
Health informatics — Genomic Sequence Variation Markup Language (GSVML)

*Informatique de santé — Langage de balisage de la variation de
séquence génomique*

STANDARDSISO.COM : Click to view the full PDF of ISO 25720:2009



PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

STANDARDSISO.COM : Click to view the full PDF of ISO 25720:2009



COPYRIGHT PROTECTED DOCUMENT

© ISO 2009

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword	iv
Introduction.....	v
1 Scope	1
2 Conformance	1
2.1 Purpose	1
2.2 Needs and general requirements.....	1
3 Normative references	2
4 Terms and definitions	2
5 GSVML specification.....	5
5.1 Specification requirements and GSVML positioning	5
5.2 GSVML structure	5
5.3 GSVML DTD and XML schema	5
6 GSVML development process.....	5
Annex A (normative) DTD of GSVML	22
Annex B (normative) XML schema of GSVML	46
Annex C (informative) Basic reference works.....	105
Bibliography.....	131

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 25720 was prepared by Technical Committee ISO/TC 215, *Health informatics*.

STANDARDSISO.COM : Click to view the full PDF of ISO 25720:2009

Introduction

In the current electronic world, there are multiple different types of data for healthcare, as shown in Figure 1. Besides clinical data and image data, as we move into this post genomic era, we are creating, internationally, overwhelming amounts of genomic data. The International Standards developing organizations are developing standards for these data; Health Level Seven develops standards for clinical data, DICOM and JPEG develop standards for image data. Genomic Sequence Variation Markup Language (GSVML) defines a standard for genomic data, especially human-related DNA variation data. The core target for the GSVML is the Single Nucleotide Polymorphism (SNP).

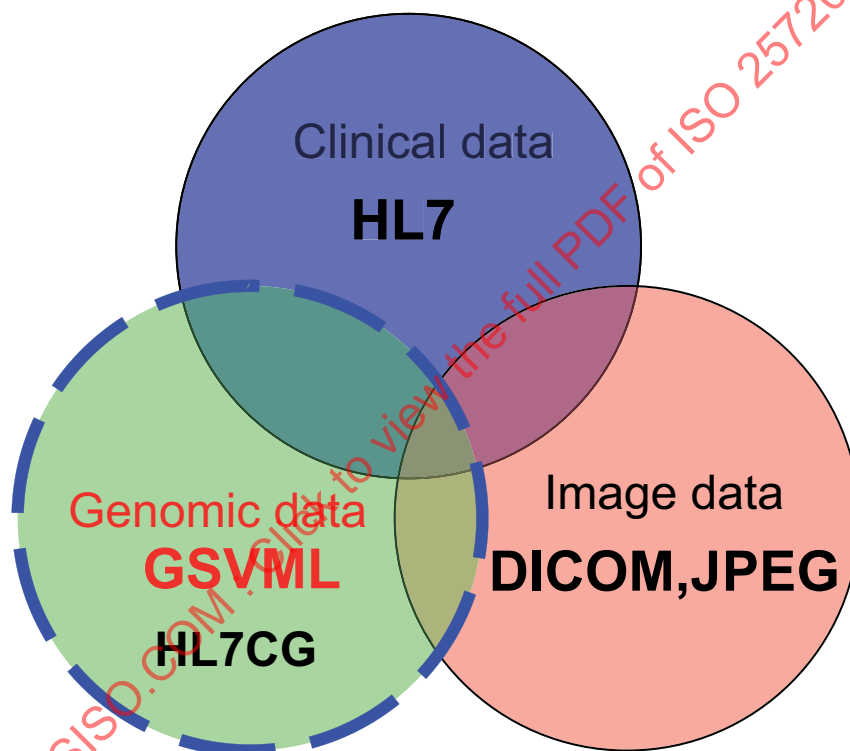


Figure 1 — Major data types of health care

In this post genomic era, the management of health-related data is becoming increasingly important to both genomic research and genome-based medicine (see reference [1]). Informational approaches to the management of clinical, image and genomic data are beginning to have as much worth as basic, bench top research. Nowadays there are many kinds of -omics data around the world awaiting effective utilization for human health. The hurdle that must be overcome to achieve this goal is the development of data format and message standards to support the interchange of -omics data. Genomic data include genome sequence, DNA sequence variation and other genome-based data such as expression data, proteomics data, molecular network, etc. As an entry point, this International Standard focuses on the DNA sequence variation. Among the DNA sequence variation, the SNP is selected as the core object because of the following three reasons.

- a) SNP is the most researched sequence variation for human health.
- b) In the current context, vast amounts of SNP data exist around the world in various types of data formats. As a result of the recent explosion in SNP research, the vast amounts of experimental data have been

accumulating in many databases in various types of data formats. These data await utilization in drug discovery, clinical diagnosis and clinical research.

- c) SNP data already have a great impact for human applications such as gene-based medicine and pharmacogenomics.

With a view to this context, the international community requires an interoperable format for the interchange of SNP data. Prior to the standardization development, we elucidated the need for data exchange among the human health-related facilities that have various types of data formats.

In the present circumstances, SNP is expected to be a key to understanding human response to external stimuli such as any kind of alien invasions, therapies, and the environmental interactions (see reference [2]). Bacterial infection is an example of alien invasion, and the responses to the infections are different amongst individuals. According to the therapy, the side effects to a drug are different amongst the patients. These responses are also different in various environments.

The Markup Language is a set of symbols and rules for their use when doing a markup of a document (see reference [3]). The first standardized markup language was Standard Generalized Markup Language (SGML), [4] which has strong similarities with troff and nroff text layout languages supplied with Unix systems. Hypertext Markup Language (HTML) is based on SGML [5]. Extensible Markup Language (XML) is a pared-down version of SGML, designed especially for Web documents (see reference [6]). XML acts as the basis for Extensible HTML (XHTML) [7] and Wireless Markup Language (WML) (see reference [8]) and for standardized definitions of system interaction such as Simple Object Access Protocol (SOAP) [9]. By contrast, text layout or semantics are often defined in a purely machine-interpretable form, as in most word processor file formats (see reference [10]).

Markup Language for the biomedical field, based on XML, has been in development for several decades to enhance the exchange data among researchers. Bioinformatic Sequence Markup Language (BSML) (see reference [11]), Systems Biology Markup Language (SBML) [12], Cell Markup Language (Cell ML) [13], and Neuro Markup Language (Neuro-ML) [14] are examples of markup languages. Polymorphism Mining and Annotation Programs (PolyMAPr) [15] is centric on SNP and tries to achieve mining, annotation and functional analysis of public databases such as dbSNP [16], the Cancer Gene Anatomy Project (CGAP) (see reference [17]), and Japanese single nucleotide polymorphisms (JSNP) (see reference [18]) through programming.

To utilize the accumulated SNP data among many facilities around the world, standards for the interchange of SNP data must be defined. The required standards include defining a data format and exchange messages. Markup Language is the reasonable choice to address this need. As for genomic data message handling, Health Level Seven Clinical Genomics Special Interest Group [19] has summarized clinical use cases for general genomic data. The GSVML project has contributed to these efforts. Additionally, this work incorporated use cases based on the Japanese Millennium Project [20]. Based on these contexts and investigations, this International Standard elucidates the needs and the requirements for GSVML and then proposes the specification of GSVML for the international standardization.

Health informatics — Genomic Sequence Variation Markup Language (GSVML)

IMPORTANT — The electronic file of this document contains colours which are considered to be useful for the correct understanding of the document. Users should therefore consider printing this document using a colour printer.

1 Scope

This International Standard is applicable to the data exchange format that is designed to facilitate the exchange of the genomic sequence variation data around the world, without forcing change of any database schema. From an informatics perspective, GSVM defines the data exchange format based on XML. The scope of this International Standard is the data exchange format, but the database schema itself is outside the scope of this International Standard. From a biological point of view, all genetic sequence variations are taken into consideration and are within the scope of this International Standard, while polymorphisms, especially SNPs, are the main focus of this International Standard. In other words, the annotations of variation as clinical concerns and -omics concerns are within the scope of this International Standard. Though SNPs exist in various biological species, the scope of this International Standard covers the human health associated species as human, cell line, and preclinical animals. The other biological species are outside the scope of this International Standard. The clinical field is within the scope of this International Standard, but the basic research fields and other scientific fields are outside the scope of this International Standard. Here, clinical research, including drug discovery, is within the scope of this International Standard. As for supposed application fields, the main focus is in human health, including clinical practice, preventive medicine, translational research and clinical researches.

2 Conformance

2.1 Purpose

This International Standard provides a data exchange format for genomic sequence variation data in human health. This International Standard provides the GSVM specification mainly for the case of SNP and Short Tandem Repeat Polymorphism (STRP). Considering that SNP and STRP are the major and simple polymorphisms in human health research, centering on them and expanding the specification to the other sequence variation data seems reasonable. This International Standard allows for the expandability of GSVM from SNP and STRP to other sequence variation data.

2.2 Needs and general requirements

The vast volume of experimental data from the recent explosion of genomic sequence variation research has produced an overwhelming amount of data stored in many databases with various types of format worldwide. Standardization of data exchange is urgent for managing, analysing, and utilizing these data. Standardizing the interoperable format is necessary for easy and convenient genomic sequence variation data exchange. Considering that genomic sequence variation, especially SNP and STRP, has its significant meaning in the gene-based medicine and the pharmacogenomics for human health, the data exchange format is the key to enhancing the gene-based clinical research and the gene-based medicine.

The management of genomic data is as critical as the basic research data in this new era. There are many kinds of -omics data around the world, and the time has come to effectively use these genomic data for human health. In order to use these data effectively and efficiently, standards must be developed to permit the interoperable interchange of genomic data globally. These standards must define the data format as well as

the messages to be used to interchange and share this data globally. This International Standard addresses those requirements, using a Markup Language.

3 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

EN 13606 (all parts), *Health informatics — Electronic healthcare record communication*

4 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

- 4.1**
actor
something or someone who supplies a stimulus to the system
- NOTE Actors include both humans and other quasi-autonomous things, such as machines, computer tasks and systems.
- 4.2**
Bioinformatic Sequence Markup Language
BSML
extensible language specification and container for bioinformatic data
- 4.3**
Cell Markup Language
Cell ML
Extensible Markup Language to provide a standard method for representing and exchanging computer-based biological models
- 4.4**
Cancer Gene Anatomy Project
CGAP
database containing genomic expression data collected for various tumorigenic tissues in both humans and mice and also providing information on methods and reagents used in deriving the genomic data
- 4.5**
dbSNP
database of SNPs provided by the US National Center for Biotechnology Information (NCBI)
- 4.6**
Digital Imaging and Communications in Medicine
DICOM
standard in the field of medical informatics for exchanging digital information between medical imaging equipment (such as radiological imaging) and other systems, ensuring interoperability
- 4.7**
deoxyribonucleic acid
DNA
molecule that encodes genetic information in the nucleus of cells
- 4.8**
DNA sequence variation
differences of DNA sequence among individuals in a population

NOTE DNA sequence variation implies **polymorphism** (4.20).

4.9**Document Type Definition****DTD**

separate document that contains formal definitions of all of the data elements in a particular type of HTML, SGML or XML document

4.10**entry point**

reference point that designate the class(es) from which the messages begin for the particular domain

4.11**gene-based medicine**

medicine based on genes or genetic science

4.12**Hypertext Markup Language****HTML**

set of markup symbols or codes inserted in a file intended for display on a World Wide Web browser page

4.13**Joint Photographic Experts Group****JPEG**

compression technique for images

4.14**Japanese single nucleotide polymorphisms****JSNP**

database of Japanese single nucleotide polymorphisms

4.15**markup language****ML**

set of symbols and rules for their uses when doing a markup of a document

4.16**Neuro Markup Language****Neuro-ML**

markup language for describing models of neurons and networks of neurons

4.17**nroff**

unix text-formatting program that is a predecessor of the Unix troff document processing system

4.18**pharmacogenomics**

branch of pharmaceuticals aiming to develop rational means to optimize drug therapy, with respect to the patient's genotype

4.19**Polymorphism Mining and Annotation Programs****PolyMAPr**

programs for polymorphism database mining, annotation and functional analysis

4.20**polymorphism**

variation in the sequence of DNA among individuals

NOTE Polymorphism implies **SNP** (4.23) and **STRP** (4.26).

4.21

Systems Biology Markup Language

SBML

markup language for simulations in systems biology

4.22

Standard Generalized Markup Language

SGML

standard for defining description of the structure of different types of electronic documents

4.23

Single Nucleotide Polymorphism

SNP

single nucleotide variation in a genetic sequence that occurs at appreciable frequency in the population

4.24

Systematized Nomenclature of Medicine – Clinical Terms

SNOMED CT

dynamic, scientifically validated clinical health care terminology and infrastructure

4.25

Simple Object Access Protocol

SOAP

lightweight protocol for exchange of information in a decentralized, distributed environment

4.26

Short Tandem Repeat Polymorphism

STRP

variable segments of DNA that are two to five bases long with numerous repeats

4.27

troff

document processing system developed by AT&T for the Unix operating system

4.28

variable number of tandem repeat

VNTR

class of polymorphism characterized by the highly variable copy number of identical or closely related sequences

4.29

Wireless Markup Language

WML

XML language used to specify content and user interface for WAP (wireless application protocol) devices

4.30

Extensible HTML

XHTML

hybrid between HTML and XML specifically designed for net device displays

4.31

Extensible Markup Language

XML

pared-down version of SGML, designed especially for web documents

4.32

XML schema

language for describing the structure and constraining the contents of XML documents

5 GSVML specification

5.1 Specification requirements and GSVML positioning

In the current context, annotative information about genomic sequence variation is increasing, which is filling in the gaps in information. The genomic sequence variation data themselves are also increasing but are stored in various databases. This trend is typical of SNP data. The pitfall of genomic sequence variation data handling is the lack of standardization of the data formats for the genomic sequence variation. Historically, the markup languages listed in Clause 4 have been used, and programs are developed to handle the genomic information. However, there have been no genomic sequence variation centric markup languages so far. GSVML is the first genomic sequence variation centric markup language and is human health centric. Considering that SNP is a highly researched polymorphism and has a great impact, especially for human health and response, we can say that GSVML has the greatest potential to be the designated markup language for human healthcare. On the other hand, setting the applications to practical human health means it must handle direct or indirect SNP annotations. Here the direct SNP annotation indicates general annotative information such as SNP associated genes and experimental preparations. The indirect SNP annotation indicates all of the -omics data and clinical data that result from SNP variation. To understand the gene-based clinical situation of each patient, we need this kind of additional information. Considering the requirement to add many kinds of additional information, the development and standardization of GSVML cannot stand alone and need harmonization with the other international standardization organizations such as Health Level Seven.

GSVML is intended to be used in data exchange messages related to human health. In the development and standardization of GSVML in this application domain, we must always keep an eye on the patient's safety, clinical efficiency and medical costs. For the patient's safety, from an informational viewpoint, the conservation and the protection of patient information are important. For the enhancement of clinical efficiency, simplicity and ease of understanding are important. For medical cost reduction, the adaptation ability and ease of installation are important. GSVML tries to respond to these basic requirements by providing the sharable XML based data exchanging format. GSVML can be used for the clinically genomic sequence variation data exchange among various types of data formats. In the greater framework of clinical data standardization, GSVML plays the part of describing the genomic sequence variation data and their necessary information.

5.2 GSVML structure

The outlined structure of GSVML is shown in Figure 3. GSVML consists of three data criteria, *viz.* variation data, direct annotation and indirect annotation. The variation data criterion describes the straightforward variation data as allele, type, position, length, region, etc. The direct annotation criterion describes the attached data of variation data as experiment analysis, epidemiology or associated gene, etc. The indirect annotation criterion describes the explanatory/higher-level information of variation data such as the -omics data, the clinical information and the environmental data. These data criteria have internal relations to each other. The detailed structure of GSVML is shown in Figures 4 to 21.

5.3 GSVML DTD and XML schema

The DTD of GSVML is shown in Annex A. The XML schema of GSVML is shown in Annex B.

6 GSVML development process

The development of GSVML followed eight steps:

- Step 1: Set the elements and needs according to the investigated use cases.

We prepared six use cases for three typical criteria. Four use cases concerned the clinical practice, and one use case for each clinical trial and translational research.

- Step 2: Construct the basic structure and DTD.

- Step 3: Investigate the existing biological ML, and its applicability to the needs (comparison with MAGE-ML, BSML, SBML, RNAMEL^[21], ProML, CellML, PolyMAPr).
- Step 4: Refine the basic structure and DTD, construct the XML Schema (XSD).
- Step 5: Investigate the existing SNP databases (their data format comparison).
- Step 6: Check the interface ability to the Health Level Seven Genotype Model.
- Step 7: Redefine the needs of GSVML and its demanded elements.
- Step 8: Refine the basic structure, DTD, and XML Schema.

Figure 2 shows the outline of the process of the development. We did design work in harmony with HL7 Clinical Genomics SIG. There were “to and fro” processes between design work and the standardization process.

Additionally, we analysed the interface between GSVML and EN 13606, SNOMED-CT

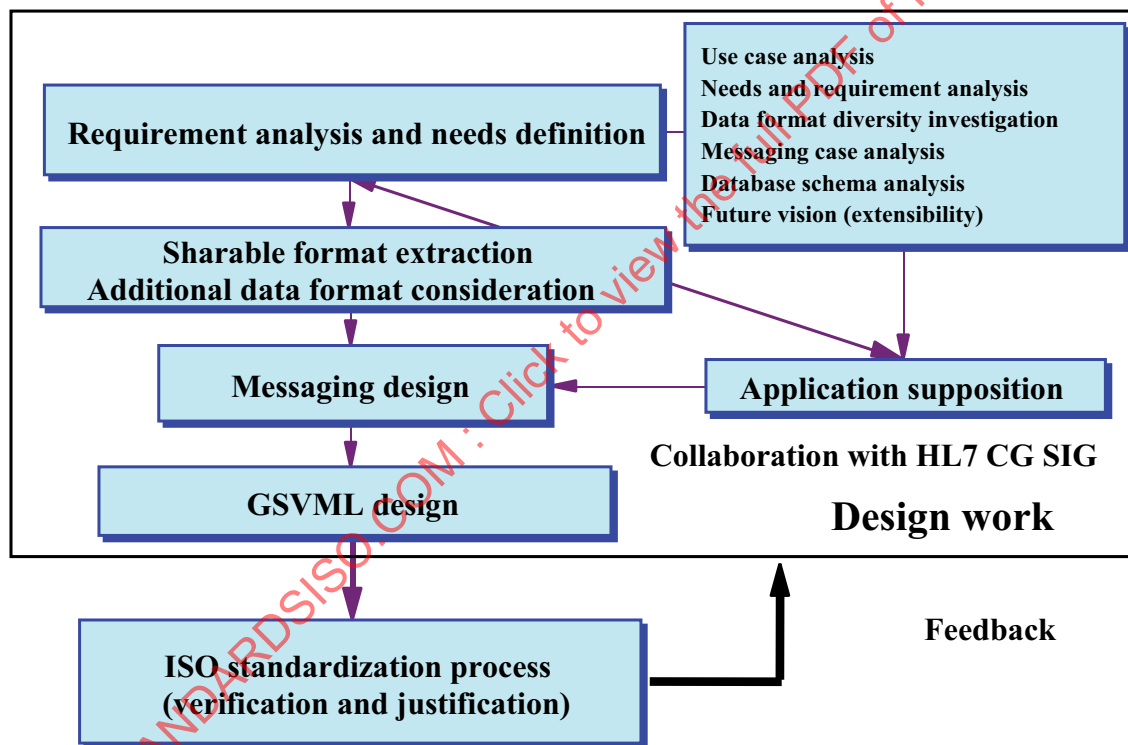
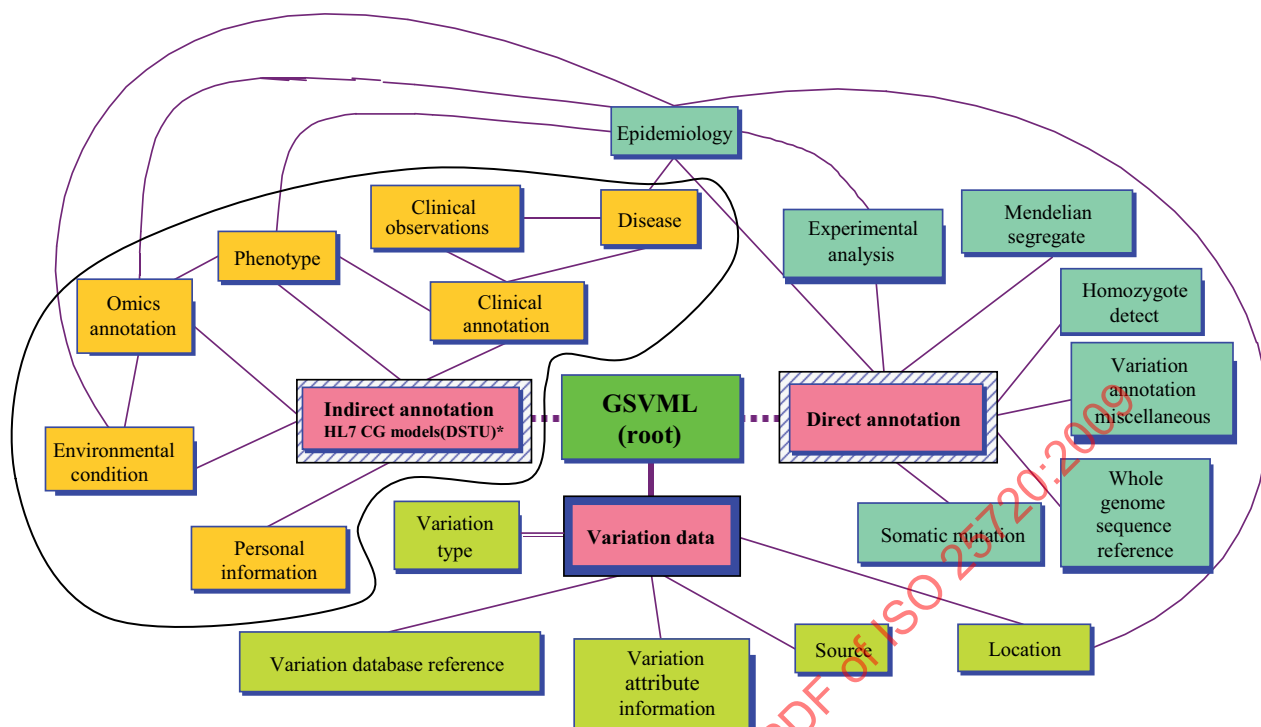


Figure 2 — Outline of the process of GSVML development



* HL7 CG models (DSTU) will be used instead of indirect annotation criterion.

Figure 3 — The outlined structure of GSVML

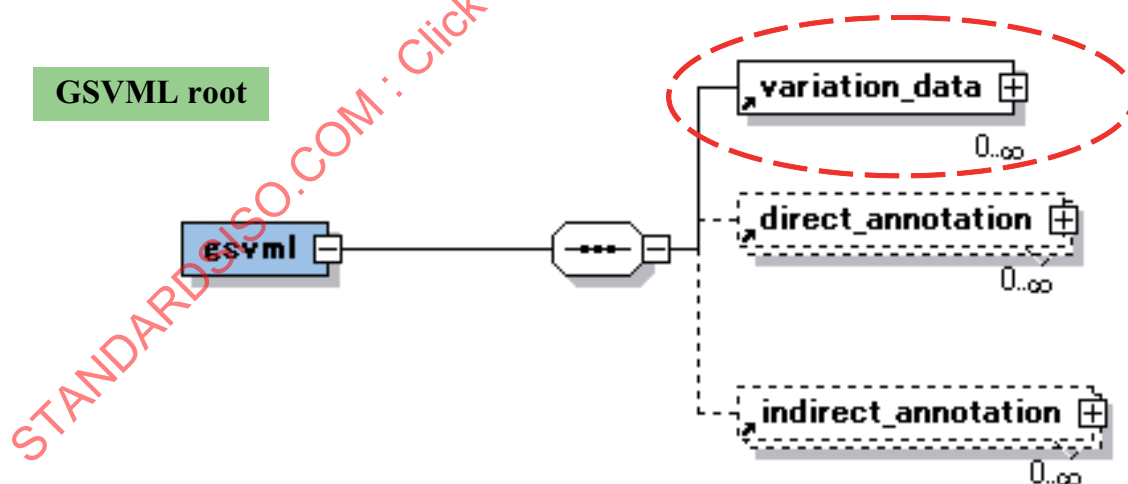


Figure 4 — Detailed structure of GSVML

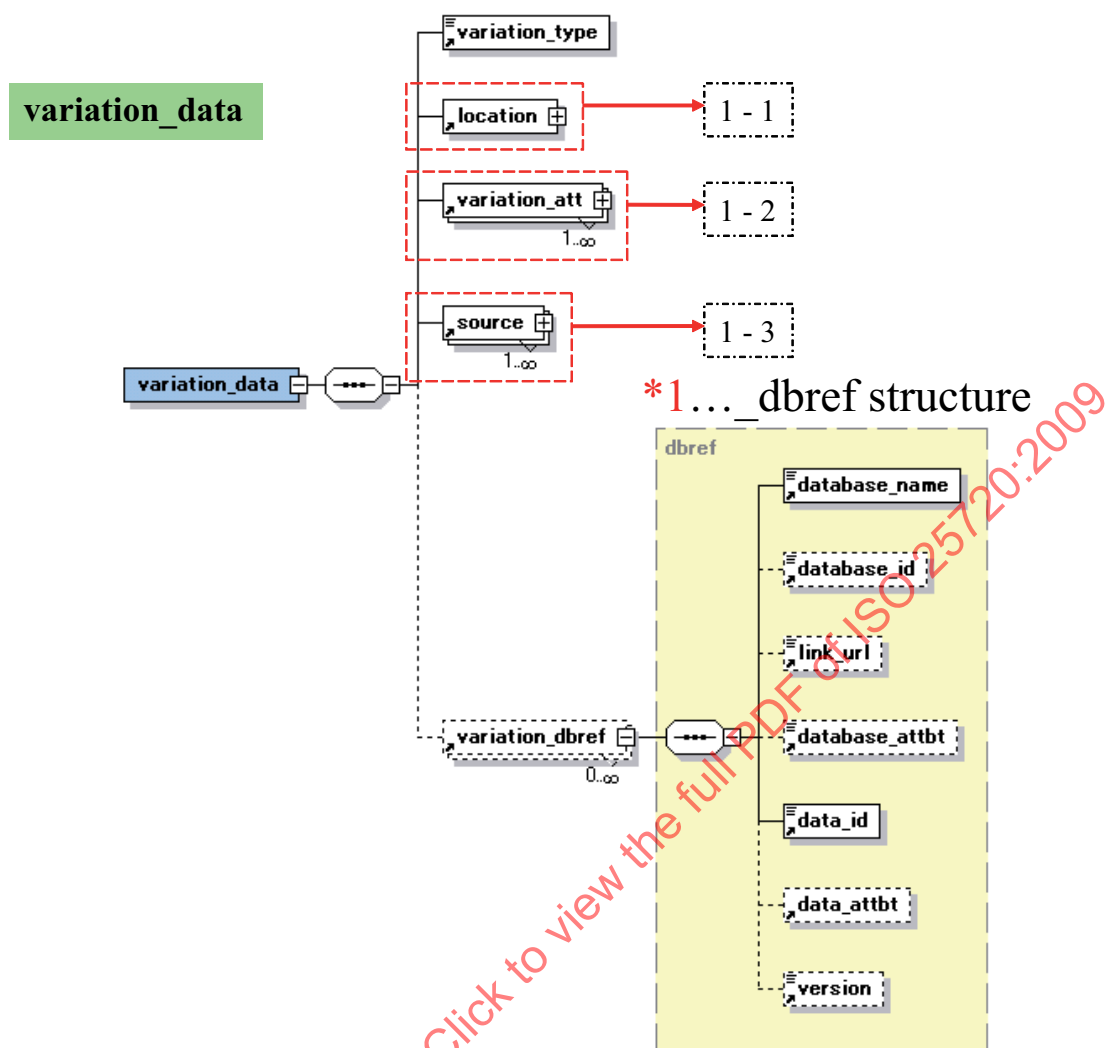


Figure 5 — Detailed structure of GSVML

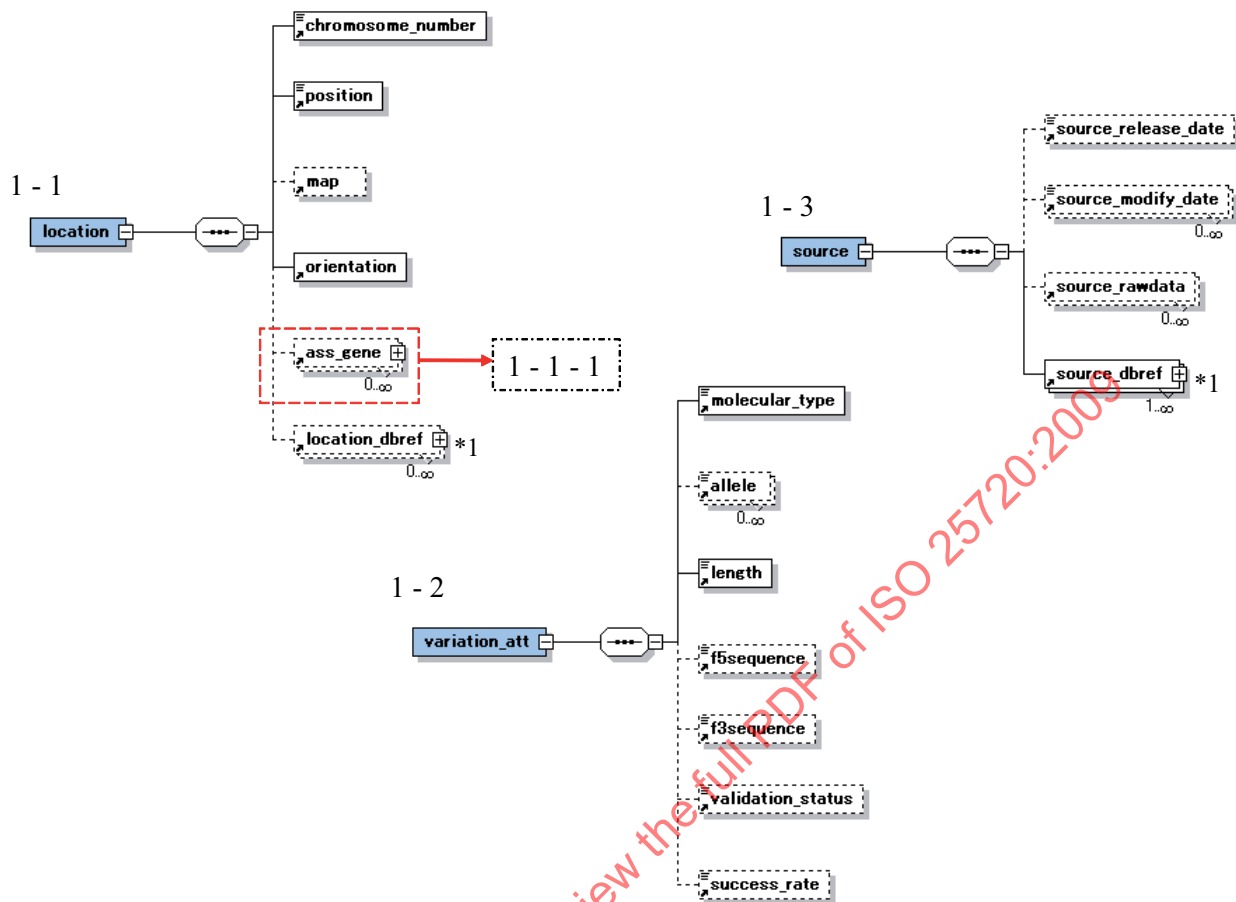


Figure 6 — Detailed structure of GSVML

ass_gene

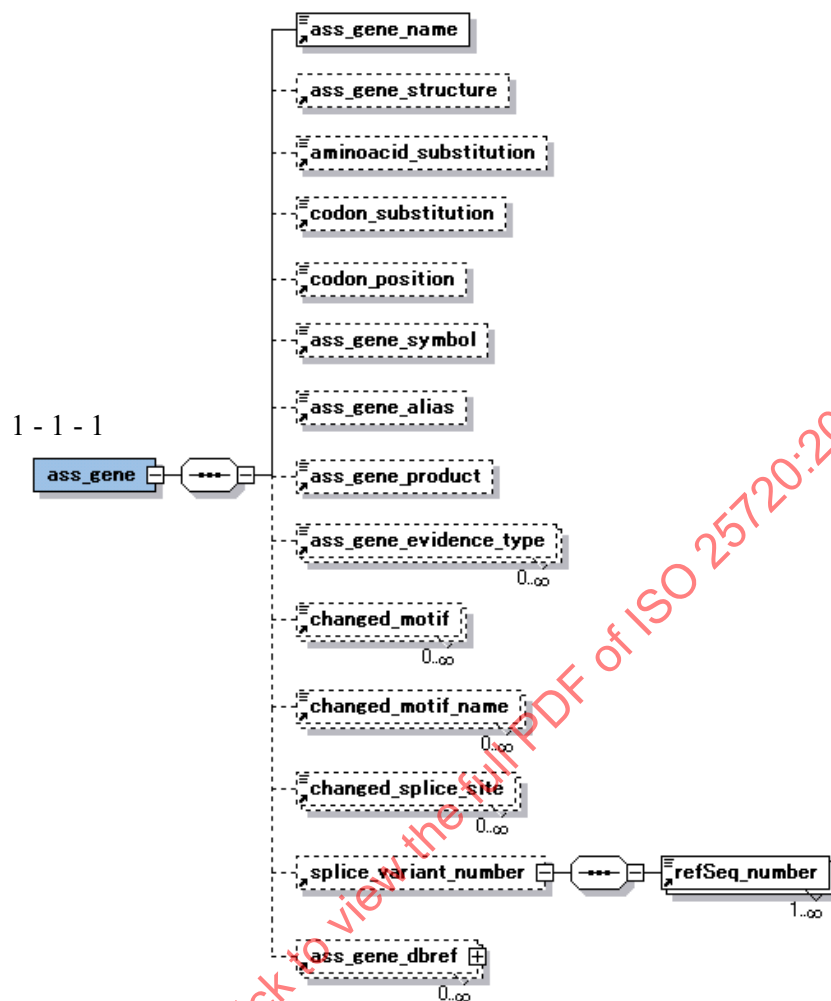


Figure 7 — Detailed structure of GSVML

GSVML root

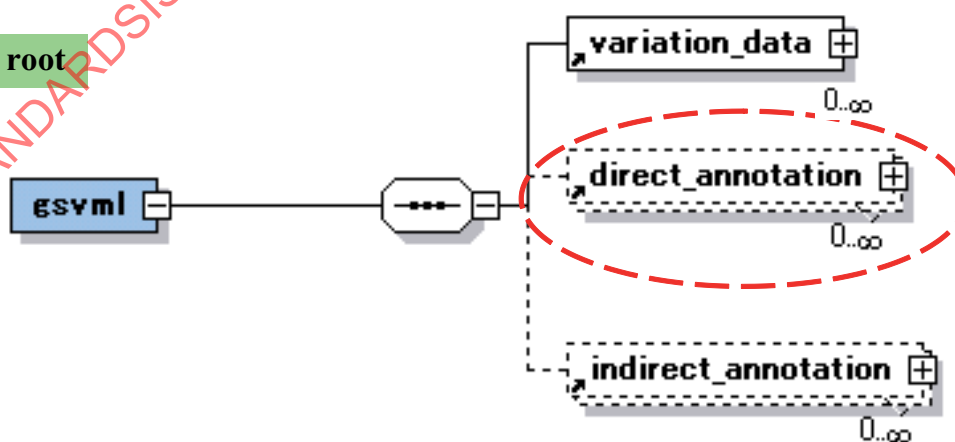


Figure 8 — Detailed structure of GSVML

direct_annotation

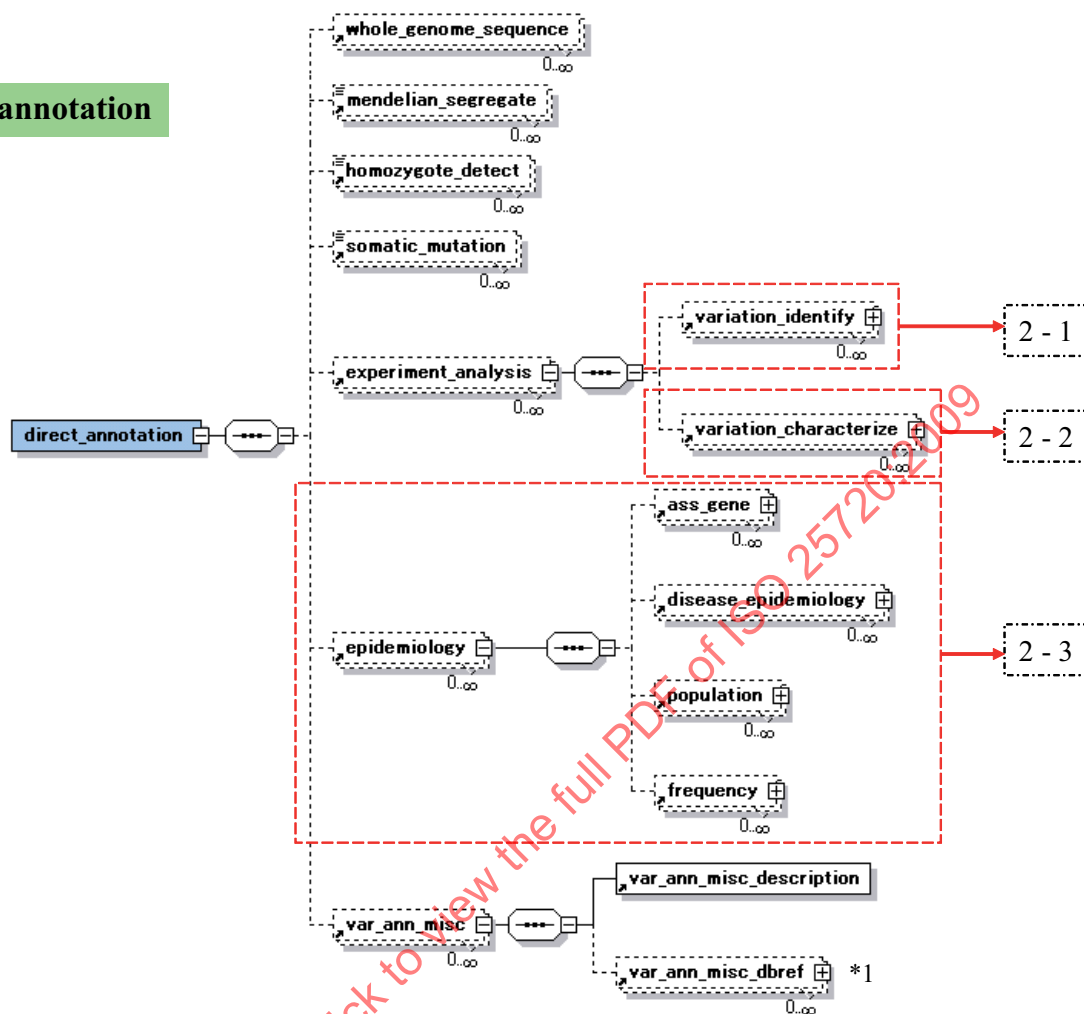


Figure 9 — Detailed structure of GSVML

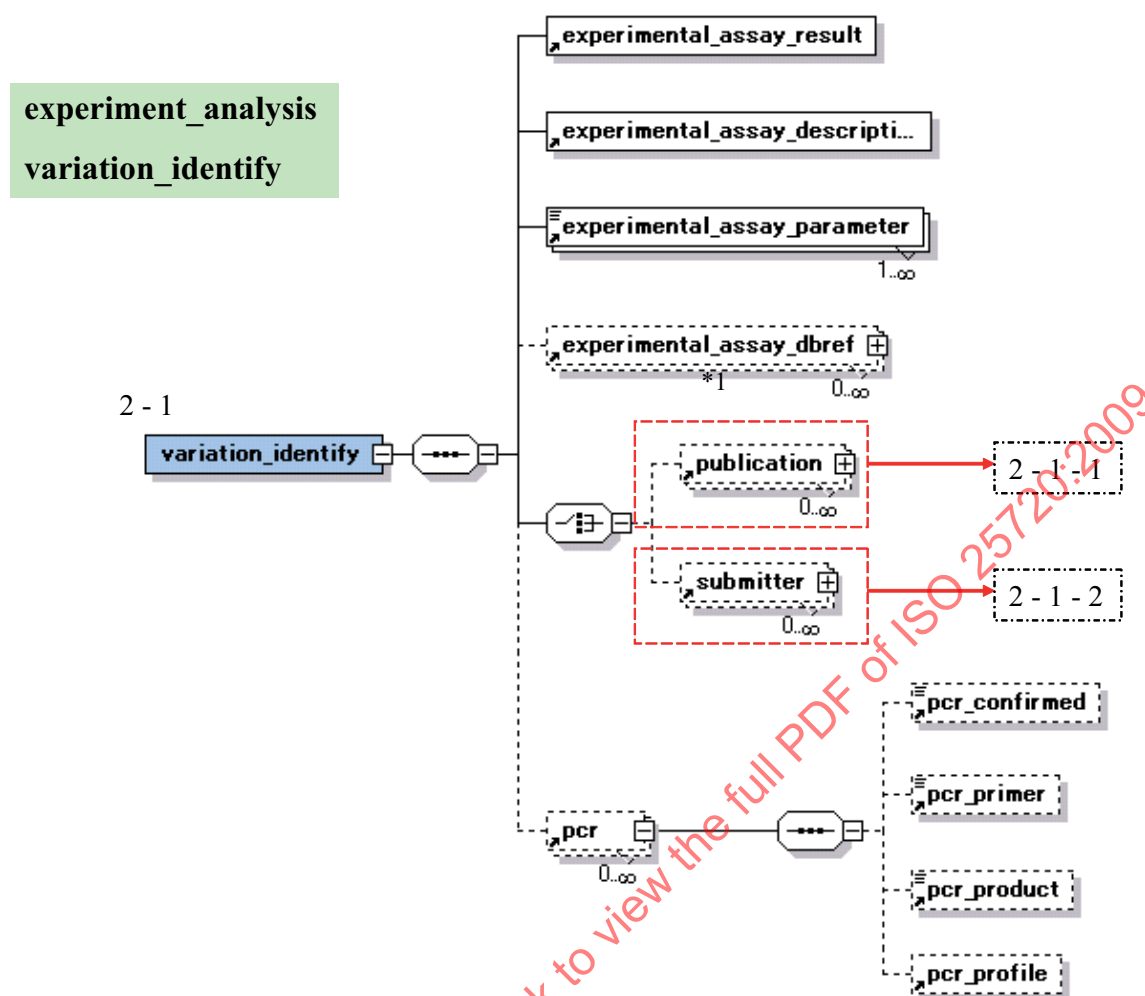


Figure 10 — Detailed structure of GSVML

publication and submitter

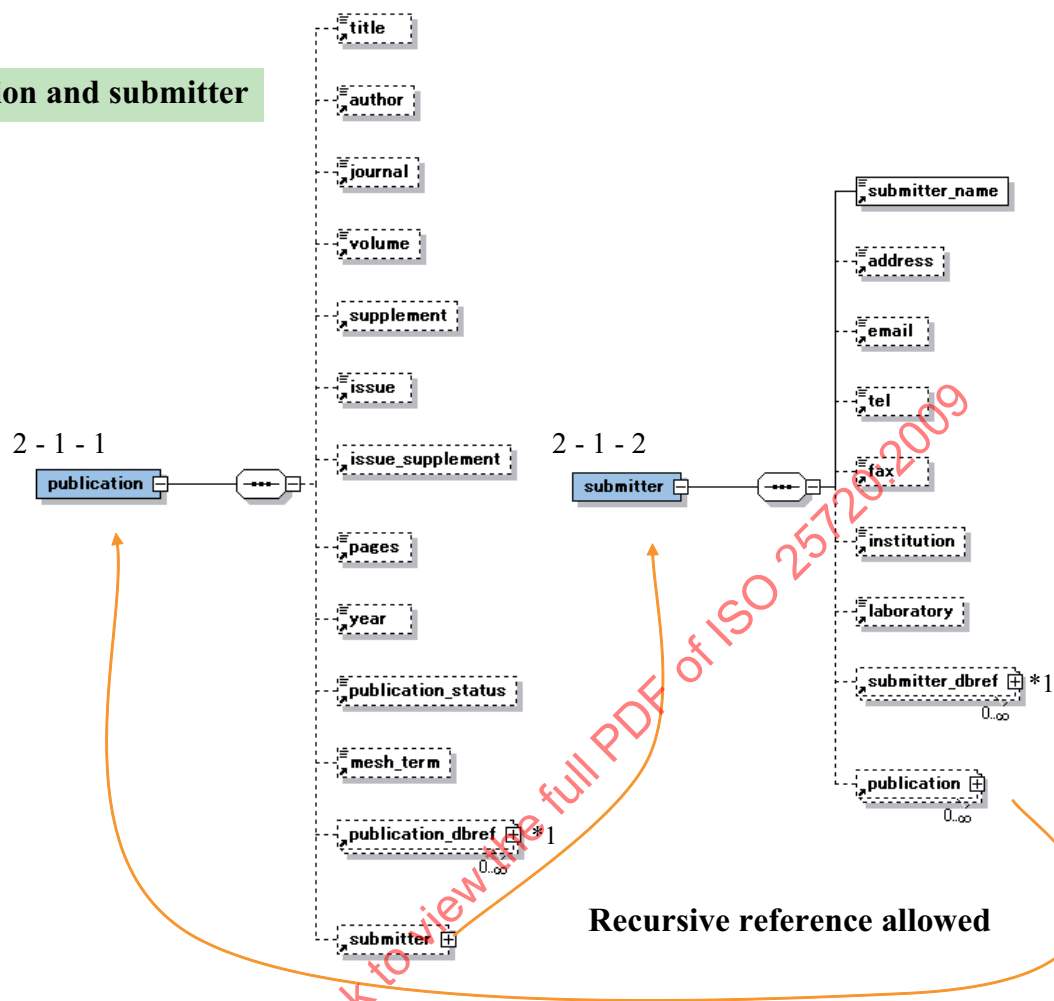


Figure 11 — Detailed structure of GSVML

experiment_analysis
variation_characterize

2 - 2

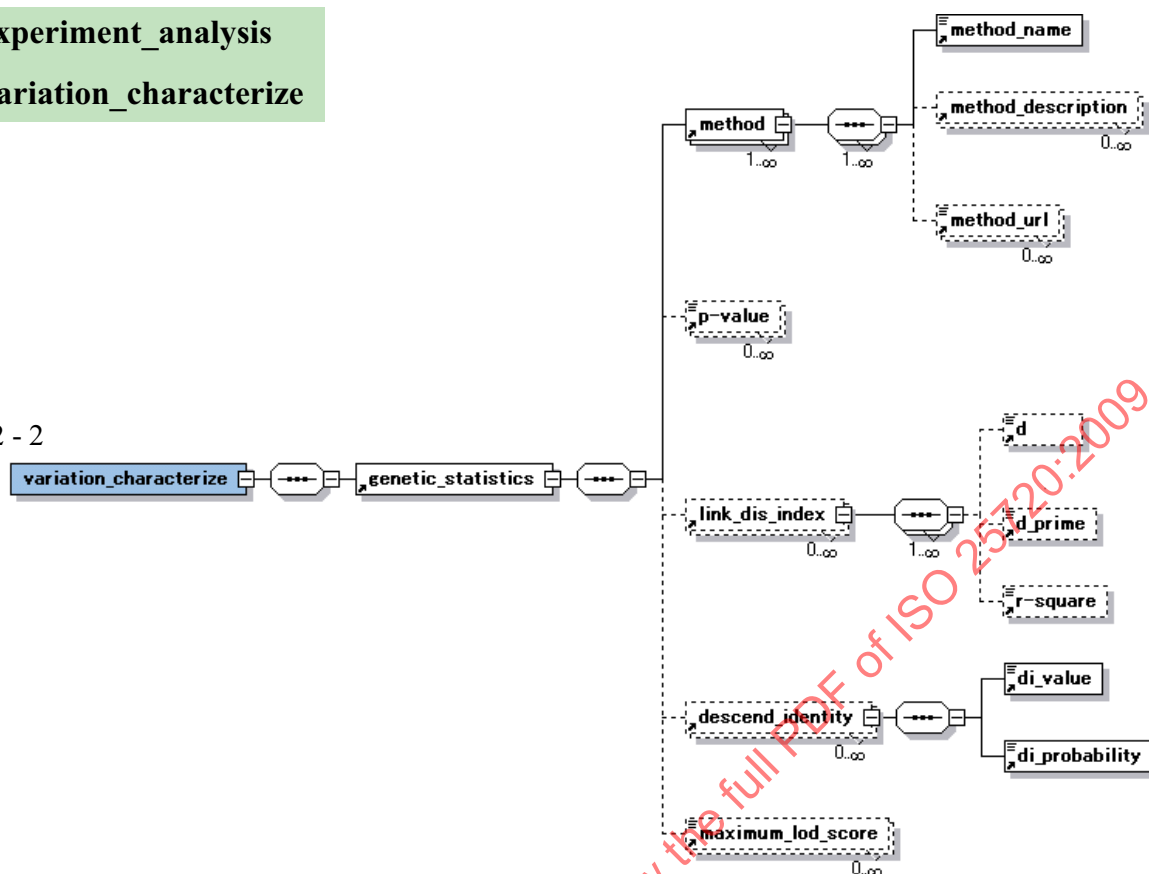


Figure 12 — Detailed structure of GSVML

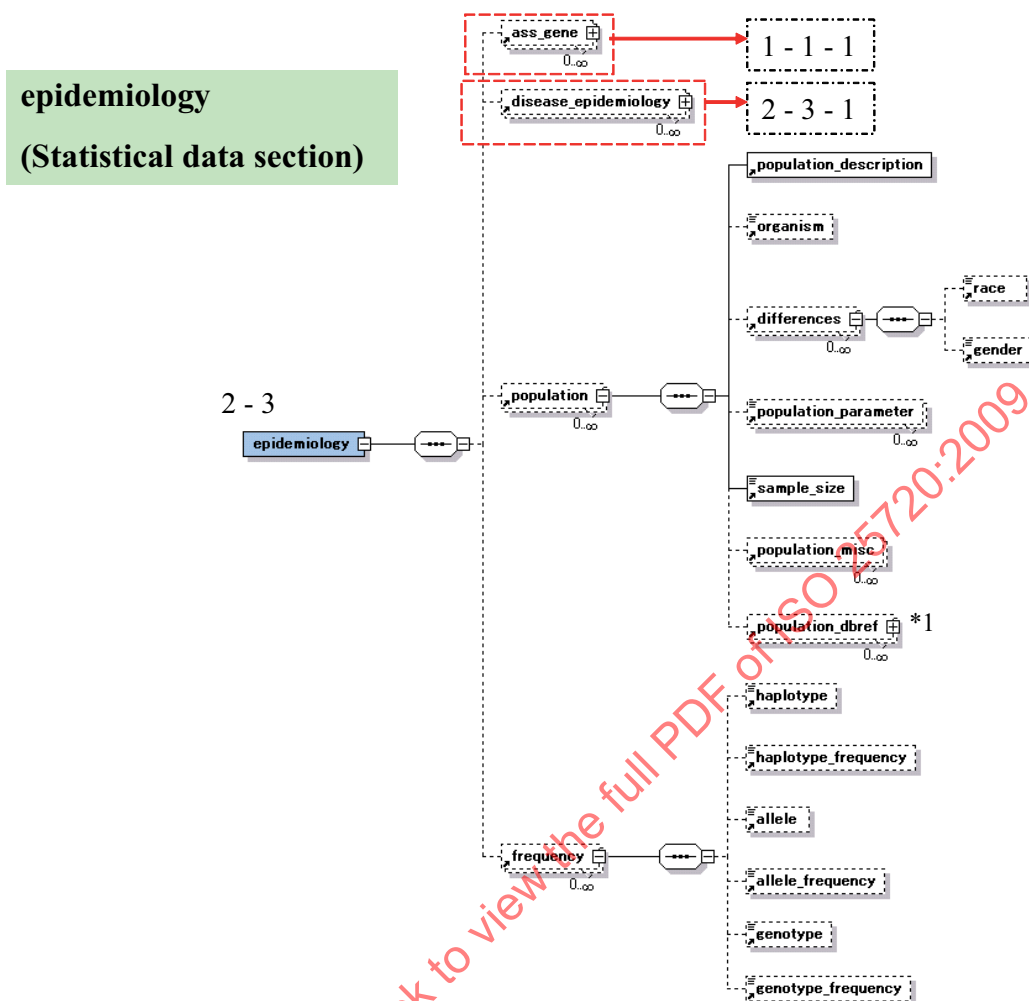


Figure 13 — Detailed structure of GSVML

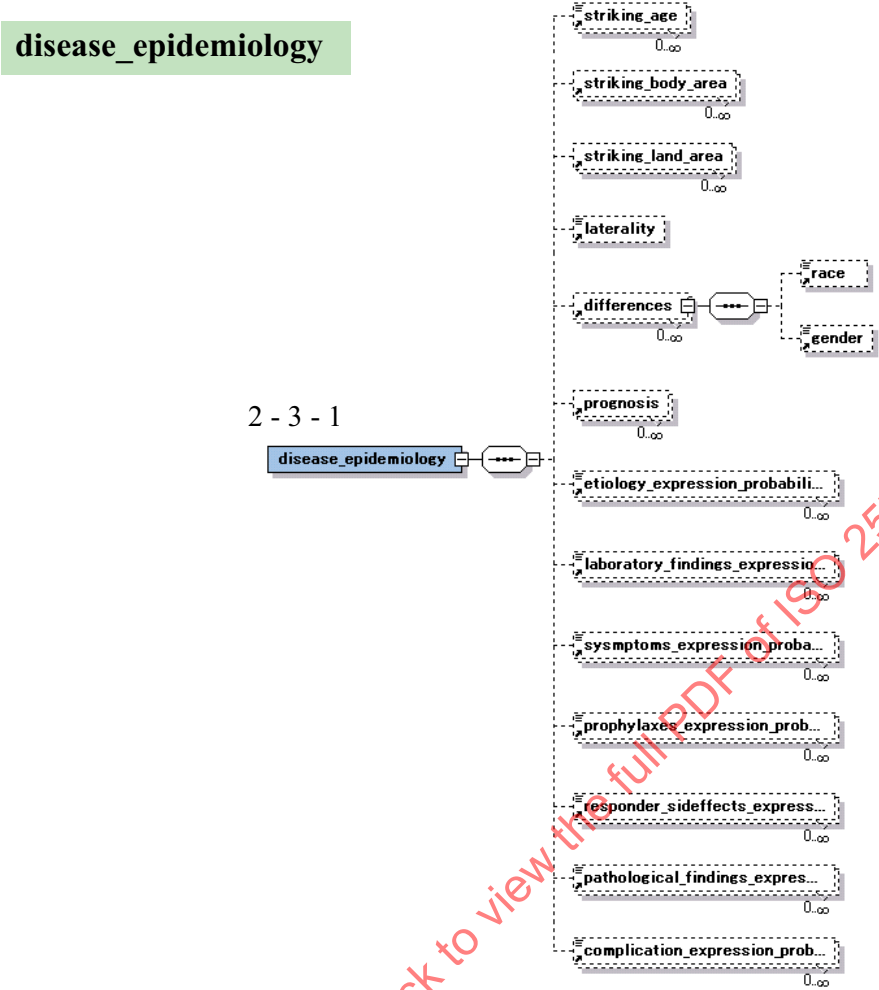


Figure 14 — Detailed structure of GSVML

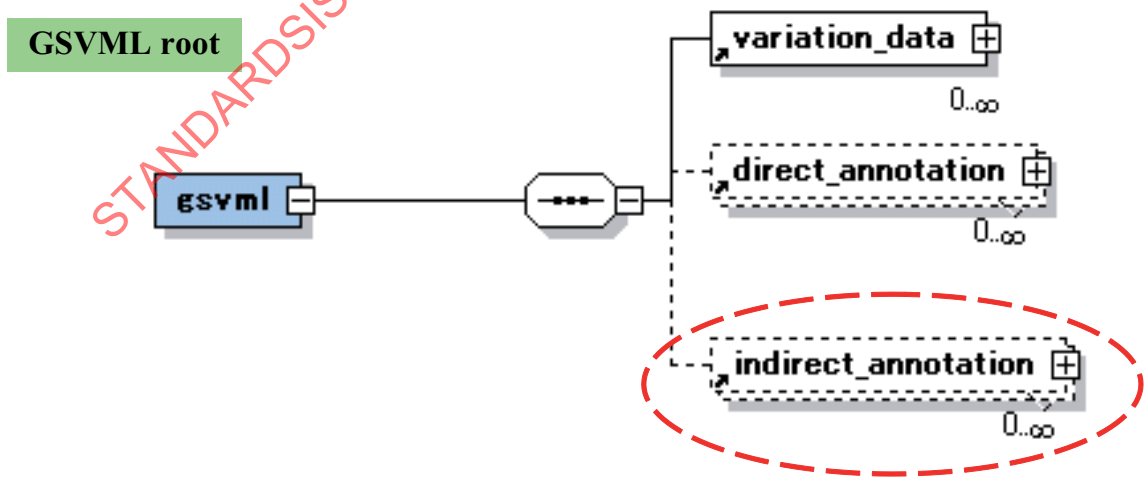


Figure 15 — Detailed structure of GSVML

indirect_annotation

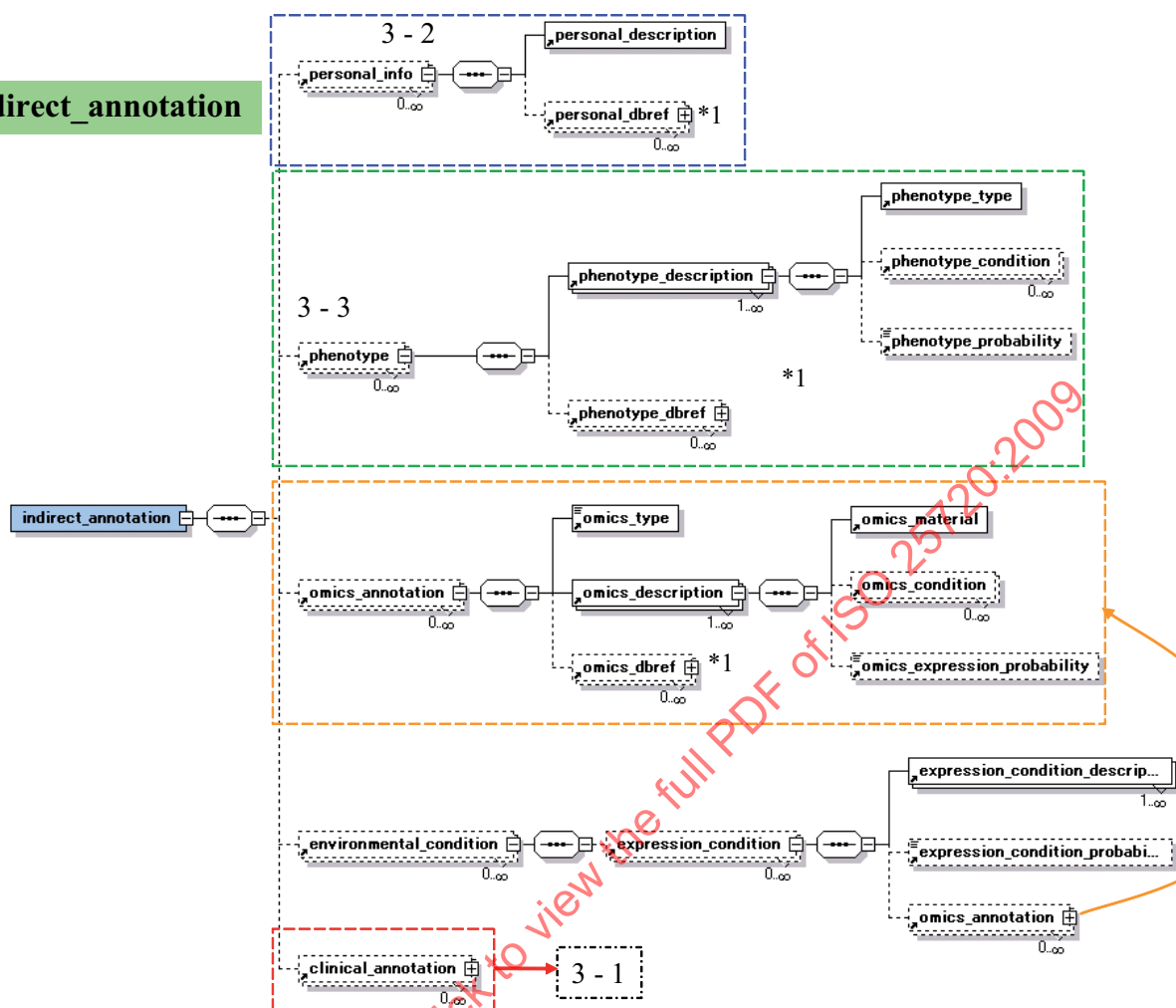


Figure 16 — Detailed structure of GSVML

clinical_annotation

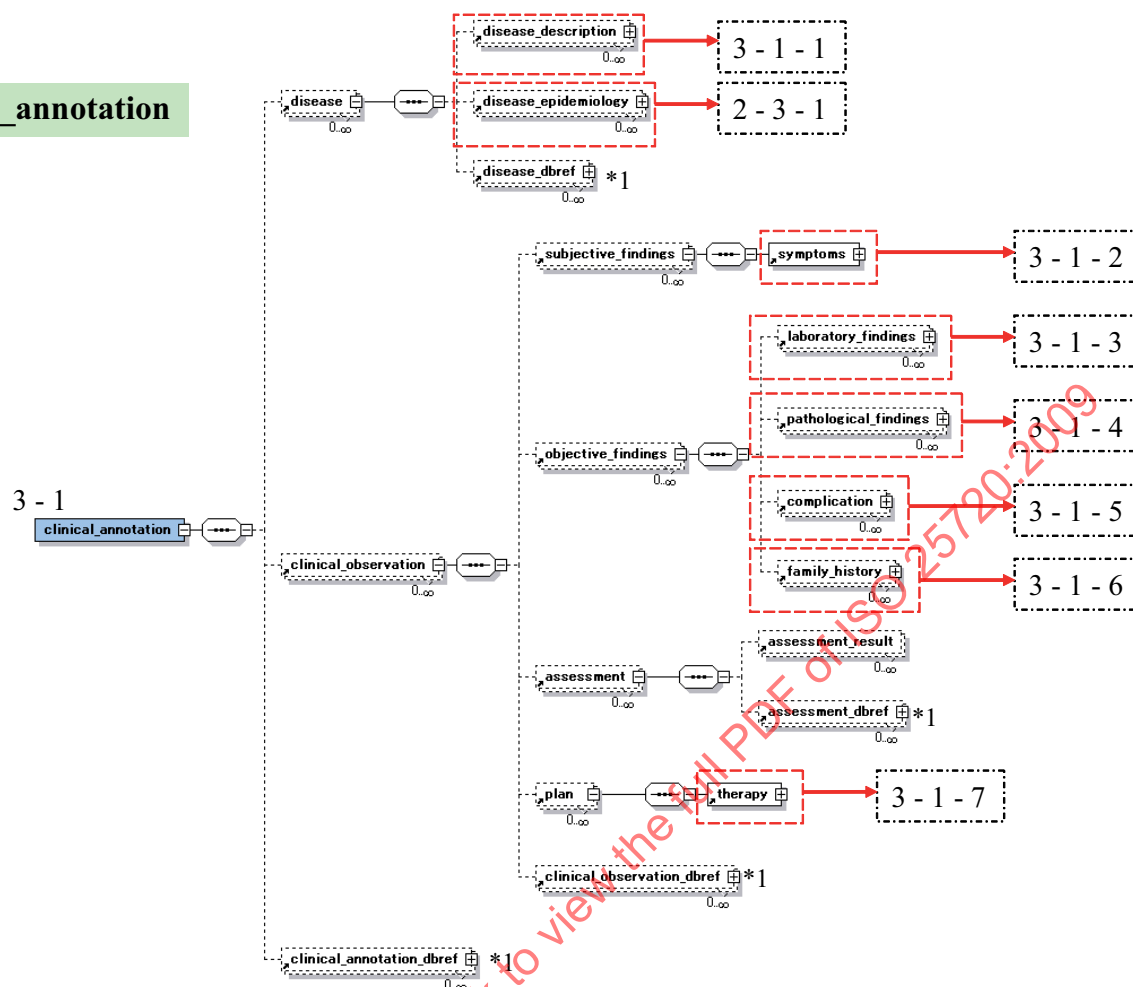


Figure 17 — Detailed structure of GSVML

disease_description

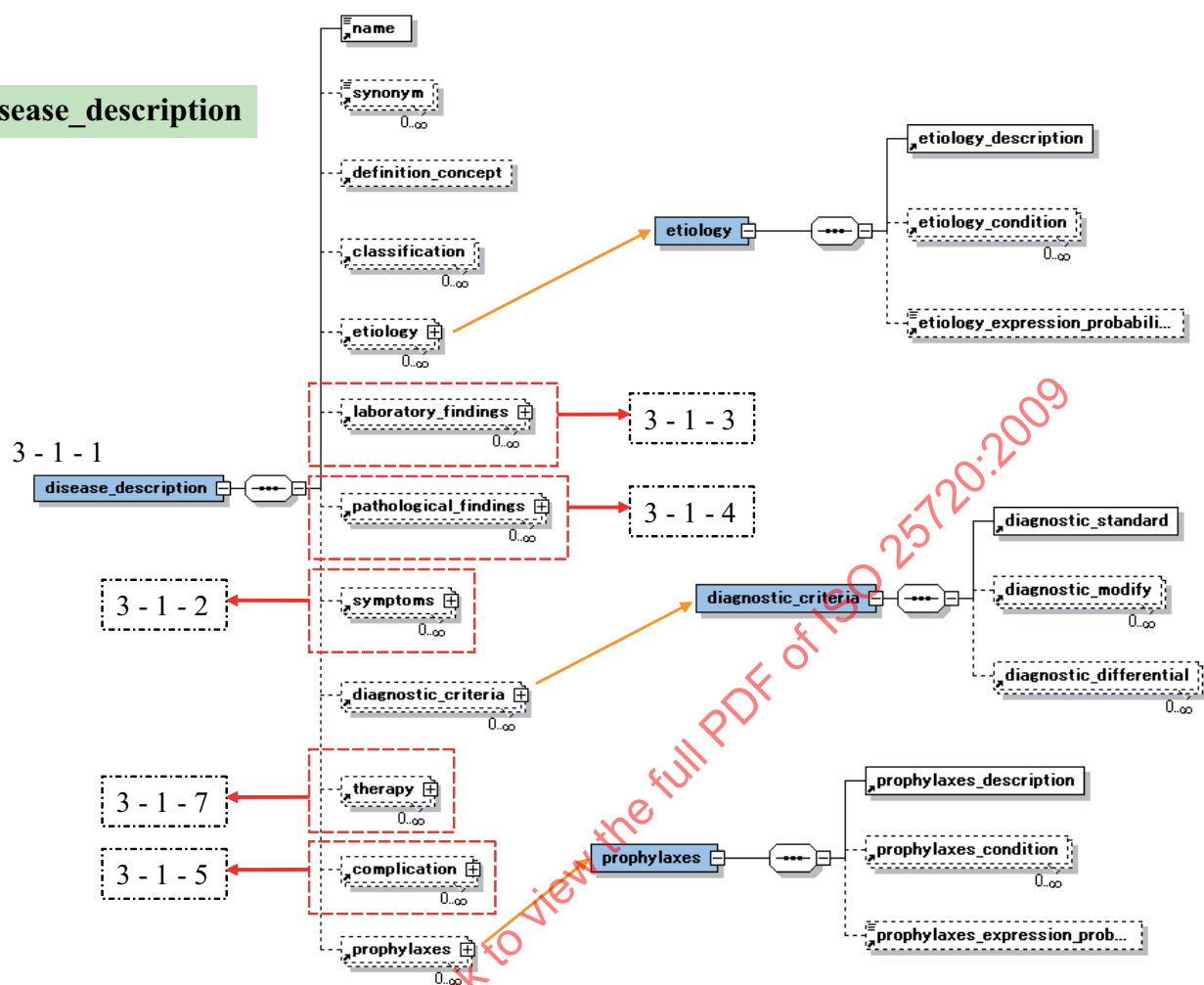


Figure 18 — Detailed structure of GSVML

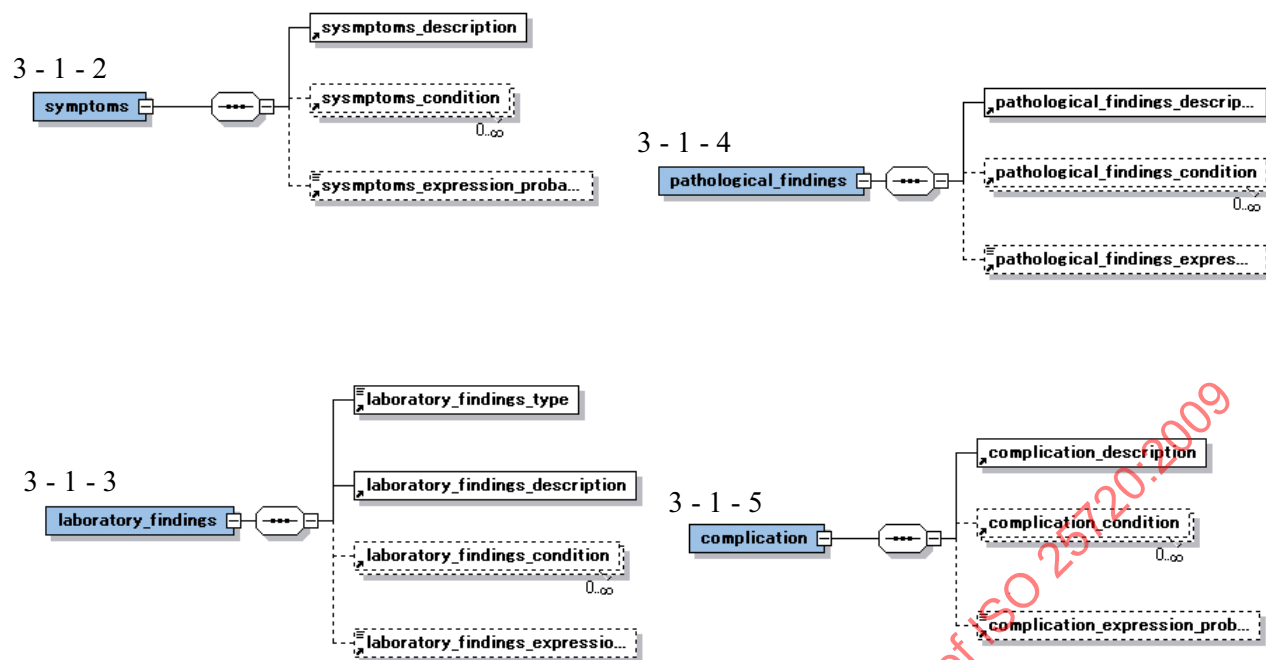


Figure 19 — Detailed structure of GSVML

family_history

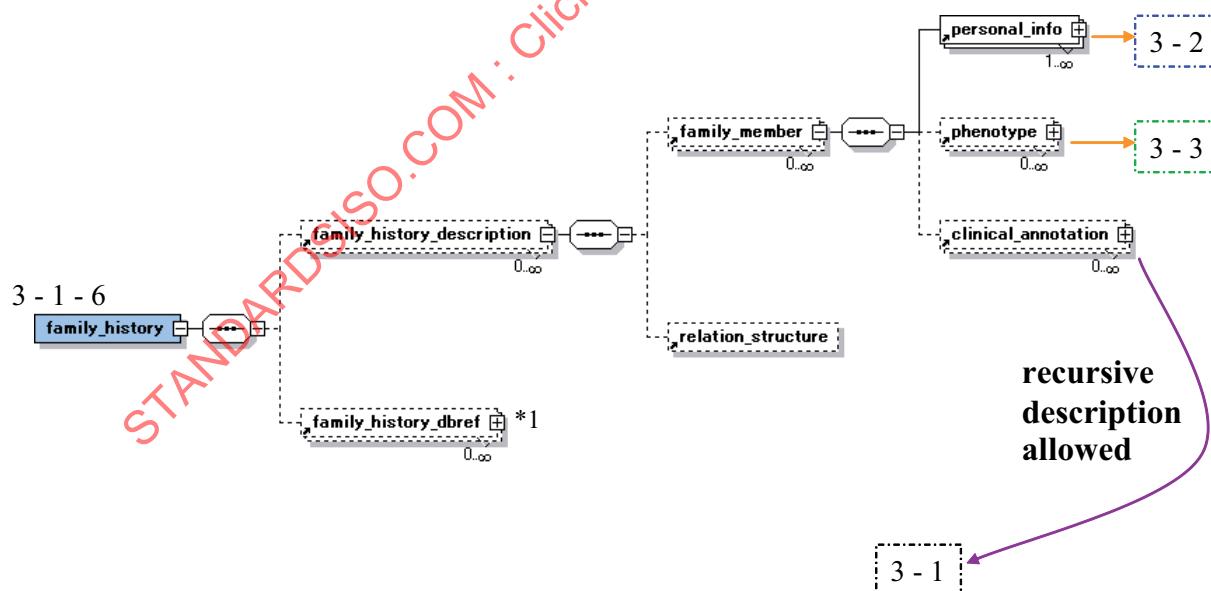


Figure 20 — Detailed structure of GSVML

therapy

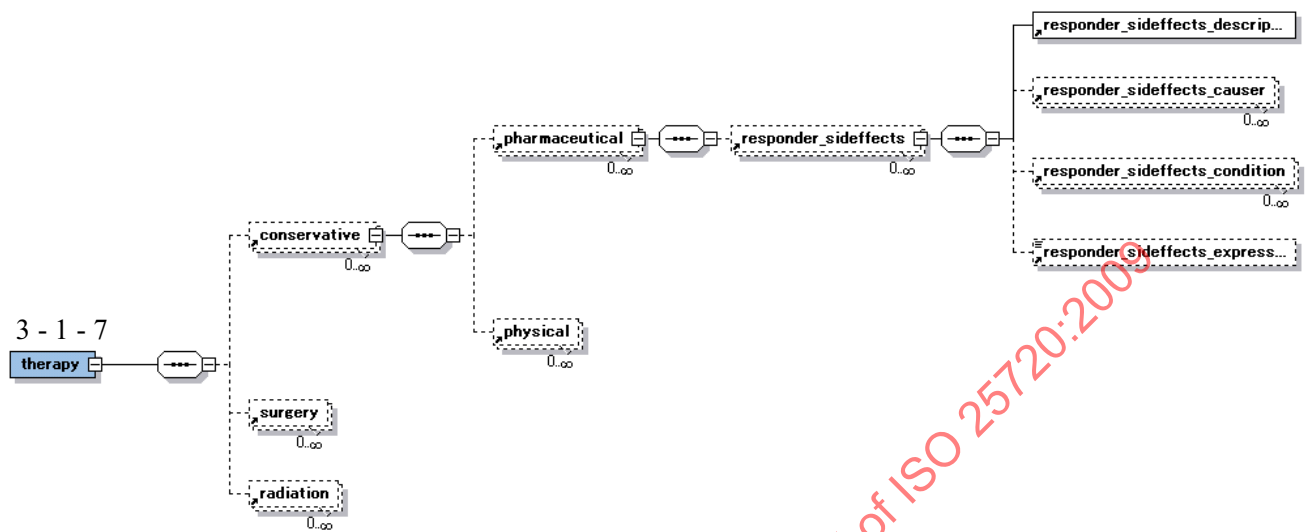


Figure 21 — Detailed structure of GSVML

Annex A (normative)

DTD of GSVML

<?xml version='1.0' encoding='Shift_JIS' ?>

<!-- comment

Jun Nakaya (Information Center for Medical Sciences, Tokyo Medical and Dental University, JAPAN)

"gsvml element: Genomic Sequence Variation Markup Language

variation_data: variation data

direct_annotation: direct annotation of variation data

indirect_annotation: indirect annotation of variation data

"assessment element

assessment_result: results of the assessment

assessment_dbref: database reference of the assessment

"ass_gene element

ass_gene_name: gene name

ass_gene_structure: category of gene structure e.g. exon, intron

aminoacid_substitution: aminoacid substitution generated by variation

codon_substitution: codon substitution generated by variation

codon_position: codon position

ass_gene_symbol: gene symbol

ass_gene_alias: gene alias

ass_gene_product: gene product

ass_gene_evidence_type: gene type e.g. functional gene, predicted EST, computational gene, Pseudogene

changed_motif: motif change exists or not

changed_motif_name: name of motif

changed_splice_site: splice site change exist or not

splice_variant_number: number of splice variant and refSeq

ass_gene_dbref: database reference information

"ass_gene_dbref element

database_name: name of the database

database_id: ID of the database

link_url: URL of the database

database_attbt: attribute of the database

data_id: ID of the datum

data_attbt: attribute of the datum

version: version of the database

"clinical_annotation element

disease: disease information

"clinical_annotation_dbref element

database_name: name of the database

database_id: ID of the database

link_url: URL of the database

database_attbt: attribute of the database

data_id: ID of the datum

data_attbt: attribute of the datum

version: version of the database

"clinical_observation element

subjective_findings: subjective findings in clinical observation

objective_findings: objective findings in clinical observation

assessment: assessment of the clinical observation

plan: plan of the clinical observation

clinical_observation_dbref: database reference of the clinical observation

"clinical_observation_dbref element

database_name: name of the database

database_id: ID of the database

link_url: URL of the database

database_attbt: attribute of the database

data_id: ID of the datum

data_atbt: attribute of the datum

version: version of the database

"complication element

complication_description: description of the complication

complication_condition: condition to express the complication

complication_: probability when the complication is expressed

"conservative element

pharmaceutical: pharmaceutical element

physical: physical treatment element

"diagnostic_criteria element

diagnostic_standard: standard diagnostic criteria

diagnostic_modify: modified diagnostic criteria

diagnostic_differential: differential diagnosis use

"differences element

race: racial difference

gender: gender difference

"direct_annotation element

whole_genome_sequence: whole genome sequence of the datum

mendelian_seggregate: known mendelization

homozygote_detect: homozygote individuals observation in sample

somatic_mutation: known somatic mutation

experimental_analysis: explanation of the experimental and the analysis

epidemiology: epidemiology of the disease and associated gene

var_ann_misc: variation annotation miscellaneous

"disease element

disease_description: description of the disease

disease_epidemiology: epidemiology of the disease

disease_dbref element: database references of the disease

attributes

_id: of the disease

_id: of the submitter

_date: created

_date: modified

"disease_dbref element

database_name: name of the database

database_id: ID of the database

link_url: URL of the database

database_attbt: attribute of the database

data_id: ID of the datum

data_attbt: attribute of the datum

version: version of the database

"disease_description element

name: name of the disease

synonym: synonym of the disease

definition_concept: definition or concept of the disease

classification: classification of the disease

etiology: etiology of the disease

laboratory_findings: element of laboratory findings of the disease

pathological_findings: pathological findings of the disease

symptoms: symptoms of the disease

diagnostic_criteria: criteria of diagnosis for the disease

therapy: therapy of the disease

complication: complications of the disease

prophylaxes: prophylaxes of the disease

"disease_epidemiology element

striking_age: striking age of the disease

striking_body_area: striking body area of the disease

striking_land_area: striking land area of the disease

laterality: laterality of the disease e.g. hemilateral, bilateral

differences: statistical differences

prognosis: prognosis of the disease

etiology_expression_probability: expression probability of the etiology

laboratory_findings_expression_probability: expression probability of the laboratory findings

symptoms_expression_probability: expression probability of the symptom

prophylaxes_expression_probability: expression probability of the prophylaxes

responder_sideeffects_expression_probability: expression probability of the responder, the side effect

pathological_findings_expression_probability: expression probability of the pathological findings

complication_expression_probability: expression probability of the complication

"environmental_condition element

expression_condition: environmental condition of the expression

"epidemiology element

ass_gene: associated gene

disease_epidemiology: epidemiology of the associated disease

population: population of the variation

frequency: frequency of the variation

"etiology element

etiology_description: description of the etiology and its mechanism

etiology_condition: conditions to express the etiology

etiology_expression_probability: expression probability of the etiology

"experiment_analysis element

variation_identify: information to identify the variation

variation_characterize: characterization of the variation datum

"experimental_assay_dbref element

database_name: name of the database

database_id: ID of the database

link_url: URL of the database

database_attr: attribute of the database

data_id: ID of the datum

data_attr: attribute of the datum

version: version of the database

"expression_condition element

expression_condition_description: of the expression condition

expression_condition_probability: of the expression condition

omics_annotation: annotation of the omics

"family_history element

family_history_description: description of the family history

family_history_dbref: database_reference of the family history

"family_history_dbref element

database_name: name of the database

database_id: ID of the database

link_url: URL of the database

database_attbt: attribute of the database

data_id: ID of the datum

data_attbt: attribute of the datum

version: version of the database

"family_history_description element

family_member: member of the family

relation_structure: relational structure of the family

"family_member element

personal_info: personal information of variation data

phenotype: phenotype of the sequence variation

clinical_annotation: clinical annotation of the sequence variation

"frequency element

haplotype: haplotype

haplotype_frequency: frequency of the haplotype

allele: observed allele

allele_frequency: frequency of the observed allele

genotype: genotype

genotype_frequency: frequency of the genotype

attributes

_id: ID of the frequency

_id: ID of the submitter

_id: ID of the population

_id: ID of the assay

_id: ID of the publication

_date: date created

_date: date modified

"indirect_annotation element

person_info: personal information of variation data

phenotype: phenotype of the sequence variation

omics_annotation: annotation type of the omics

environmental_condition: environmental conditions of the expression

clinical_annotation: clinical annotation of the sequence variation

"laboratory_findings element

laboratory_findings_type: type of laboratory findings e.g. urine, blood, ECG, image

laboratory_findings_description: description of the laboratory findings

laboratory_findings_condition: condition of the laboratory findings

laboratory_findings_expression_probability: expression probability of the laboratory findings with its degree

"link_dis_index element

d: d value for Linkage Disequilibrium test

d_prime: d prime for LD test

r-square: r square for LD test

"location element

chromosome_number: the number of the chromosome

position: position of the variation in the chromosome

map: chromosome map on which the variation is

orientation: chromosome orientation on which the variation is

ass_gene: associated gene

location_dbref: database reference of location

"method element

method_name: statistical method name

method_description: description or explanation of the method

method_url: URL of the method

"objective_findings element

laboratory_findings: laboratory findings of the disease

pathological_findings: pathological findings of the disease

complication: complications of the disease

family_history: family history

"omics_annotation element

omics_type element: type of omics

omics_description element: description of the omics

omics_dbref element: database reference of the omics

"pathological_findings element

pathological_findings_description: description of the pathological finding

pathological_findings_condition: condition of the pathological finding

pathological_findings_expression_probability: expression probability of the pathological finding

"pcr element

pcr_confirmed: artifact verification e.g. variation found on repeat PCR sample

pcr_primer: primer sequence

pcr_product: PCR product e.g. single band, multi band

pcr_profile: PCR profile

"personal_info element

personal_description: description of the personal information

personal_dbref: database reference of the personal information

"pharmaceutical element

responder_sideeffects: responder and/or side effects

"phenotype element

phenotype_description: description of the phenotype

phenotype_dbref: database references of the phenotype

attributes

_id: ID of the phenotype

_id: ID of the submitter

_date: date created

_date: date modified

"phenotype_dbref element

database_name: name of the database

database_id: ID of the database

link_url: URL of the database

database_attr: attribute of the database

data_id: ID of the datum

data_attr: attribute of the datum

version: version of the database

"phenotype_description element

phenotype_type: type of the phenotype

phenotype_condition: condition to express the phenotype

phenotype_probability: probability to express the phenotype on the conditions

"plan element

therapy: therapy of the disease

"population element

population_description: description of the population

organism: organism

differences: statistical differences

population_parameter: parameter of population

sample_size: sample size of population

population_misc: population miscellaneous

population_dbref: database references of the population

attributes

_id: ID of the population

_id: ID of the submitter

_date: date created

_date: date modified

"population_dbref element

database_name: name of the database

database_id: ID of the database

link_url: URL of the database

database_attbt: attribute of the database

data_id: ID of the datum

data_attbt: attribute of the datum

version: version of the database

"prophylaxes element

prophylaxes_description: description of the prophylaxis

prophylaxes_condition: condition of the prophylaxis

prophylaxes_expression_probability: expression probability of the prophylaxis

"publication element

title: title of the publication

author: author of the publication

journal: journal of the publication

volume: volume of the publication

supplement: supplement of the publication

issue: issue of the publication

issue_supplement: issue supplement of the publication

pages: page of the publication

year: year of the publication

publication_status: status of the publication

mesh_term: mesh term of the publication

publication_dbref: database references of the publications

submitter: submitter of the publication

attributes

_id: ID of the publication

_id: ID of the submitter

_date: date created

_date: date modified

"publication_dbref element

database_name: name of the database

database_id: ID of the database

link_url: URL of the database

database_attbt: attribute of the database

data_id: ID of the datum

data_attbt: attribute of the datum

version: version of the database

"responder_sideeffects element

responder_sideeffects_causer: causer of the responder and/or side effect

responder_sideeffects_description: description of the responder and/or side effect

responder_sideeffects_condition: condition of the responder and/or side effect

responder_sideeffects_expression_probability: expression probability of the responder and/or side effect

"source element

source_release_date: date released

source_modify_date: date modified

source_rawdata: rawdatum of the source

source_dbref: database reference of source

"splice_variant_number element

refSeq_number: reference sequence number

"submitter element

submitter_name: name of the submitter

address: address of the submitter

email: email of the submitter

tel: telephone of the submitter

fax: Fax of the submitter

institution: Institution of the submitter

laboratory: Laboratory of the submitter

submitter_dbref: database references of the submitter

publication: publication of the experiment

"subjective_findings element

symptoms: laboratory findings of the disease

"submitter_dbref element

database_name: name of the database

database_id: ID of the database

link_url: URL of the database

database_attbt: attribute of the database

data_id: ID of the datum

data_attbt: attribute of the datum

version: version of the database

"symptoms element

symptoms_description: description of the symptom

symptoms_condition: condition of the symptom

symptoms_expression_probability: expression probability of the symptom with its degree (table)

"therapy element

conservative: conservative treatment

surgery: surgical treatment

radiation: radiation therapy

"var_ann_misc element

var_ann_misc_description: description of the variation annotation miscellaneous

var_ann_misc_dbref: database references of the variation annotation miscellaneous

attributes

_ann_misc_id: ID of the variation annotation miscellaneous

_id: ID of the submitter

_date: date created

_date: date modified

"var_ann_misc_dbref element

database_name: name of the database

database_id: ID of the database

link_url: URL of the database

database_attbt: attribute of the database

data_id: ID of the datum

data_attbt: attribute of the datum

version: version of the database

"variation_att element

molecular_type: type of molecule e.g. DNA, RNA

allele: observed allele

length: sequence length including flanking sequence

f5sequence: 5' flanking sequence

f3sequence: 3' flanking sequence

validation_status: status of validation as (Proven, Suspected)

success_rate: certainty of variation information

"variation_characterize element

method: statistical method

p-value: p value for significance (Association study)

link_dis_index: linkage disequilibrium index for LD test

"variation_data element

variation_type: type of variation

location: location of the variation

variation_att: attribute information of the variation

source: source of the sequence variation

variation_dbref: database references of the variation

"variation_dbref element

database_name: name of the database

database_id: ID of the database

link_url: URL of the database

database_attbt: attribute of the database

data_id: ID of the datum

data_attbt: attribute of the datum

version: version of the database

"variation_identify element

experimental_assay_result: result of the experimental assay

experimental_assay_description: description of the experimental assay

experimental_assay_parameter: parameter of the experimental assay

experimental_assay_dbref: database reference information

publication: publication of the experiment

submitter: submitter of the publication

pcr: PCR

-->

<!ELEMENT gsvml (variation_data , direct_annotation* , indirect_annotation*)>

<!ELEMENT variation_data (variation_type, location, variation_att+, source*, variation_dbref*)>

<!-- variation_type comment

scope of variation_type are: SNP, rSNP, cSNP, iSNP, uSNP, gSNP, RFLP, MS, STRP, VNTR, Insertion, Deletion, Sustitution, Other

-->

<!ELEMENT variation_type (#PCDATA)>

<!ELEMENT location (chromosome_number , position , map? , orientation , ass_gene* , location_dbref*)>

<!ELEMENT chromosome_number (#PCDATA)>

<!ELEMENT position (#PCDATA)>

<!ELEMENT map (#PCDATA)>

<!ELEMENT orientation (#PCDATA)>

<!ELEMENT ass_gene (ass_gene_name , ass_gene_structure? , aminoacid_substitution? , codon_substitution? , codon_position? ,

ass_gene_symbol? , ass_gene_alias? , ass_gene_product? , ass_gene_evidence_type* , changed_motif* ,
changed_motif_name* ,

changed_splice_site* , splice_variant_number? , ass_gene_dbref*)>

<!ELEMENT ass_gene_name (#PCDATA)>

<!ELEMENT ass_gene_structure (#PCDATA)>

<!ELEMENT aminoacid_substitution (#PCDATA)>

<!ELEMENT codon_substitution (#PCDATA)>

<!ELEMENT codon_position (#PCDATA)>

<!ELEMENT ass_gene_symbol (#PCDATA)>

<!ELEMENT ass_gene_alias (#PCDATA)>

<!ELEMENT ass_gene_product (#PCDATA)>

<!ELEMENT ass_gene_evidence_type (#PCDATA)>

<!ELEMENT changed_motif (#PCDATA)>

<!ELEMENT changed_motif_name (#PCDATA)>

<!ELEMENT changed_splice_site (#PCDATA)>

<!ELEMENT splice_variant_number (refSeq_number+)>

<!ELEMENT refSeq_number (#PCDATA)>

<!ELEMENT ass_gene_dbref(database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

<!ELEMENT database_name (#PCDATA)>

<!ELEMENT database_id (#PCDATA)>

<!ELEMENT link_url (#PCDATA)>

<!ELEMENT database_attbt (#PCDATA)>

<!ELEMENT data_id (#PCDATA)>

<!ELEMENT data_attbt (#PCDATA)>

<!ELEMENT version (#PCDATA)>

<!ELEMENT location_dbref(database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

<!ELEMENT variation_att(molecular_type , allele* , length , f5sequence? , f3sequence? , validation_status? , success_rate?)>

<!ATTLIST variation

variation_id CDATA #REQUIRED

submitter_id CDATA #REQUIRED

population_id CDATA #REQUIRED

experimental_assay_id CDATA #REQUIRED

publication_id CDATA #REQUIRED

create_date CDATA #IMPLIED

modify_date CDATA #IMPLIED

```

>

<!ELEMENT molecular_type (#PCDATA)>

<!ELEMENT allele (#PCDATA)>

<!ELEMENT length (#PCDATA)>

<!ELEMENT f5sequence (#PCDATA)>

<!ELEMENT f3sequence (#PCDATA)>

<!-- validation_status comment

scope of validation_status are: Proven, Suspected

-->

<!ELEMENT validation_status (#PCDATA)>

<!ELEMENT success_rate (#PCDATA)>

<!ELEMENT source (source_release_date? , source_modify_date* , source_rawdata* , source_dbref+)>

<!ELEMENT source_release_date (#PCDATA)>

<!ELEMENT source_modify_date (#PCDATA)>

<!ELEMENT source_rawdata (#PCDATA)>

<!ELEMENT source_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

<!ELEMENT variation_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

<!ELEMENT direct_annotation (whole_genome_sequence* , mendelian_seggregate* , homozygote_detect* ,
                             somatic_mutation* , experiment_analysis* , epidemiology* , var_ann_misc*)>

<!ELEMENT whole_genome_sequence (#PCDATA)>

<!ELEMENT mendelian_seggregate (#PCDATA)>

<!-- validation_status comment

scope of homozygote_detect are: homo, hetero, unknown

-->

<!ELEMENT homozygote_detect (#PCDATA)>

<!ELEMENT somatic_mutation (#PCDATA)>

<!ELEMENT experiment_analysis (variation_identify* , variation_characterize*)>

<!ELEMENT variation_identify (experimental_assay_result , experimental_assay_description , experimental_assay_parameter+ ,
                             experimental_assay_dbref* , (publication* | submitter*) , pcr*)>

```

<!ATTLIST variation_identify

experimental_assay_id CDATA #REQUIRED

submitter_id CDATA #REQUIRED

create_date CDATA #IMPLIED

modify_date CDATA #IMPLIED

>

<!ELEMENT experimental_assay_result (#PCDATA)>

<!ELEMENT experimental_assay_description (#PCDATA)>

<!ELEMENT experimental_assay_parameter (#PCDATA)>

<!ELEMENT experimental_assay_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

<!ELEMENT publication (title? , author? , journal? , volume? , supplement? , issue? , issue_supplement? ,
pages? , year? , publication_status? , mesh_term? , publication_dbref* , submitter?)>

<!ATTLIST publication

publication_id CDATA #REQUIRED

submitter_id CDATA #REQUIRED

create_date CDATA #IMPLIED

modify_date CDATA #IMPLIED

>

<!ELEMENT title (#PCDATA)>

<!ELEMENT author (#PCDATA)>

<!ELEMENT journal (#PCDATA)>

<!ELEMENT volume (#PCDATA)>

<!ELEMENT supplement (#PCDATA)>

<!ELEMENT issue (#PCDATA)>

<!ELEMENT issue_supplement (#PCDATA)>

<!ELEMENT pages (#PCDATA)>

<!ELEMENT year (#PCDATA)>

<!ELEMENT publication_status (#PCDATA)>

<!ELEMENT mesh_term (#PCDATA)>

<!ELEMENT publication_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

```

<!ELEMENT submitter (submitter_name , address? , email? , tel? , fax? , institution? , laboratory? , submitter_dbref* , publication)>

<!-- submitter

    submitter_id CDATA #REQUIRED

    create_date CDATA #IMPLIED

    modify_date CDATA #IMPLIED

-->

<!ELEMENT submitter_name (#PCDATA)>

<!ELEMENT address (#PCDATA)>

<!ELEMENT email (#PCDATA)>

<!ELEMENT tel (#PCDATA)>

<!ELEMENT fax (#PCDATA)>

<!ELEMENT institution (#PCDATA)>

<!ELEMENT laboratory (#PCDATA)>

<!ELEMENT submitter_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

<!ELEMENT pcr (pcr_confirmed? , pcr_primer? , pcr_product? , pcr_profile?)>

<!ELEMENT pcr_confirmed (#PCDATA)>

<!ELEMENT pcr_primer (#PCDATA)>

<!ELEMENT pcr_product (#PCDATA)>

<!ELEMENT pcr_profile (#PCDATA)>

<!ELEMENT variation_characterize (genetic_statistics)>

<!ELEMENT genetic_statistics (method+ , p-value* , link_dis_index* , descend_identity* , maximum_lod_score*)>

<!ELEMENT method (method_name , method_description* , method_url*)>

<!-- method_name comment

scope of method_name are: sib pair, linkage disequilibrium, association study and any other string

-->

<!ELEMENT method_name (#PCDATA)>

<!ELEMENT method_description (#PCDATA)>

<!ELEMENT method_url (#PCDATA)>

<!ELEMENT p-value (#PCDATA)>

<!ELEMENT link_dis_index (d? , d_prime? , r_square?)>

```

<!ELEMENT d (#PCDATA)>

<!ELEMENT d_prime (#PCDATA)>

<!ELEMENT r_square (#PCDATA)>

<!ELEMENT descend_identity (di_value , di_probability)>

<!ELEMENT di_value (#PCDATA)>

<!ELEMENT di_probability (#PCDATA)>

<!ELEMENT maximum_lod_score (#PCDATA)>

<!ELEMENT epidemiology (ass_gene* , disease_epidemiology* , population* , frequency*)>

<!ELEMENT population (population_description , organism? , differences* , population_parameter* , sample_size , population_misc* , population_dbref*)>

<!ATTLIST population

population_id CDATA #REQUIRED

submitter_id CDATA #REQUIRED

create_date CDATA #IMPLIED

modify_date CDATA #IMPLIED

>

<!ELEMENT population_description (#PCDATA)>

<!ELEMENT organism (#PCDATA)>

<!ELEMENT differences (race? , gender*)>

<!ELEMENT race (#PCDATA)>

<!ELEMENT gender (#PCDATA)>

<!ELEMENT population_parameter (#PCDATA)>

<!ELEMENT sample_size (#PCDATA)>

<!ELEMENT population_misc (#PCDATA)>

<!ELEMENT population_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version*)>

<!ELEMENT frequency (haplotype? , haplotype_frequency? , allele? , allele_frequency? , genotype? , genotype_frequency*)>

<!ATTLIST frequency

frequency_id CDATA #REQUIRED

submitter_id CDATA #REQUIRED

population_id CDATA #REQUIRED

```

    assay_id    CDATA #REQUIRED

    publication_id CDATA #REQUIRED

    create_date  CDATA #IMPLIED

    modify_date  CDATA #IMPLIED

>

<!ELEMENT haplotype (#PCDATA)>

<!ELEMENT haplotype_frequency (#PCDATA)>

<!ELEMENT allele_frequency (#PCDATA)>

<!ELEMENT genotype (#PCDATA)>

<!ELEMENT genotype_frequency (#PCDATA)>

<!ELEMENT var_ann_misc (var_ann_misc_description , var_ann_misc_dbref*)>

!ATTLIST var_ann_misc

    var_ann_misc_id    CDATA #REQUIRED

    submitter_id    CDATA #REQUIRED

    create_date    CDATA #IMPLIED

    modify_date    CDATA #IMPLIED

>

<!ELEMENT var_ann_misc_description (#PCDATA)>

<!ELEMENT var_ann_misc_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

<!ELEMENT indirect_annotation (personal_info* , phenotype* , omics_annotation* , environmental_condition* ,
clinical_annotation*)>

<!ELEMENT personal_info (personal_description , personal_dbref*)>

<!ELEMENT personal_description (#PCDATA)>

<!ELEMENT personal_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

<!ELEMENT phenotype (phenotype_description+ , phenotype_dbref*)>

<!ATTLIST phenotype

    phenotype_id CDATA #REQUIRED

    submitter_id CDATA #REQUIRED

    create_date CDATA #IMPLIED

    modify_date CDATA #IMPLIED

```

```

>

<!ELEMENT phenotype_description (phenotype_type , phenotype_condition* , phenotype_probability?)>

<!ELEMENT phenotype_type (#PCDATA)>

<!ELEMENT phenotype_condition (#PCDATA)>

<!ELEMENT phenotype_probability (#PCDATA)>

<!ELEMENT phenotype_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

<!ELEMENT omics_annotation (omics_type , omics_description+ , omics_dbref*)>

<!-- omics_type element comment

scope of omics_type are: transcriptomics, proteomics, metabolomics, signalomics, organomics and any other string

-->

<!ELEMENT omics_type (#PCDATA)>

<!ELEMENT omics_description (omics_material, omics_condition* , omics_expression_probability?)>

<!ELEMENT omics_material (#PCDATA)>

<!ELEMENT omics_condition (#PCDATA)>

<!ELEMENT omics_expression_probability (#PCDATA)>

<!ELEMENT omics_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

<!ELEMENT environmental_condition (expression_condition*)>

<!ELEMENT expression_condition (expression_condition_description+ , expression_condition_probability? , omics_annotation*)>

<!ELEMENT expression_condition_description (#PCDATA)>

<!ELEMENT expression_condition_probability (#PCDATA)>

<!ELEMENT clinical_annotation (disease* , clinical_observation* , clinical_annotation_dbref*)>

<!ELEMENT disease (disease_description* , disease_epidemiology* , disease_dbref*)>

<!-- ATTLIST disease

disease_id CDATA #REQUIRED

submitter_id CDATA #REQUIRED

create_date CDATA #IMPLIED

modify_date CDATA #IMPLIED

-->

<!ELEMENT disease_description (name , synonym* , definition_concept? , classification* , etiology* , laboratory_findings* ,
pathological_findings* ,

symptoms* , diagnostic_criteria* , therapy* , complication* , prophylaxes*)>

```

<!ELEMENT name (#PCDATA)>

<!ELEMENT synonym (#PCDATA)>

<!ELEMENT definition_concept (#PCDATA)>

<!ELEMENT classification (#PCDATA)>

<!ELEMENT etiology (etiology_description , etiology_condition* , etiology_expression_probability?)>

<!ELEMENT etiology_description (#PCDATA)>

<!ELEMENT etiology_condition (#PCDATA)>

<!ELEMENT etiology_expression_probability (#PCDATA)>

<!ELEMENT laboratory_findings (laboratory_findings_type , laboratory_findings_description , laboratory_findings_condition* , laboratory_findings_expression_probability?)>

<!ELEMENT laboratory_findings_type (#PCDATA)>

<!ELEMENT laboratory_findings_description (#PCDATA)>

<!ELEMENT laboratory_findings_condition (#PCDATA)>

<!ELEMENT laboratory_findings_expression_probability (#PCDATA)>

<!ELEMENT pathological_findings (pathological_findings_description , pathological_findings_condition* , pathological_findings_expression_probability?)>

<!ELEMENT pathological_findings_description (#PCDATA)>

<!ELEMENT pathological_findings_condition (#PCDATA)>

<!ELEMENT pathological_findings_expression_probability (#PCDATA)>

<!ELEMENT symptoms (symptoms_description , symptoms_condition* , symptoms_expression_probability?)>

<!ELEMENT symptoms_description (#PCDATA)>

<!ELEMENT symptoms_condition (#PCDATA)>

<!ELEMENT symptoms_expression_probability (#PCDATA)>

<!ELEMENT diagnostic_criteria (diagnostic_standard , diagnostic_modify* , diagnostic_differential*)>

<!ELEMENT diagnostic_modify (#PCDATA)>

<!ELEMENT diagnostic_standard (#PCDATA)>

<!ELEMENT diagnostic_differential (#PCDATA)>

<!ELEMENT therapy (conservative* , surgery* , radiation*)>

<!ELEMENT conservative (pharmaceutical* , physical*)>

<!ELEMENT pharmaceutical (responder_sideeffects*)>

<!ELEMENT responder_sideeffects (responder_sideeffects_description , responder_sideeffects_causer* , responder_sideeffects_condition* , responder_sideeffects_expression_probability?)>

<!ELEMENT responder_sideeffects_causer (#PCDATA)>

<!ELEMENT responder_sideeffects_description (#PCDATA)>

<!ELEMENT responder_sideeffects_condition (#PCDATA)>

<!ELEMENT responder_sideeffects_expression_probability (#PCDATA)>

<!ELEMENT physical (#PCDATA)>

<!ELEMENT surgery (#PCDATA)>

<!ELEMENT radiation (#PCDATA)>

<!ELEMENT complication (complication_description , complication_condition* , complication_expression_probability?)>

<!ELEMENT complication_description (#PCDATA)>

<!ELEMENT complication_condition (#PCDATA)>

<!ELEMENT complication_expression_probability (#PCDATA)>

<!ELEMENT prophylaxes (prophylaxes_description , prophylaxes_condition* , prophylaxes_expression_probability?)>

<!ELEMENT prophylaxes_description (#PCDATA)>

<!ELEMENT prophylaxes_condition (#PCDATA)>

<!ELEMENT prophylaxes_expression_probability (#PCDATA)>

<!ELEMENT disease_epidemiology (striking_age* , striking_body_area* , striking_land_area* , laterality? , differences* , prognosis* , etiology_expression_probability* ,

laboratory_findings_expression_probability* , symptoms_expression_probability* ,

prophylaxes_expression_probability* ,

responder_sideeffects_expression_probability* , pathological_findings_expression_probability* ,

complication_expression_probability*)>

<!ELEMENT striking_age (#PCDATA)>

<!ELEMENT striking_body_area (#PCDATA)>

<!ELEMENT striking_land_area (#PCDATA)>

<!-- laterality element comment

scope of laterality are: hemi, bi, right, left, other

-->

<!ELEMENT laterality (#PCDATA)>

<!ELEMENT prognosis (#PCDATA)>

<!ELEMENT disease_dbref (database_name , database_id? , link_url? , database_attrib? , data_id , data_attrib? , version?)>

<!ELEMENT clinical_observation (subjective_findings* , objective_findings* , assessment* , plan* , clinical_observation_dbref*)>

<!ELEMENT subjective_findings (symptoms)>

<!ELEMENT objective_findings (laboratory_findings* , pathological_findings* , complication* , family_history*)>

<!ELEMENT family_history (family_history_description* , family_history_dbref*)>

<!ELEMENT family_history_description (family_member* , relation_structure?)>

<!ELEMENT family_member (personal_info+ , phenotype* , clinical_annotation*)>

<!ELEMENT relation_structure (#PCDATA)>

<!ELEMENT family_history_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

<!ELEMENT assessment (assessment_result* , assessment_dbref*)>

<!ELEMENT assessment_result (#PCDATA)>

<!ELEMENT assessment_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

<!ELEMENT plan (therapy)>

<!ELEMENT clinical_observation_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

<!ELEMENT clinical_annotation_dbref (database_name , database_id? , link_url? , database_attbt? , data_id , data_attbt? , version?)>

Annex B (normative)

XML schema of GSVML

```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

  <!--===== root element =====>

  <xs:element name="gsvml">

    <xs:annotation>

      <xs:documentation>Genomic Sequence Variation Markup Language</xs:documentation>

    </xs:annotation>

    <xs:complexType>

      <xs:sequence>

        <xs:element ref="variation_data" minOccurs="0" maxOccurs="unbounded"/>

        <xs:element ref="direct_annotation" minOccurs="0" maxOccurs="unbounded"/>

        <xs:element ref="indirect_annotation" minOccurs="0" maxOccurs="unbounded"/>

      </xs:sequence>

    </xs:complexType>

  </xs:element>

  <!--===== variation data =====>

  <xs:element name="variation_data">

    <xs:annotation>

      <xs:documentation>variation data</xs:documentation>

    </xs:annotation>

    <xs:complexType>

      <xs:sequence>

        <xs:element ref="variation_type"/>

        <xs:element ref="location"/>

        <xs:element ref="variation_att" maxOccurs="unbounded"/>

        <xs:element ref="source" minOccurs="0" maxOccurs="unbounded"/>

      </xs:sequence>

    </xs:complexType>

  </xs:element>

</xs:schema>
```

```

    <xs:element ref="variation_dbref" minOccurs="0" maxOccurs="unbounded"/>

  </xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="variation_type">

  <xs:annotation>

    <xs:documentation>type of variation</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

    <xs:restriction base="xs:string">

      <xs:enumeration value="SNP"/>

      <xs:enumeration value="rSNP"/>

      <xs:enumeration value="cSNP"/>

      <xs:enumeration value="iSNP"/>

      <xs:enumeration value="uSNP"/>

      <xs:enumeration value="gSNP"/>

      <xs:enumeration value="RFLP"/>

      <xs:enumeration value="MS"/>

      <xs:enumeration value="STRP"/>

      <xs:enumeration value="VNTR"/>

      <xs:enumeration value="Insertion"/>

      <xs:enumeration value="Deletion"/>

      <xs:enumeration value="Sustitution"/>

      <xs:enumeration value="Other"/>

      <xs:enumeration value=""/>

    </xs:restriction>

  </xs:simpleType>

</xs:element>

<xs:element name="location">

  <xs:annotation>

```

```

    <xs:documentation>location of the variation</xs:documentation>

</xs:annotation>

<xs:complexType>

  <xs:sequence>

    <xs:element ref="chromosome_number"/>

    <xs:element ref="position"/>

    <xs:element ref="map" minOccurs="0"/>

    <xs:element ref="orientation"/>

    <xs:element ref="ass_gene" minOccurs="0" maxOccurs="unbounded"/>

    <xs:element ref="location_dbref" minOccurs="0" maxOccurs="unbounded"/>

  </xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="chromosome_number" type="xs:string">

  <xs:annotation>

    <xs:documentation>the number of the chromosome</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="position" type="xs:string">

  <xs:annotation>

    <xs:documentation>position of the variation in the chromosome</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="map">

  <xs:annotation>

    <xs:documentation>chromosome map on which the variation is</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="orientation">

  <xs:annotation>

```

```

    <xs:documentation>chromosome orientation on which the variation is</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="ass_gene">

  <xs:annotation>

    <xs:documentation>associated gene</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="ass_gene_name"/>

      <xs:element ref="ass_gene_structure" minOccurs="0"/>

      <xs:element ref="aminoacid_substitution" minOccurs="0"/>

      <xs:element ref="codon_substitution" minOccurs="0"/>

      <xs:element ref="codon_position" minOccurs="0"/>

      <xs:element ref="ass_gene_symbol" minOccurs="0"/>

      <xs:element ref="ass_gene_alias" minOccurs="0"/>

      <xs:element ref="ass_gene_product" minOccurs="0"/>

      <xs:element ref="ass_gene_evidence_type" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="changed_motif" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="changed_motif_name" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="changed_splice_site" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="splice_variant_number" minOccurs="0"/>

      <xs:element ref="ass_gene_dbref" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="ass_gene_name" type="xs:string">

  <xs:annotation>

    <xs:documentation>gene name</xs:documentation>

  </xs:annotation>

```

</xs:element>

<xs:element name="ass_gene_structure">

<xs:annotation>

<xs:documentation>category of gene structure e.g. exon, intron</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="aminoacid_substitution">

<xs:annotation>

<xs:documentation>aminoacid sustitution generated by variation</xs:documentation>

</xs:annotation>

<xs:simpleType>

<xs:restriction base="xs:string">

<xs:enumeration value="A"/>

<xs:enumeration value="R"/>

<xs:enumeration value="N"/>

<xs:enumeration value="D"/>

<xs:enumeration value="C"/>

<xs:enumeration value="Q"/>

<xs:enumeration value="E"/>

<xs:enumeration value="G"/>

<xs:enumeration value="H"/>

<xs:enumeration value="I"/>

<xs:enumeration value="L"/>

<xs:enumeration value="K"/>

<xs:enumeration value="M"/>

<xs:enumeration value="F"/>

<xs:enumeration value="P"/>

<xs:enumeration value="S"/>

<xs:enumeration value="W"/>

<xs:enumeration value="T"/>

```

    <xs:enumeration value="Y"/>

    <xs:enumeration value="V"/>

  </xs:restriction>

</xs:simpleType>

</xs:element>

<xs:element name="codon_substitution">

  <xs:annotation>

    <xs:documentation>codon substitution generated by variation</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

    <xs:restriction base="xs:string">

      <xs:enumeration value="A"/>

      <xs:enumeration value="T"/>

      <xs:enumeration value="G"/>

      <xs:enumeration value="C"/>

    </xs:restriction>

  </xs:simpleType>

</xs:element>

<xs:element name="codon_position" type="xs:string">

  <xs:annotation>

    <xs:documentation>codon position</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="ass_gene_symbol" type="xs:string">

  <xs:annotation>

    <xs:documentation>gene symbol</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="ass_gene_alias" type="xs:string">

  <xs:annotation>

```

```

    <xs:documentation>gene alias</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="ass_gene_product" type="xs:string">

  <xs:annotation>

    <xs:documentation>gene product</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="ass_gene_evidence_type" type="xs:string">

  <xs:annotation>

    <xs:documentation>gene type e.g. functional gene, predicted EST, computational gene, Pseudogene</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="changed_motif" type="xs:boolean">

  <xs:annotation>

    <xs:documentation>motif change exists or not</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="changed_motif_name" type="xs:string">

  <xs:annotation>

    <xs:documentation>name of motif</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="changed_splice_site" type="xs:boolean">

  <xs:annotation>

    <xs:documentation>splice site change exist or not</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="splice_variant_number">

  <xs:annotation>

```

```

    <xs:documentation>number of splice variant and refSeq</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="refSeq_number" maxOccurs="unbounded"/>

    </xs:sequence>

    <xs:attribute name="number" type="xs:integer"/>

  </xs:complexType>

</xs:element>

<xs:element name="refSeq_number" type="xs:string">

  <xs:annotation>

    <xs:documentation>reference sequence number</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="ass_gene_dbref" type="dbref">

  <xs:annotation>

    <xs:documentation>database reference of the associated gene</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="location_dbref" type="dbref">

  <xs:annotation>

    <xs:documentation>database reference of location</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="variation_att">

  <xs:annotation>

    <xs:documentation>attribute information of the variation</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

```

```

<xs:element ref="molecular_type"/>

<xs:element ref="allele" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="length"/>

<xs:element ref="f5sequence" minOccurs="0"/>

<xs:element ref="f3sequence" minOccurs="0"/>

<xs:element ref="validation_status" minOccurs="0"/>

<xs:element ref="success_rate" minOccurs="0"/>

</xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="molecular_type" type="xs:string">

  <xs:annotation>

    <xs:documentation>type of molecule e.g. DNA, RNA</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="allele" type="xs:string">

  <xs:annotation>

    <xs:documentation>observed allele</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="length" type="xs:float">

  <xs:annotation>

    <xs:documentation>sequence length including flanking sequence</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="f5sequence" type="xs:string">

  <xs:annotation>

    <xs:documentation>5' flanking sequence</xs:documentation>

  </xs:annotation>

</xs:element>

```

```

<xs:element name="f3sequence" type="xs:string">

  <xs:annotation>

    <xs:documentation>3' flanking sequence</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="validation_status">

  <xs:annotation>

    <xs:documentation>status of validation as (Proven, Suspected)</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

    <xs:union>

      <xs:simpleType>

        <xs:restriction base="xs:string">

          <xs:enumeration value="Proven"/>

          <xs:enumeration value="Suspected"/>

        </xs:restriction>

      </xs:simpleType>

      <xs:simpleType>

        <xs:restriction base="xs:string"/>

      </xs:simpleType>

    </xs:union>

  </xs:simpleType>

</xs:element>

<xs:element name="success_rate" type="xs:float">

  <xs:annotation>

    <xs:documentation>certainty of variation information</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="source">

  <xs:annotation>

```

```

    <xs:documentation>source of the sequence variation</xs:documentation>

</xs:annotation>

<xs:complexType>

  <xs:sequence>

    <xs:element ref="source_release_date" minOccurs="0"/>

    <xs:element ref="source_modify_date" minOccurs="0" maxOccurs="unbounded"/>

    <xs:element ref="source_rawdata" minOccurs="0" maxOccurs="unbounded"/>

    <xs:element ref="source_dbref" maxOccurs="unbounded"/>

  </xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="source_release_date" type="xs:dateTime">

  <xs:annotation>

    <xs:documentation>date released</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="source_modify_date" type="xs:dateTime">

  <xs:annotation>

    <xs:documentation>date modified</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="source_rawdata">

  <xs:annotation>

    <xs:documentation>rawdatum of the source</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="source_dbref" type="dbref">

  <xs:annotation>

    <xs:documentation>database reference of source</xs:documentation>

  </xs:annotation>

```

```

</xs:element>

<xs:element name="variation_dbref" type="dbref">

  <xs:annotation>

    <xs:documentation>database references of the variation</xs:documentation>

  </xs:annotation>

</xs:element>

<!--===== direct annotation =====-->

<xs:element name="direct_annotation">

  <xs:annotation>

    <xs:documentation>direct annotation of variation data</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="whole_genome_sequence" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="mendelian_segregate" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="homozygote_detect" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="somatic_mutation" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="experiment_analysis" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="epidemiology" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="var_ann_misc" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="whole_genome_sequence">

  <xs:annotation>

    <xs:documentation>whole genome sequence of the datum</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="mendelian_segregate" type="xs:string">

  <xs:annotation>

```

```

    <xs:documentation>known mendelization</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="homozygote_detect">

    <xs:annotation>

        <xs:documentation>homozygote individuals observation in sample</xs:documentation>

    </xs:annotation>

    <xs:simpleType>

        <xs:restriction base="xs:string">

            <xs:enumeration value="homo"/>

            <xs:enumeration value="hetero"/>

            <xs:enumeration value="unknown"/>

        </xs:restriction>

    </xs:simpleType>

</xs:element>

<xs:element name="somatic_mutation" type="xs:string">

    <xs:annotation>

        <xs:documentation>known somatic mutation</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="experiment_analysis">

    <xs:annotation>

        <xs:documentation>explanation of the experimental and the analysis</xs:documentation>

    </xs:annotation>

    <xs:complexType>

        <xs:sequence>

            <xs:element ref="variation_identify" minOccurs="0" maxOccurs="unbounded"/>

            <xs:element ref="variation_characterize" minOccurs="0" maxOccurs="unbounded"/>

        </xs:sequence>

    </xs:complexType>

```

</xs:element>

<xs:element name="variation_identify">

<xs:annotation>

<xs:documentation>information to identify the variation</xs:documentation>

</xs:annotation>

<xs:complexType>

<xs:sequence>

<xs:element ref="experimental_assay_result"/>

<xs:element ref="experimental_assay_description"/>

<xs:element ref="experimental_assay_parameter" maxOccurs="unbounded"/>

<xs:element ref="experimental_assay_dbref" minOccurs="0" maxOccurs="unbounded"/>

<xs:choice>

<xs:element ref="publication" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="submitter" minOccurs="0" maxOccurs="unbounded"/>

</xs:choice>

<xs:element ref="pcr" minOccurs="0" maxOccurs="unbounded"/>

</xs:sequence>

<xs:attribute name="experimental_assay_id" use="required"/>

<xs:attribute name="submitter_id" use="required"/>

<xs:attribute name="create_date"/>

<xs:attribute name="modify_date"/>

</xs:complexType>

</xs:element>

<xs:element name="experimental_assay_result">

<xs:annotation>

<xs:documentation>result of the experimental assay</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="experimental_assay_description">

<xs:annotation>

```

    <xs:documentation>description of the experimental assay</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="experimental_assay_parameter" type="xs:string">

    <xs:annotation>

        <xs:documentation>parameter of the experimental assay</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="experimental_assay_dbref" type="dbref">

    <xs:annotation>

        <xs:documentation>database reference information</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="publication">

    <xs:annotation>

        <xs:documentation>publication of the experiment</xs:documentation>

    </xs:annotation>

    <xs:complexType>

        <xs:sequence>

            <xs:element ref="title" minOccurs="0"/>

            <xs:element ref="author" minOccurs="0"/>

            <xs:element ref="journal" minOccurs="0"/>

            <xs:element ref="volume" minOccurs="0"/>

            <xs:element ref="supplement" minOccurs="0"/>

            <xs:element ref="issue" minOccurs="0"/>

            <xs:element ref="issue_supplement" minOccurs="0"/>

            <xs:element ref="pages" minOccurs="0"/>

            <xs:element ref="year" minOccurs="0"/>

            <xs:element ref="publication_status" minOccurs="0"/>

            <xs:element ref="mesh_term" minOccurs="0"/>

```

```

    <xs:element ref="publication_dbref" minOccurs="0" maxOccurs="unbounded"/>

    <xs:element ref="submitter" minOccurs="0"/>

  </xs:sequence>

  <xs:attribute name="publication_id" use="required"/>

  <xs:attribute name="submitter_id" use="required"/>

  <xs:attribute name="create_date"/>

  <xs:attribute name="modify_date"/>

</xs:complexType>

</xs:element>

<xs:element name="title" type="xs:string">

  <xs:annotation>

    <xs:documentation>title of the publication</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="author" type="xs:string">

  <xs:annotation>

    <xs:documentation>author of the publication</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="journal" type="xs:string">

  <xs:annotation>

    <xs:documentation>journal of the publication</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="volume" type="xs:string">

  <xs:annotation>

    <xs:documentation>volume of the publication</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="supplement">

```

```

<xs:annotation>

  <xs:documentation>supplement of the publication</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="issue" type="xs:string">

  <xs:annotation>

    <xs:documentation>issue of the publication</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="issue_supplement">

  <xs:annotation>

    <xs:documentation>issue supplement of the publication</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="pages" type="xs:string">

  <xs:annotation>

    <xs:documentation>page of the publication</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="year" type="xs:string">

  <xs:annotation>

    <xs:documentation>year of the publication</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="publication_status" type="xs:string">

  <xs:annotation>

    <xs:documentation>status of the publication</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="mesh_term" type="xs:string">

```

```

<xs:annotation>

  <xs:documentation>mesh term of the publication</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="publication_dbref" type="dbref">

  <xs:annotation>

    <xs:documentation>database references of the publications</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="submitter">

  <xs:annotation>

    <xs:documentation>submitter of the publication</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="submitter_name"/>

      <xs:element ref="address" minOccurs="0"/>

      <xs:element ref="email" minOccurs="0"/>

      <xs:element ref="tel" minOccurs="0"/>

      <xs:element ref="fax" minOccurs="0"/>

      <xs:element ref="institution" minOccurs="0"/>

      <xs:element ref="laboratory" minOccurs="0"/>

      <xs:element ref="submitter_dbref" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="publication" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

    <xs:attribute name="submitter_id" use="required"/>

    <xs:attribute name="create_date"/>

    <xs:attribute name="modify_date"/>

  </xs:complexType>

</xs:element>

```

```
<xs:element name="submitter_name" type="xs:string">

  <xs:annotation>

    <xs:documentation>name of the submitter</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="address" type="xs:string">

  <xs:annotation>

    <xs:documentation>address of the submitter</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="email" type="xs:string">

  <xs:annotation>

    <xs:documentation>email of the submitter</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="tel" type="xs:string">

  <xs:annotation>

    <xs:documentation>telephone of the submitter</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="fax" type="xs:string">

  <xs:annotation>

    <xs:documentation>Fax of the submitter</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="institution" type="xs:string">

  <xs:annotation>

    <xs:documentation>Institution of the submitter</xs:documentation>

  </xs:annotation>

</xs:element>
```

```

<xs:element name="laboratory" type="xs:string">

  <xs:annotation>

    <xs:documentation>Laboratory of the submitter</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="submitter_dbref" type="dbref">

  <xs:annotation>

    <xs:documentation>database references of the submitter</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="pcr">

  <xs:annotation>

    <xs:documentation>PCR (polymerase chain reaction)</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="pcr_confirmed" minOccurs="0"/>

      <xs:element ref="pcr_primer" minOccurs="0"/>

      <xs:element ref="pcr_product" minOccurs="0"/>

      <xs:element ref="pcr_profile" minOccurs="0"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="pcr_confirmed">

  <xs:annotation>

    <xs:documentation>artifact verification e.g. variation found on repeat PCR sample</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

    <xs:union>

      <xs:simpleType>

```

```

    <xs:restriction base="xs:string">

        <xs:enumeration value="Yes"/>

        <xs:enumeration value="No"/>

        <xs:enumeration value="Unknown"/>

    </xs:restriction>

</xs:simpleType>

<xs:simpleType>

    <xs:restriction base="xs:string"/>

</xs:simpleType>

</xs:union>

</xs:simpleType>

</xs:element>

<xs:element name="pcr_primer" type="xs:string">

    <xs:annotation>

        <xs:documentation>primer sequence</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="pcr_product" type="xs:string">

    <xs:annotation>

        <xs:documentation>PCR product, e.g. single band, multi band</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="pcr_profile">

    <xs:annotation>

        <xs:documentation>PCR profile</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="variation_characterize">

    <xs:annotation>

        <xs:documentation>characterization of the variation datum</xs:documentation>

```

```

</xs:annotation>

<xs:complexType>

  <xs:sequence>

    <xs:element ref="genetic_statistics"/>

  </xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="genetic_statistics">

  <xs:annotation>

    <xs:documentation>genetic statistics</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="method" maxOccurs="unbounded"/>

      <xs:element ref="p-value" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="link_dis_index" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="descend_identity" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="maximum_lod_score" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="method">

  <xs:annotation>

    <xs:documentation>statistical method</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence maxOccurs="unbounded">

      <xs:element ref="method_name"/>

      <xs:element ref="method_description" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="method_url" minOccurs="0" maxOccurs="unbounded"/>

```

```

</xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="method_name">

  <xs:annotation>

    <xs:documentation>statistical method name</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

    <xs:union>

      <xs:simpleType>

        <xs:restriction base="xs:string">

          <xs:enumeration value="sib pair"/>

          <xs:enumeration value="linkage disequilibrium"/>

          <xs:enumeration value="association study"/>

        </xs:restriction>

      </xs:simpleType>

      <xs:simpleType>

        <xs:restriction base="xs:string"/>

      </xs:simpleType>

    </xs:union>

  </xs:simpleType>

</xs:element>

<xs:element name="method_description">

  <xs:annotation>

    <xs:documentation>description or explanation of the method</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="method_url" type="xs:string">

  <xs:annotation>

    <xs:documentation>URL of the method</xs:documentation>

```

```

</xs:annotation>

</xs:element>

<xs:element name="p-value" type="xs:float">

  <xs:annotation>

    <xs:documentation>p value for significance (Association study)</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="link_dis_index">

  <xs:annotation>

    <xs:documentation>linkage disequilibrium index for LD test</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence maxOccurs="unbounded">

      <xs:element ref="d" minOccurs="0"/>

      <xs:element ref="d_prime" minOccurs="0"/>

      <xs:element ref="r-square" minOccurs="0"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="d" type="xs:float">

  <xs:annotation>

    <xs:documentation>d value for Linquage Disequilibrium Test</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="d_prime" type="xs:float">

  <xs:annotation>

    <xs:documentation>d prime for LD Test</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="r-square" type="xs:float">

```

```

<xs:annotation>

  <xs:documentation>r square for LD test</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="descend_identity">

  <xs:annotation>

    <xs:documentation>Identity By Descent (IBD) for Sib-Pair Test</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="di_value"/>

      <xs:element ref="di_probability"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="di_value" type="xs:float">

  <xs:annotation>

    <xs:documentation>IBD value</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="di_probability">

  <xs:annotation>

    <xs:documentation>IBD probability</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

    <xs:restriction base="xs:float">

      <xs:minInclusive value="0"/>

      <xs:maxInclusive value="1"/>

    </xs:restriction>

  </xs:simpleType>

```

```
</xs:element>
```

```
<xs:element name="maximum_lod_score" type="xs:float">
```

```
  <xs:annotation>
```

```
    <xs:documentation>maximum LOD score for sib-pair analysis</xs:documentation>
```

```
  </xs:annotation>
```

```
</xs:element>
```

```
<xs:element name="epidemiology">
```

```
  <xs:annotation>
```

```
    <xs:documentation>epidemiology of the disease and associated gene</xs:documentation>
```

```
  </xs:annotation>
```

```
  <xs:complexType>
```

```
    <xs:sequence>
```

```
      <xs:element ref="ass_gene" minOccurs="0" maxOccurs="unbounded"/>
```

```
      <xs:element ref="disease_epidemiology" minOccurs="0" maxOccurs="unbounded"/>
```

```
      <xs:element ref="population" minOccurs="0" maxOccurs="unbounded"/>
```

```
      <xs:element ref="frequency" minOccurs="0" maxOccurs="unbounded"/>
```

```
    </xs:sequence>
```

```
  </xs:complexType>
```

```
</xs:element>
```

```
<xs:element name="disease_epidemiology">
```

```
  <xs:annotation>
```

```
    <xs:documentation>epidemiology of the disease</xs:documentation>
```

```
  </xs:annotation>
```

```
  <xs:complexType>
```

```
    <xs:sequence>
```

```
      <xs:element ref="striking_age" minOccurs="0" maxOccurs="unbounded"/>
```

```
      <xs:element ref="striking_body_area" minOccurs="0" maxOccurs="unbounded"/>
```

```
      <xs:element ref="striking_land_area" minOccurs="0" maxOccurs="unbounded"/>
```

```
      <xs:element ref="laterality" minOccurs="0"/>
```

```
      <xs:element ref="differences" minOccurs="0" maxOccurs="unbounded"/>
```

```

<xs:element ref="prognosis" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="etiology_expression_probability" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="laboratory_findings_expression_probability" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="symptoms_expression_probability" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="prophylaxes_expression_probability" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="responder_sideeffects_expression_probability" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="pathological_findings_expression_probability" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="complication_expression_probability" minOccurs="0" maxOccurs="unbounded"/>

</xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="striking_age" type="xs:string">

  <xs:annotation>

    <xs:documentation>striking age of the disease</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="striking_body_area">

  <xs:annotation>

    <xs:documentation>striking body area of the disease</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="striking_land_area">

  <xs:annotation>

    <xs:documentation>striking land area of the disease</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="laterality">

  <xs:annotation>

    <xs:documentation>laterality of the disease e.g. hemilateral, bilateral</xs:documentation>

  </xs:annotation>

```

```

<xs:simpleType>

  <xs:restriction base="xs:string">

    <xs:enumeration value="hemi"/>

    <xs:enumeration value="bi"/>

    <xs:enumeration value="right"/>

    <xs:enumeration value="left"/>

    <xs:enumeration value="other"/>

  </xs:restriction>

</xs:simpleType>

</xs:element>

<xs:element name="prognosis">

  <xs:annotation>

    <xs:documentation>prognosis of the disease</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="population">

  <xs:annotation>

    <xs:documentation>population of the variation</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="population_description"/>

      <xs:element ref="organism" minOccurs="0"/>

      <xs:element ref="differences" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="population_parameter" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="sample_size"/>

      <xs:element ref="population_misc" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="population_dbref" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

    <xs:attribute name="population_id" use="required"/>

```

```

    <xs:attribute name="submitter_id" use="required"/>

    <xs:attribute name="create_date"/>

    <xs:attribute name="modify_date"/>

</xs:complexType>

</xs:element>

<xs:element name="population_description">

    <xs:annotation>

        <xs:documentation>description of the population</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="organism" type="xs:string">

    <xs:annotation>

        <xs:documentation>organism</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="differences">

    <xs:annotation>

        <xs:documentation>statistical differences</xs:documentation>

    </xs:annotation>

    <xs:complexType>

        <xs:sequence>

            <xs:element ref="race" minOccurs="0"/>

            <xs:element ref="gender" minOccurs="0"/>

        </xs:sequence>

    </xs:complexType>

</xs:element>

<xs:element name="race" type="xs:string">

    <xs:annotation>

        <xs:documentation>racial difference</xs:documentation>

    </xs:annotation>

```

```

</xs:element>

<xs:element name="gender">

  <xs:annotation>

    <xs:documentation>gender difference</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

    <xs:union>

      <xs:simpleType>

        <xs:restriction base="xs:string">

          <xs:enumeration value="male"/>

          <xs:enumeration value="female"/>

          <xs:enumeration value="unknown"/>

          <xs:enumeration value="other"/>

        </xs:restriction>

      </xs:simpleType>

      <xs:simpleType>

        <xs:restriction base="xs:string"/>

      </xs:simpleType>

    </xs:union>

  </xs:simpleType>

</xs:element>

<xs:element name="population_parameter" type="xs:string">

  <xs:annotation>

    <xs:documentation>parameter of population</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="sample_size" type="xs:float">

  <xs:annotation>

    <xs:documentation>sample size of population</xs:documentation>

  </xs:annotation>

```

```

</xs:element>

<xs:element name="population_misc">

  <xs:annotation>

    <xs:documentation>population miscellaneous</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="population_dbref" type="dbref">

  <xs:annotation>

    <xs:documentation>database references of the population</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="frequency">

  <xs:annotation>

    <xs:documentation>frequency of the variation</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="haplotype" minOccurs="0"/>

      <xs:element ref="haplotype_frequency" minOccurs="0"/>

      <xs:element ref="allele" minOccurs="0"/>

      <xs:element ref="allele_frequency" minOccurs="0"/>

      <xs:element ref="genotype" minOccurs="0"/>

      <xs:element ref="genotype_frequency" minOccurs="0"/>

    </xs:sequence>

    <xs:attribute name="frequency_id" use="required"/>

    <xs:attribute name="submitter_id" use="required"/>

    <xs:attribute name="population_id" use="required"/>

    <xs:attribute name="assay_id" use="required"/>

    <xs:attribute name="publication_id" use="required"/>

    <xs:attribute name="create_date"/>

```

```

    <xs:attribute name="modify_date"/>

  </xs:complexType>

</xs:element>

<xs:element name="haplotype" type="xs:string">

  <xs:annotation>

    <xs:documentation>haplotype</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="haplotype_frequency" type="xs:float">

  <xs:annotation>

    <xs:documentation>frequency of the haplotype</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="allele_frequency" type="xs:float">

  <xs:annotation>

    <xs:documentation>frequency of the observed allele</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="genotype" type="xs:string">

  <xs:annotation>

    <xs:documentation>genotype</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="genotype_frequency" type="xs:float">

  <xs:annotation>

    <xs:documentation>frequency of the genotype</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="var_ann_misc">

  <xs:annotation>

```

```

    <xs:documentation>variation annotation miscellaneous</xs:documentation>

</xs:annotation>

<xs:complexType>

    <xs:sequence>

        <xs:element ref="var_ann_misc_description"/>

        <xs:element ref="var_ann_misc_dbref" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

    <xs:attribute name="var_ann_misc_id" use="required"/>

    <xs:attribute name="submitter_id" use="required"/>

    <xs:attribute name="create_date"/>

    <xs:attribute name="modify_date"/>

</xs:complexType>

</xs:element>

<xs:element name="var_ann_misc_description">

    <xs:annotation>

        <xs:documentation>description of the variation annotation miscellaneous</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="var_ann_misc_dbref" type="dbref">

    <xs:annotation>

        <xs:documentation>database references of the variation annotation miscellaneous</xs:documentation>

    </xs:annotation>

</xs:element>

<!--===== indirect_annotation =====-->

<xs:element name="indirect_annotation">

    <xs:annotation>

        <xs:documentation>indirect annotation of variation data</xs:documentation>

    </xs:annotation>

</xs:complexType>

    <xs:sequence>

```

```

<xs:element ref="personal_info" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="phenotype" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="omics_annotation" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="environmental_condition" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="clinical_annotation" minOccurs="0" maxOccurs="unbounded"/>

</xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="personal_info">

  <xs:annotation>

    <xs:documentation>personal information of variation data</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="personal_description"/>

      <xs:element ref="personal_dbref" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="personal_description">

  <xs:annotation>

    <xs:documentation>description of the personal information e.g. Name, Job</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="personal_dbref" type="dbref">

  <xs:annotation>

    <xs:documentation>database reference of personal information</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="phenotype">

```

```

<xs:annotation>

  <xs:documentation>phenotype of the sequence variation</xs:documentation>

</xs:annotation>

<xs:complexType>

  <xs:sequence>

    <xs:element ref="phenotype_description" maxOccurs="unbounded"/>

    <xs:element ref="phenotype_dbref" minOccurs="0" maxOccurs="unbounded"/>

  </xs:sequence>

  <xs:attribute name="phenotype_id" use="required"/>

  <xs:attribute name="submitter_id" use="required"/>

  <xs:attribute name="create_date"/>

  <xs:attribute name="modify_date"/>

</xs:complexType>

</xs:element>

<xs:element name="phenotype_description">

  <xs:annotation>

    <xs:documentation>description of the phenotype</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="phenotype_type"/>

      <xs:element ref="phenotype_condition" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="phenotype_probability" minOccurs="0"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="phenotype_type">

  <xs:annotation>

    <xs:documentation>type of the phenotype</xs:documentation>

  </xs:annotation>

```

</xs:element>

<xs:element name="phenotype_condition">

<xs:annotation>

<xs:documentation>condition to express the phenotype</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="phenotype_probability">

<xs:annotation>

<xs:documentation>probability to express the phenotype on the conditions</xs:documentation>

</xs:annotation>

<xs:simpleType>

<xs:restriction base="xs:float">

<xs:minInclusive value="0"/>

<xs:maxInclusive value="1"/>

</xs:restriction>

</xs:simpleType>

</xs:element>

<xs:element name="phenotype_dbref" type="dbref">

<xs:annotation>

<xs:documentation>database references of the phenotype</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="omics_annotation">

<xs:annotation>

<xs:documentation>annotation type of the omics</xs:documentation>

</xs:annotation>

<xs:complexType>

<xs:sequence>

<xs:element ref="omics_type"/>

<xs:element ref="omics_description" maxOccurs="unbounded"/>

```

    <xs:element ref="omics_dbref" minOccurs="0" maxOccurs="unbounded"/>

  </xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="omics_type">

  <xs:annotation>

    <xs:documentation>type of omics</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

    <xs:union>

      <xs:simpleType>

        <xs:restriction base="xs:string">

          <xs:enumeration value="genomics"/>

          <xs:enumeration value="transcriptomics"/>

          <xs:enumeration value="proteomics"/>

          <xs:enumeration value="signalomics"/>

          <xs:enumeration value="metabolomics"/>

          <xs:enumeration value="organomics"/>

        </xs:restriction>

      </xs:simpleType>

      <xs:simpleType>

        <xs:restriction base="xs:string"/>

      </xs:simpleType>

    </xs:union>

  </xs:simpleType>

</xs:element>

<xs:element name="omics_description">

  <xs:annotation>

    <xs:documentation>description of the omics</xs:documentation>

  </xs:annotation>

```

```

<xs:complexType>

  <xs:sequence>

    <xs:element ref="omics_material"/>

    <xs:element ref="omics_condition" minOccurs="0" maxOccurs="unbounded"/>

    <xs:element ref="omics_expression_probability" minOccurs="0"/>

  </xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="omics_material">

  <xs:annotation>

    <xs:documentation>material of the omics</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="omics_condition">

  <xs:annotation>

    <xs:documentation>condition of the omics</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="omics_expression_probability">

  <xs:annotation>

    <xs:documentation>expression probability of the omics</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

    <xs:restriction base="xs:float">

      <xs:minInclusive value="0"/>

      <xs:maxInclusive value="1"/>

    </xs:restriction>

  </xs:simpleType>

</xs:element>

<xs:element name="omics_dbref" type="dbref">

```

```

<xs:annotation>

  <xs:documentation>database reference of the omics</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="environmental_condition">

  <xs:annotation>

    <xs:documentation>environmental conditions of the expression</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="expression_condition" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="expression_condition">

  <xs:annotation>

    <xs:documentation>environmental condition of the expression</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="expression_condition_description" maxOccurs="unbounded"/>

      <xs:element ref="expression_condition_probability" minOccurs="0"/>

      <xs:element ref="omics_annotation" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="expression_condition_description">

  <xs:annotation>

    <xs:documentation>description of the expression condition</xs:documentation>

  </xs:annotation>

```

```

</xs:element>

<xs:element name="expression_condition_probability">

  <xs:annotation>

    <xs:documentation>probability of the expression condition</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

    <xs:restriction base="xs:float">

      <xs:minInclusive value="0"/>

      <xs:maxInclusive value="1"/>

    </xs:restriction>

  </xs:simpleType>

</xs:element>

<xs:element name="clinical_annotation">

  <xs:annotation>

    <xs:documentation>clinical annotation of the sequence variation</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="disease" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="clinical_observation" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="clinical_annotation_dbref" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="disease">

  <xs:annotation>

    <xs:documentation>disease information</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

```

```

<xs:element ref="disease_description" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="disease_epidemiology" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="disease_dbref" minOccurs="0" maxOccurs="unbounded"/>

</xs:sequence>

<xs:attribute name="disease_id" use="required"/>

<xs:attribute name="submitter_id" use="required"/>

<xs:attribute name="create_date"/>

<xs:attribute name="modify_date"/>

</xs:complexType>

</xs:element>

<xs:element name="disease_description">

  <xs:annotation>

    <xs:documentation>description of the disease</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="name"/>

      <xs:element ref="synonym" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="definition_concept" minOccurs="0"/>

      <xs:element ref="classification" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="etiology" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="laboratory_findings" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="pathological_findings" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="symptoms" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="diagnostic_criteria" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="therapy" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="complication" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="prophylaxes" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

  </xs:complexType>

```

</xs:element>

<xs:element name="name" type="xs:string">

<xs:annotation>

<xs:documentation>name of the disease</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="synonym" type="xs:string">

<xs:annotation>

<xs:documentation>synonym of the disease</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="definition_concept">

<xs:annotation>

<xs:documentation>definition or concept of the disease</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="classification">

<xs:annotation>

<xs:documentation>classification of the disease</xs:documentation>

</xs:annotation>

</xs:element>

<xs:element name="etiology">

<xs:annotation>

<xs:documentation>etiology of the disease</xs:documentation>

</xs:annotation>

<xs:complexType>

<xs:sequence>

<xs:element ref="etiology_description"/>

<xs:element ref="etiology_condition" minOccurs="0" maxOccurs="unbounded"/>

<xs:element ref="etiology_expression_probability" minOccurs="0"/>

```

    </xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="etiology_description">

    <xs:annotation>

        <xs:documentation>description of the etiology and its mechanism</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="etiology_condition">

    <xs:annotation>

        <xs:documentation>conditions to express the etiology</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="etiology_expression_probability">

    <xs:annotation>

        <xs:documentation>expression probability of the etiology</xs:documentation>

    </xs:annotation>

    <xs:simpleType>

        <xs:restriction base="xs:float">

            <xs:minInclusive value="0"/>

            <xs:maxInclusive value="1"/>

        </xs:restriction>

    </xs:simpleType>

</xs:element>

<xs:element name="laboratory_findings">

    <xs:annotation>

        <xs:documentation>laboratory findings of the disease</xs:documentation>

    </xs:annotation>

    <xs:complexType>

        <xs:sequence>

```

```

    <xs:element ref="laboratory_findings_type"/>

    <xs:element ref="laboratory_findings_description"/>

    <xs:element ref="laboratory_findings_condition" minOccurs="0" maxOccurs="unbounded"/>

    <xs:element ref="laboratory_findings_expression_probability" minOccurs="0"/>

  </xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="laboratory_findings_type" type="xs:string">

  <xs:annotation>

    <xs:documentation>type of laboratory findings e.g. urine, blood, ECG, image</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="laboratory_findings_description">

  <xs:annotation>

    <xs:documentation>description of the laboratory findings</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="laboratory_findings_condition">

  <xs:annotation>

    <xs:documentation>condition of the laboratory findings</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="laboratory_findings_expression_probability">

  <xs:annotation>

    <xs:documentation>expression probability of the laboratory findings</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

    <xs:restriction base="xs:float">

      <xs:minInclusive value="0"/>

      <xs:maxInclusive value="1"/>

```

```

</xs:restriction>

</xs:simpleType>

</xs:element>

<xs:element name="pathological_findings">

  <xs:annotation>

    <xs:documentation>pathological findings of the disease</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="pathological_findings_description"/>

      <xs:element ref="pathological_findings_condition" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="pathological_findings_expression_probability" minOccurs="0"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="pathological_findings_description">

  <xs:annotation>

    <xs:documentation>description of the pathological finding</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="pathological_findings_condition">

  <xs:annotation>

    <xs:documentation>condition of the pathological finding</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="pathological_findings_expression_probability">

  <xs:annotation>

    <xs:documentation>expression probability of the pathological findings</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

```

```

    <xs:restriction base="xs:float">

        <xs:minInclusive value="0"/>

        <xs:maxInclusive value="1"/>

    </xs:restriction>

</xs:simpleType>

</xs:element>

<xs:element name="symptoms">

    <xs:annotation>

        <xs:documentation>laboratory findings of the disease</xs:documentation>

    </xs:annotation>

    <xs:complexType>

        <xs:sequence>

            <xs:element ref="symptoms_description"/>

            <xs:element ref="symptoms_condition" minOccurs="0" maxOccurs="unbounded"/>

            <xs:element ref="symptoms_expression_probability" minOccurs="0"/>

        </xs:sequence>

    </xs:complexType>

</xs:element>

<xs:element name="symptoms_description">

    <xs:annotation>

        <xs:documentation>description of the symptom</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="symptoms_condition">

    <xs:annotation>

        <xs:documentation>condition of the symptom</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="symptoms_expression_probability">

    <xs:annotation>

```

```

    <xs:documentation>expression probability of the symptom</xs:documentation>

</xs:annotation>

<xs:simpleType>

    <xs:restriction base="xs:float">

        <xs:minInclusive value="0"/>

        <xs:maxInclusive value="1"/>

    </xs:restriction>

</xs:simpleType>

</xs:element>

<xs:element name="diagnostic_criteria">

    <xs:annotation>

        <xs:documentation>criteria of diagnosis for the disease</xs:documentation>

    </xs:annotation>

    <xs:complexType>

        <xs:sequence>

            <xs:element ref="diagnostic_standard"/>

            <xs:element ref="diagnostic_modify" minOccurs="0" maxOccurs="unbounded"/>

            <xs:element ref="diagnostic_differential" minOccurs="0" maxOccurs="unbounded"/>

        </xs:sequence>

    </xs:complexType>

</xs:element>

<xs:element name="diagnostic_standard">

    <xs:annotation>

        <xs:documentation>standard diagnostic criteria</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="diagnostic_modify">

    <xs:annotation>

        <xs:documentation>modified diagnostic criteria</xs:documentation>

    </xs:annotation>

```

```
</xs:element>
```

```
<xs:element name="diagnostic_differential">
```

```
  <xs:annotation>
```

```
    <xs:documentation>differential diagnosis use</xs:documentation>
```

```
  </xs:annotation>
```

```
</xs:element>
```

```
<xs:element name="therapy">
```

```
  <xs:annotation>
```

```
    <xs:documentation>therapy of the disease</xs:documentation>
```

```
  </xs:annotation>
```

```
  <xs:complexType>
```

```
    <xs:sequence>
```

```
      <xs:element ref="conservative" minOccurs="0" maxOccurs="unbounded"/>
```

```
      <xs:element ref="surgery" minOccurs="0" maxOccurs="unbounded"/>
```

```
      <xs:element ref="radiation" minOccurs="0" maxOccurs="unbounded"/>
```

```
    </xs:sequence>
```

```
  </xs:complexType>
```

```
</xs:element>
```

```
<xs:element name="conservative">
```

```
  <xs:annotation>
```

```
    <xs:documentation>conservative treatment</xs:documentation>
```

```
  </xs:annotation>
```

```
  <xs:complexType>
```

```
    <xs:sequence>
```

```
      <xs:element ref="pharmaceutical" minOccurs="0" maxOccurs="unbounded"/>
```

```
      <xs:element ref="physical" minOccurs="0" maxOccurs="unbounded"/>
```

```
    </xs:sequence>
```

```
  </xs:complexType>
```

```
</xs:element>
```

```
<xs:element name="pharmaceutical">
```

```

<xs:annotation>

  <xs:documentation>pharmaceutical element</xs:documentation>

</xs:annotation>

<xs:complexType>

  <xs:sequence>

    <xs:element ref="responder_sideeffects" minOccurs="0" maxOccurs="unbounded"/>

  </xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="responder_sideeffects">

  <xs:annotation>

    <xs:documentation>responder and/or side effects</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="responder_sideeffects_description"/>

      <xs:element ref="responder_sideeffects_causer" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="responder_sideeffects_condition" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="responder_sideeffects_expression_probability" minOccurs="0"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="responder_sideeffects_causer">

  <xs:annotation>

    <xs:documentation>causer of the responder and/or side effect</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="responder_sideeffects_description">

  <xs:annotation>

    <xs:documentation>description of the responder and/or side effect</xs:documentation>

```

```

</xs:annotation>

</xs:element>

<xs:element name="responder_sideeffects_condition">

  <xs:annotation>

    <xs:documentation>condition of the responder and/or side effect</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="responder_sideeffects_expression_probability">

  <xs:annotation>

    <xs:documentation>expression probability of the responder, the side effect</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

    <xs:restriction base="xs:float">

      <xs:minInclusive value="0"/>

      <xs:maxInclusive value="1"/>

    </xs:restriction>

  </xs:simpleType>

</xs:element>

<xs:element name="physical">

  <xs:annotation>

    <xs:documentation>physical treatment element</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="surgery">

  <xs:annotation>

    <xs:documentation>surgical treatment</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="radiation">

  <xs:annotation>

```

```

    <xs:documentation>radiation therapy</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="complication">

  <xs:annotation>

    <xs:documentation>complications of the disease</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="complication_description"/>

      <xs:element ref="complication_condition" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="complication_expression_probability" minOccurs="0"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="complication_description">

  <xs:annotation>

    <xs:documentation>description of the complication</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="complication_condition">

  <xs:annotation>

    <xs:documentation>condition to express the complication</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="complication_expression_probability">

  <xs:annotation>

    <xs:documentation>probability when the complication is expressed</xs:documentation>

  </xs:annotation>

  <xs:simpleType>

```

```

    <xs:restriction base="xs:float">

        <xs:minInclusive value="0"/>

        <xs:maxInclusive value="1"/>

    </xs:restriction>

</xs:simpleType>

</xs:element>

<xs:element name="prophylaxes">

    <xs:annotation>

        <xs:documentation>prophylaxes of the disease</xs:documentation>

    </xs:annotation>

    <xs:complexType>

        <xs:sequence>

            <xs:element ref="prophylaxes_description"/>

            <xs:element ref="prophylaxes_condition" minOccurs="0" maxOccurs="unbounded"/>

            <xs:element ref="prophylaxes_expression_probability" minOccurs="0"/>

        </xs:sequence>

    </xs:complexType>

</xs:element>

<xs:element name="prophylaxes_description">

    <xs:annotation>

        <xs:documentation>description of the prophylaxis</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="prophylaxes_condition">

    <xs:annotation>

        <xs:documentation>condition of the prophylaxis</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="prophylaxes_expression_probability">

    <xs:annotation>

```

```

    <xs:documentation>expression probability of the prophylaxes</xs:documentation>

</xs:annotation>

<xs:simpleType>

    <xs:restriction base="xs:float">

        <xs:minInclusive value="0"/>

        <xs:maxInclusive value="1"/>

    </xs:restriction>

</xs:simpleType>

</xs:element>

<xs:element name="disease_dbref" type="dbref">

    <xs:annotation>

        <xs:documentation>database references of the disease</xs:documentation>

    </xs:annotation>

</xs:element>

<xs:element name="clinical_observation">

    <xs:annotation>

        <xs:documentation>clinical observation of the person</xs:documentation>

    </xs:annotation>

    <xs:complexType>

        <xs:sequence>

            <xs:element ref="subjective_findings" minOccurs="0" maxOccurs="unbounded"/>

            <xs:element ref="objective_findings" minOccurs="0" maxOccurs="unbounded"/>

            <xs:element ref="assessment" minOccurs="0" maxOccurs="unbounded"/>

            <xs:element ref="plan" minOccurs="0" maxOccurs="unbounded"/>

            <xs:element ref="clinical_observation_dbref" minOccurs="0" maxOccurs="unbounded"/>

        </xs:sequence>

    </xs:complexType>

</xs:element>

<xs:element name="subjective_findings">

    <xs:annotation>

```

```

    <xs:documentation>subjective findings in clinical observation</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="symptoms"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="objective_findings">

  <xs:annotation>

    <xs:documentation>objective findings in clinical observation</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="laboratory_findings" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="pathological_findings" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="complication" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="family_history" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="family_history">

  <xs:annotation>

    <xs:documentation>family history</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="family_history_description" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="family_history_dbref" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

```

```

</xs:complexType>

</xs:element>

<xs:element name="family_history_description">

  <xs:annotation>

    <xs:documentation>description of the family history</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="family_member" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="relation_structure" minOccurs="0"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="family_member">

  <xs:annotation>

    <xs:documentation>member of the family</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="personal_info" maxOccurs="unbounded"/>

      <xs:element ref="phenotype" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="clinical_annotation" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="relation_structure">

  <xs:annotation>

    <xs:documentation>relational structure of the family</xs:documentation>

  </xs:annotation>

</xs:element>

```

```

<xs:element name="family_history_dbref" type="dbref">

  <xs:annotation>

    <xs:documentation>database reference of the family history</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="assessment">

  <xs:annotation>

    <xs:documentation>assessment of the clinical observation</xs:documentation>

  </xs:annotation>

  <xs:complexType>

    <xs:sequence>

      <xs:element ref="assessment_result" minOccurs="0" maxOccurs="unbounded"/>

      <xs:element ref="assessment_dbref" minOccurs="0" maxOccurs="unbounded"/>

    </xs:sequence>

  </xs:complexType>

</xs:element>

<xs:element name="assessment_result">

  <xs:annotation>

    <xs:documentation>results of the assessment</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="assessment_dbref" type="dbref">

  <xs:annotation>

    <xs:documentation>database reference of the assessment</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="plan">

  <xs:annotation>

    <xs:documentation>plan of the clinical observation</xs:documentation>

  </xs:annotation>

```

```

<xs:complexType>

  <xs:sequence>

    <xs:element ref="therapy"/>

  </xs:sequence>

</xs:complexType>

</xs:element>

<xs:element name="clinical_observation_dbref" type="dbref">

  <xs:annotation>

    <xs:documentation>database reference of the clinical observation </xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="clinical_annotation_dbref" type="dbref">

  <xs:annotation>

    <xs:documentation>database reference of the clinical annotation</xs:documentation>

  </xs:annotation>

</xs:element>

<!--===== dbref complexType =====-->

<xs:complexType name="dbref">

  <xs:annotation>

    <xs:documentation>database reference</xs:documentation>

  </xs:annotation>

  <xs:sequence>

    <xs:element ref="database_name"/>

    <xs:element ref="database_id" minOccurs="0"/>

    <xs:element ref="link_url" minOccurs="0"/>

    <xs:element ref="database_attbt" minOccurs="0"/>

    <xs:element ref="data_id"/>

    <xs:element ref="data_attbt" minOccurs="0"/>

    <xs:element ref="version" minOccurs="0"/>

  </xs:sequence>

```

```

</xs:complexType>

<xs:element name="database_name" type="xs:string">

  <xs:annotation>

    <xs:documentation>name of the database</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="database_id" type="xs:string">

  <xs:annotation>

    <xs:documentation>ID of the database</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="link_url" type="xs:string">

  <xs:annotation>

    <xs:documentation>URL of the database</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="database_attbt" type="xs:string">

  <xs:annotation>

    <xs:documentation>attribute of the database</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="data_id" type="xs:string">

  <xs:annotation>

    <xs:documentation>ID of the datum</xs:documentation>

  </xs:annotation>

</xs:element>

<xs:element name="data_attbt">

  <xs:annotation>

    <xs:documentation>attribute of the datum</xs:documentation>

  </xs:annotation>

```

```
<xs:complexType/>

</xs:element>

<xs:element name="version" type="xs:string">

  <xs:annotation>

    <xs:documentation>version of the database</xs:documentation>

  </xs:annotation>

</xs:element>

</xs:schema>
```

STANDARDSISO.COM : Click to view the full PDF of ISO 25720:2009

Annex C (informative)

Basic reference works

C.1 Introduction

This annex gives the reference works for development of GSVML (see Clause 6) and a non-exhaustive list of the most generally applicable basic reference works. Details concerning currently valid ISO documents are available in the ISO catalogue. Copies can be obtained from the ISO national bodies. For specific subjects, the provisions of other, less generally applicable, documents will be relevant.

C.2 Use case analysis

C.2.1 General

We summarised typical use cases of genomic sequence variation data exchanging in human health. Six use cases for three types of criteria for the SNP data exchange are summarised as follows.

Type I: Clinical practice use

The first type of use case is clinical practice. In this use case, SNP data are exchanged amongst the facilities and the clinical specialists.

Genetic diagnosis or genetic counselling

For genetic diagnosis or genetic counselling, the SNP data are exchanged amongst MDs, lab technicians and counsellors. In this case, individual SNP data are sent/received with individual clinical data. For more advanced diagnosis, individual genomic data including -omics data are demanded.

Prescription derived from pharmacogenomics

For prescription derived from pharmacogenomics, the SNP data will not be exchanged in most cases. The exchange data will be the prescription, reasons and its annotations. In this case, individual SNP data are not sent/received with individual prescription results.

Gene therapy

For gene therapy, the SNP data are exchanged amongst hospitals, other facilities, MDs and patients. In this case, individual SNP data are sent/received with individual clinical data and individual genomic data.

Disease prevention based on the individual polymorphism

For disease prevention based on the individual polymorphism, the SNP data are exchanged between MDs and lab technicians. In this case, individual SNP data are sent/received with individual clinical data.

Type II: Clinical trial use

In the case of clinical trials, the SNP data are exchanged amongst hospitals, research institutes, MDs and the pharmaceutical company. In this case, individual SNP data are sent/received with individual clinical data and other data that are needed to specify the experiment. The clinical data required depend on the clinical phase. Early phases do not need many individuals but need many parameters, while in late phases the opposite is true.

Type III: Translational research use

In translational research, the SNP data are exchanged amongst hospitals, research institutes, MDs, researchers and the pharmaceutical company. In this case, individual SNP data are sent/received with individual clinical data along with other additional data that are needed to specify the experiment. The number of clinical data elements required will be several dozen, while the parameters for each individual are many.

C.2.2 Overview

Figure C.1 is the general use case of GSVML in the clinical scene.

Through GSVML, every actor can exchange data smoothly without requiring a change in their existing database schema. In the same way, the researchers can exchange their genomic sequence variation data without any pain.

As an example, in the case of genetic diagnosis, the individual SNP data are exchanged amongst facilities such as hospitals and medical laboratories. These data are also exchanged amongst persons such as MDs, laboratory analysts, counsellors and, in some cases, the patient him/herself. Here individual SNP data are encapsulated with the individual clinical data and his/her -omics data, in some cases, for further examination. To analyse this individual SNP data, the individual SNP data need to be compared with the database derived SNP data that have various types of data formats.

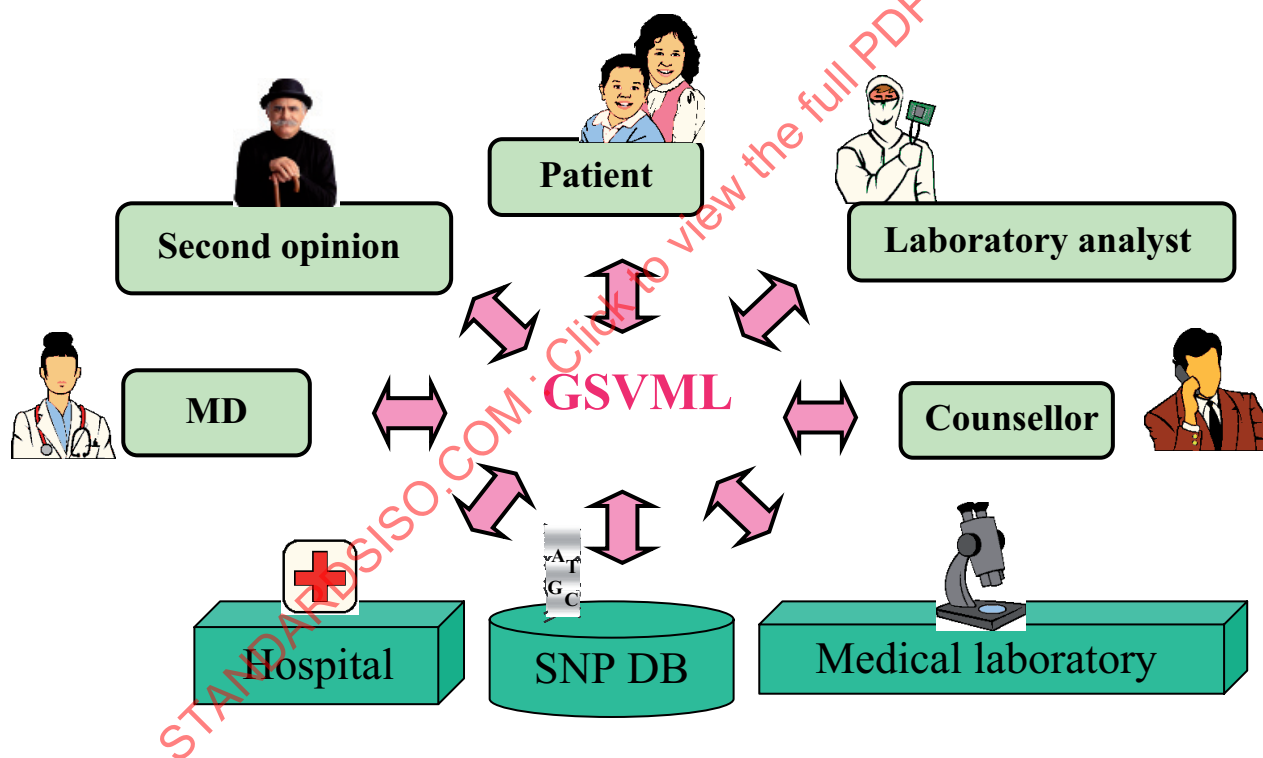


Figure C.1 — An outlined example of general use case

C.2.3 Use case of SNP analysis

Figure C.2 is an example work flow of the SNP analysis.

This is the case of the “Japanese National Millennium Project”.

This project tries to find the relevant SNPs or genes for the five lifestyle diseases.

This is the case for the diabetes mellitus.