WITH STANDARDS – UNLOCK THE POWER OF DATA

cdisc

2022
US
INTERCHANGE
26-27 OCTOBER | AUSTIN

SDTM Standards for Genomics: Genomics Findings (GF) and Future Directions

Presented by Christine Connolly, Head of Standards Projects, CDISC
Dr. Erin Muhlbradt, Clinical/Biomedical Information Specialist, CDISC Terminology Lead,
US NCI-EVS [c], MSC, a Guidehouse company

Meet the Speakers

Christine Connolly

Title: Head, Standards Projects

Organization: CDISC

Christine is an advocate for data standardization given its potential to expedite development of preventive approaches, harm reduction strategies, and quality therapies to improve health outcomes. She has over twenty years of experience working in clinical trials in academic and pharmaceutical settings. She co-leads the Genomics Subteam and other CDISC initiatives.

Dr. Erin Muhlbradt, PhD

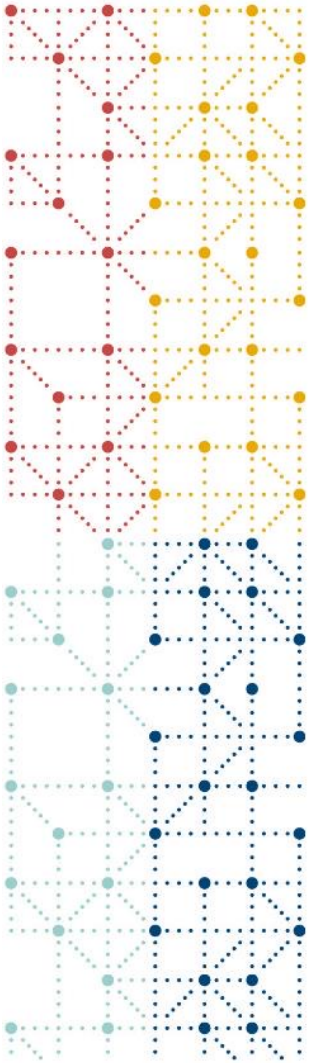
Title: Clinical/Biomedical Information Specialist; CDISC Terminology Lead

Organization: US NCI-EVS [c], MSC, a Guidehouse company

Erin is the CDISC and EVS lead for the CDISC terminology development program with nearly 15 years of experience as a clinical/biomedical terminology specialist. She co-leads the Genomics Subteam and other CDISC initiatives.

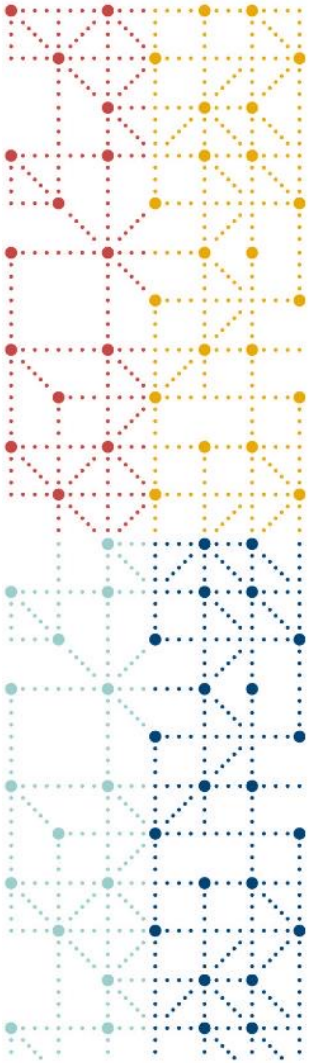
Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*



Agenda

1. Standards for Genomics
2. Genomics Findings (GF)
3. Future Directions
4. How you can be involved!



Standards for Genomics

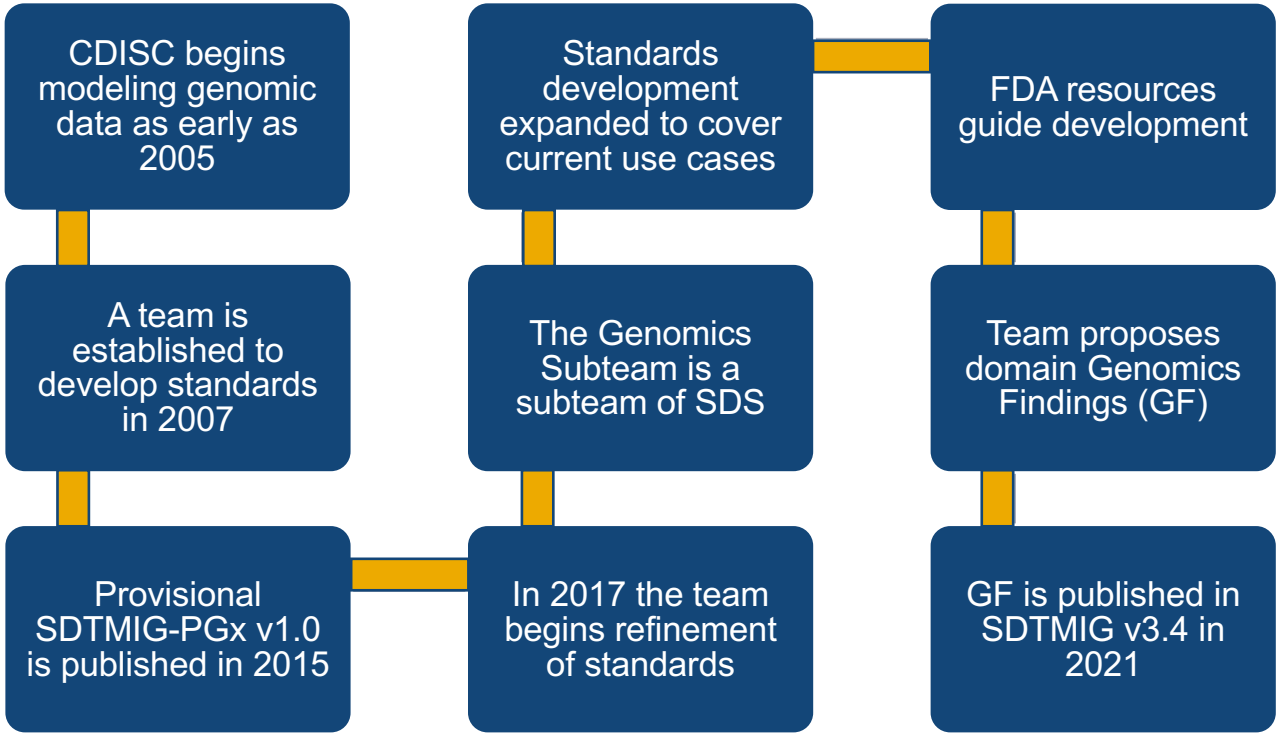
Genomics

Genomics refers to the structure, function, evolution, mapping, and editing of an organism's genome.

Genomic data collected as part of clinical research supports both development of quality patient care and improvements in patient outcomes.

Genomic analysis of subject samples continues to become a standard practice and the methodology for generating these data continues to evolve.

Development History



Refined Standards for Genomic Data

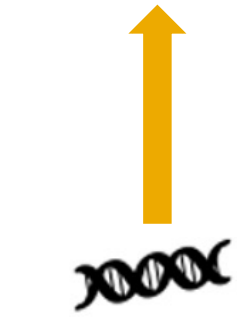
A single domain, Genomics Findings (GF), published in the SDTMIG v3.4 in 2021

Deprecation of SDTMIG-PGx v1.0 with:

- Provisional PF domain deprecated and superseded by the GF domain
- Biospecimens domains BE, BS, and RELSPEC published in the SDTMIG v3.4 as is and pending updates in future versions
- Provisional PG, PB, and SB domains deprecated with re-instantiation considered if valid use cases are found

Refined Standards for Genomic Data

A *single domain*, Genomics Findings (GF), published in the SDTMIG v3.4



Findings related to the structure, function, evolution, mapping, and editing of subject and non-host organism genomic material of interest

PF
Pharmacogenomics Findings

GF is continuous improvement standards

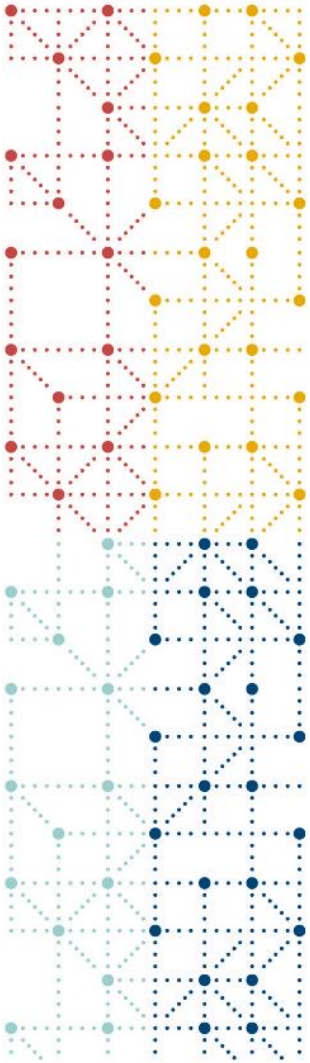
- Renamed Genomics Findings (GF) with domain clarifications
- New use cases modeled for GF
- Eighteen variable concepts and definitions clarified
- Two new concepts added
- Five established SDTM variables added
- Two outdated concepts retired

Renaming PF to Genomics Findings (GF)

- *Pharmacogenomics/Genetics Findings (PF)* name and definition do not accurately describe data represented in the domain

- Pharmacogenomics and pharmacogenetics are use cases for genomic data
  - **Pharmacogenomics** - Science that examines inherited variations in genes that dictate ***drug response*** and explores the ways such variations can be used to predict **whether a person will respond** favorably, adversely, or not at all **to an investigational product**.
  - **Pharmacogenetics** - Study of **the way drugs interact with genetic makeup** or the study of **genetic response to a drug**.

The terms above describe use cases and do not describe genomic data
Additionally, genomic data have many use cases beyond drug response



Genomics Findings (GF)

- Domain Walkthrough
- Terminology Considerations

Domain Walkthrough

Let's walkthrough Genomics Findings (GF) with SDTMIG v3.4 Example 2

GF Ex 2 - Single Nucleotide Variation

Created by Dana Booth, last modified on Oct 19, 2021

This example shows findings from an assessment of a known single nucleotide variant in gene ABCG2 using wet laboratory methodology real-time polymerase chain reaction. Findings from this assessment show the genotypes from DNA extracted from the blood of 3 individuals, each with a different genotype at the genetic locus of interest. Because the DNA specimen was extracted from normal blood, the inheritability of the variation is considered to be in the germline.

gf.xpt

Row 1: Shows a subject genotype which is homozygous for the variant nucleotide in the reference sequence.
Row 2: Shows a subject genotype which is heterozygous for the nucleotide in the reference sequence.
Row 3: Shows a subject genotype which is homozygous for the nucleotide in the reference sequence.

gf.xpt

Row	STUDYID	DOMAIN	USUBJID	GFSEQ	GFREFID	GFTESTCD	GFTEST	GFTSTDTL	GFORRES	GFORREF	GFSTRESC	GFSTREFC	GFINHERT	GFGENREF	GFCHROM	GFSYM
1	C12345	GF	C12345-001	1	NA18537	SNV	Single Nucleotide Variation	GENOTYPE	T/T	G/G	T/T	G/G	GERMLINE VARIATION	GRCh38.p13	4	ABCG2

GF Identifier Variables

GF Ex 2 - Single Nucleotide Variation

STUDYID	DOMAIN	USUBJID	GFSEQ	GFREFID	GFTESTCD	GFTEST	GFTSTDTL	GFORRES
C12345	GF	C12345-001	1	NA18537	SNV	Single Nucleotide Variation	GENOTYPE	T/T

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format[1]	Role	CDISC Notes	Core
STUDYID ★	Study Identifier	Char		Identifier	Unique identifier for a study.	Req
DOMAIN ★	Domain Abbreviation	Char	GF	Identifier	Two-character abbreviation for the domain.	Req
USUBJID ★	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.	Req
SPDEVID	Sponsor Device Identifier	Char		Identifier	Sponsor-defined identifier for a device.	Perm
NHOID	Non-Host Organism Identifier	Char		Identifier	Sponsor-defined identifier for a non-host organism which should only be used when the organism is the subject of the TEST. This variable should be populated with an intuitive name based on the identity of the non-host organism as reported by a lab (e.g., "A/California/7/2009 (H1N1)"). It is not to be used as a qualifier of the result in the record on which it appears.	Perm
GFSEQ ★	Sequence Number	Num		Identifier	Sequence number to ensure uniqueness of records within a dataset for a subject. May be any valid number (including decimals) and does not have to start at 1.	Req
GFGRPID	Group ID	Char		Identifier	Used to link together a block of related records within a subject in a domain.	Perm
GFREFID ★	Reference ID	Char		Identifier	A unique identifier for the assayed genetic specimen.	Exp
GFSPID	Sponsor-Defined Identifier	Char		Identifier	Sponsor-defined identifier.	Perm
GFLNKID	Link ID	Char		Identifier	Identifier used to link related records across domains. This may be a one-to-one or a one-to-many relationship.	Perm
GFLNKGRP	Link Group ID	Char		Identifier	Identifier used to link related records across domains. This will usually be a many-to-one relationship.	Perm

Platform used to detect the finding may be represented here

GF Topic Variable

GF Ex 2 - Single Nucleotide Variation

STUDYID	DOMAIN	USUBJID	GFSEQ	GFREFID	GFTESTCD	GFTEST	GFTSTDTL	GFORRES
C12345	GF	C12345-001	1	NA18537	SNV	Single Nucleotide Variation	GENOTYPE	T/T

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format[1]	Role	CDISC Notes	Core
GFTESTCD ★	Short Name of Genomic Measurement	Char	(GFTESTCD)	Topic	Short name of the measurement, test, or examination described in GFTEST. It can be used as a column name when converting a dataset from a vertical to a horizontal format. The value in GFTESTCD cannot be longer than 8 characters, nor can it start with a number (e.g., "1TEST" is not valid). GFTESTCD cannot contain characters other than letters, numbers, or underscores.	Req

GF Qualifier Variables

GF Ex 2 - Single Nucleotide Variation

STUDYID	DOMAIN	USUBJID	GFSEQ	GFREFID	GFTESTCD	GFTEST	GFTSTDTL	GFORRES
C12345	GF	C12345-001	1	NA18537	SNV	Single Nucleotide Variation	GENOTYPE	T/T

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format[1]	Role	CDISC Notes	Core
GFTEST ★	Name of Genomic Measurement	Char	(GFTEST)	Synonym Qualifier	Long name for GFTESTCD. The value in GFTEST cannot be longer than 40 characters.	Req
GFTSTDTL ★	Measurement, Test, or Examination Detail	Char	(GFTSDTL)	Variable Qualifier	Description of a reportable qualifying the assessment in GFTESTCD and GFTEST.	Perm
GFCAT	Category for Genomic Finding	Char		Grouping Qualifier	Used to define a category of topic-variable values.	Perm
GFSCAT	Subcategory for Genomic Finding	Char		Grouping Qualifier	Used to define a further categorization of GFCAT values.	Perm

GF Qualifier Variables

GF Ex 2 - Single Nucleotide Variation

GFTESTCD	GFTEST	GFTSTDTL	GFORRES	GFORREF	GFSTRESC	GFSTREFC	GFINHERT	GFGENREF	GFCHROM	GFSYM	GFSYMTYP	GFGENLOC	GFSEQID	GFPVRID	GFNAM	GFSPEC	GFMETHOD
SNV	Single Nucleotide Variation	GENOTYPE	T/T	G/G	T/T	G/G	GERMLINE VARIATION	GRCh38.p13	4	ABCG2	GENE WITH PROTEIN PRODUCT	4:88131171	ENSG00000118777	rs2231142	ACME LABS	DNA	REAL_TIME POLYMERASE CHAIN REACTION

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format[1]	Role	CDISC Notes	Core
GFORRES ★	Result or Finding in Original Units	Char		Result Qualifier	Result of the measurement or finding as originally received or collected.	Exp
GFORRESU	Original Units	Char	(UNIT)	Variable Qualifier	Unit for GFORRES.	Perm
GFORREF ★	Reference Result in Original Units	Char		Variable Qualifier	Reference value for the result or finding as originally received or collected. GFORREF uses the same units as GFORRES, if applicable.	Perm
GFSTRESC ★	Result or Finding in Standard Format	Char		Result Qualifier	Contains the result value for all findings, copied or derived from GFORRES, in a standard format or in standard units. GFSTRESC should store all results or findings in character format; if results are numeric, they should also be stored in numeric format in GFSTRESN.	Exp
GFSTRESN	Numeric Result/Finding in Standard Units	Num		Result Qualifier	Used for continuous or numeric results or findings in standard format; copied in numeric format from GFSTRESC. GFSTRESN should store all numeric test results or findings.	Perm
GFSTRESU	Standard Units	Char	(UNIT)	Variable Qualifier	Standardized units used for GFSTRESC, GFSTRESN, GFSTREFC, and GFSTREFN.	Perm
GFSTREFC ★	Reference Result in Standard Format	Char		Variable Qualifier	Reference value for the result or finding copied or derived from GFORREF in a standard format.	Perm
GFSTREFN	Numeric Reference Result in Std Units	Num		Variable Qualifier	Reference value for continuous or numeric results or findings in standard format or in standard units. GFSTREFN uses the same units as GFSTRESN, if applicable.	Perm
GFRESCAT	Result Category	Char		Variable Qualifier	Used to categorize the result of a finding.	Perm

GF Qualifier Variables

GF Ex 2 - Single Nucleotide Variation

GFTESTCD	GFTEST	GFTSTDTL	GFORRES	GFORREF	GFSTRESC	GFSTREFC	GFINHERT	GFGENREF	GFCHROM	GFSYM	GFSYMTYP	GFGENLOC	GFSEQID	GFPVRID	GFNAM	GFSPEC	GFMETHOD
SNV	Single Nucleotide Variation	GENOTYPE	T/T	G/G	T/T	G/G	GERMLINE VARIATION	GRCh38.p13	4	ABCG2	GENE WITH PROTEIN PRODUCT	4:88131171	ENSG00000118777	rs2231142	ACME LABS	DNA	REAL_TIME POLYMERASE CHAIN REACTION

New variables for genomics in SDTM v2.0

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format[1]	Role	CDISC Notes	Core
GFINHERT ★	Inheritability	Char	(INHERTGF)	Variable Qualifier	Identifies whether the variation can be passed to the next generation.	Perm
GFGENREF ★	Genome Reference	Char		Variable Qualifier	An identifier for the genome reference used to generate the reported result. For example, Genome Reference Consortium Human Build 38 patch release 13 may be represented as "GRCh38.p13".	Perm
GFCHROM ★	Chromosome Identifier	Char		Variable Qualifier	The designation (name or number) of the chromosome or contig on which the variant or other feature appears (e.g., "17"; "X").	Perm
GFSYM ★	Genomic Symbol	Char	*	Variable Qualifier	A published symbol for the portion of the genome serving as a locus for the experiment/test.	Perm
GFSYMTYP ★	Genomic Symbol Type	Char	(SYMTYPGF)	Variable Qualifier	A description of the type of genomic entity that is represented by the published symbol in GFSYM.	Perm
GFGENLOC ★	Genetic Location	Char		Variable Qualifier	Specifies the location within a sequence for the observed value in GFORRES.	Perm
GFGENSR	Genetic Sub-Region	Char		Variable Qualifier	The portion of the locus in which the variation was found. Examples: "Exon 15", "Kinase domain".	Perm
GFSEQID ★	Sequence Identifier	Char		Variable Qualifier	A unique identifier for the sequence used as the reference to identify the genetic variation in the result. Examples: "NM_001234", "ENSG00000182533", "ENST00000343849.2".	Perm
GFPVRID ★	Published Variant Identifier	Char		Variable Qualifier	A unique identifier for the variation that has been publicly characterized in an external database. Examples: "rs2231142", "COSM41596".	Perm
GFCOPYID	Copy Identifier	Char		Variable Qualifier	An arbitrary identifier used to differentiate between copies of a genetic target of interest present on homologous chromosomes.	Perm

GF Qualifier Variables

GF Ex 2 - Single Nucleotide Variation

GFTESTCD	GFTEST	GFTSTDTL	GFORRES	GFORREF	GFSTRESC	GFSTREFC	GFINHERT	GFGENREF	GFCHROM	GFSYM	GFSYMTYP	GFGENLOC	GFSEQID	GFPVRID	GFNAM	GFSPEC	GFMETHOD
SNV	Single Nucleotide Variation	GENOTYPE	T/T	G/G	T/T	G/G	GERMLINE VARIATION	GRCh38.p13	4	ABCG2	GENE WITH PROTEIN PRODUCT	4:88131171	ENSG00000118777	rs2231142	ACME LABS	DNA	REAL_TIME POLYMERASE CHAIN REACTION

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format[1]	Role	CDISC Notes	Core
GFSTAT	Completion Status	Char	(ND)	Record Qualifier	Used to indicate that a question was not asked or a test was not done, or a test was attempted but did not generate a result. Should be null or have a value of "NOT DONE".	Perm
GFREASND	Reason Test Not Done	Char		Record Qualifier	Reason not done. Used in conjunction with GFSTAT when value is "NOT DONE".	Perm
GFXFN	External File Path	Char		Record Qualifier	The filename and/or path to external data not stored in the same format and possibly not the same location as the other data for a study.	Perm
GFNAM ★	Laboratory/Vendor Name	Char		Record Qualifier	Name or identifier of the vendor that provided the test result. When more than 1 vendor is involved in the generation of the result, additional vendors should be represented as supplemental qualifiers.	Perm
GFSPEC ★	Specimen Material Type	Char	(GENSMP)	Record Qualifier	Identifies the type of genetic material used for the measurement.	Perm

GF Qualifier Variables

GF Ex 2 - Single Nucleotide Variation

GFTESTCD	GFTEST	GFTSTDTL	GFORRES	GFORREF	GFSTRESC	GFSTREFC	GFINHERT	GFGENREF	GFCHROM	GFSYM	GFSYMTYP	GFGENLOC	GFSEQID	GFPVRID	GFNAM	GFSPEC	GFMETHOD
SNV	Single Nucleotide Variation	GENOTYPE	T/T	G/G	T/T	G/G	GERMLINE VARIATION	GRCh38.p13	4	ABCG2	GENE WITH PROTEIN PRODUCT	4:88131171	ENSG00000118777	rs2231142	ACME LABS	DNA	REAL_TIME POLYMERASE CHAIN REACTION

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format[1]	Role	CDISC Notes	Core
GFMETHOD ★	Method of Test or Examination	Char	(METHOD)	Record Qualifier	The test method by which the examination is performed by the wet lab in order to yield the result reported in the dataset.	Exp
GFRUNID	Run ID	Char		Record Qualifier	A unique identifier for a particular run of a test performed by the wet lab on a particular batch of samples. This identifier can be used to distinguish between records for the same test performed at different times.	Perm
GFANMETH	Analysis Method	Char	(GFANMET)	Record Qualifier	The method of secondary processing performed by the dry lab to yield the result reported in the dataset.	Perm
GFBLFL	Baseline Flag	Char	(NY)	Record Qualifier	Indicator used to identify a baseline value. Should be "Y" or null.	Perm
GFDRVFL	Derived Flag	Char	(NY)	Record Qualifier	Used to indicate a derived record (e.g., a record that represents the average of other records such as a computed baseline). Should be "Y" or null.	Perm
GFLLOQ	Lower Limit of Quantitation	Num		Variable Qualifier	Indicates the lower limit of quantitation for an assay. Units will be those used for GFSTRESU.	Perm
GFREPNUM	Repetition Number	Num		Record Qualifier	The instance number of a test that is repeated within a given timeframe for the same test performed by the wet lab.	Perm

GF Timing Variables

GF Ex 2 - Single Nucleotide Variation

VISITNUM	VISIT	VISITDY	GFDTC	GFDY
1	SCREENING	-1	2020-06-25	-3

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format[1]	Role	CDISC Notes	Core
VISITNUM ★	Visit Number	Num		Timing	Clinical encounter number. Numeric version of VISIT, used for sorting.	Exp
VISIT ★	Visit Name	Char		Timing	Protocol-defined description of clinical encounter.	Perm
VISITDY ★	Planned Study Day of Visit	Num		Timing	Planned study day of VISIT. Should be an integer.	Perm
GFDTC ★	Date/Time of Specimen Collection	Char	ISO 8601 datetime or interval	Timing	Date and time of specimen collection.	Exp
GFDY ★	Study Day of Specimen Collection	Num		Timing	Actual study day of visit/collection/exam expressed in integer days relative to the sponsor-defined RFSTDTC in Demographics.	Perm
GFTPT	Planned Time Point Name	Char		Timing	Text description of time when a measurement or observation should be taken as defined in the protocol. This may be represented as an elapsed time relative to a fixed reference point, such as time of last dose. See GFTPTNUM and GFTPTREF.	Perm
GFTPTNUM	Planned Time Point Number	Num		Timing	Numerical version of GFTPT used in sorting.	Perm
GFELTM	Planned Elapsed Time from Time Point Ref	Char	ISO 8601 duration	Timing	Elapsed time relative to a planned fixed reference (GFTPTREF). This variable is useful where there are repetitive measures. Not a clock time or a date time variable, but an interval, represented as ISO duration.	Perm
GFTPTREF	Time Point Reference	Char		Timing	Name of the fixed reference point referred to by GFELTM, GFTPTNUM, and GFTPT. Examples: "PREVIOUS DOSE", "PREVIOUS MEAL".	Perm
GFRFTDTC	Date/Time of Reference Time Point	Char	ISO 8601 datetime or interval	Timing	Date/time for a fixed reference time point defined by GFTPTREF.	Perm

Terminology Considerations

GF variables with Controlled Terminology

Variable Name	Variable Label	Description	Associated Controlled Terminology?
GFTEST/CD	Name/Short Name of Genomic Measurement	Long/short name of the measurement, test, or examination described in GFTEST.	CDISC CT
GFTSTDTL	Measurement, Test, or Examination Detail	Description of a reportable qualifying the assessment in GFTESTCD and GFTEST.	CDISC CT
GFINHERT	Inheritability	Identifies whether the variation can be passed to the next generation.	CDISC CT
GFGENREF	Genome Reference	An identifier for the genome reference used to generate the reported result.	External
GFSYM	Genomic Symbol	A published symbol for the portion of the genome serving as a locus for the experiment/test.	External-HGNC
GFSYMTYP	Genomic Symbol Type	A description of the type of genomic entity that is represented by the published symbol in --SYM.	CDISC-CT
GFSEQID	Sequence Identifier	A unique identifier for the sequence used as the reference to identify the genetic variation in the result.	External
GFPRVID	Published Variant Identifier	A unique identifier for the variation that has been publicly characterized in an external database.	External

Terminology Considerations for GFTEST/CD and GFTSTDTL

This example shows findings from an assessment of a known single nucleotide variant in gene ABCG2 using wet laboratory methodology real-time polyerase chain reaction. Findings from this assessment show the genotypes from DNA extracted from the blood of 3 individuals, each with a different genotype at the genetic locus of interest. Because the DNA specimen was extracted from normal blood, the inheritabilty of the variation is considered to be in the germline.

gf.xpt

Row 1: Shows a subject genotype which is homozygous for the variant nucleotide in the reference sequence.
Row 2: Shows a subject genotype which is heterozygous for the nucleotide in the reference sequence.
Row 3: Shows a subject genotype which is homozygous for the nucleotide in the reference sequence.

GF Ex 2 - Single Nucleotide Variation

gf.xpt

Row	STUDYID	DOMAIN	USUBJID	GFSEQ	GFREFID	GFTESTCD	GFTEST	GFTSTDTL	GFORRES	GFORREF	GFSTRESC	GFSTREFC	GFINHERT	GFGENREF	GFCHROM	GFSYM	GFSYMTYP	GFGENLOC	GFSEQID	GFPVRID	GFNAM	GFSPEC	GFMETHOD	VISITNUM	VISIT	VISITDY	GFDTC	GFDY
1	C12345	GF	C12345-001	1	NA1853	SNV	Single Nucleotide Variation	GENOTYPE	T/T	G/G	T/T	G/G	GERMLINE VARIATION	GRCh38.p13	4	ABCG2	GENE WITH PROTEIN PRODUCT	4:88131171	ENSG00000118777	rs2231142	ACME LABS	DNA	REAL_TIME POLYMERASE CHAIN REACTION	1	SCREENING	-1	2020-06-25	-3
2	C12345	GF	C12345-002	2	NA0700	SNV	Single Nucleotide Variation	GENOTYPE	G/T	G/G	G/T	G/G	GERMLINE VARIATION	GRCh38.p13	4	ABCG2	GENE WITH PROTEIN PRODUCT	4:88131171	ENSG00000118777	rs2231142	ACME LABS	DNA	REAL_TIME POLYMERASE CHAIN REACTION	1	SCREENING	-1	2020-06-25	-3
3	C12345	GF	C12345-003	3	NA0013	SNV	Single Nucleotide Variation	GENOTYPE	G/G	G/G	G/G	G/G	GERMLINE VARIATION	GRCh38.p13	4	ABCG2	GENE WITH PROTEIN PRODUCT	4:88131171	ENSG00000118777	rs2231142	ACME LABS	DNA	REAL_TIME POLYMERASE CHAIN REACTION	1	SCREENING	-1	2020-06-25	-3

gf.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	GFSEQ	GFGRPID	GFREFID	GFTESTCD	GFTEST	GFTSTDTL	GFORRES	GFORRESU	GFSTRESC	GFSTRESN	GFSTRESU	GFINHERT	GFGENREF	GFCHROM	GFSYM	GFSYMTYP	GFGENLOC
1	ABC-123	GF	123101	ACME GenePanel 500	1	1	TRF001338	SNV	Single Nucleotide Variation	PREDICTED AMINO ACID CHANGE	D1853N		D1853N			SOMATIC VARIATION	GRCh37.75	11	ATM	GENE WITH PROTEIN PRODUCT	108175462
2	ABC-123	GF	123101	ACME GenePanel 500	2	1	TRF001338	SNV	Single Nucleotide Variation	PREDICTED CODING SEQUENCE CHANGE	5557G>A		5557G>A			SOMATIC VARIATION	GRCh37.75	11	ATM	GENE WITH PROTEIN PRODUCT	108175462
3	ABC-123	GF	123101	ACME GenePanel 500	3	1	TRF001338	SNV	Single Nucleotide Variation	VARIANT IMPACT CLASSIFICATION	ambiguous		ambiguous			SOMATIC VARIATION	GRCh37.75	11	ATM	GENE WITH PROTEIN PRODUCT	108175462
4	ABC-123	GF	123101	ACME GenePanel 500	4	1	TRF001338	SNV	Single Nucleotide Variation	READ DEPTH	501		501	501		SOMATIC VARIATION	GRCh37.75	11	ATM	GENE WITH PROTEIN PRODUCT	108175462
5	ABC-123	GF	123101	ACME GenePanel 500	5	1	TRF001338	SNV	Single Nucleotide Variation	VARIANT READ DEPTH/READ DEPTH	51	%	51	51	%	SOMATIC VARIATION	GRCh37.75	11	ATM	GENE WITH PROTEIN PRODUCT	108175462

Terminology Considerations for GFSYM

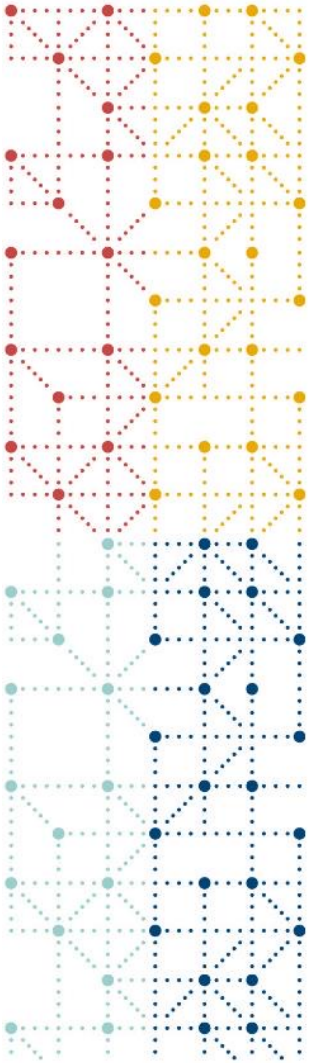
Where to put the Gene Name?

SDTMIGv3.4, GF Domain Assumption 5

- "For human genetic data, standard nomenclature populated in variable GFSYM must be obtained from the genomic symbol list maintained in the HUGO Gene Nomenclature Committee (HGNC) database (www.genenames.org)."
- Gene Symbols do not belong in GFTEST/CD – *Request will be denied.*

Terminology Considerations for GFANMETH

- GFANMETH variable is supported by the GFANMET codelist
  - Contains a list of named formulas or gene signatures
  - Codelist is extensible
- The definition for each value will contain a text description of the formula.
- The actual mathematical formula can be placed in the Define-XML file, owing to character constraints in the dataset.
- When submitting a CDISC change request for a new GFANMETH value, a paper citation for the formula as well as the related GFTEST and GFTSTDTL values should be submitted with the request for better understanding by the team.



Future Directions

Today

The CDISC Genomics Subteam goal for 2022 is to:

- Support stakeholder implementation of genomics standards through outreach and development/publication of resources and new standards

To achieve this goal, we are working toward deliverables related to:

- Communication of Standards
- Implementation Support
- Standards Development
- Refinements to Genomics Findings (GF)

In progress for 2022

Communication of Standards

- Introduction to the SDTM Genomics Findings (GF) Webinar (24 March)
- CDISC Europe Interchange (28 April)
- CDISC US Interchange (27 October)
- Training course (possible 2023)

Implementation Support

- CDISC Website Landing page updated: Genomics | CDISC
- Introduction to GF Knowledge Base Article (possible 2023)
- Concept Map for GF domain (in progress)
- GF domain examples in Examples Collection (Public Review completed)

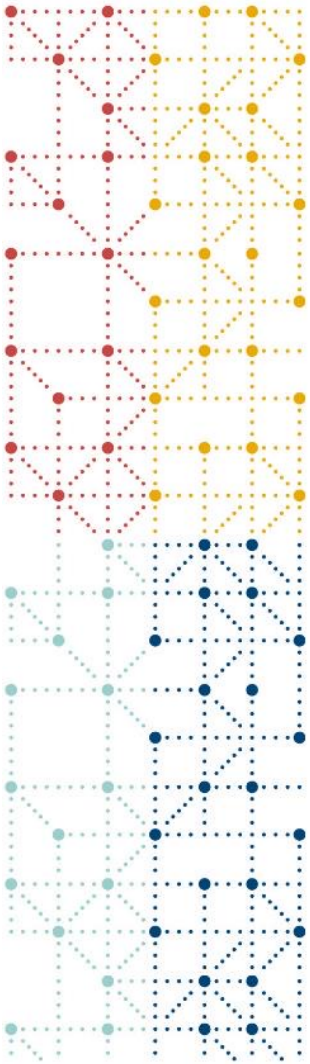
In progress for 2022

Standards Development

- GF Codetable Mapping File (published 25 March)
- Controlled Terminology Rules for GF (published 24 June)
- CDASH collaboration for genomic data collection (scoping 2023)

Refinements to Genomics Findings (GF)

- Pending development work

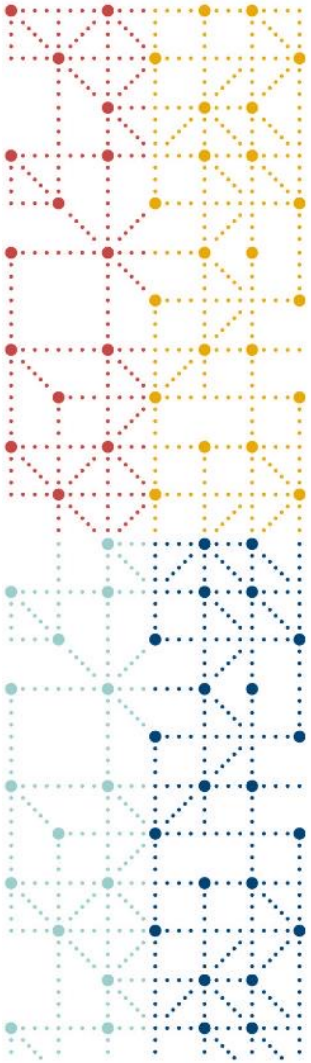


How you can be involved!

How you can be involved!

We invite you to contribute to continuous improvement of genomic standards:

- Become a CDISC Genomics Subteam volunteer
  - www.cdisc.org/volunteer
    - Click link to *Become a Volunteer*
- Contribute FAQs and use case examples for modeling
  - Use case examples should be real-life, de-identified, and submission related
  - We would like to discuss your use case examples with you
  - Reach out to Christine Connolly, CDISC Head of Standards Projects (cconnolly@cdisc.org)
- Review draft standards as they are released



Thank You!

cdisc