TECHNICAL SPECIFICATION

ISO/TS 19844

First edition
2015-12-15

Health informatics — Identification of medicinal products — Implementation guidelines for data elements and structures for the unique identification and exchange of regulated information on substances

Informatique de santé — Identification des médicaments — Lignes directrices pour la mise en œuvre des éléments de données et structures pour l'identification unique et l'échange d'informations réglementées sur les substances

Reference number
ISO/TS 19844:2015(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword viii
Introduction ix
1 Scope 1
2 Normative references 2
3 General background and history 2
4 Substance (Mandatory) 3
4.1 Introduction 3
4.2 Defining Substances 5
4.3 Substance Types (Mandatory) 7
4.4 Substance ID (Mandatory) 10
4.5 Substance Names (Mandatory) 11
4.5.1 Substance Name 12
4.5.2 Substance Name Type 12
4.5.3 Language 13
4.5.4 Official Name (repeat as necessary) 13
4.6 Reference Sources (Mandatory) 15
4.6.1 Public Domain 16
4.6.2 Reference Source Type 16
4.6.3 Reference Source Class 16
4.6.4 Reference Source ID 17
4.6.5 Reference Source Citation 17
4.6.6 Reference Source Document (new class to be included in the second edition of ISO 11238) 17
4.6.7 Reference Source Document Type (new class to be included in the second edition of ISO 11238) 18
4.6.8 Reference Source Document Classification (new class to be included in the second edition of ISO 11238) 18
4.6.9 Reference Source URL (new class to be included in the second edition of ISO 11238) 18
4.7 Substance Code (Conditional) 18
4.7.1 Code 19
4.7.2 Code System 19
4.7.3 Code System ID 20
4.7.4 Code System Status 20
4.7.5 Code System Status Change Date 20
4.7.6 Comment 20
4.7.7 Reference Source 21
4.7.8 Substance Classification (repeat as necessary) 21
4.7.9 Target 23
4.7.10 Gene 25
4.7.11 Gene Elements 26
4.7.12 Substance Relationship 27
4.8 Structure (repeat as necessary) (Conditional) 29
4.8.1 Structural Representation Type 34
4.8.2 Structural Representation 35
4.8.3 Structural Representation Attachment 35
4.8.4 Stereochemistry 35
4.8.5 Optical Activity 36
4.8.6 Molecular Formula 36
4.8.7 Molecular Formula by Moieties (new class to be included in the second edition of ISO 11238) 37
4.8.8 Isotope (repeat as necessary) 37

4.9 Amount (Conditional) 38
4.9.1 Average 38
4.9.2 Low Limit 38
4.9.3 High Limit 39
4.9.4 Unit 39
4.9.5 Non-numeric Value 39
4.10 Source Material (Conditional) 39
4.10.1 Source Material Class 40
4.10.2 Source Material Type 41
4.10.3 Source Material state 41
4.10.4 Organism ID 41
4.10.5 Organism Name 41
4.10.6 Development Stage 42
4.10.7 Part Description (repeat as necessary) 42
4.10.8 Fraction (repeat as necessary) 42
4.10.9 Organism 43
4.11 Modification (repeat as necessary) (Conditional) 49
4.11.1 Modification Type 50
4.11.2 Residue Modified 51
4.11.3 Residue Site 51
4.11.4 Structural Modification 51
4.12 Property (Conditional) 55
4.12.1 Property Type 56
4.12.2 Property Name 56
4.12.3 Property Parameters (new class to be included in the second edition of ISO 11238) 56
4.12.4 Substance Name 57
4.12.5 Substance ID 57
4.12.6 Amount type 57
4.13 Version (repeat as necessary) (Mandatory) 57
4.13.1 Version Number 57
4.13.2 Effective date 58
4.13.3 Change Made 58

5 Substance definitions 58
5.1 Chemical Substance 58
5.1.1 Structure 59
5.1.2 Stoichiometric 59
5.1.3 Stoichiometric Chemicals 59
5.1.4 Comment 62
5.1.5 Non- Stoichiometric Chemicals 62
5.2 Proteins/ Peptides 64
5.2.1 Microheterogeneity 65
5.2.2 Sequence Type 66
5.2.3 Number of subunits 66
5.2.4 Disulfide Linkage 66
5.2.5 Comment 67
5.2.6 Protein Subunit (repeat as necessary) 67
5.2.7 Molecular Weight (repeat as necessary) 69
5.2.8 Glycosylation 69
5.2.9 Structure 71
5.2.10 Modification 71
5.2.11 Property 71
5.2.12 Molecular Weight 71
5.3 Nucleic Acids 72
5.3.1 Structure 73
5.3.2 Sequence Type 73
5.3.3 Number of Subunits 73
5.3.4 Area of hybridisation 74

5.3.5 Comment ... 74
5.3.6 Nucleic Acid Subunit (repeat as necessary) ... 74
5.3.7 Modification ... 77
5.3.8 Property ... 77
5.3.9 Molecular Weight ... 77
5.4 Polymers –To be addressed in more detail in the next edition of this Technical Specification ... 78
5.4.1 Substance Name ... 79
5.4.2 Structure ... 79
5.4.3 Polymer Class ... 79
5.4.4 Polymer Geometry ... 80
5.4.5 Copolymer Sequence type ... 80
5.4.6 Comment ... 80
5.4.7 Monomer Description (repeat as necessary) ... 80
5.4.8 Structural Repeat (repeat as necessary) ... 81
5.4.9 Molecular Weight (repeat as necessary) ... 83
5.4.10 Property (repeat as necessary) ... 83
5.4.11 Reference Source (repeat as necessary) ... 83
5.5 Structurally-Diverse Substances ... 83
5.5.1 herbals and Substances Used in the Preparation of Plant-Based Allergenic Extracts ... 84
5.5.2 Vaccines — Annex addressing this will be included in the next edition of this Technical Specification ... 93
5.5.3 Purified Blood Products and Polyclonal Antibodies — Annex addressing this will be included in the next edition of this Technical Specification ... 93
5.5.4 Cells and Tissues — Annex addressing this will be included in the next edition of this Technical Specification ... 93
5.5.5 Minerals — Annex addressing this will be included in the next edition of this Technical Specification ... 93
5.6 Mixture Substance (repeat as necessary) ... 94
5.6.1 Mixture Type ... 94
5.6.2 Mixture Constituent (repeat as necessary) ... 94
6 specified substance (Optional) ... 95
6.1 specified substance Group 1 (repeat as necessary) ... 96
6.1.1 specified substance Group 1 ID ... 97
6.1.2 specified substance Group1 Name ... 97
6.1.3 Substance Name (repeat as necessary) ... 97
6.1.4 Substance Code ... 97
6.1.5 Version (repeat as necessary) ... 97
6.1.6 Reference Sources ... 97
6.1.7 Property (repeat as necessary) ... 97
6.1.8 Fraction (new class to be included in the second edition of ISO 11238) ... 98
6.1.9 Modification ... 98
6.1.10 Reference Information (repeat as necessary) ... 98
6.1.11 Constituent (repeat as necessary) ... 98
6.1.12 Physical Form (repeat as necessary) ... 99
6.2 specified substance Group 1 intended for herbal Substance and herbal Preparation ... 100
6.2.1 specified substance Group 1 ID ... 101
6.2.2 specified substance Group1 Name ... 101
6.2.3 Reference Sources ... 101
6.2.4 Fraction (new class to be included in the second edition of ISO 11238) ... 101
6.2.5 Modification (new classes to be included in the second edition of ISO 11238) ... 102
6.2.6 Constituent (repeat as necessary) ... 102
6.2.7 Physical Form (repeat as necessary) ... 103
6.3 specified substance Group 2 (repeat as necessary) ... 104
6.3.1 specified substance Group2 ID ... 107
6.3.2 specified substance Group2 Name ... 107
6.3.3 Parent Substance ID ... 107
6.3.4 Manufacturing ... 107

6.4 specified substance Group 2 for herbal preparations....111
6.4.1 specified substance Group2 ID....111
6.4.2 specified substance Group2 Name....111
6.4.3 Parent Substance ID....111
6.4.4 Manufacturing....112
6.4.5 Version (repeat as necessary)....113
6.5 specified substance Group 3 (repeat as necessary)....113
6.5.1 specified substance Group3 ID....114
6.5.2 specified substance Group3 Name....114
6.5.3 Parent Substance ID....114
6.5.4 Grade....115
6.5.5 Reference Source (repeat as necessary)....115
6.5.6 Version (repeat as necessary)....115
6.5.7 Reference Source (repeat as necessary)....116
6.5.8 Version (repeat as necessary)....116

Annex A (normative) Choosing a Substance ID....117
A.1 Requesting a Substance ID and providing information....117

Annex B (normative) Chemical Substance....119
B.0 Introduction....119
B.0.1 Proposal for the update of the ISO 11238 Substance standard....120
B.0.2 Outline of Annex B....122
B.1 Scope....123
B.2 Terms and definitions....123
B.3 Chemical Substance subtypes and Mixture Substance....132
B.3.1 Substance type, Chemical substance....132
B.3.2 Solid state forms of the Substance....132
B.3.3 Need to substantiate the chemical structure, molecular formula and molecular weight....135
B.3.4 Polymorphism....136
B.3.5 Non-Stoichiometric chemical substances....137
B.3.6 Mixture substance....138
B.3.7 Multi substance material....139
B.4 Discussion of the key elements of a chemical substance....142
B.4.1 Identity of material....142
B.4.2 Nomenclature....142
B.4.3 Molecular formula....142
B.4.4 Molecular weight....143
B.4.5 Substance Structure....143
B.4.6 Geometric Isomerism....146
B.4.7 Stereo-descriptors in systematic nomenclature: Substance with one centre of Asymmetry....147
B.4.8 Substance with two centres of Asymmetry, Epimers, Diastereomers....148
B.4.9 Anomers....148
B.4.10 Substance with more than two centres of Asymmetry (Mixture of stereoisomers)....151
B.4.11 Conclusion for the Key elements....152
B.4.12 Decision tree for a new Substance ID....152
B.5 Discussing other elements of importance regarding the characteristics of a substance....153
B.5.1 Introduction....153
B.5.2 Naming Vegetable Oils....153
B.5.3 Castor Oil and related products....169
B.5.4 Properties to be captured, related to liquids (Gas), Nitrous oxide....172
B.6 Examples....177
B.6.1 Example 1: Amlodipine besilate....178
B.6.2 Example 2: Ponatinib hydrochloride....185
B.6.3 Example 3: Benzathine Benzylpenicillin tetrahydrate, sterilised.....201
B.6.4 specified substance Group 2 information level....206
B.6.5 specified substance Group 3 information level....207
B.7 Radiopharmaceutical substance....207
B.7.1 Introduction....207

B.7.2 Example: Florbetapir ^{18}F 208
B.7.3 Identity of material, combining the elements for Florbetapir ^{18}F, Substance and specified substance information level. 209
B.7.4 specified substance Group 2 information level 214

Annex C (normative) Protein Substance 215
C.1 Scope 215
C.2 Introduction 215
C.3 Peptide Substances 216
C.3.1 Example IIIa: Protein Substance: Vasopressin 216
C.3.2 Example IIIb: Desmopressin 219
C.3.3 Example IIIc: Desmopressin Acetate 223
C.3.4 Example IIId: Calcitonin Salmon 227
C.3.5 Example IIIe: Human Insulin 233
C.4 Element Group Protein, specified substance Group 1 235
C.4.1 Example IVa: Insulin Human Zinc Suspension (Amorphous), taken from EP/USP Monographs for Zinc Suspensions 235
C.4.2 Example IVb: Insulin Human Zinc Suspension (Crystalline), taken from EP/USP Monographs for Zinc Suspensions 236
C.4.3 Example IVc: Insulin Human Zinc Suspension taken from EP/USP Monographs for Zinc Suspensions, including discussion for specified substance Group 3 information 237
C.5 Element Group Protein, specified substance Group 2 240
C.5.1 Example Va: Synthetic Calcitonin Salmon-Manufacturer Company Acme 240
C.5.2 specified substance Group 2 information level 240
C.5.3 Example Vb: Recombinant Calcitonin Salmon-- Company Acme 241
C.6 Element Group Protein, specified substance Group 3 241
C.6.1 Example: Calcitonin Salmon (Synthetic) specified substance Group 3 242
C.6.2 Example: Calcitonin Salmon (Recombinant) specified substance Group 3 242
C.7 Example Protein Substance Example of a Monoclonal Antibody conjugated Toxin 243
C.8 Addendum: Microheterogeneity 248

Annex D (normative) Nucleic Acid Substance 249
D.1 Scope 249
D.2 Introduction 249
D.3 Examples 250
D.3.1 Example: 5-Methylcytosine 250
D.3.2 Example: A 5'-phosphate ribonucleotide 251
D.4 Substances 252
D.4.1 Example: Mipomersen Sodium information 252
D.4.2 Example: Oligonucleotide Elements 252
D.4.3 Example: Anti-sense RNA 257

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword - Supplementary information

The committee responsible for this document is ISO/TC 215, *Health informatics*.

Introduction

This Technical Specification is a guide for implementing ISO 11238, *Health informatics — Identification of medicinal products — Data elements and structures for the unique identification and exchange of regulated information on substances*. This Technical Specification was developed in response to a worldwide demand for guidance on the implementation of internationally harmonised specifications for medicinal products. It is one of a group of four implementation guides for a total of five ISO standards which together provide the basis for the unique identification of medicinal products. The other standards in this group are:

- ISO 11615, *Health informatics — Identification of medicinal products — Data elements and structures for the unique identification and exchange of regulated medicinal product information*

- ISO 11616, *Health informatics — Identification of medicinal products — Data elements and structures for the unique identification and exchange of regulated pharmaceutical product information*

- ISO 11239, *Health informatics — Identification of medicinal products — Data elements and structures for the unique identification and exchange of regulated information on pharmaceutical dose forms, units of presentation, routes of administration and packaging*

- ISO 11240, *Health informatics — Identification of medicinal products — Data elements and structures for the unique identification and exchange of units of measurement*

The standards for the Identification of Medicinal Products (IDMP) support the activities of medicines regulatory agencies worldwide by jurisdiction. These include a variety of regulatory activities related to development, registration and life cycle management of medicinal products as well as pharmacovigilance and risk management.

The business objective of this implementation guide is to provide a means for exchanging regulatory substance information. To meet the primary objectives of the regulation of medicines and pharmacovigilance, it is necessary to exchange medicinal product information in a robust and reliable manner.

For the purposes of this Technical Specification, all conditions (e.g. mandatory, conditional, optional) correspond to the necessary requirements to uniquely and unambiguously identify a substance. Implementation of the ISO IDMP standards may dictate that mandatory elements for identification be tagged as conditional or optional, based on regional requirements. If a section is identified as 'optional' but is implemented in a specific region, conformance described within that section is applicable. The scope of this Technical Specification is to identify the scientifically necessary elements for the unique identification of substances/specified substances.

Health informatics — Identification of medicinal products — Implementation guidelines for data elements and structures for the unique identification and exchange of regulated information on substances

1 Scope

This Technical Specification is used in the implementation of ISO 11238. This Technical Specification defines substances based on their scientific identity (i.e. what they are) rather than on their use or method of production.

ISO 11238 provides the conceptual framework for defining substances and specified substances and for assigning unique identifiers in the context of the ISO IDMP standards. ISO 11238 describes general concepts for defining and distinguishing substances and a high level model for the structuring of information for substances. This Technical Specification provides detailed explanations of each type or grouping of substance information, an element-by-element description for implementation of ISO 11238, and examples for a variety of substances and specified substances.

This first edition of the Technical Specification will only address substances, and Groups 1 to 3 of the specified substances as defined in ISO 11238 and Annexes A, B, C, and D. It is anticipated that specified substances Group 4, as defined in ISO 11238, will be addressed in a subsequent edition of this Technical Specification. Some information that would typically fall under specified substances Group 4 may be covered in the Annexes of this Technical Specification. This information, although not defining of either a substance or a specified substance Group 1, may be essential to distinguishing substances.

This Technical Specification addresses the following:

- Data elements necessary for defining substances and specified substances Groups 1 to 3;
- The logical use of data elements as defined in ISO 11238;
- Substances and specified substances Groups 1 to 3 business rules for
  - determining necessary data elements,
  - distinguishing and defining materials according to ISO 11238,
  - triggering the assignment of identifiers.

This Technical Specification does not address the following:

- Business processes for data management;
- Implementation of a specific data information system (e.g. a relational database schema);
- Normative messaging standards for substances;
- The maintenance of controlled vocabularies;
- The specific global identifier system that should be used;
- Nomenclature standards for substances.

2 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.

ISO 8601, *Data elements and interchange formats — Information interchange — Representation of dates and times*

ISO 11238, *Health informatics — Identification of medicinal products — Data elements and structures for the unique identification and exchange of regulated information on substances*

3 General background and history

Due to the lack of a common and harmonised approach to define substances, regulators and pharmaceutical industry are faced with the inability to:

1) effectively exchange medicinal substance information in a structured and efficient way;

2) ensure data consistency and evaluate/compare information across regions, which especially impairs pharmacovigilance and compliance activities;

3) develop consistent terminology for use throughout the healthcare community.

The objectives of the IDMP standards are to address the issues outlined above by developing harmonised standards that build on the regulatory and technical processes already established and to support the population and maintenance of existing systems/applications with fully reliable regulatory medicinal product information.

Harmonised standards will stimulate vendors to develop “off-the-shelf” tools (that are interoperable due to the standard itself). Harmonised standards will also help to maximise forward compatibility of data and minimise the complexities of backward compatibility.

This implementation guide is intended to assist reporters (including pharmaceutical companies, regulatory authorities and non-commercial sponsors) in constructing messages or transmitting information that allows substances to be defined unambiguously and assigned unique IDs. It also provides guidance to help choose the correct Substance ID from a public data source that provides unique substance and specified substance identifiers. It is anticipated that an extensive list of substance identifiers as well as the definitional elements upon which the ID was based will be provided. This Technical Specification is not intended to be a guide for a maintenance organisation. The maintenance organisation may also create alternative methods to submit information consistent with the ISO model.

Table 1 is an example table for class and elements description.

Table 1 — Example table for class and element description

User Guidance	
Example(s)	
Conformance	
Data Type	
Values Allowed	
Business Rule(s)	

In contrast to other parts of the guide, conformance will refer to whether an element is required for a given substance type or a specified substance group. Conformance is not meant to be applied globally to all types of messages.

Definition: Conformance will be expressed based on the following terminology: Mandatory, Conditional and Optional.

Mandatory: refers to data elements that are required and shall therefore be implemented.

Conditional: refers to data elements that are subject to business rules and may became required by:
- data rules;
- process rules;
- regional rules.

Optional: refers to data elements that are informative but not definitional.

The description on whether a data element is conditional by data, by process or by regional rule is out of scope of this Technical Specification and will be defined within regional implementation guides.

The information provided in the table refers to the global guidance. When there is no information in the conformance table row (e.g. information on business rule is not provided), please refer to the regional implementation guide.

4 Substance (Mandatory)

4.1 General

All medicinal products consist of substances; these substances can be active ingredients, excipients, or packaging materials. There are two fundamental levels of information described in ISO 11238, a “substance level” and a “specified substance” level. Both levels are included in the more generic concept of an ingredient. At the substance level, substances are defined based on inherent attributes rather than use or method of manufacture. At the specified substance level, four separate groups of elements provide additional information.

In order to define or distinguish material either at a substance or specified substance level, a number of attributes should be taken into consideration:

- For chemicals, the molecular structure is captured at the substance level;

- For proteins, the amino acid sequence, sites and type of glycosylation, and the presence and position of disulfide bonds will be captured at the substance level;

- For nucleic acids, the sequence, type of sugar and linkage will be captured at the substance level;

- For other polymers, the monomers used to synthesize the polymer, the structural repeating units, the molecular weight and/or a property related to molecular weight (e.g. viscosity), the source of naturally derived polymers and any modifications that irreversibly alter the molecular structure will be captured at the substance level;

- For structurally diverse material, taxonomic, anatomical and fractionation information, properties related to the underlying molecular structure of the material, and modifications that alter the underlying molecular structure will be captured at the substance level;

- Mixture substance consists of a simple combination of single substances that are either isolated together or are the result of the same synthetic process. The biological source of the mixture is also captured where relevant at the substance level. Proportions are not captured at the substance level. It should be noted that a mixture

substance description should only include the substances that are generally or consistently present in the material. This excludes impurities and degradants.

Other attributes will be specific to the specified substance levels:

- Constituent substances in a multi-substance material;
- Proportions of constituent substances in a multi-substance material;
- Physical state;
- Grade or purity of material;
- Manufacturing information;
- Analytical data.

There are four groups of elements that are used to further define and specify substances. Specified substances are always composed of at least one substance.

Specified substance Group 1 is typically used to define:

- Multi-substance materials consisting of multiple substances, which are not defined as mixture substances;
- Additional information regarding herbal and allergenic extracts;
- Physical state, including polymorphic forms;
- Detailed glycosylation information.

Specified substance Group 2 is typically used to define:

- Manufacturer and the overall manufacturing process and critical process version number;
- In addition there is the possibility to make use of the reference source document class to store Release specifications of the intended manufactured substance.

Specified substance Group 3 is typically used to define:

- Grade or level of purity (Pharmacopoeial Specifications) and In house specification used to cover a set of specifications of all approved manufacturers for the substance.

Specified substance Group 4 is typically used to define:

- Detailed Analytical Data;
- Detailed Manufacturing Information.

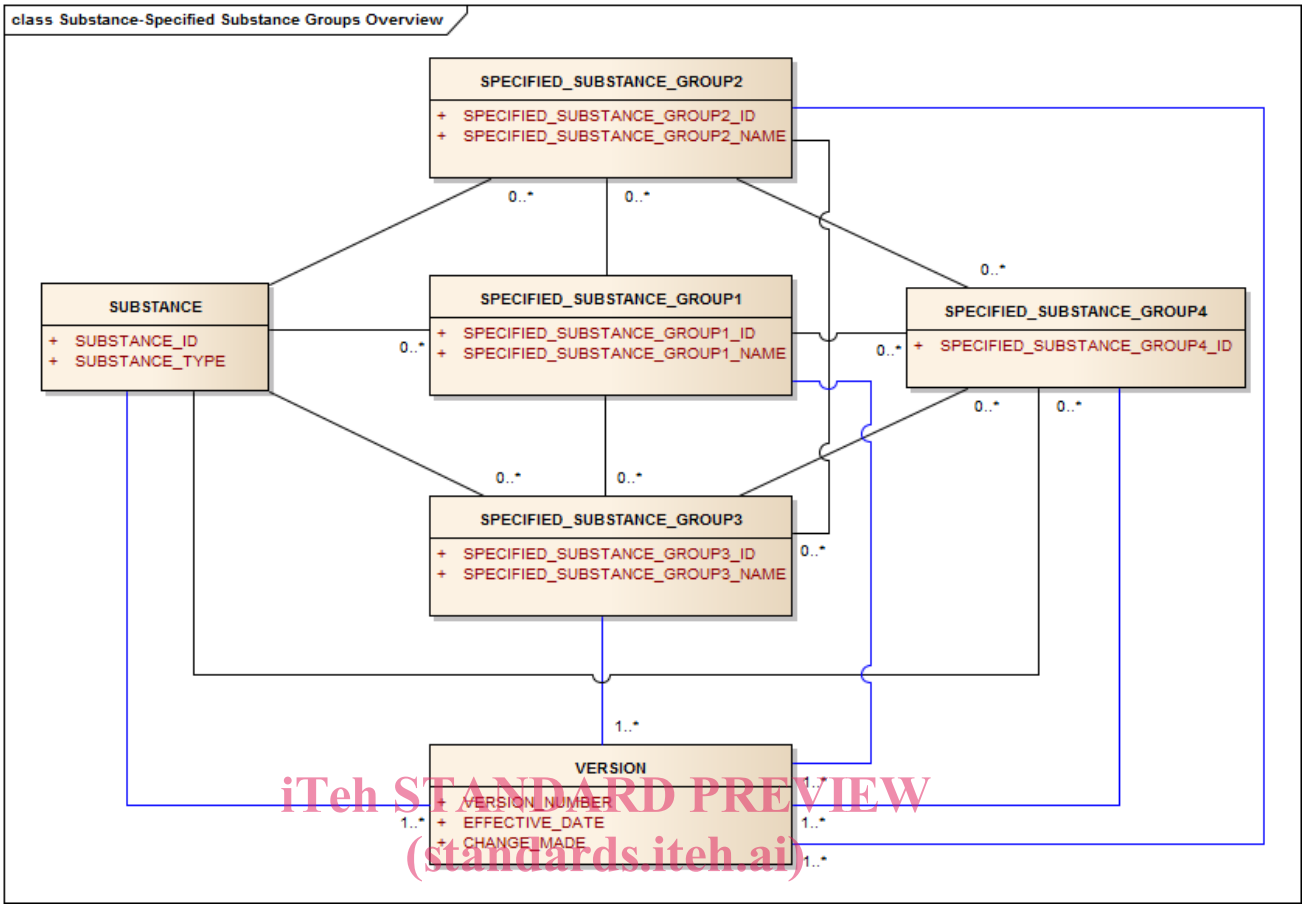

Figure 1 — High-level Substance-Specified substance information model

4.2 Defining substances

A substance is any matter that has a discrete existence, irrespective of origin, which may be biological or chemical. Substances can be single well-defined chemical entities containing a definite molecular structure, synthetic (e.g. isomeric mixtures) or naturally-occurring (e.g. conjugated oestrogens) mixtures of chemicals, or materials derived from plants, animals, microorganisms or inorganic matrices that are not definable by a single or limited number of molecular structures. Substances may be active moieties, salts, solvates, stoichiometric complexes or mixtures of compounds that are isolated or synthesised/ obtained or produced in a process together. Materials that are combined from multiple sources to form a product are not considered substances.

A substance is generally defined by what it is, and not by how it is made or used. Substance definitions are typically based on the immutable properties of a given material. These properties include the molecular structure, or structures of a given material, taxonomic, anatomical or fractionation information for material that cannot be represented by molecular structures. Purity, physical form, and method of production are typically not considered when defining substances.

In addition to defining information there is also information that is essential to validate the defining information and this information should be submitted if available. Validation may be performed based on relevant provided documents (e.g. regulatory dossier) or information available in recognised source (e.g. pharmacopoeias).

The primary goal of the ISO IDMP standard - *Data Elements and Structures for the Unique Identification and Exchange of Regulated Information on Substances* is to define unambiguously all substances present in regulated products. Once a substance has been defined, a unique identifier that is permanently associated with that substance will be assigned. This Technical Specification describes the necessary information for this registration process. Reference information, names, codes and IDs that can be associated with a substance are also described. The document is not comprehensive in regards to reference and definitional information. Other substance