

Don't Be Afraid to Touch the PGx Standards

Ipppei Akiya (ipppei.aki@datadriven.jp)

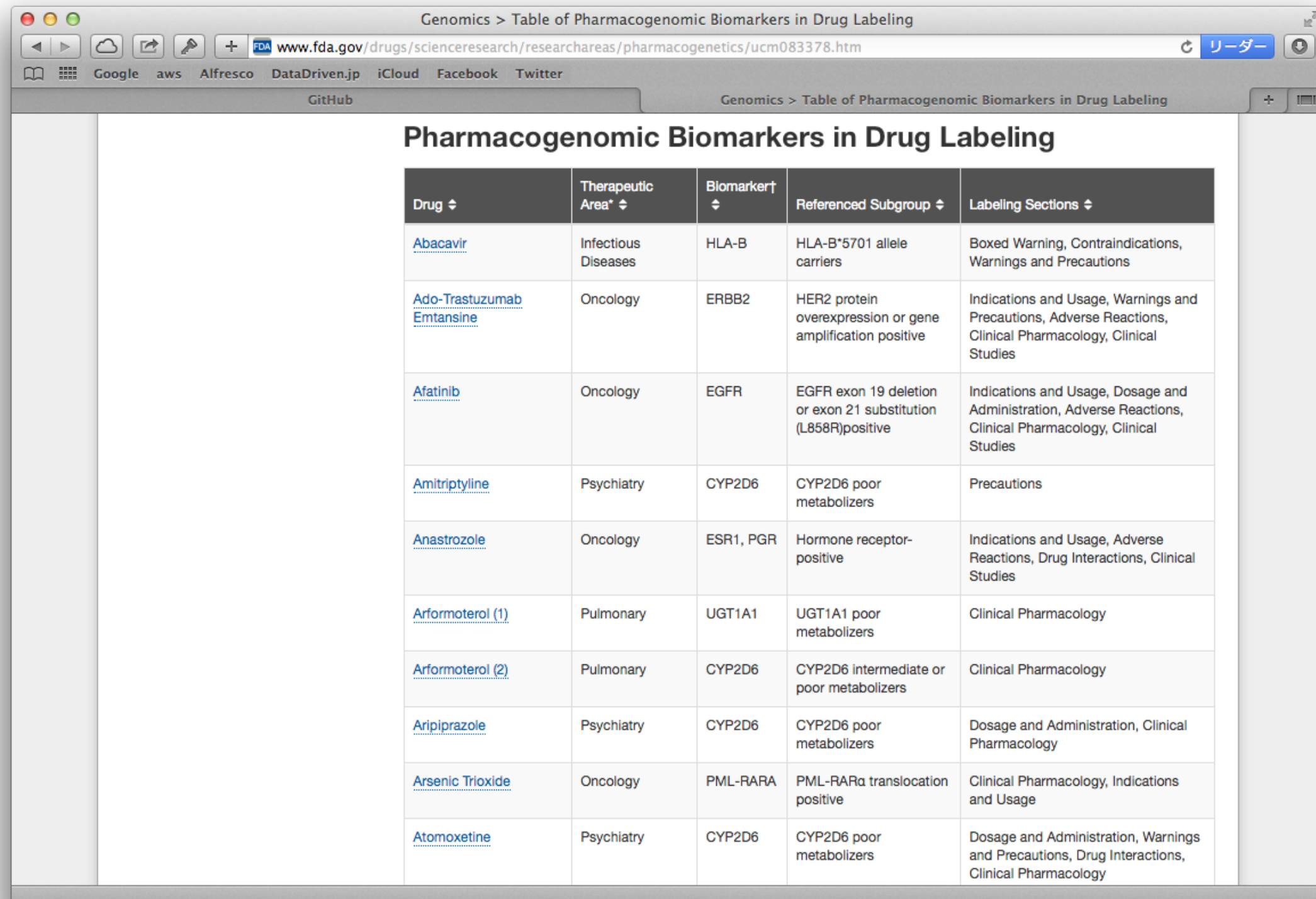
DataDriven inc.

Background

It is recognized that human genome is one of the most important factors to develop the specific diseases, and specific genome variation induces a perturbation of treatment effect.

ex) CYP2D6, EGFR, HLA-B

Background



The screenshot shows a web browser window with the address bar displaying the URL: www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm. The browser's address bar also shows the text 'Genomics > Table of Pharmacogenomic Biomarkers in Drug Labeling'. The browser's search bar contains the text 'リーダー'. The browser's tabs show 'GitHub' and 'Genomics > Table of Pharmacogenomic Biomarkers in Drug Labeling'. The main content area displays the title 'Pharmacogenomic Biomarkers in Drug Labeling' and a table with the following data:

Drug	Therapeutic Area*	Biomarker†	Referenced Subgroup	Labeling Sections
Abacavir	Infectious Diseases	HLA-B	HLA-B*5701 allele carriers	Boxed Warning, Contraindications, Warnings and Precautions
Ado-Trastuzumab Emtansine	Oncology	ERBB2	HER2 protein overexpression or gene amplification positive	Indications and Usage, Warnings and Precautions, Adverse Reactions, Clinical Pharmacology, Clinical Studies
Afatinib	Oncology	EGFR	EGFR exon 19 deletion or exon 21 substitution (L858R)positive	Indications and Usage, Dosage and Administration, Adverse Reactions, Clinical Pharmacology, Clinical Studies
Amitriptyline	Psychiatry	CYP2D6	CYP2D6 poor metabolizers	Precautions
Anastrozole	Oncology	ESR1, PGR	Hormone receptor-positive	Indications and Usage, Adverse Reactions, Drug Interactions, Clinical Studies
Arformoterol (1)	Pulmonary	UGT1A1	UGT1A1 poor metabolizers	Clinical Pharmacology
Arformoterol (2)	Pulmonary	CYP2D6	CYP2D6 intermediate or poor metabolizers	Clinical Pharmacology
Aripiprazole	Psychiatry	CYP2D6	CYP2D6 poor metabolizers	Dosage and Administration, Clinical Pharmacology
Arsenic Trioxide	Oncology	PML-RARA	PML-RARa translocation positive	Clinical Pharmacology, Indications and Usage
Atomoxetine	Psychiatry	CYP2D6	CYP2D6 poor metabolizers	Dosage and Administration, Warnings and Precautions, Drug Interactions, Clinical Pharmacology

<http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm>

Background

Genomic data is helpful to predict a treatment effect and to prevent a serious adverse reaction.

All stakeholders in pharmaceutical and healthcare industry expected to develop new drugs based on genomic and related information.

However, clinical data standard for genomic data had not published.



Study Data Tabulation Model Implementation Guide: Pharmacogenomics/Genetics Version 1.0 (Provisional)



Study Data Tabulation Model Implementation Guide: Pharmacogenomics/Genetics Version 1.0 (Provisional)

Prepared by the
CDISC PGx Team

Notes to Readers

- This is the provisional version 1.0 of the implementation guide for Pharmacogenomics and Pharmacogenetics. It is intended to correspond to version 1.5 of the CDISC Study Data Tabulation Model.
- Because SDTM v1.5 has not yet been published, this document remains a provisional release only.

Revision History

Date	Version	Summary of Changes
2015-05-26	1.0	Provisional release. See Section 1.7 for major revisions based on comments from public review.
2014-10-10	1.0 Draft	Draft for Public Review

See [Appendix E](#) for Representations and Warranties, Limitations of Liability, and Disclaimers.

July 2015, CDISC officially released SDTM Implementation Guide related to genomics data.

We are ready to start to the genomics data implementation with SDTM.

Really?

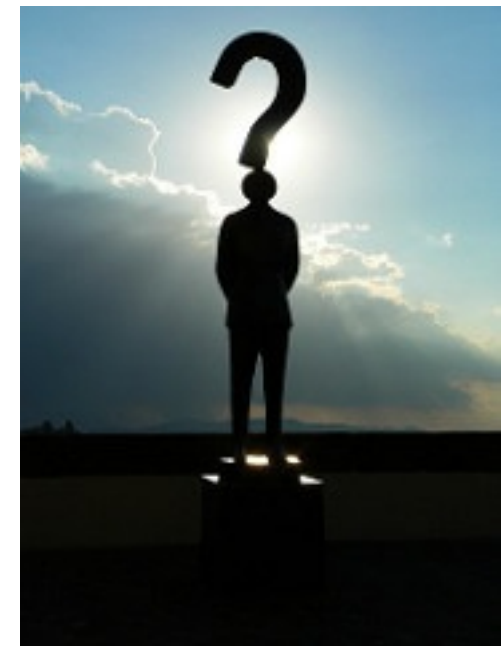


Photo by Marco Bellucci

Notice of SDTMIG-PGx

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.

Scope of SDTMIG-PGx

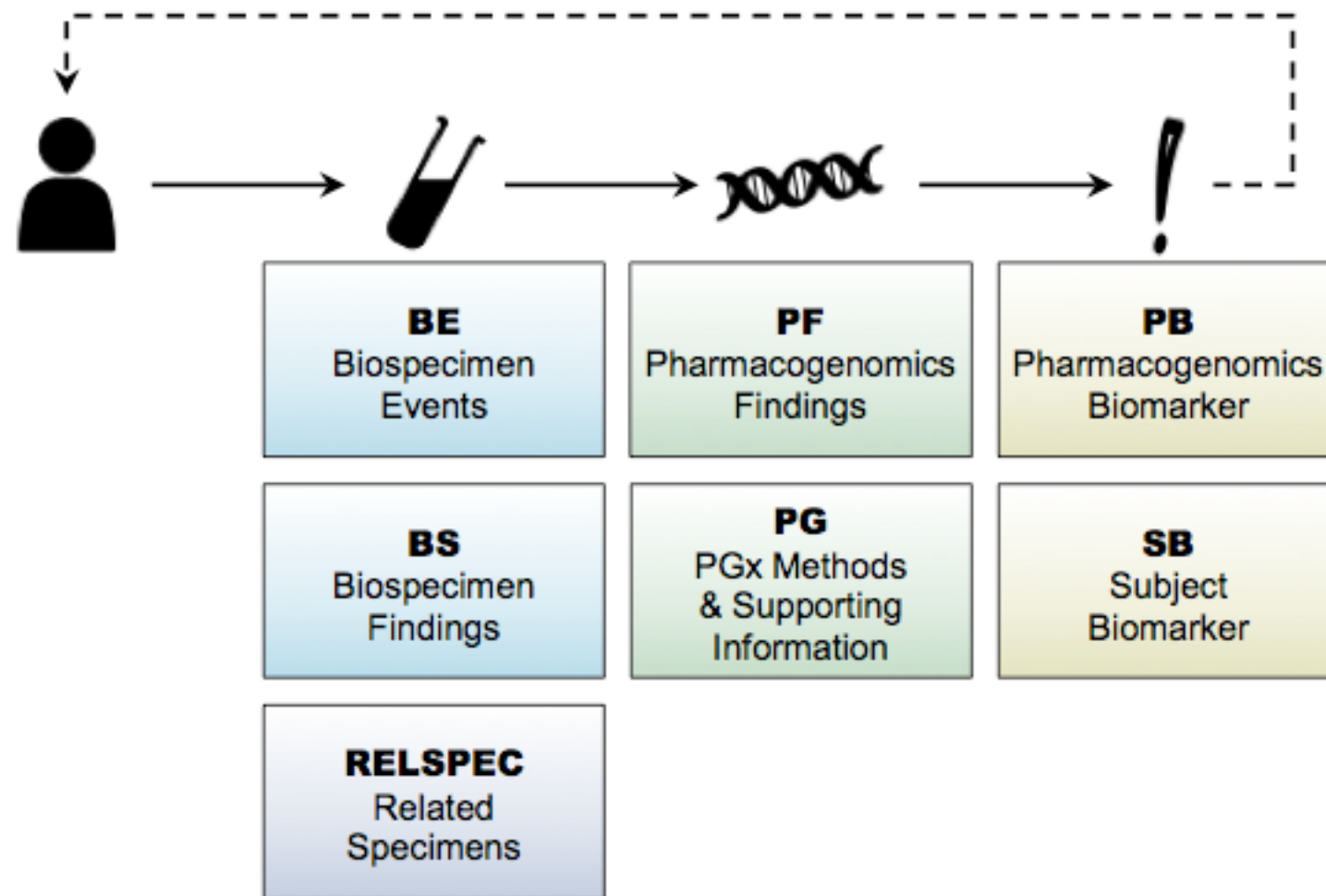
- Biospecimen Collection and Handling
- Quality Data
- Genetic Mutation
- Genotyping
- Gene Expression
- Cytogenetics
- Viral Genetics
- Proteomics

Dataset Level Metadata of SDTMIG-PGx

Dataset	Description	Class	Structure	Purpose	Keys*	Location
PB	Pharmacogenomics/ Genetics Biomarker	Special Purpose Domains	One record per biomarker used in the study	Tabulation	STUDYID, PBMKRID, PBMKR	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 Study Data Tabulation Model Implementation
Guide: Pharmacogenomics/Genetics Version 1.0

Classification of SDTMIG-PGx Domains



From Study Data Tabulation Model Implementation Guide: Pharmacogenomics/Genetics Version 1.0

BE, BS, and RELSPEC

Biospecimen Events(BE)

The Biospecimen Events domain is used to capture information about actions taken that affect a specimen or alter its status.

Biospecimen Findings(BS)

The Biospecimen Findings domain contains the details regarding the characteristics of biospecimens and extracted samples (e.g., RNA, DNA) such as specimen volume, quantity of extracted sample, specimen condition and the integrity of the DNA or RNA samples.

RELSPEC Dataset(RELSPEC)

The Related Specimens dataset holds the hierarchy of specimen relationships.

Biospecimen Collection and Handling

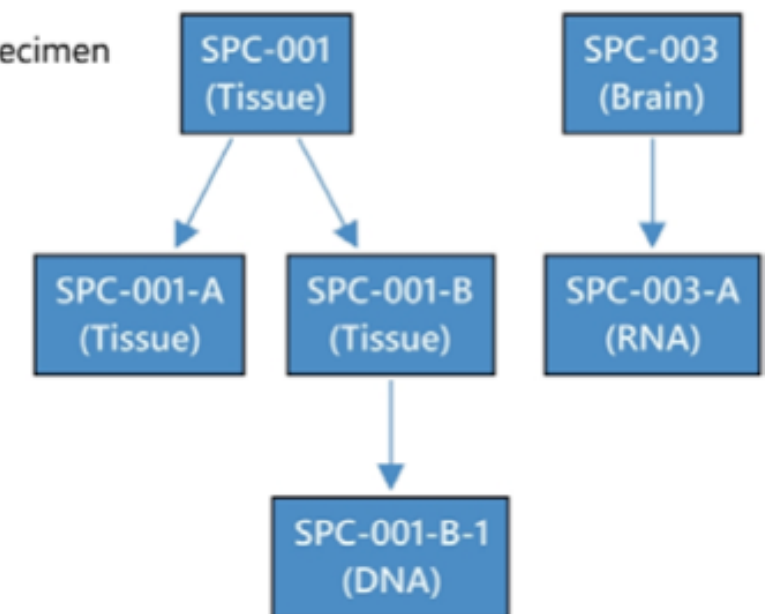
Row	STUDYID	USUBJID	REFID	SPEC	PARENT	LEVEL
1	ABC-123	001-01	SPC-001	TISSUE		1
2	ABC-123	001-01	SPC-001-A	TISSUE	SPC-001	2
3	ABC-123	001-01	SPC-001-B	TISSUE	SPC-001	2
4	ABC-123	001-01	SPC-001-B-1	DNA	SPC-001-B	3
5	ABC-123	001-01	SPC-003	TISSUE		1
6	ABC-123	001-01	SPC-003-A	RNA	SPC-003	2



Originally Collected Specimen
Level 1

Child Specimen
Level 2

Child Specimen
Level 3



From Study Data Tabulation Model Implementation Guide: Pharmacogenomics/Genetics Version 1.0

Example of BE datasets

be.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	BESEQ	BEREFID	BETERM	BEDECOD	BEPARTY	BEPRTYID	BECAT	VISITNUM
1	3441271	BE	MU-298		1	298B1	Collected	COLLECTED	SITE	05	COLLECTION	2
2	3441271	BE	MU-298	293USHE8	2	298B1-1	Extracted	EXTRACTED	SITE	05	EXTRACTION	2
3	3441271	BE	MU-298		3	298B1-1	Shipped	SHIPPED	ABC LAB		TRANSPORT	2
4	3441271	BE	MU-298	PURKIT	4	298R1-1R0	Extracted	EXTRACTED	ABC LAB		EXTRACTION	2
5	3441271	BE	MU-298		5	298R1-1R1	Aliquoted	ALIQUOTED	ABC LAB		PREPARATION	2
6	3441271	BE	MU-298		6	298R1-1R2	Aliquoted	ALIQUOTED	ABC LAB		PREPARATION	2
7	3441271	BE	MU-298		7	298R1-1R3	Aliquoted	ALIQUOTED	ABC LAB		PREPARATION	2
8	3441271	BE	MU-298		8	298R1-1R1	Sequenced	SEQUENCED	ABC LAB		PREPARATION	2
9	3441271	BE	MU-298		9	298R1-1R2	Frozen	FROZEN	ABC LAB		PREPARATION	2
10	3441271	BE	MU-298		10	298R1-1R3	Frozen	FROZEN	ABC LAB		PREPARATION	2

Row	BEDTC	BESTDTC	BEENDTC
1 (cont)	2010-04-01T11:50	2010-04-01T11:50	
2 (cont)	2010-04-01T11:50	2010-04-01T12:10	
3 (cont)	2010-04-01T11:50	2010-04-01T15:00	2010-04-02T8:00
4 (cont)	2010-04-01T11:50	2010-04-02T9:00	2010-04-05T13:50
5 (cont)	2010-04-01T11:50	2010-04-05T13:50	
6 (cont)	2010-04-01T11:50	2010-04-05T13:50	
7 (cont)	2010-04-01T11:50	2010-04-05T13:50	
8 (cont)	2010-04-01T11:50	2010-04-05T13:50	2010-04-06T10:30
9 (cont)	2010-04-01T11:50	2010-04-05T13:50	
10 (cont)	2010-04-01T11:50	2010-04-05T13:50	

suppbe.xpt

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL	QORIG	QEVAL
1	3441271	BE	MU-298	BEREFID	298B1	BESPEC	Specimen Type	BLOOD	CRF	
2	3441271	BE	MU-298	BEREFID	298B1-1	BESPEC	Specimen Type	PLASMA	CRF	
3	3441271	BE	MU-298	BEREFID	298R1-1R0	BESPEC	Specimen Type	RNA	eDT	
4	3441271	BE	MU-298	BEREFID	298R1-1R1	BESPEC	Specimen Type	RNA	eDT	
5	3441271	BE	MU-298	BEREFID	298R1-1R2	BESPEC	Specimen Type	RNA	eDT	
6	3441271	BE	MU-298	BEREFID	298R1-1R3	BESPEC	Specimen Type	RNA	eDT	

From Study Data Tabulation Model Implementation
Guide: Pharmacogenomics/Genetics Version 1.0

Example of BS and RELSPEC dataset

bs.xpt

Row	STUDYID	DOMAIN	USUBJID	BSSEQ	BSREFID	BSTESTCD	BSTEST	BSCAT	BSORRES	BSORRESU	BSSTRESC
1	3441271	BS	MU-298	1	298B1	VOLUME	Volume	SPECIMEN MEASUREMENT	12	mL	6
2	3441271	BS	MU-298	2	298B1-1	VOLUME	Volume	SPECIMEN MEASUREMENT	7	mL	6
3	3441271	BS	MU-298	3	298B1-1R0	VOLUME	Volume	SPECIMEN MEASUREMENT	6	mL	6
4	3441271	BS	MU-298	4	298B1-1R0	RIN	RNA Integrity Number	QUALITY CONTROL	9.3		9.3
5	3441271	BS	MU-298	5	298R1-1R1	VOLUME	Volume	SPECIMEN MEASUREMENT	2	mL	2
6	3441271	BS	MU-298	6	298R1-1R2	VOLUME	Volume	SPECIMEN MEASUREMENT	2	mL	2
7	3441271	BS	MU-298	7	298R1-1R3	VOLUME	Volume	SPECIMEN MEASUREMENT	2	mL	2

Row	BSSTRESN	BSSTRESU	BSSPEC	VISITNUM	BSDTC
1 (cont)	6	mL	BLOOD	2	2010-04-01T11:50
2 (cont)	6	mL	PLASMA	2	2010-04-01T11:50
3 (cont)	6	mL	RNA	2	2010-04-01T11:50
4 (cont)	9.3		RNA	2	2010-04-01T11:50
5 (cont)	2	mL	RNA	2	2010-04-01T11:50
6 (cont)	2	mL	RNA	2	2010-04-01T11:50
7 (cont)	2	mL	RNA	2	2010-04-01T11:50

relspec.xpt

Row	STUDYID	USUBJID	REFID	SPEC	PARENT	LEVEL
1	3441271	MU-298	298B1	BLOOD		1
2	3441271	MU-298	298B1-1	PLASMA	298B1	2
3	3441271	MU-298	298R1-1R0	RNA	298B1-1	3
4	3441271	MU-298	298R1-1R1	RNA	298R1-1R0	4
5	3441271	MU-298	298R1-1R2	RNA	298R1-1R0	4
6	3441271	MU-298	298R1-1R3	RNA	298R1-1R0	4

From Study Data Tabulation Model Implementation
Guide: Pharmacogenomics/Genetics Version 1.0

PF and PG

Pharmacogenomics/Genetics Findings(PF)

The Pharmacogenomics/Genetics Findings domain contains the results of genetic variation and gene expression tests. For genetic variation tests, the results may include portions of the genetic sequence and comparisons with reference gene sequences.

Pharmacogenomics/Genetics Methods and Supporting Information(PG)

The Pharmacogenomics/Genetics Methods and Supporting Information domain contains additional data that may affect interpretation of PF data, such as setup processes, test or sequencing parameters, and QC observations.

Examples of PF dataset

Row	STUDYID	DOMAIN	USUBJID	PFSEQ	PFREFID	PFTESTCD	PFTEST	PFGENRI	PFREFSEQ	PFCAT	PFSCAT	PFORRES	PFORREF
1	PAAAAA	PF	AA66-666-6666	1	12-708-1501C	NUC	Nucleotide	NOS3	NM_000603.4	GENETIC VARIATION	GENOTYPE		
2	PAAAAA	PF	AA66-666-6666	2	12-708-0579A	NUC	Nucleotide	VEGFA	NM_001025366.2	GENETIC VARIATION	GENOTYPE	G/A	A
3	PAAAAA	PF	AA66-666-6666	3	12-708-0579A	NUC	Nucleotide	VEGFA	NM_001025366.2	GENETIC VARIATION	GENOTYPE	A/A	A
4	PAAAAA	PF	AA66-666-6666	4	12-708-0579A	NUC	Nucleotide	VEGFA	NM_001025366.2	GENETIC VARIATION	GENOTYPE	G/G	C
5	PAAAAA	PF	AA66-666-6666	5	12-708-0579A	NUC	Nucleotide	FCGR3a	NM_001127596.1	GENETIC VARIATION	GENOTYPE	T/T	T

Row	PFGENLOC	PFSTRESC	PFRSNUM	PFSTAT	PFREASND	PFXFN	PFNAM	PFSPEC	PFMUTYP	PFMETHOD	PFRUNID	VISITNUM	PFDTC
1 (cont)			rs2070744	NOT DONE	NO AMP	snp_database.csv	ABC-LAB	DNA	GERMLINE	qPCR	UC-1202	-1	2005-02-20
2 (cont)	-614	c.[-614A>G];[=]	rs1570360			snp_database.csv	ABC-LAB	DNA	GERMLINE	qPCR	UC-1202	-1	2005-02-20
3 (cont)	-2055	c.[=];[=]	rs699947			snp_database.csv	ABC-LAB	DNA	GERMLINE	qPCR	UC-1202	-1	2005-02-20
4 (cont)	-94	c.[-94C>G];[-94C>G]	rs2010963			snp_database.csv	ABC-LAB	DNA	GERMLINE	qPCR	UC-1202	-1	2005-02-20
5 (cont)	526	c.[=];[=]	rs396991			snp_database.csv	ABC-LAB	DNA	GERMLINE	qPCR	UC-1202	-1	2005-02-20

From Study Data Tabulation Model Implementation
Guide: Pharmacogenomics/Genetics Version 1.0

c.[-614A>G];[=]

Position

- c. coding DNA reference sequence
- g. genomic reference sequence
- p. protein reference sequence

Examples of PF dataset

Gene Expression

pf.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	PFSEQ	PFGRPID	PFREFID	PFTESTCD	PFTEST	PFGENRI	PFGENTYP	PFCAT	PFSCAT
1	A12345	PF	43871	AGS-G4900DA	1	1	2287.09443	NINT1VAL	Normalized Intensity 1 Value	EGFR	GENE	GENE EXPRESSION	Pre-Analytic
2	A12345	PF	43871	AGS-G4900DA	2	1	2287.09443	NINT2VAL	Normalized Intensity 2 Value	EGFR	GENE	GENE EXPRESSION	Pre-Analytic
3	A12345	PF	43871	MANAN03	3	1	2287.09443	PVAL	P Value	EGFR	GENE	GENE EXPRESSION	Analytic
4	A12345	PF	43871	MANAN03	4	1	2287.09443	FOLDCHG	Fold Change	EGFR	GENE	GENE EXPRESSION	Analytic
5	A12345	PF	43871		5	1	2287.09443	NTASMN	Net Assessment	EGFR	GENE	GENE EXPRESSION	Interpretation

Row	PFORRES	PFSTRESC	PFSTRESN	PFXFN	PFNAM	PFSPEC	PFMETHOD	PFRUNID	PFANMETH	VISITNUM	PFDTC
1 (cont)	1.16279	1.16279	1.16279	2.16.090.1.135764.3.4:7280912	Deluxe Central Labs	RNA	Microarray	1000450001	LOWESS	2	2005-03-21T11:28:17
2 (cont)	0.96469	0.96469	0.96469	2.16.090.1.135764.3.4:7280912	Deluxe Central Labs	RNA	Microarray	1000450001	LOWESS	2	2005-03-21T11:28:17
3 (cont)	0.05391	0.05391	0.05391	2.16.090.1.135764.3.4:7280912	Deluxe Central Labs	RNA	Microarray	1000450001		2	2005-03-21T11:28:17
4 (cont)	1.8	1.8	1.8	2.16.090.1.135764.3.4:7280912	Deluxe Central Labs	RNA	Microarray	1000450001		2	2005-03-21T11:28:17
5 (cont)	Over-expressed	Over-expressed		2.16.090.1.135764.3.4:7280912	Deluxe Central Labs	RNA	Microarray	1000450001		2	2005-03-21T11:28:17

From Study Data Tabulation Model Implementation
Guide: Pharmacogenomics/Genetics Version 1.0

Examples of PG dataset

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

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

From Study Data Tabulation Model Implementation Guide: Pharmacogenomics/Genetics Version 1.0

PB and SB

Pharmacogenomics/Genetics Biomarker (PB)

The Pharmacogenomics/Genetics Biomarker domain documents known associations between observed variations and medical conclusions (e.g., disease diagnosis, resistance of a virus to a particular drug).

Subject Biomarker (SB)

The Subject Biomarker domain holds data about any genetic biomarkers (as defined in PB) that a subject may have.

Examples of PB and SB datasets

pb.xpt

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	L10I+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

sb.xpt

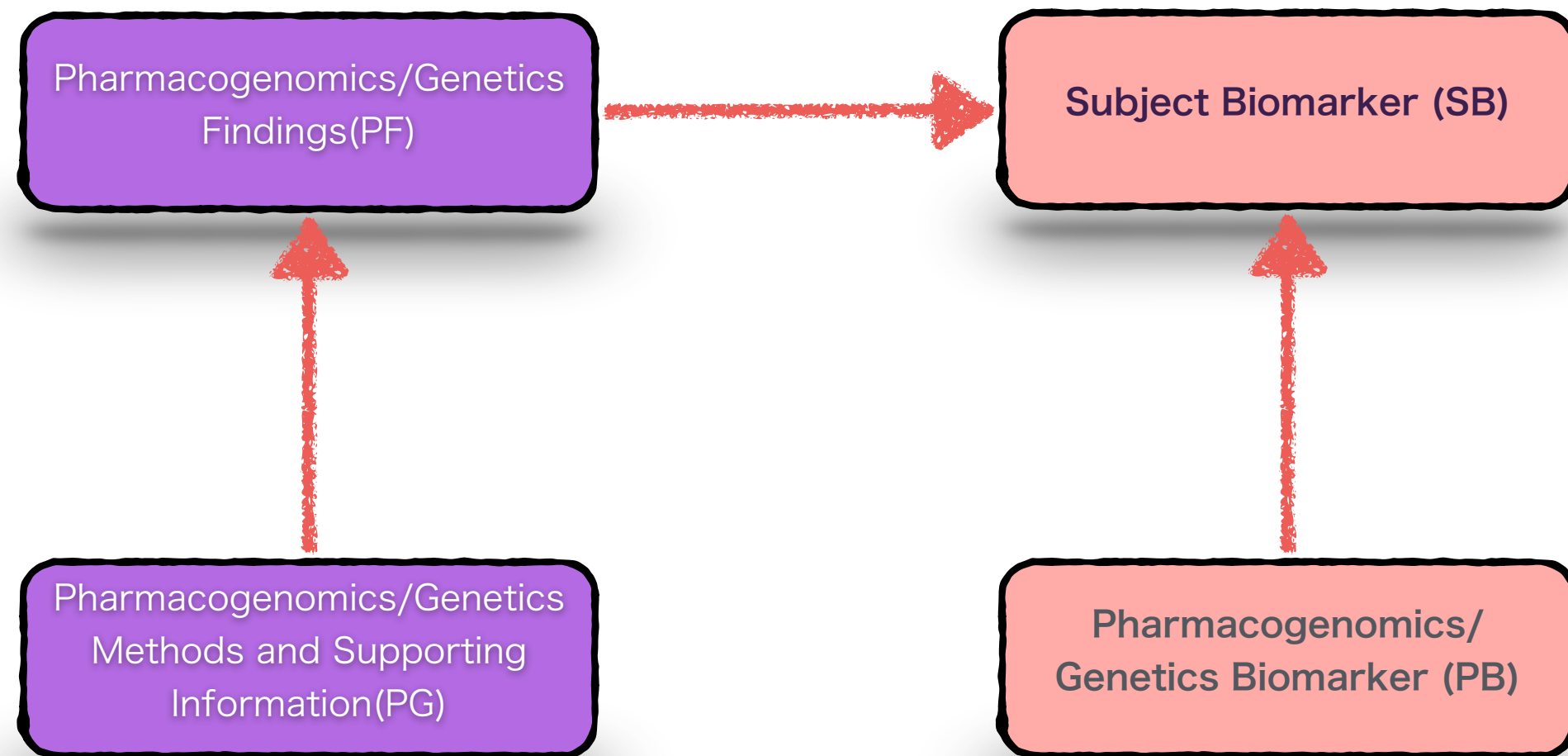
Row	STUDYID	DOMAIN	USUBJID	SBSEQ	SBGRPID	SBREFID	SBMRKRID	SBGENRI	SBGENTYP	SPNSPCES	SBNSTRN	SBNAM
1	STDY-505357	SB	521298	1	RESIST	D391395	G48V	Protease	Sector	HIV	1a	BIOTECHA
2	STDY-505357	SB	521298	2	RESIST	D391395	L10I+K20R+M36I+A71V+V82T	Protease	Sector	HIV	1a	BIOTECHA

Row	VISITNUM	VISIT	VISITDY	SBDTC
1 (cont)	1	BASELINE	-1	2010-03-01
2 (cont)	1	BASELINE	-1	2010-03-01

relrec.xpt

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
1	STDY-505357	PB		PBMRKRID		MANY	1
2	STDY-505357	SB		SBMRKRID		MANY	1

From Study Data Tabulation Model Implementation Guide: Pharmacogenomics/Genetics Version 1.0



Collaboration for a PGx Study

PGx Domains

Genomic Analysis Lab

Wet Lab Researcher



Photo by U.S. Army RDECOM

Dry Lab Researcher



Photo by Héctor García

Other Domains

Data Science Center

Clinical Programmer



Christopher Porter

Thank you!