – the science and the data

- Pharmacogenomics (PGx) is a branch of pharmacology
- explains the relationship between genetic variations and drug response in patients by correlating gene expression or SNPs with a drug's efficacy and toxicity
- develop rational means to optimize drug therapy, with respect to the patients' genotype, to ensure maximum efficacy with minimal adverse effects.
- promise the advent of "personalized medicine"; in which drugs and drug combinations are optimized for each individual's unique genetic makeup

Pharmacogenomics – the science and the data

- Pharmacological tests are used to identify which patient will have toxicity from commonly used cancer drugs and identify which patient will not respond to commonly used cancer drug.

- The tests most commonly include gene expression analysis using microarrays, which can be performed on specific tissues specimens collected from the patients.

- Other tests may include SNP or any genetic polymorphism analysis, genotyping and a few more.

- The data one would expect out of such PGx tests are mainly biospecimen and genetic (DNA, RNA) samples along with their date and time captures including clinical significance information

PGx Data

- CDISC published the clinical data standard for genomics and biomarker data, Study Data Tabulation Model Implementation Guide: Pharmacogenomics/Genetics (SDTMIG-PGx) in May 2015.

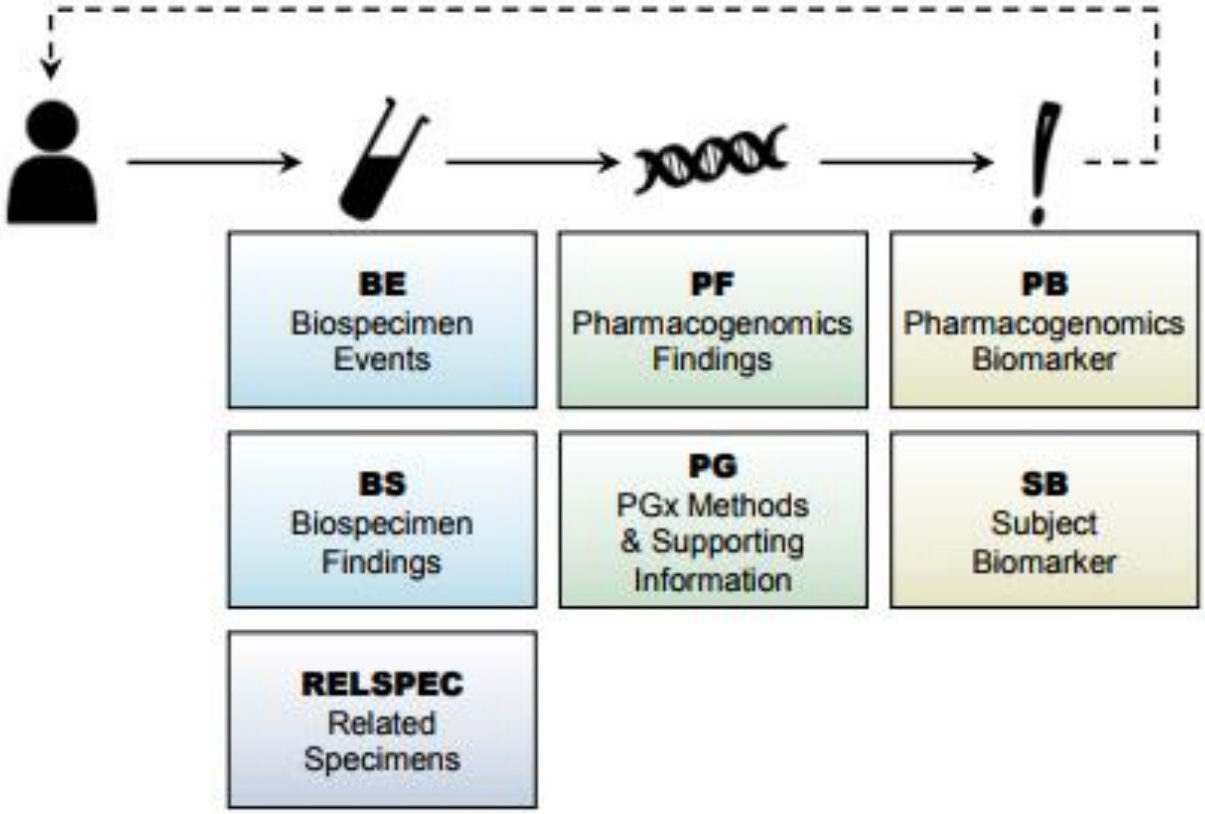
- The development of these CDISC PGx domains was done in parallel with the work being done by the HL7 Clinical Genomics Work Group (CG), which was initiated jointly several years ago by CDISC and HL7

- Since PGx is a part of biomedical studies, the data collected within PGx study has many characteristics of biological data. Actually, the PGx data is multiple dimensional and highly complex. PGxIG1.0 is structured to accommodate these characteristics.

Scope of SDTMIG-PGx

- Biospecimen Collection and Handling
- Quality data
- Genetic Mutation
- Genotyping
- Gene Expression
- Cytogenetics
- Viral Genetics
- Proteomics

CDISC SDTMIG-PGx Domains and Datasets



From SDTMIG-PGx v1.0

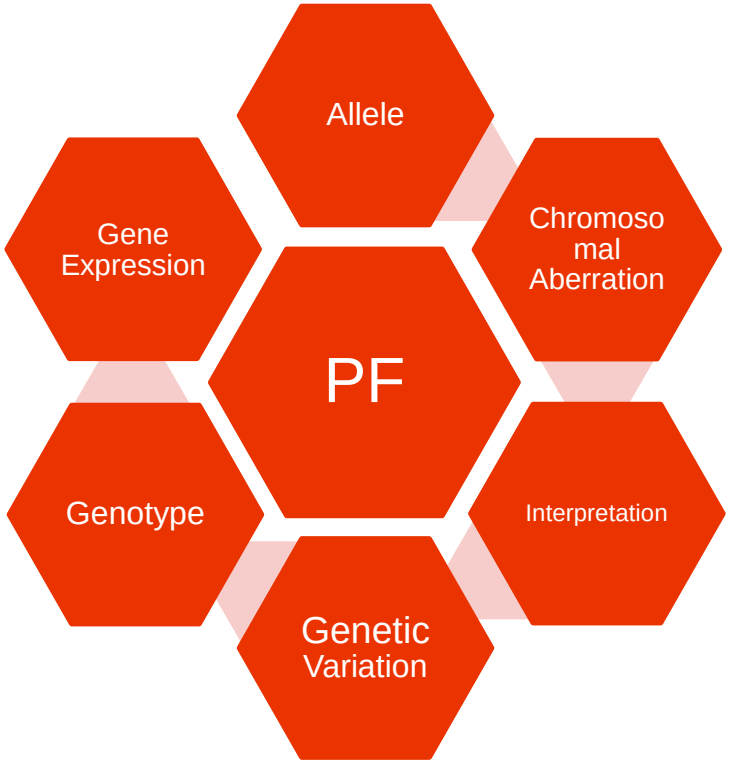
CDISC SDTMIG-PGx Domains and Datasets

Dataset	Description	Class	Structure	Purpose	Keys*	Location
PB	Pharmacogenomics/ Genetics Biomarker	Special Purpose Domains	One record per biomarker used in the study	Tabulation	STUDYID, PBMRKRID, PBMRKR	pb.xpt
SB	Subject Biomarker	Special Purpose Domains	One record per subject per observed biomarker in the study	Tabulation	STUDYID, USUBJID, SBMRKRID	sb.xpt
BE	Biospecimen Events	Events	One record per instance per biospecimen event per biospecimen identifier per subject	Tabulation	STUDYID, USUBJID, BEREFID, BETERM, BESDTC	be.xpt
BS	Biospecimen Findings	Findings	One record per measurement per biospecimen identifier per subject	Tabulation	STUDYID, USUBJID, BSREFID, BSTESTCD	bs.xpt
PF	Pharmacogenomics/ Genetics Findings	Findings	One record per finding per observation per biospecimen per subject	Tabulation	STUDYID, USUBJID, PFREFID, PFTESTCD, PFGENRI, PFSTRESC	pf.xpt
PG	Pharmacogenomics/ Genetics Methods and Supporting Information	Findings	One record per observation per biospecimen per run per subject	Tabulation	STUDYID, USUBJID, PGREFID, PGTESTCD, PGGENRI, PGNSPCES, PGNSTRN, PGDTC	pg.xpt
RELSPEC	Related Specimens	Special-Purpose Datasets	One record per specimen identifier per subject	Tabulation	STUDYID, USUBJID, REFID	relspec.xpt

From SDTMIG-PGx v1.0

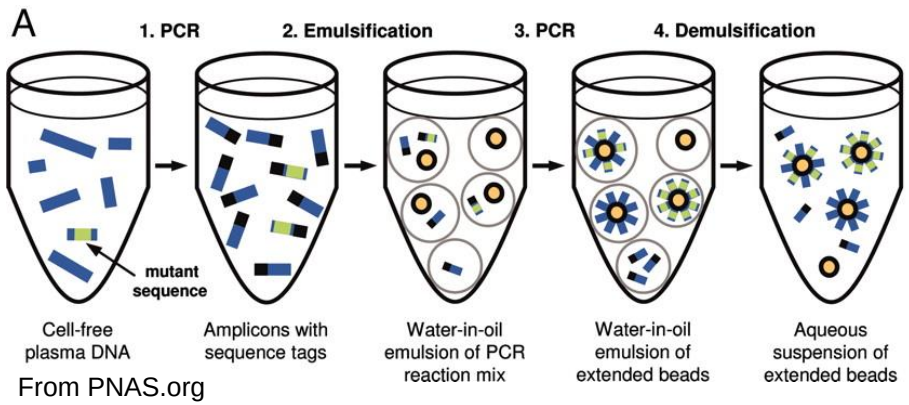
Pharmacogenomics/Genetics Findings domain

- PF domain captures results for both genetic variation and gene expression
- PFTESTCD and PFTEST generally identify the type of test performed by specifying the type of material assessed, such as nucleotides or amino acid, or the level of granularity, such as codon or allele
- PFNSPCES and PFNSTRN –used to differentiate between subject and microbial records
- PFMETHOD lists techniques for the execution of genomics or genetic testing.
- PFGENRI –portion of genome serving as locus for the test.



An example for SDTM PF

- Below is a test that determines whether certain somatic variations are present in a tumor sample
- BEAMing* -involves transformation of a population of DNA molecules into a population of beads each coated with thousands of copies of identical sequences
- The results are a count of fluorescent magnetic particles representing each genetic variation being tested.



From PNAS.org

Row	STUDYID	DOMAIN	USUBJID	PFSEQ	PFREFID	PFTESTCD	PFTEST	PFGENRI	PFGENLI	PFGENTRG	PFCAT
1	AAA	PF	AAA-XXX1	1	L1066774-3-S0072248	OBSLVL	Observed Level	KRAS	34	A	GENETIC VARIATION
2	AAA	PF	AAA-XXX1	2	L1066774-3-S0072248	OBSLVL	Observed Level	EGFR	2235_2249	del	GENETIC VARIATION

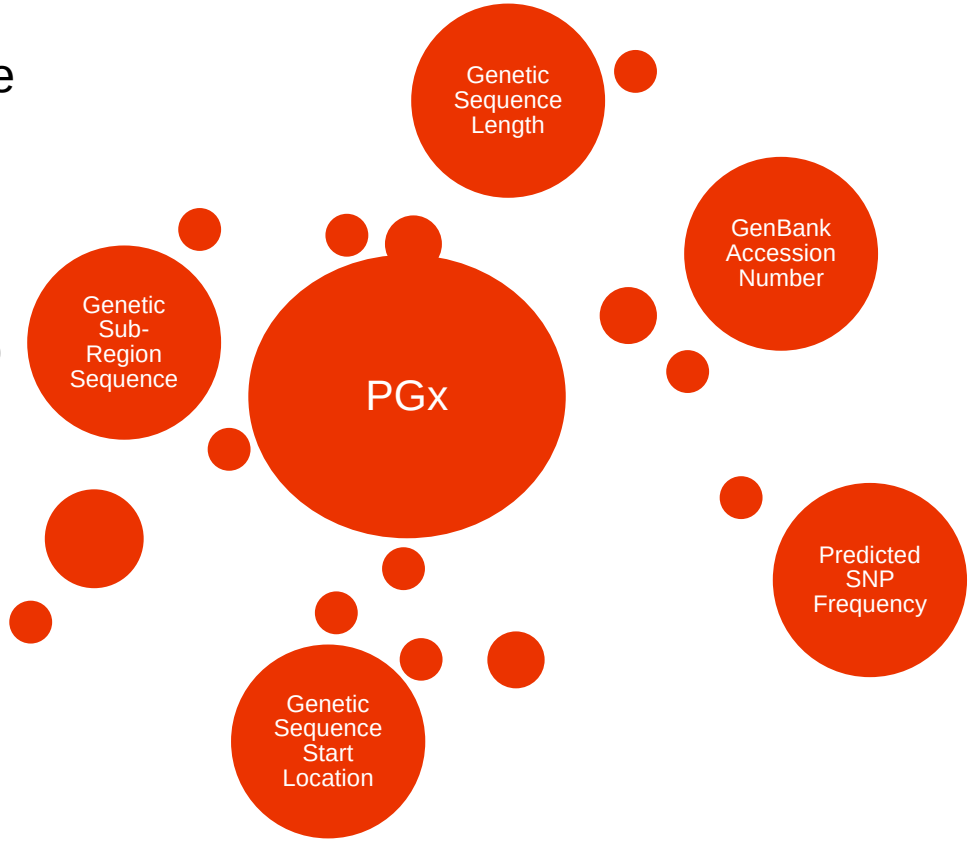
Row	PFORRES	PFORRESU	PFSTRESC	PFSTRESN	PFSTRESU	PFRESCAT	PFMUTYP	PFNAM	PFMETHOD	PFRUNID	VISITNUM	PFDTC
1 (cont)	29.675	%	29.675	29.675	%	Mutant	SOMATIC	Vendor C	BEAMing	797350234879	1	2005-05-05
2 (cont)	0.0072	%	0.0072	0.0072	%	Wildtype	SOMATIC	Vendor C	BEAMing	797350234879	1	2005-05-05

*Dressman D. et al (2003) *Proc Natl Acad Sci USA* 100(15);8817-22

From SDTMIG-PGx v1.0

Pharmacogenomics/Genetics Methods and Supporting Information

- PGx domain holds the methods, algorithms and parameters used in the analysis
- PGTEST/PGTESTCD include Exons sequenced, exons length etc.
- PGNSPCES and PGNSTRN –used to differentiate between subject and microbial records
- PGGENRI will be used to collect the gene of interest in the results record
- PGMETHOD contains the method of sequencing the genome



Example for SDTM PG

- Below are test parameters and details for a qRT-PCR run performed to determine gene expression.

Row	STUDYID	DOMAIN	USUBJID	PGSEQ	PGGRPID	PGREFID	PGTESTCD	PGTEST	PGGENRI	PGCAT
1	AAA	PG	AAA-XXX1	1	1	L1046058-3-S0072248	SRSEQ	Sub-Region Sequenced	MYC	GENE EXPRESSION
2	AAA	PG	AAA-XXX1	2	2	L1046058-3-S0072248	SRSEQ	Sub-Region Sequenced	MFNG	GENE EXPRESSION

PGORRES	PGSTRESC	PGSTRESN	PGSPEC	PGMETHOD	PGRUNID	VISITNUM	PGDTC
EXON Boundary 2-3	EXON Boundary 2-3		RNA	QUANTITATIVE REVERSE TRANSCRIPTASE POLYMERASE CHAIN REACTION	Hs00153408_m1	-1	2013-01-30T19:56:00
EXON Boundary 4-5	EXON Boundary 4-5		RNA	QUANTITATIVE REVERSE TRANSCRIPTASE POLYMERASE CHAIN REACTION	Hs00153408_m1	-1	2013-01-30T19:56:00

From SDTMIG-PGx v1.0

Pharmacogenomics/Genetics Biomarker

- Special purpose domain that serves as a reference to associate observed genetic mutations with medical conclusions
- This dataset employs a special structure and does not conform to any of the CDISC three observational classes of findings, events or interventions.
- PBSTMT -medical conclusion associated with genetic biological marker.
- PBMRKR – individual variant/marker

Row	STUDYID	DOMAIN	PBSEQ	PBMRKRID	PBMRKR	PBGENRI	PBGENTYP	PBNSPCES	PBSTRN	PBDRUG	PBDIAG	PBSTMT
1	STDY-505357	PB	1	G48V	p.G48V	Protease	SECTOR	HIV	1a	Saquinavir		RESISTANCE
2	STDY-505357	PB	2	L10I+K20R+M36I+A71V+V82T	p.L10I	Protease	SECTOR	HIV	1a	Indinavir		RESISTANCE
3	STDY-505357	PB	3	L10I +K20R+M36I+A71V+V82T	p.K20R	Protease	SECTOR	HIV	1a	Indinavir		RESISTANCE
4	STDY-505357	PB	4	L101+K20R+M36I+A71V+V82T	p.M36I	Protease	SECTOR	HIV	1a	Indinavir		RESISTANCE
5	STDY-505357	PB	5	L10I +K20R+M36I+A71V+V82T	p.A71V	Protease	SECTOR	HIV	1a	Indinavir		RESISTANCE
6	STDY-505357	PB	6	L10I +K20R+M36I+A71V+V82T	p.V82T	Protease	SECTOR	HIV	1a	Indinavir		RESISTANCE

From SDTMIG-PGx v1.0

Example of EGFR gene

pb.xpt

Row	STUDYID	DOMAIN	PBSEQ	PBMRKRID	PBMRKR	PBGENRI	PBGENTYP	PBDRUG	PBDIAG	PBSTMT
1	ABX-01256	PB	1	2073A>T	c.2073A>T	EGFR	GENE		Diffusely Infiltrating Astrocytoma	Decreased risk
2	ABX-01256	PB	2	G719A	c.2156G>C	EGFR	GENE	EGFR TKIs		Increased sensitivity
3	ABX-01256	PB	3	T790M	c.2369C>T	EGFR	GENE	EGFR TKIs		Decreased sensitivity

pf.xpt

Row	STUDYID	DOMAIN	USUBJID	PFSEQ	PFREFID	PFTESTCD	PFTEST	PFGENRI	PFGENTYP	PFREFSEQ	PFCAT	PFORRES	PFORREF
1	ABX-01256	PF	XX7-154	1	5493283	NUC	Nucleotide	EGFR	GENE	NM_005228.3	GENETIC VARIATION	C	G
2	ABX-01256	PF	XX7-212	1	8970343	NUC	Nucleotide	EGFR	GENE	NM_005228.3	GENETIC VARIATION	T	C
3	ABX-01256	PF	XX7-220	1	7629230	NUC	Nucleotide	EGFR	GENE	NM_005228.3	GENETIC VARIATION	T	A
4	ABX-01256	PF	XX7-237	1	9092742	NUC	Nucleotide	EGFR	GENE	NM_005228.3	GENETIC VARIATION	C	G

Row	PFGENLOC	PFGENSR	PFSTRESC	PFXFN	PFNAM	PFMETHOD	PFDRVFL	VISITNUM	PFDTC	PFRUNID
1 (cont)	2156	Exon 18	c.2156G>C	5.23.445.1.4.165008.1.8:86175	Biotech ABC	Massively Parallel Sequencing		1	2012-10-23T10:06	8970723
2 (cont)	2369	Exon 20	c.2369C>T	5.23.445.1.4.165008.1.8:87952	Biotech ABC	Massively Parallel Sequencing		1	2012-10-23T12:50	8925000
3 (cont)	2073	Exon 16	c.2073A>T	5.23.445.1.4.165008.1.8:87970	Biotech ABC	Massively Parallel Sequencing		1	2012-10-23T13:03	8925018
4 (cont)	2293	Exon 20	c.2293G>C	5.23.445.1.4.165008.1.8:88100	Biotech ABC	Massively Parallel Sequencing		1	2012-10-23T13:55	8928003

pg.xpt

Row	STUDYID	DOMAIN	USUBJID	PGSEQ	PGTESTCD	PGTEST	PGGENRI	PGGENTYP	PGCAT	PGORRES	PGSTRESC
1	ABX-01256	PG	XX7-154	1	EXON	Exons Sequenced	EGFR	GENE	GENETIC VARIATION	13-21	13-21
2	ABX-01256	PG	XX7-154	2	SEQSTART	Sequence Start	EGFR	GENE	GENETIC VARIATION	1499	1499
3	ABX-01256	PG	XX7-154	3	SEQLONG	Sequence Length	EGFR	GENE	GENETIC VARIATION	1127	1127
4	ABX-01256	PG	XX7-212	1	EXON	Exons Sequenced	EGFR	GENE	GENETIC VARIATION	13-21	13-21
5	ABX-01256	PG	XX7-212	2	SEQSTART	Sequence Start	EGFR	GENE	GENETIC VARIATION	1499	1499
6	ABX-01256	PG	XX7-212	3	SEQLONG	Sequence Length	EGFR	GENE	GENETIC VARIATION	1127	1127
7	ABX-01256	PG	XX7-220	1	EXON	Exons Sequenced	EGFR	GENE	GENETIC VARIATION	13-21	13-21
8	ABX-01256	PG	XX7-220	2	SEQSTART	Sequence Start	EGFR	GENE	GENETIC VARIATION	1499	1499
9	ABX-01256	PG	XX7-220	3	SEQLONG	Sequence Length	EGFR	GENE	GENETIC VARIATION	1127	1127
10	ABX-01256	PG	XX7-237	1	EXON	Exons Sequenced	EGFR	GENE	GENETIC VARIATION	13-21	13-21
11	ABX-01256	PG	XX7-237	2	SEQSTART	Sequence Start	EGFR	GENE	GENETIC VARIATION	1499	1499
12	ABX-01256	PG	XX7-237	3	SEQLONG	Sequence Length	EGFR	GENE	GENETIC VARIATION	1127	1127

From SDTMIG-PGx v1.0

Challenges for SAS Programmer

- The PGx Findings domain stores key results such as intensity values (both raw and normalized), P-Value, fold-change, ratio, genetic change, amino-acid change, etc. Such data are significantly different from the currently available findings module (Lab data, Vital Signs, Pharmacokinetic data, etc.).
- CDISC plans to create vocabulary for CDISC TESTCD and TEST which will reside in the NCI EVS (National Cancer Institute's Enterprise Vocabulary Services) and be a counterpart to the LOINC (Logical Observation Identifiers Names and Codes) codes.
- A genetic variation data could be anything from complex arrangements of strings that looks like random character strings (generally A T G (U) C in case of a DNA/RNA sequence information like a gene substring) to strings or number arranged in an ambiguous array which actually holds some hidden meaning that need to be decoded again by some complex algorithm which a programmer has to implement.

Challenges for SAS Programmer

- Mapping codes likely will involve databases like NCBI and GenBank along with medical dictionaries like MeddRA and WHODRUG.
- Unlike the other finding variables, which contains derived or assigned- text, numeric values and discrete values (Yes/No, 0/1, Male/Female), the variables in the PGx domain may not be necessarily simple and a programmer may need to think of rules and very specific programming to derive certain variables from the captured raw data in order to make them CDISC data structure complaint.
- Programming may involve developing macros to automate standardized algorithms across the domains and major operations effective to handle complex strings such as using regular expressions and string functions.
- All these new kind of data collected in PGx domains brings new opportunities for more flexible SDTM and ADaM programming for CDISC

What are we waiting for?

- SDTMIG-PGx includes variables of SDTM v1.5 which has not yet officially published.

- FDA still has not officially announced that sponsor should submit genomic data with SDTMIG-PGx

Summary

- Since the completion of the Human Genome Project in 2003, there has been a significant rise in the use of Pharmacogenomic data for both research and clinical care.
- The use of this new type of data can explain the relationship between genetic variations and drug response in patients, thereby adjusting therapies on a patient-by-patient basis.
- Understanding SDTM PGx domains will help us map the crucial genomic data adequately and efficiently thereby standardizing the data.
- Precision medicine in oncology will play a major role in reducing the adverse events experienced by the patients.
- Precision medicine is not the future but the present of Oncology clinical trials.

References

- Study Data Tabulation Model Implementation Guide: Pharmacogenomics/Genetics Version 1.0 (Provisional)
- http://www.cdisc.org/pgx-review-article
- Transforming Biomarker Data into an SDTM based Dataset - Kiran Cherukuri
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4846136/
- Canadian Cancer Society
- www.cancer.net

Shortening the Distance From Lab to Life™

Syneos™ Health