PD CEN ISO/TS 19844:2017

BSI Standards Publication

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

Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

National foreword

This Published Document is the UK implementation of CEN ISO/TS 19844:2017. It is identical to ISO/TS 19844:2016. It supersedes PD CEN ISO/TS 19844:2015, which is withdrawn.

The UK participation in its preparation was entrusted to Technical Committee IST/35, Health informatics.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

© The British Standards Institution 2018
Published by BSI Standards Limited 2018

ISBN 978 0 580 93300 4

ICS 35.240.80

Compliance with a British Standard cannot confer immunity from legal obligations.

This Published Document was published under the authority of the Standards Policy and Strategy Committee on 31 August 2018.

Amendments/corrigenda issued since publication

Date	Text affected

Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

TECHNICAL SPECIFICATION

SPÉCIFICATION TECHNIQUE

TECHNISCHE SPEZIFIKATION

CEN ISO/TS 19844

August 2017

ICS 35.240.80

Supersedes CEN ISO/TS 19844:2015

English Version

Health informatics - Identification of medicinal products - Implementation guidelines for data elements and structures for the unique identification and exchange of regulated information on substances (ISO/TS 19844:2016)

Informatique de santé - Identification des médicaments - Lignes directrices pour la mise en oeuvre des éléments de données et structures pour l'identification unique et l'échange d'informations réglementées sur les substances (ISO/TS 19844:2016)

Medizinische Informatik - Identifikation von Arzneimitteln - Implementierungsleitfaden für Datenelemente und Strukturen zur eindeutigen Identifikation und zum Austausch von vorgeschriebenen Informationen von Substanzen (ISO/TS 19844:2016)

This Technical Specification (CEN/TS) was approved by CEN on 26 July 2017 for provisional application.

The period of validity of this CEN/TS is limited initially to three years. After two years the members of CEN will be requested to submit their comments, particularly on the question whether the CEN/TS can be converted into a European Standard.

CEN members are required to announce the existence of this CEN/TS in the same way as for an EN and to make the CEN/TS available promptly at national level in an appropriate form. It is permissible to keep conflicting national standards in force (in parallel to the CEN/TS) until the final decision about the possible conversion of the CEN/TS into an EN is reached.

CEN members are the national standards bodies of Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Former Yugoslav Republic of Macedonia, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey and United Kingdom.

EUROPEAN COMMITTEE FOR STANDARDIZATION
COMITÉ EUROPÉEN DE NORMALISATION
EUROPÄISCHES KOMITEE FÜR NORMUNG

CEN-CENELEC Management Centre: Avenue Marnix 17, B-1000 Brussels

© 2017 CEN All rights of exploitation in any form and by any means reserved worldwide for CEN national Members.

Ref. No. CEN ISO/TS 19844:2017 E

Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

European foreword

The text of ISO/TS 19844:2016 has been prepared by Technical Committee ISO/TC 215 "Health informatics" of the International Organization for Standardization (ISO) and has been taken over as CEN ISO/TS 19844:2017 by Technical Committee CEN/TC 251 "Health informatics" the secretariat of which is held by NEN.

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

This document supersedes CEN ISO/TS 19844:2015.

According to the CEN-CENELEC Internal Regulations, the national standards organizations of the following countries are bound to announce this Technical Specification: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Former Yugoslav Republic of Macedonia, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey and the United Kingdom.

Endorsement notice

The text of ISO/TS 19844:2016 has been approved by CEN as CEN ISO/TS 19844:2017 without any modification.

AVA Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

Contents

1 Scope 1
2 Normative references 2
3 General background and history 2
4 Substance (Mandatory) 3
4.1 General 3
4.2 Defining substances 5
4.2.1 Substance type (Mandatory) 7
4.2.2 Substance ID (Mandatory) 10
4.3 Substance names (Mandatory) 11
4.3.1 Substance name 12
4.3.2 Substance name type 13
4.3.3 Language 14
4.3.4 Official name (Conditional 14
4.4 Reference source (Conditional) 17
4.4.1 Public domain 17
4.4.2 Reference source type 18
4.4.3 Reference source class 18
4.4.4 Reference source ID 19
4.4.5 Reference source citation 19
4.5 Reference source document (Conditional) 19
4.5.1 Public domain 19
4.5.2 Reference source document 20
4.5.3 Reference source document type 20
4.5.4 Reference source document ID 21
4.5.5 Reference source document classification 21
4.5.6 Reference source document URL 21
4.6 Substance code (Conditional) 21
4.6.1 Code 22
4.6.2 Code system 22
4.6.3 Code system ID 23
4.6.4 Code system status 23
4.6.5 Code change date 24
4.6.6 Comment 24
4.6.7 Reference source 24
4.7 Reference information (Conditional) 24
4.7.1 Comment 25
4.7.2 Substance classification (Conditional) 25
4.7.3 Substance relationship (Conditional) 28
4.7.4 Target (Conditional) 30
4.7.5 Gene (Conditional) 33
4.7.6 Gene element (Conditional) 35
4.8 Structure 36
4.8.1 Structural Representation (Conditional) 36
4.8.2 Stereochemistry 42
4.8.3 Optical activity 43

© ISO 2016 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

4.8.4 Molecular Formula ... 44
4.8.5 Molecular Formula by Moiety ... 44
4.8.6 Molecular weight (Mandatory) ... 44
4.8.7 Isotope (Conditional) ... 44
4.9 Amount (Conditional) ... 46
4.9.1 Average ... 46
4.9.2 Low limit ... 47
4.9.3 High limit ... 47
4.9.4 Unit ... 47
4.9.5 Non-numeric Value ... 48
4.9.6 Reference Source (Conditional) ... 48
4.9.7 Reference source document (Conditional) ... 48
4.10 Source material (Conditional) ... 48
4.10.1 Source material class ... 49
4.10.2 Source material type ... 50
4.10.3 Source material state ... 50
4.10.4 Organism ID ... 50
4.10.5 Organism name ... 51
4.10.6 Parent substance ID ... 51
4.10.7 Parent substance name ... 51
4.10.8 Development stage ... 52
4.10.9 Part Description (CONDITIONAL) ... 52
4.10.10 Fraction (Conditional) ... 54
4.10.11 Organism (Conditional) ... 57
4.11 Modification (Conditional) ... 64
4.11.1 Modification type ... 66
4.11.2 Residue modified ... 66
4.11.3 Residue sites ... 66
4.11.4 Structural modification (Conditional) ... 67
4.11.5 Agent modification (Conditional) ... 69
4.11.6 Physical Modification (Conditional) ... 70
4.12 Property (Conditional) ... 72
4.12.1 Property type ... 72
4.12.2 Property name ... 73
4.12.3 Property parameters ... 73
4.12.4 Substance ID ... 73
4.12.5 Substance name ... 74
4.12.6 Amount type (Mandatory) ... 74
4.13 Version (Mandatory) ... 74
4.13.1 Version number ... 74
4.13.2 Effective date ... 75
4.13.3 Change made ... 75

5 Substance definitions ... 75
5.1 Chemical substance ... 75
5.1.1 Comment ... 76
5.1.2 Structure ... 77
5.1.3 Stoichiometric/Non-stoichiometric chemicals ... 77
5.1.4 Stoichiometric chemicals ... 78
5.1.5 Non-stoichiometric chemicals (Conditional) ... 81
5.1.6 Substance Name (Mandatory) ... 83
5.1.7 Substance Code (Conditional) ... 83
5.1.8 Version (Mandatory) ... 83

© ISO 2016 - All rights reserved
Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

5.1.9 Reference information ... 83
5.1.10 Reference source (Conditional) ... 83
5.1.11 Reference source document (Conditional) ... 83
5.2 Proteins/peptides ... 83
5.2.1 Microheterogeneity ... 84
5.2.2 Sequence type ... 86
5.2.3 Number of subunits ... 86
5.2.4 Disulfide linkage ... 86
5.2.5 Comment ... 87
5.2.6 Protein subunit (Mandatory) ... 87
5.2.7 Molecular weight (Conditional) ... 90
5.2.8 Glycosylation (Conditional) ... 91
5.2.9 Property (Conditional) ... 92
5.2.10 Structure (Mandatory) ... 93
5.2.11 Substance name (Mandatory) ... 93
5.2.12 Modification (Conditional) ... 93
5.2.13 Substance code (Conditional) ... 93
5.2.14 Source material (Conditional) ... 93
5.2.15 Version (Mandatory) ... 93
5.2.16 Reference information (Conditional) ... 93
5.2.17 Reference source (Conditional) ... 93
5.2.18 Reference source document (Conditional) ... 93
5.3 Nucleic acids ... 93
5.3.1 Structure (Conditional) ... 94
5.3.2 Sequence type ... 95
5.3.3 Number of subunits ... 95
5.3.4 Area of hybridisation ... 96
5.3.5 Comment ... 96
5.3.6 Nucleic acid subunit (Mandatory) ... 96
5.3.7 Modification (Conditional) ... 100
5.3.8 Property (Conditional) ... 100
5.3.9 Molecular weight (Conditional) ... 101
5.3.10 Substance Name (Mandatory) ... 101
5.3.11 Substance Code (Conditional) ... 101
5.3.12 Version (Mandatory) ... 101
5.3.13 Reference information (Conditional) ... 101
5.3.14 Reference source (Conditional) ... 101
5.3.15 Reference source document (Conditional) ... 101
5.4 Polymers ... 101
5.4.1 Polymer class ... 103
5.4.2 Polymer geometry ... 103
5.4.3 Copolymer sequence type ... 103
5.4.4 Comment ... 103
5.4.5 Substance name (Mandatory) ... 103
5.4.6 Structure (Mandatory) ... 104
5.4.7 Monomer description (Conditional) ... 104
5.4.8 Structural repeat (Conditional) ... 105
5.4.9 Molecular weight (Mandatory) ... 108
5.4.10 Property (Conditional) ... 108
5.4.11 Substance code (Conditional) ... 108
5.4.12 Version (Mandatory) ... 108
5.4.13 Reference information (Conditional) ... 108

© ISO 2016 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

5.4.14 Modification (Conditional) ... 108
5.4.15 Source material (Conditional) ... 109
5.4.16 Reference source (Conditional) ... 109
5.4.17 Reference source document (Conditional) ... 109
5.5 Structurally diverse substances ... 109
5.5.1 Comment ... 110
5.5.2 Substance name (Mandatory) ... 110
5.5.3 Structure (Mandatory) ... 110
5.5.4 Property (Conditional) ... 110
5.5.5 Molecular weight ... 111
5.5.6 Glycosylation (Conditional) ... 111
5.5.7 Modification (Conditional) ... 111
5.5.8 Source material (Conditional) ... 111
5.5.9 Substance code (Conditional) ... 111
5.5.10 Reference information (Conditional) ... 111
5.5.11 Version (Mandatory) ... 111
5.5.12 Reference source (Conditional) ... 111
5.5.13 Reference source document (Conditional) ... 111
5.5.14 Herbals and substances used in the preparation of plant-based allergenic extracts ... 111
5.5.15 Vaccines ... 114
5.5.16 Plasma-derived substance for human blood products and polyclonal antibodies ... 114
5.5.17 Allergens ... 114
5.5.18 Advance Therapies and Advanced Vaccines (Genes, Modified Viruses, Cells and Tissues as Substances) ... 115
5.5.19 Minerals ... 115
5.6 Mixture substance ... 116
5.6.1 Mixture type ... 116
5.6.2 Mixture constituent (Mandatory) ... 116
5.6.3 Modification (Conditional) ... 117
5.6.4 Source material (Conditional) ... 117
5.6.5 Substance name (Mandatory) ... 117
5.6.6 Substance code (Conditional) ... 117
5.6.7 Reference information (Conditional) ... 118
5.6.8 Version (Mandatory) ... 118

6 Specified substance (Optional) ... 118
6.1 Specified Substance Group 1 (repeat as necessary) ... 118
6.1.1 Specified substance Group 1 ID ... 119
6.1.2 Specified substance Group1 Name ... 120
6.1.3 Substance Name (Mandatory) ... 120
6.1.4 Substance Code (Conditional) ... 120
6.1.5 Version (Mandatory) ... 120
6.1.6 Reference source (Conditional) ... 120
6.1.7 Reference source document (Conditional) ... 120
6.1.8 Property (Conditional) ... 120
6.1.9 Fraction (Conditional) ... 121
6.1.10 Modification (Conditional) ... 125
6.1.11 Reference Information (Conditional) ... 125
6.1.12 Constituent (Mandatory) ... 125
6.1.13 Physical form (Conditional) ... 127
6.1.14 Specified substance particulars ... 128
6.2 Specified substance Group 2 ... 133
6.2.1 Specified Substance Group2 ID ... 135

© ISO 2016 - All rights reserved
Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

6.2.2 Specified Substance Group2 Name ... 136
6.2.3 Parent Substance ID ... 136
6.2.4 Reference source (Conditional) ... 136
6.2.5 Reference source document (Conditional) ... 136
6.2.6 Manufacturing (Mandatory) ... 136
6.3 Specified Substance Group 2 for Herbal preparations ... 142
6.3.1 Specified Substance Group2 ID ... 142
6.3.2 Specified substance Group2 Name ... 142
6.3.3 Parent Substance ID ... 142
6.3.4 Manufacturing ... 143
6.3.5 Version ... 144
6.4 Specified Substance Group 3 ... 144
6.4.1 Specified Substance Group 3 ID ... 145
6.4.2 Specified Substance Group3 Name ... 145
6.4.3 Parent Substance ID ... 145
6.4.4 Grade (Mandatory) ... 145
6.4.5 Version (Mandatory) ... 146
6.4.6 Reference source (Conditional) ... 146
6.4.7 Reference source document (Conditional) ... 147
6.4.8 Substance name (Mandatory) ... 147
6.4.9 Substance code (Conditional) ... 147
6.4.10 Version (Mandatory) ... 147

7 Description of the information modelling principles and practices ... 147

Annex A (normative) Choosing a Substance ID ... 148

Annex B (normative) Chemical substance ... 150

Annex C (normative) Protein substance ... 270

Annex D (normative) Nucleic acid substance ... 329

Annex E (normative) Structurally Diverse Substance – Herbal Substance/Herbal Specified Substance ... 348

Annex F (normative) Structurally Diverse Substance, Homeopathic substance ... 466

Annex G (normative) Structurally Diverse Substance – Plasma-derived substances ... 509

Annex H (normative) Polymer Substance ... 581

© ISO 2016 – All rights reserved

Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

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*.

This second edition cancels and replaces the first edition (ISO/TS 19844:2015), which has been technically revised.

© ISO 2016 - All rights reserved
Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

Introduction

This document provides guidelines for implementing ISO 11238. This document is 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 document, 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 subclause is identified as 'optional' but is implemented in a specific region, conformance described within that subclause is applicable. The scope of this document is to identify the scientifically necessary elements for the unique identification of Substances/Specified Substances.

© ISO 2016 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

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 document is used in the implementation of ISO 11238. This document 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 document 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 second edition of the document addresses substances, Groups 1 to 3 of the Specified Substances as defined in ISO 11238 and Annexes A, B, C, D, E, F, G and H. It is anticipated that Specified Substances Group 4, as defined in ISO 11238, will be addressed in a subsequent edition of this document. Some information that would typically fall under Specified Substances Group 4 may be covered in the Annexes of this document. This information, although not defining of either a Substance or a Specified Substance Group 1, may be essential to distinguishing substances. This document 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 document 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.

© ISO 2015 – All rights reserved

AVA Documents Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements 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 document 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.

© ISO 2016 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

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 refers 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 become 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 document 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:

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

- 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 are captured at the substance level;

- For nucleic acids, the sequence, type of sugar and linkage are captured at the Substance level;

- For other polymers, the monomers used to synthesize the polymer, the structural repeat 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 are 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 are 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;

© ISO 2016 – All rights reserved
AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

- 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.

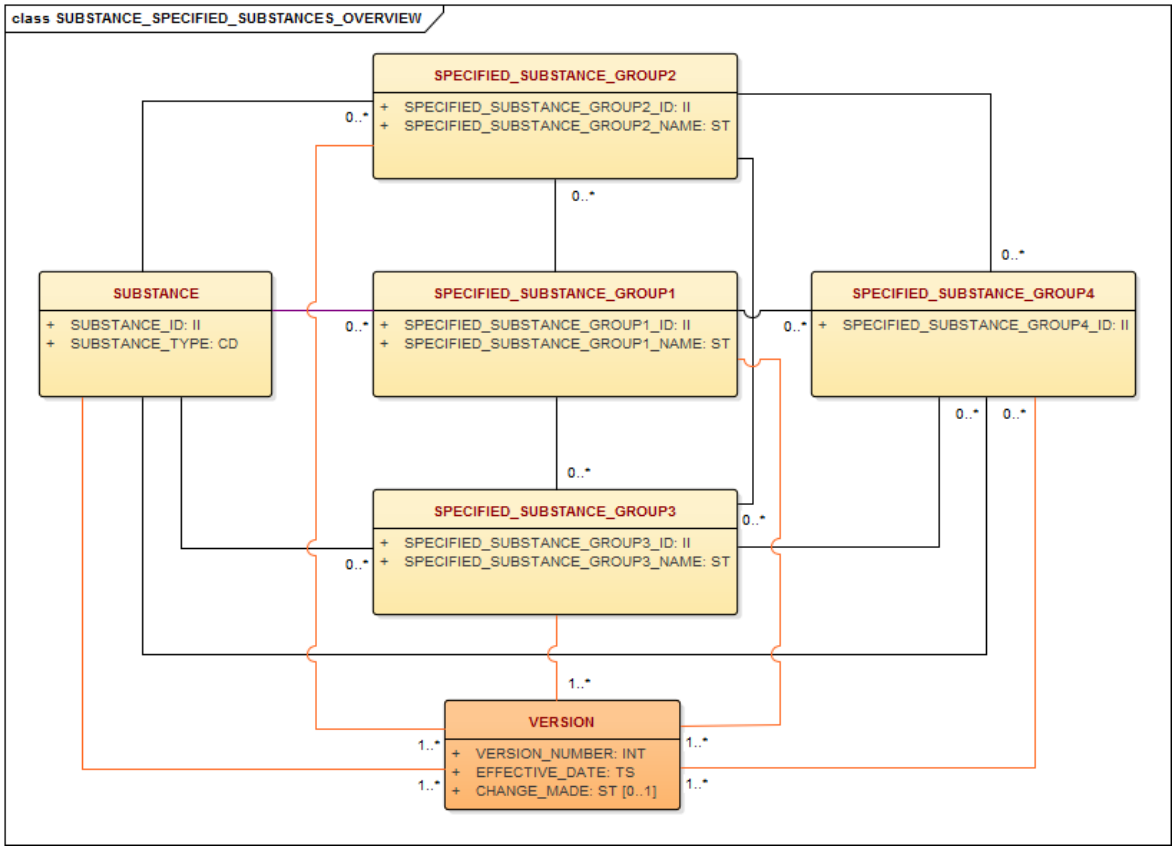


Figure 1 — High-level Substance-Specified Substance information model

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

© ISO 2016 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

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 ISO 11238 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 document 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 attributes may be associated for both reference and definitional information as new scientific standards or information becomes available.

Substances that are mixed together to form a product, even if there are physical interactions between the Substances will be described as a Specified Substance or a product.

EXAMPLE 1 FD&C Blue No. 1 Aluminium Lake is not defined as a mixture substance but rather as two separate substances FD&C Blue No. 1 and aluminium oxide. Also, simethicone is not a substance because it contains two substances, namely dimethicone and silicon dioxide. These materials would be described as Group 1 Specified Substances.

Conversely, if new covalent bonds and/or simple salts or solvates are formed, separate Substance IDs for each molecular entity are assigned.

EXAMPLE 2 Fluticasone propionate would have a separate ID from fluticasone; atorvastatin calcium trihydrate would have a separate ID from atorvastatin calcium anhydrous.

Differences at the Specified Substance level should trigger a new Specified Substance ID.

EXAMPLE 3 Insulin Human Crystalline and Insulin Human Amorphous have different Specified Substances Group 1 IDs and map to the substance Human Insulin.

© ISO 2016 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

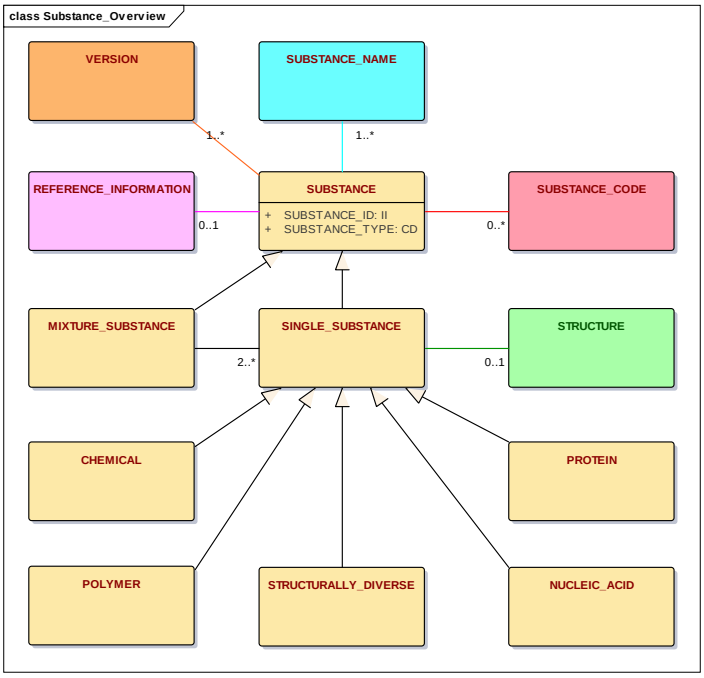


Figure 2 — High-level information model for substances

4.2.1 Substance type (Mandatory)

At the basic level, all material present in medicinal products will either be described as a Substance and/or a Specified Substance Group 1. Additional information, such as manufacturer, grade and manufacturing information on a given material will be associated with Specified Substance Groups 2 to 4.

Substances shall be defined using one or more of the five types of single substances or a mixture. Mixtures consist of a combination of related single substances where the source material and essential properties are also captured but proportions are not. For the purposes of definition, all single substances are of one of the following types:

- (simple) chemical;
- protein;
- nucleic acid;
- polymer;
- structurally-diverse.

Each of these different types of substances has distinct elements essential for definition. Substances can also be a **mixture of single substances,** typically related and of the same type. Some materials may appear to have elements from multiple types of substances e.g. PEGylated proteins, drugs covalently linked to polymeric matrices or proteins, etc., to fully define the substance. Each of these substances will

© ISO 2016 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

be described as a single type, although it is possible for the same substance to be defined by using two or more substance types of which one is the predominant one.

EXAMPLE 1: Proteins modified with a polymer or drug will be described as a protein with the polymer or drug captured as a modification. Drugs covalently linked to a polymer will be described as a polymer with the drug either captured as an end group or as structural modification. Another example can be structurally diverse substances such as vaccines containing a protein fraction, which should also be described using elements from the protein model. Detailed examples will be provided in the Annex to address these complex materials. Although every substance definition will be defined with a single type it is possible that an alternative definition of a substance will be of a different type.

EXAMPLE 2: Desmopressin can either be defined as a modified peptide or as a chemical substance.

Mixtures consist of a combination of related single substances where the source material and essential properties are also captured but proportions are not. At the Specified Substance Group 1 level proportions can be captured.

Additional information, such as manufacturer, grade, and manufacturing information on a given material will be associated with Specified Substances Groups 2-4. The first step in defining a material is to determine whether the material can be defined as a Single Substance, a Mixture of single substances or a Specified Substance Group 1. A mixture substance is a substance that cannot be defined using a single set of substance elements or structural representations. Racemic mixtures, most polymers and structurally diverse substances, although inherently mixtures, can be defined using a single set of elements and therefore are not described as mixtures. Material that contains multiple substances that are not related and not synthesised or isolated together will be described as Specified Substance Group 1. Specified substance Group 1 would also include crystalline polymorphs, substances in varied physical states along with the capture of additional properties regarding sterility or the presence of endotoxin.

Examples of Specified Substance Group 1 include simethicone, adjuvanted vaccine antigens, isophane insulins and additional herbal and allergenic extract information not captured at the Substance level. Sterile water, water for injection, ice, and steam would be different Specified Substances Group 1 that would map to the substance water.

Mixtures are substances like Gentamicin sulfate, USP or Gentamicin sulfate, Ph.Eur. which consists of mixtures of distinct but related individual aminoglycosides. In USP Gentamicin sulfate is described as a mixture of Gentamicin C1A sulfate, Gentamicin C1 sulfate, and Gentamicin C2 sulfate. In Ph.Eur. it is described as a mixture of five components Gentamicin C1, C1a, C2, C2a, and C2b respectively which are typically isolated together from a bacterial culture. At the substance level all related substances typically present at a level greater than 1% would be listed as components of the mixture regardless of whether the substance is listed in a given pharmacopoeia. Mixture Substances are defined in the ISO 11238 standard as a simple combination of single substances that are either isolated together or are the result of the same synthetic process.

© ISO 2016 – All rights reserved
AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

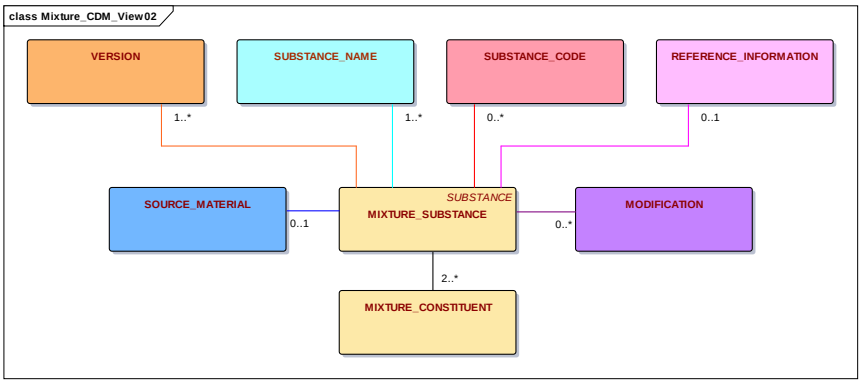
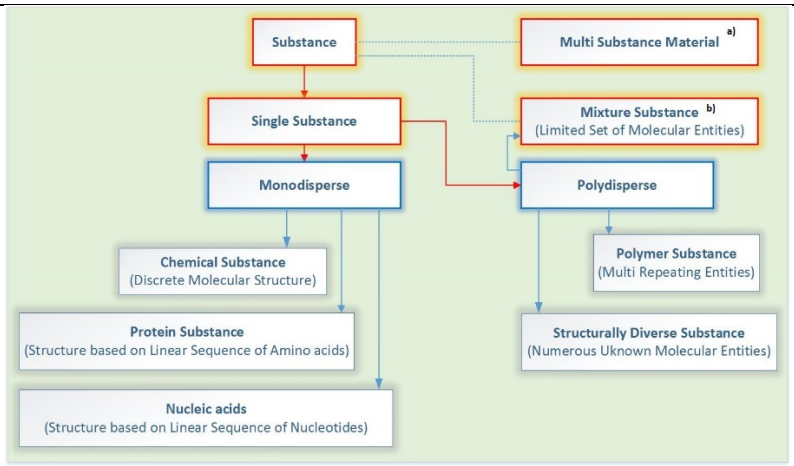


Figure 3 — High-level information model for Mixture Substance

For mixture substances in which multiple related substances are isolated together from a single source, the source will be a defining element. Starches are typically a mixture substance of the two distinct polymers amylose and amylopectin. Corn, whexat, potato and tapioca starches are all distinct substances distinguished by the source.

Most other polymers will be described as single substances with the definition based on the structural repeat unit and a representation of the average molecular weight of the polymer as well as the biological source for biopolymers or polymers derived from biological matrices.

Figure 4 represents a decision tree or process by which the type of substance is determined. Monodisperse substances are substances that can be described as a single molecular entity. Polydisperse substances are substances that typically have multiple molecular entities that are too numerous or too diverse to be captured as a mixture (e.g. cells or tissues) or where the production of the substance inherently results in polydispersity (e.g. polymers).



a) Multi-substance material: Multi Substances and/or Specified Substances of diverse origin.

b) Mixture Substance: Type of polydisperse substance that is a combination of single substances isolated together or synthesised/obtained or produced in the same process.

Figure 4 — Decision tree or process by which the type of substance is determined

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

Table 2 — Substance types

User Guidance	Information about the substance type as described in and in accordance with ISO 11238 definitions
Example(s)	"Chemical ", "Protein" "Nucleic acid, "Polymer", "Structurally Diverse"
Conformance	MANDATORY
Data Type	CD
Values Allowed	Chemical, Protein, Nucleic Acid, Polymer, Structurally Diverse, Mixture, Specified Substance Group 1, Specified Substance Group 2, Specified Substance Group 3, Specified Substance Group 4

4.2.2 Substance ID (Mandatory)

Every substance will be identified by an ID. An ID will be generated following the initial submission of substance data. The ID should be unique to each substance, non-semantic, non-chronological, and of fixed length with an integrity check. Once an ID is assigned to a given physical or conceptual material the ID will not change. Changes or additions to defining information result in a new version of the substance definition but would not result in a new Substance ID. Although an effort should always be made to obtain complete information prior to assignment there will be instances when additional definitional information on a given substance is transmitted after assignment. When incorrect or incomplete information may be transmitted and used to define a substance the information will be changed or added but the ID will not change. In rare cases when the same substance is found to have two IDs one of the IDs will remain a primary ID and the use of the other ID will be deprecated and captured in the Substance Code class. For some materials there may be multiple IDs that a material could map to. A submitter should always submit the ID that is the most descriptive of the material.

EXAMPLE Glucose, when a solid crystalline material can exist with an alpha or beta pyranose structure in an anhydrous state, or as a monohydrate with an alpha pyranose structure. If a solid crystalline form of glucose exists in a medicinal product the ID specific to the structure should be used. Glucose in a liquid or an amorphous solid state would use the Substance ID that does not specify the specific form.

Both Substances and Specified Substances will be identified with an ID. IDs will be released to the public if the substance is in a licensed medicinal product. The IDs for substances in the investigational stage will only be released if an official name exists or a company code is associated with defining information and is found together in at least one single reference that is from a reputable source in the public domain (i.e. scientific journal, presentations or posters at scientific conferences, company publication, patent or published patent application, public databases such as STN from CAS). Substance IDs will always be issued to the requester if sufficient information is provided to unambiguously define the substance.

© ISO 2016 – All rights reserved
AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

Table 3 — Substance ID

User Guidance	ID to be used in all electronic submissions to identify a substance. Generated when sufficient information is available to unambiguously define a substance. ID will be permanently associated with a given substance and each substance at the substance level shall have one and only one ID. NOTE If a unique "Substance ID" has been assigned, this "Substance ID" is specified based on the Substance Name controlled vocabulary. In the absence of a unique "Substance ID", e. g. for the initial submission of the substance, this data element is not mandatory
Example(s)	UNII Code, EV-Code, Global Substance ID[1] (DEVTYS543H)
Conformance	MANDATORY
Data Type	II
Values Allowed	Value could be a code associated with a preferred term, or a specific type of data. If a code is transmitted the preferred term associated with that code should also be transmitted.
Business Rule(s)	All substances will be identified by a single ID. NOTE The ID will only be released to the public if the defining information is in the public domain or if a company that provides the defining information requests public release or releases the code in public marketing materials. An ID will always be issued to an organisation that requests an ID and supplies the information necessary to define a substance. A flag to control the release of the ID to the general public is part of the Reference Source information 4.4. Defining information found in patents will usually not be sufficient to release the ID to the public.

4.3 Substance names (Mandatory)

Although ISO 11238 does not provide any guidance on substance nomenclature, it does provide a structure for the capture of names and codes that are used to refer to a substance. For the purpose of ISO 11238 there are basically eight types of names, Official names, systematic names, other names, brand names, company codes, homeopathic names, specified homeopathic names and plasma-derived names.

At least one name or company code should be associated with each substance. Names for a given substance and/or standardised names submitted by another party may be available in the substance terminology. Indication that a name is the preferred name or term by which a substance will be referred to in a given jurisdiction and domain may be presented in the terminology. This preferred term may vary depending on jurisdiction or domain or the state of development of a given substance.

EXAMPLE During the early stages of clinical development, the preferred term for a substance may be its company code. In later stages an INN name may be associated with the substance and become the preferred term for that substance.

[1] Global Substance ID is corresponding to the ISO IDMP identifier of a Substance intended to be used globally.

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

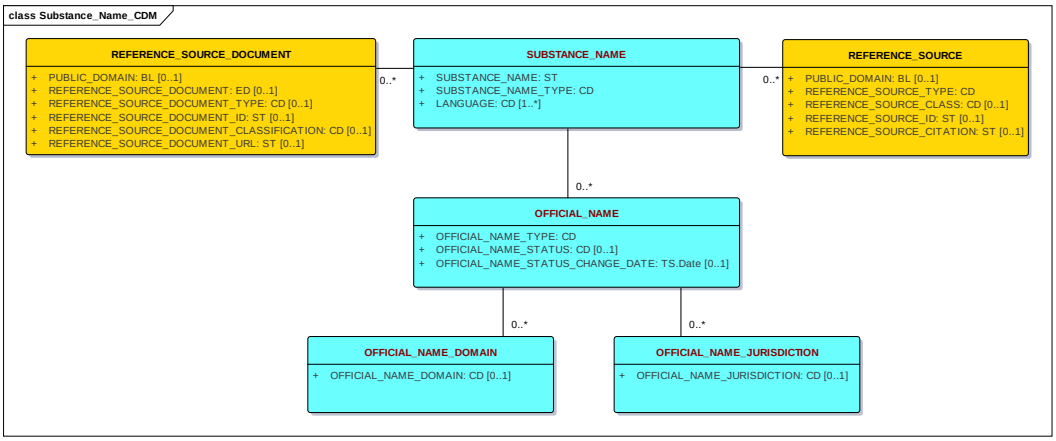


Figure 5 — Information model for Substance Names

4.3.1 Substance name

When requesting an ID, a submitter should submit all names and company codes which have been associated with the substance. A submitter may also want to add a name to a substance that is already in the database. Defining information for a given substance should also be transmitted when requesting that a name is added to the database to ensure the correct substance identity. A submitter may also submit names in multiple languages. Although there is no requirement that names are unique, other or common names should also explicitly distinguish related substances from one another. Distinguishing characteristics for many polymers, such as molecular weight, degree of polymerisation, and viscosity, should be indicated in the names of these substances. Standardisation on such names to facilitate searching and identification of substances may be adopted within regions.

EXAMPLE Polyethylene glycol is a name for a family of polymers, so it would not be an acceptable name for an individual substance. Polyethylene Glycol 1000 is an official name in the US for a polyethylene glycol polymer with an average molecular weight of 1000. In the European Union the INN name Macrogol 1000 would be an official name for the same substance. In the personal care products domain PEG-20 is the official name for the same substance and 20 refers to the average degree of polymerisation.

© ISO 2016 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

Table 4 — Substance name

User Guidance	The Name or company code associated with the Substance shall be specified. For substances originated from Plants the Substance is always a single species, or infraspecies, or is occasionally a single genus. It will have a Name which is composed of: [Scientific genus/binomial/trinomial with Author] [Part(s)] The scientific element of the name is usually a binomial, but may be a trinomial or a genus, as appropriate, together with the author of the name. The Name is not an official name in any pharmacopoeia or other legislation, and the inclusion of the author is to ensure that it is not confused with the official names for specified substances that are derived from a single species substance. It can be written in any language and its jurisdiction can be restricted using Name Status.
Example(s)	"Amoxicillin"; Harpagophytum procumbens (Burch.) DC. ex Meisn., Root
Conformance	MANDATORY
Data Type	ST
Values Allowed	Free text, multiple names allowed, each in its own <asNamedEntity> element.
Business Rule(s)	Mandatory all substances shall have at least one name or company code (also considered to be a Substance Name) associated with the substance.

4.3.2 Substance name type

Every name shall have one and only one name type. Official names are typically non-proprietary names used in a given jurisdiction and domain to refer to a specific substance. The domain, jurisdiction, and authority that assigned the name (USAN, INN, JAN etc.) and the language of the name are also captured.

Systematic names will typically be used for simple chemicals and structurally diverse materials where the definitions are based on a chemical structure or systematic taxonomic information. For chemicals, these names are typically derived from IUPAC or CAS systems of nomenclature. From the systematic names of chemicals, a molecular structure can typically be derived and the name can be checked by a number of chemical drawing programmes that convert a given name to a molecular structure. Brand names are names by which a company identifies a given substance typically for marketing purposes. Other names associated with a given substance are for instance common names. Company codes (also known as lab/laboratory codes or clinical trial codes) are also treated as names. They are assigned by a given company to substances in clinical or preclinical development.

Table 5 — Substance name type

User Guidance	Each name shall be associated with a type.
Example(s)	Official name, systematic name, scientific name, other name, brand name, company code, homeopathic name, specified homeopathic name, plasma-derived name, preferred name
Conformance	MANDATORY
Data Type	CD

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

Table 5 *(continued)*

Values Allowed	OFFICIAL, SYSTEMATIC, SCIENTIFIC, COMPANY CODE, OTHER, BRAND, HERBAL, HOMEOPATHIC, SPECIFIED HOMEOPATHIC, PLASMA-DERIVED, PREFERRED NAME
Business Rule(s)	A Substance Name shall always correspond to at least one and only one Substance Name Type. The Preferred Name is the name that describes at least all moieties of a substance and shall be used in all EU languages, e.g. Benzathine Benzylpenicillin tetrahydrate (Dutch: Benzathinebenzylpenicillinetetrahydraat). For the Substance (fresh) (e.g. Harpagophytum procumbens (Burch.) DC. ex Meisn., Root) the Substance Name Type shall be 'Other'.

4.3.3 Language

All names shall be associated with at least one language. Often the name, particularly INN names, will be identical in multiple languages. The language for company codes that use a Latin character set shall be described as English (en). The language for the systematic name of plant, microorganism or animal derived substances shall also be described as English since many of these will be a hybrid of a Latin Name for the species and an English name for the part and/or the subspecies information.

EXAMPLE The INN name alisertib is identical in English, French and Spanish.

Table 6 — Language

User Guidance	The language used to provide the structured substance information shall be specified according to the XML standard's xml:lang attribute specification, which is based on ISO 639-1, alpha-2 codes
Example(s)	en – English, de – German, fr –French
Conformance	MANDATORY
Data Type	CD
Values Allowed	ISO 639-1, alpha-2 codes
Business Rule(s)	While xml:lang allows country-specification of language, e.g. en_US vs. en_UK or pt_PT vs. pt_BR, such specificity is to be avoided except if the country variant of the language exceptionally does give rise to an alternative spelling. In that case, only the exceptional spelling needs to be tagged with the country-specific language code, not the regular spelling for that language in all other countries. If the same name spelling exists for multiple languages of interest, each language shall be specified.

4.3.4 Official name (Conditional

Each official name shall have one or more official name types. The official name type reflects the organisation that assigns or recognises the name associated with the substance. These names are typically non-proprietary names that are used in the labelling of pharmaceuticals. The domains and jurisdictions in which the official name is used are also captured, tracked and maintained within the

© ISO 2016 – All rights reserved
AVA Documents Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

terminology. Although most official names will refer to material at the substance level there will be instances where an official name will refer to material at the Specified Substances Group 1 level [e.g. Simeticone (BAN, INN) or Simethicone (USAN)]. The official name class is optional and shall be repeated as necessary.

4.3.4.1 Official name type

Any authority which assigns official names will typically do so in a publication, and as such, referencing the publication is sufficient indication for the name being "official" and sufficient representation for the "official name type". When new official names are issued by official naming bodies, the naming bodies will be reflected in this terminology.

Table 7 — Official name type

User Guidance	Designation of which authority assigned the Official Name. All official names need to have at least one such designation. The name type is the name of the authority or authorities that have assigned or have adopted the name.
Example(s)	BAN, COSING, Ph.Eur., FCC, INCI, INN, JAN, JECFA, MARTINDALE, USAN, USP, PHF, HAB, PhF (Pharmacopée française), IUIS
Conformance	MANDATORY — all official names require a designation of the naming authority
Data Type	CD
Values Allowed	All values in the name assigning authority identifier system.
Business Rule(s)	All official names shall have a naming authority designated. If multiple authorities have assigned the same name, a separate <asNamedEntity> is specified for each authority that needs to be listed.

4.3.4.2 Official name status

The status of the official name is associated with a given jurisdiction. It is possible that an official name could be official in one jurisdiction and superseded in another. With some official name types an initial name will be proposed and the name will later become recommended. Recommended INN names would be considered to be the Official name in many jurisdictions.

Table 8 — Official name status

User Guidance	The status of the official name.
Example(s)	Current, Alternate, Superseded, Proposed, Primary
Conformance	OPTIONAL
Data Type	CD
Values Allowed	Current, Alternate, Superseded, Proposed, Primary
Business Rule(s)	Each official name shall have a single status. The fact that a name is alternate is evident by having more than one name assigned to the same substance by the same authority for the same language, domain, etc.

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

4.3.4.3 Official name status change date

The change date of the status of the official name shall be specified according to the ISO 8601 date format (the variant without any delimiters).

Table 9 — Official name status change date

User Guidance	Date of official name change if known.
Example(s)	20110601
Conformance	OPTIONAL
Data Type	TS.DATE
Business Rule(s)	If official name status changes the date of the official change will be captured if known.

4.3.4.4 Official name domain (CONDITIONAL)

Official names will also be associated with a domain. The domain and the actual name of a domain will vary according to jurisdiction.

EXAMPLE The terms DRUG and BIOLOGIC could be used in Japan and the US as official name domain. In the EU the term MEDICINE may cover the name of many substances that would be either DRUGS or BIOLOGICS in the US and Japan.

Table 10 — Official name domain

User Guidance	Specifies when and for what purpose the official name is to be used. All official names can have at least one domain. This is useful to differentiate different names for the same substance as used for a drug active ingredient as opposed to a food colour additive.
Example(s)	BIOLOGIC, COSMETIC, DRUG, FOOD, MEDICINE, DIETARY SUPPLEMENT
Conformance	CONDITIONAL
Data Type	CD
Business Rule(s)	Used if different names exist that are not appropriate to be used across all domains.

4.3.4.5 Official name jurisdiction (CONDITIONAL)

When a product is on the market and contains pharmaceutically active substances at least one jurisdiction should be associated with the name of the active substance. For inactive substances or excipients, the jurisdiction may not be apparent or captured. Pharmacopoeias will often be consulted to determine the jurisdiction of an official name. The Official Name Jurisdiction class is optional; when necessary the following conformances apply:

© ISO 2016 – All rights reserved
AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

Table 11 — Official name jurisdiction

User Guidance	The jurisdiction of the official name shall be provided.
Example(s)	US, EU, JA
Conformance	CONDITIONAL
Data Type	CD
Values Allowed	ISO 3166-1 alpha-2 codes. For the European Union EU shall be used.
Business Rule(s)	All the jurisdictions in which the Official Name is to be used should be indicated.

4.4 Reference source (Conditional)

All data elements associated with a Substance or Specified Substance should have a reference source. These sources could be regulatory submissions, publications by naming authorities such as INN or USAN, pharmacopoeias, books, journals, web sites, meeting materials of public databases such as chemical abstracts. A source can refer to multiple data elements and a given element can have multiple sources.

4.4.1 Public domain

Most reference sources other than regulatory submissions will be in the public domain. All regulatory submissions will typically be considered to not be in the public domain. If the reference source is in the public domain, the data elements referenced by the source could be made public. There may be a separate determination whether any portion of the record can be made public. This determination may vary depending on type of substance. Regardless submitters of information should always explicitly indicate whether data should be considered confidential.

EXAMPLE Particular care should be taken in the release of company codes. Although the company codes are frequently mentioned in the public domain the release of a company code associated with a Substance ID and the defining information should only be made if all of the defining elements are associated with the company code in a single source in the public domain. For chemicals, the chemical structure should be associated with the company code in the public domain. For proteins such as monoclonal antibodies, names or company codes could be released to the public only if the target of the antibody is mentioned in the same public source as the name or code. A separate determination will be made as to whether the protein sequence is in the public domain.

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

Table 12 — Public domain

User Guidance	Indication of whether the source is in the public domain. Company codes will only be released if the defining information is associated with the code in a public source such as chemical abstracts (CAS), a journal article, or a poster at scientific meeting attributable to the company.
Example(s)	Yes, No
Conformance	OPTIONAL
Data Type	BL
Values Allowed	Yes, No
Business Rule(s)	All regulatory submissions are considered not to be in the public domain. There may be an additional flag that determines whether any information can be released on a given substance.

4.4.2 Reference source type

All reference sources will be associated with a type. The type may actually identify the source of the information if it is a public database or naming authority. For official names and defining information, the primary source should be used if available. A single source can be used for many data elements. Public databases particularly those of official naming bodies, regulatory submissions, manufacturer‘s technical specifications, scientific journal articles and pharmacopoeias are all sources that can be used.

Table 13 — Reference source type

User Guidance	The reference source type in which the data elements were actually found.
Example(s)	BP, CHEMID (NLM), Ph.Eur., USP/NF, IND (Investigational New Drug), INN, ITIS, JAN, JOURNAL, JP, KEGG, MARTINDALE, NDA, PERSONAL CARE PRODUCTS COUNCIL (PCPC), PUBCHEM, USAN, USP, WEB PAGE, CAS Registry, FDA-(G)-SRS, WHO Collaborating Centre for Drug Statistics methodology (WHOCC), CHEMICAL BOOK, ORANGE BOOK, JAVA APPLICATION, MAA (Marketing Authorisation Application), Allegome-A database of allergenic molecules, Uniprot
Conformance	MANDATORY
Data Type	CD
Business Rule(s)	All names should have at least one source and hence source type.

4.4.3 Reference source class

Each reference information type will be associated with a class. The class will be automatically associated with each type of information.

© ISO 2016 – All rights reserved
AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

Table 14 — Reference source class

User Guidance	Useful for classification of sources.
Example(s)	Regulatory submission, Public Database
Conformance	OPTIONAL
Data Type	CD
Values Allowed	LITERATURE, OFFICIAL NAME SOURCE, OTHER NAME SOURCE, REGULATORY SUBMISSION, PUBLIC DATABASE, WEB.
Business Rule(s)	Need not be entered as the value is dependent on the Reference Source Type.

4.4.4 Reference source ID

Many sources such as regulatory submissions, INN lists and others will be associated with a number or ID. The ID should be captured.

Table 15 — Reference source ID

User Guidance	An ID associated with reference source.
Example(s)	IND number, INN list; PMID (PubMed Identifier).
Conformance	CONDITIONAL
Data Type	ST

4.4.5 Reference source citation

A citation should be captured for a literature reference. The specific format for a given citation may be defined at regional level. In case the Reference source citation is a Web Page or Public Database, the URL should be provided.

Table 16 — Reference source citation

User Guidance	The actual reference source or citation in which the name was found.
Example(s)	Literature reference: Encycl. 3: 429 (1792) WHO Drug Information Vol. 25, no 3, 2011 Recommend INN: list 66
Conformance	CONDITIONAL
Data Type	ST

4.5 Reference source document (Conditional)

4.5.1 Public domain

Most sources other than regulatory submissions will be in the public domain. All regulatory submissions will typically be considered to not be in the public domain. If the reference source is in the public domain, the data elements referenced by the source could be made public. There may be a

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

separate determination whether any portion of the record can be made public. This determination may vary depending on type of substance. Regardless submitters of information should always explicitly indicate whether data should be considered confidential.

Table 17 — Public domain

User Guidance	Indication of whether the source is in the public domain.
Example(s)	Yes, No
Conformance	OPTIONAL
Data Type	BL
Values Allowed	Yes, No
Business Rule(s)	All regulatory submissions are considered not to be in the public domain. This flag determines whether any information can be released on a given substance.

4.5.2 Reference source document

The document should be concise and contain information specific for given data elements.

Table 18 — Reference source document

User Guidance	Documents that contain relevant information on data elements and are referenced should be stored if available.
Example(s)	PDF, TEXT
Conformance	OPTIONAL
Data Type	ED
Values Allowed	The documents can be saved as PDF files or TEXT files. If the documents or data source contains structured data such as spectra, it can be transmitted in a standard format that can be visualised and searched.

4.5.3 Reference source document type

Table 19 — Reference source document type

User Guidance	The type of the document
Example(s)	Pharmacopoeial monograph, Regulatory submission, Journal article, Release specification, In House specification, Text file
Conformance	OPTIONAL
Data Type	CD

© ISO 2016 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

4.5.4 Reference source document ID

Table 20 — Reference source document ID

User Guidance	An ID associated with reference source document.
Example(s)	FR4563 (Artificial ID)
Conformance	OPTIONAL
Data Type	ST

4.5.5 Reference source document classification

A classification system will be also developed for documents. The classification system will indicate the type of information contained within the reference document based on a terminology for the classification of reference source documents.

Table 21 — Reference source document classification

User Guidance	The reference document should be classified according to the information contained within
Example(s)	Chemical structure, Name, NMR spectra, Mass Spectra, In House/Batch release specification
Conformance	OPTIONAL
Data Type	CD

4.5.6 Reference source document URL

Table 22 — Reference source document URL

User Guidance	The reference source document URL in which the document was found. The URL can be captured here.
Example(s)	http://www.who.int/medicines/publications/druginformation/issues/RL_66.pdf
Conformance	CONDITIONAL
Data Type	ST

4.6 Substance code (Conditional)

This subclause of reference information allows the capture of codes that are typically associated with a substance. The capture of these codes typically facilitates the defining and mapping of substances and linking of substances to a variety of information sources. Codes should be communicated if known but are not necessary or mandatory. The terminology should have the ability to capture, store and distribute codes as necessary. All the codes that are captured should be associated with a publicly recognised code system and map directly to a given substance. It should be noted that company codes are not captured in this subclause but are considered a type of name for a given substance. Commonly

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

used public codes such as CAS registry numbers with a given substance should be provided and associated to the substance.

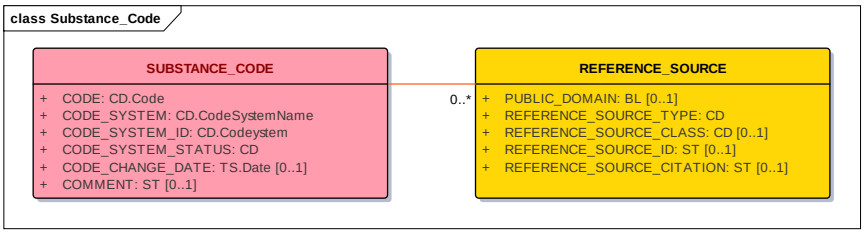


Figure 6 — Information model for substance codes

4.6.1 Code

The actual code shall be captured using the same format that is used in the code system. Only codes associated with a code system shall be captured. The code shall be specifically associated with a given substance. Many public and non-public databases identify substances with a code and these codes can be very helpful in mapping substances to various systems. Codes shall always be verified against the source system. Different jurisdictions may require a code from a code system or multiple code systems to be associated and submitted with a substance.

Table 23 — Code

User Guidance	The specific code (the value) with regard to a referenced code system. NOTE InChI is not considered a code and is specified within the structure description.
Example(s)	11-22-356; ab02sh
Conformance	MANDATORY
Data Type	CD.Code
Business Rule(s)	Only codes from recognised code systems will be captured. The code should specifically link to a substance but need not uniquely link to substance. This code can also be a deprecated Substance ID generated based on ISO 11238.

4.6.2 Code system

Every code shall be associated with a code system and a list of names of code systems expected to be submitted shall be maintained by the maintenance organisation.

Examples CAS registry numbers, EC numbers, FDA UNII codes, EMA XEVMPD codes, ASK numbers, EPA Pesticide codes.

© ISO 2016 – All rights reserved
AVA Documents Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

Table 24 — Code system

User Guidance	Name of the code system.
Example(s)	CAS, EINECS, NSC, ASK, ATC, NDF-RT; RX-CUI ; INN BQ ; ETC.
Conformance	MANDATORY
Data Type	CD.CodeSystemName
Business Rule(s)	If the Code class applies, the code system name is implicit and derived from the Code System ID.

4.6.3 Code system ID

The code system ID shall be an OID or an ID as applicable. The ID is associated with code system and need not be submitted.

Table 25 — Code system ID

User Guidance	The ID of the code system.
Example(s)	CAS Registry code: 0049; FDA Substance Registration System (UNII) 0050; INN Biological Qualifier (INN BQ) code: 0051
Conformance	MANDATORY
Data Type	CD.Codeystem
Business Rule(s)	The code system ID is required as all codes shall be linked to a code system.

4.6.4 Code system status

The code system status is typically the status provided by the system for a given code. Many code systems can end up with multiple codes representing the same substance. Often one code will be the preferred code and others will either be deleted or deprecated. Ad-hoc terms may also be developed to indicate a code status.

Table 26 — Code system status

User Guidance	The status of the code assignment
Example(s)	Current, Alternate, Superseded, Proposed
Conformance	MANDATORY
Data Type	CD
Values Allowed	Current, Alternate, Superseded, Proposed

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

4.6.5 Code change date

Table 27 — Code change date

User Guidance	The date at which the code status is changed as part of the terminology maintenance.
Example(s)	20110601
Conformance	OPTIONAL
Data Type	TS.DATE

4.6.6 Comment

Table 28 — Comment

User Guidance	Any comment can be provided in this field, if necessary.
Conformance	OPTIONAL
Data Type	ST

4.6.7 Reference source

The reference source information should be captured as described in 4.4.

4.7 Reference information (Conditional)

Reference information should be captured for all types of substances and specified Group 1 substances. According to ISO 11238, this subclause captures additional types of informative reference information for each type of substance. This information shall not affect the generation of a new unique identifier; it may include both classification and target information for active substances but it does not provide any guidance on the classification of pharmacological effects or the determination of the putative targets for any Substance or Specified Substance.

NOTE Genes from which proteins are derived, target information and codes from code systems also constitute reference information to capture and link to a substance. Classification systems such as the WHO ATC and NDF-RT, which code classification information to substances, are particularly important. Target information is important for monoclonal and polyclonal antibodies and small molecules directed against specific molecular targets.

© ISO 2016 – All rights reserved
AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

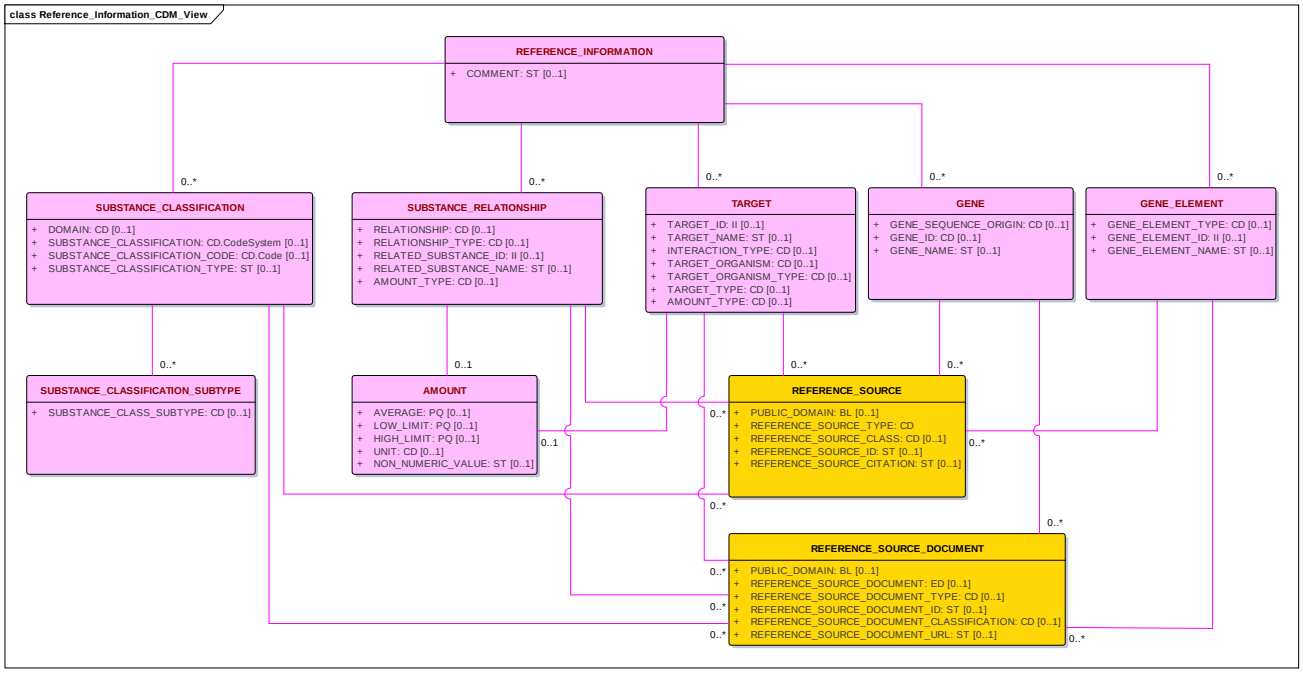


Figure 7 — Information model for Reference Information

4.7.1 Comment

Table 29 — Comment

User Guidance	Any comment can be provided in this field, if necessary.
Conformance	OPTIONAL
Data Type	ST

4.7.2 Substance classification (Conditional)

The standard should have the ability to capture a variety of classification systems. Classification systems are typically based on molecular structure, chemical properties, pharmacological effects, mechanism of action, therapeutic targets or indication.

The standard shall have the ability to capture multiple classifications and variable levels of classification. Although most classification will be associated with an external classification system, ad-hoc classification of substances may be developed within this terminology as needed.

The information should be captured as shown in Figure 8:

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

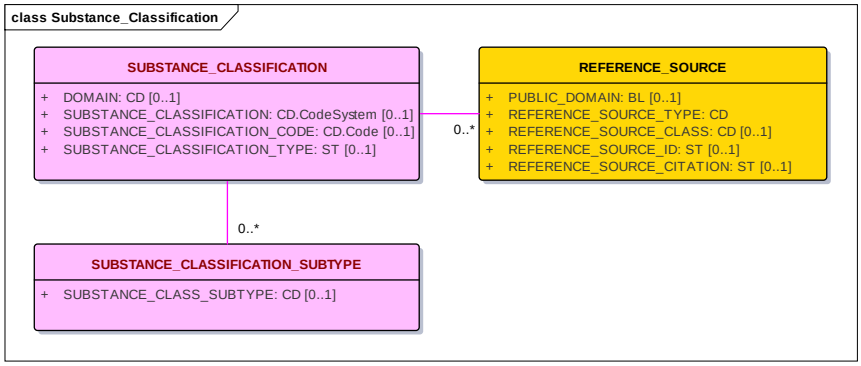


Figure 8 — Information model for substance classification

4.7.2.1 Domain

Classification systems will often only be used within a specific domain. The domain will be associated with the classification system.

Table 30 — Domain

User Guidance	The domain of the substance classification shall be provided.
Example(s)	Human pharmaceuticals, human vaccine, animal drug, animal vaccine, food, food additive, colorant, pesticide; tobacco additive; flavour, excipient
Conformance	OPTIONAL
Data Type	CD

4.7.2.2 Substance classification

There are a number of systems that classify substances based on chemical structure, pharmacological effect, organ system affected, mechanism of action or molecular targets. These classification systems are useful and this information, if available, should be provided and maintained as applicable.

Table 31 — Substance classification

User Guidance	Substance Classification terms based on a controlled vocabulary that is controlled by the system or type of substance classification. The system should be capable of capturing a variety of classification systems and each system should be identified. All active ingredients are typically classified in multiple systems. Ad-hoc classification may be created as necessary.
Example(s)	ATC, NDF-RT, NCI, AHPA, MESH; Therapeutic Category of Drugs in Japan
Conformance	OPTIONAL
Data Type	CD.CodeSystem

© ISO 2016 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

4.7.2.3 Substance classification code

Classification code with map to the active moiety of a given substance.

Table 32 — Substance classification code

User Guidance	Code from the classifying code system. The code should be tied to the substance...
Example(s)	N0000175684 NDF-RT code for Oxycodone that maps to the mechanism of action classification. N02AA05 ATC code that maps to Oxycodone.
Conformance	OPTIONAL
Data Type	CD.Code
Business Rule(s)	The classification code shall be part of an internationally recognised classification system.

4.7.2.4 Substance classification type

Nearly all classification systems contain multiple levels. Type is the highest level or the initial level of the classification system.

Table 33 — Substance classification type

User Guidance	The initial type or highest level of a classification system.
Example(s)	Organ system in ATC.
Conformance	OPTIONAL
Data Type	ST
Business Rule(s)	The value is IMPLICIT and can be derived by the “Substance Classification — Subtype “if available.

4.7.2.5 Substance classification subtype

Table 34 — Substance classification subtype

User Guidance	Many classification systems will have multiple subtypes of classification this field will attempt to capture this information.
Example(s)	NDF-RT Mechanism of Action (MOA) or Physiological Effect.
Conformance	OPTIONAL
Data Type	ST

4.7.2.6 Reference source (Conditional)

The reference source information should be captured as described in 4.4.

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

4.7.2.7 Reference source document (Optional)

The reference source document information should be captured as described in 4.5.

4.7.3 Substance relationship (Conditional)

Table 35 — Substance Relationship

User Guidance	Relationships between substances such as the relationship between a salt form and its active moiety or parent substance should be captured. These relationships are often essential to the description of medicinal products, the basis of strength and the classification of substances. These relationships are often obvious and rules will be developed for specifying substances involved in each type of relationship. For example, the active moiety of all sodium salts would be the free acid, conversely the active moiety of a hydrochloride salt would be a free base.
Example(s)	N/A
Conformance	CONDITIONAL
Data Type	N/A

4.7.3.1 Relationship

The relationship is not a mere code but the actual connection between the two substances which may be different for different kinds of relationships:

There are a variety of different relationships that are possible but for the identification of medicinal products the most important relationships are between a substance, its active moieties, the basis of strength and constituents in herbal and other complex substances. Parent and salt-solvate relationships and relationship of drug substance to its metabolites and impurities are also important although not essential for the description of a medicinal product. If there are known relationships between substances the relationships should be provided and maintained. The parent-active moiety relationship may be essential in describing medicinal products.

Table 36 — Relationship

User Guidance	Relationship between substances. NOTE: Target and gene elements are not to be specified in this subclause.
Example(s)	parent, active moiety, salt, radiolabelled, prodrug, metabolite, enantiomer, starting material, processing material, resultant material
Conformance	CONDITIONAL
Data Type	CD

© ISO 2016 – All rights reserved
AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

4.7.3.1 Relationship type

Table 37 — Relationship type

User Guidance	The Type of interaction between the substances.
Example(s)	IONIC, Van der Waals interaction, Impurity, adjuvant absorption agent, active mooety
Conformance	OPTIONAL
Data Type	CD

4.7.3.2 Related substance ID

Table 38 — Related substance ID

User Guidance	Substance ID associated with the related substance. All related substances should have a Substance ID.
Conformance	CONDITIONAL
Data Type	II
Values Allowed	ISO 11238 Substance ID

4.7.3.3 Related substance name

Table 39 — Related substance name

User Guidance	Preferred term or primary name of related substance.
Conformance	CONDITIONAL
Data Type	CD
Values Allowed	Based on ISO 11238 substance name
Business Rule(s)	The preferred term of the related substance is captured; this information is required when there is a “Substance Relationship”; The substance name is implicit and derived from the Substance ID.

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

4.7.3.4 Amount type

Table 40 — Amount type

User Guidance	Most elements that require a quantitative value will also have a field called amount type. Amount type should always be specified because the actual value of the amount is often dependent on it. EXAMPLE: In capturing the actual relative amounts of substances or molecular fragments it is essential to indicate whether the amount refers to a mole ratio or weight ratio. For any given element an effort should be made to use same the amount type for all related definitional elements.
Example(s)	Mole per cent; Weight per cent; Mole ratio; Weight ratio, Weight, Moles, Exact, Non-Numeric value, SRU/Polymer Ratio, Monomer/Polymer Ratio
Conformance	CONDITIONAL
Data Type	CD
Business Rule(s)	SRU/Polymer Ratio and Monomer/Polymer Ratio are applicable to polymers.

4.7.3.5 Amount (Conditional)

For some substance relationships it may be desirable to capture a quantitative or semi-quantitative amount that may further define the relationship (i.e. major metabolite, degradants etc.). Amount shall be captured according to 4.9.

Substance Relationships which refer to reactions (e.g. metabolism, prodrug) are quantified by specifying the quantity of the reactants:

4.7.3.6 Reference source (Optional)

The reference source information should be captured as described in 4.4.

4.7.3.7 Reference source document (Optional)

The reference source document information should be captured as described in 4.5.

4.7.4 Target (Conditional)

The target subclause allows the capture of information related to the targets that a substance is known to interact with and the interaction results in a biological effect or transformation. These targets can be therapeutic, metabolic, interactions postulated to result in toxicity or interactions of unknown effect. Targets are typically proteins, or complexes of proteins. The target may be a substance in its own right and ID shall be assigned and/or existing ID of a known system used (i.e. GENE, Uniprot) to refer to a target.

For monoclonal antibodies the target shall always be provided. For other types substances the target may be provided if known. Ideally the target is a substance with its own Substance ID. However, in specific circumstances, the target may be handled as a relationship between two substances.

© ISO 2016 – All rights reserved
AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

4.7.4.1 Target ID

Typically, the Target ID is either a Substance or Specified Substance ID.

Table 41 — Target ID

User Guidance	ID of the target relation molecular entity upon which the substance acts. One Target ID should be specified. The target should be a molecular entity or a family of molecular entities. The target is usually a protein entity that the drug actually interacts with. Both therapeutic and metabolic targets and targets associated with toxicity should be captured.
Conformance	CONDITIONAL
Data Type	II
Values Allowed	ISO 11238 Substance ID
Business Rule(s)	The ISO 11238 Substance ID shall be referenced.

4.7.4.2 Target name

At least one name of the target shall be provided when target information is submitted. Target names will typically be handled no differently than other substance names. If an active ingredient targets more than one target each target should be listed. In addition to names a list of codes associated with the target may also be used. Gene bank and Uniprot codes can be particularly useful for identifying and tracking targets. The ISO 11238 standard will be updated to rename Target Gene Name element as Target Name.

Table 42 — Target name

User Guidance	Name of the target molecular entity upon which the substance acts. Each target should be associated with a target name.
Conformance	CONDITIONAL
Data Type	ST
Values Allowed	ISO 11238 substance name
Business Rule(s)	ISO 11238 substance name implicitly derived from Target ID. A variety of names can be captured. A single name will be a preferred name and associated with the Target ID.

4.7.4.3 Interaction type

The type of interactions between a substance and its target should also be captured. A controlled vocabulary should be developed to capture these types of interactions.

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

Table 43 — Interaction type

User Guidance	Type of interaction between the substance and the target. The interaction type will be captured for all substrate-target interactions. A substance can have multiple types of interactions.
Example(s)	AGONIST, INDUCER, INHIBITOR, PARTIAL ANTAGONIST, SUBSTRATE, PRODUCT
Conformance	CONDITIONAL
Data Type	CD

4.7.4.4 Target organism

When drugs are targeted against a specific organism the organism should be captured. For antivirals and antibacterials there is typically no need to give the strain level unless the organism is resistant to more commonly used drugs.

Table 44 — Target organism

User Guidance	The organism for which the active substance is targeted.
Example(s)	HOMO SAPIENS ; M. TB, HIV-1 etc.
Conformance	CONDITIONAL
Data Type	CD
Business Rule(s)	At least one organism type should be associated.

4.7.4.5 Target organism type

Every organism should have a type that is maintained within this terminology. The organism type may also be used independent of the organism particularly for substances such as antibacterial which may be targeted against a wide range of organisms.

Table 45 — Target organism type

User Guidance	Every target organism should have a type associated with the organism.
Example(s)	HUMAN, VIRAL, BACTERIAL, FUNGAL, INSECT, HELMINTH, MALARIAL
Conformance	CONDITIONAL
Data Type	CD
Business Rule(s)	At least one organism type should be associated.

© ISO 2016 – All rights reserved
AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

4.7.4.6 Target type

Each target-substance interaction should be assigned a type. The type is dependent on the substance and also to some extent the use of the substance. A given target-substance interaction can be of more than one type.

EXAMPLE The therapeutic target would be HIV-1 protease when ritonavir is used in HIV-related therapy. The metabolic target is Cyp 3A4 and ritonavir is an inhibitor of this target. Ritonavir is also an inhibitor of P-glycoprotein a drug transporter. In many combination products ritonavir is used strictly to inhibit Cyp 3A4/5 and transporter enzymes.

Table 46 — Target type

User Guidance	Type is a limited form of classification associated with a target of an active substance. Some targets can be both toxic and therapeutic; the choice is dependent on the dose (Km value) of a given substance.
Example(s)	THERAPEUTIC, METABOLIC, TOXIC; TRANSPORTER
Conformance	OPTIONAL
Data Type	CD
Business Rule(s)	Each target should be associated with a target type when the target can be classified.

4.7.4.7 Amount type (Conditional)

Amount type should be captured as described in 4.7.3.4.

4.7.4.8 Amount (Conditional)

Amount should be captured as described in 4.9.

4.7.4.9 Reference source (Optional)

The reference source information should be captured as described in 4.4.

4.7.4.10 Reference source document (Optional)

The reference source document information should be captured as described in 4.5.

4.7.5 Gene (Conditional)

The gene subclause will only be captured for proteins and nucleic acids. It does not apply for chemicals, polymers or structurally diverse substances. It is only used as reference information and refers to the gene that is the origin of the substance (typically a protein). Genes may also be substances and would be described as nucleic acids.

4.7.5.1 Gene sequence origin

For monoclonal antibodies the gene sequence origin will be used to capture species from which the antibody was derived.

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

EXAMPLES For monoclonal antibodies the typical values can be HUMAN, HUMANISED MOUSE; HUMANISED RAT; CHIMERIC MOUSE-HUMAN; CHIMERIC RAT-HUMAN.

Table 47 — Gene sequence origin

User Guidance	Common name for the organism from which the final gene sequence originated.
Example(s)	Human, Bovine, Human; Humanised; Mouse-Human Chimeric; Mouse
Conformance	CONDITIONAL
Data Type	CD
Values Allowed	Typically, from the NCBI Gene database.
Business Rule(s)	NOTE it applies primarily to proteins.

4.7.5.2 Gene ID

Table 48 — Gene ID

User Guidance	ID associated with the gene of origin
Example(s)	YP_299723.1
Conformance	CONDITIONAL
Data Type	CD
Values Allowed	Typically, from the NCBI Gene database.

4.7.5.3 Gene name

Table 49 — Gene name

User Guidance	Complete name given for a gene. Every gene which has an ID should also have a name.
Example(s)	hIL-10 gene
Conformance	CONDITIONAL
Data Type	ST
Business Rule(s)	Implicit from the Gene ID from the applicable Gene database.

4.7.5.4 Reference source (Conditional)

The reference source information should be captured as described in 4.4.

4.7.5.5 Reference source document (Conditional)

The reference source document information should be captured as described in 4.5.

© ISO 2016 – All rights reserved
AVA Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

4.7.6 Gene element (Conditional)

Gene elements will be captured for genes that are used in gene therapy. When the gene element subclause applies, the following information shall be provided:

4.7.6.1 Gene element type

Table 50 — Gene element type

User Guidance	Type of the gene element. Should be captured for all genes that are transduced into cells and are intended to be expressed.
Example(s)	enhancer, promoter, silencer, terminator, repressor
Conformance	CONDITIONAL
Data Type	CD

4.7.6.2 Gene element ID

Table 51 — Gene element ID

User Guidance	Unique identifier for gene element.
Conformance	CONDITIONAL
Data Type	II
Values Allowed	A unique identifier will be associated with each gene element.

4.7.6.3 Gene element name

Table 52 — Gene element name

User Guidance	Specific gene element name.
Example(s)	SV40 Enhancer
Conformance	CONDITIONAL
Data Type	ST
Values Allowed	A unique identifier will be associated with each gene element.
Business Rule(s)	Implicit from the Gene element ID from the applicable Gene element controlled vocabulary.

4.7.6.4 Reference source (Conditional)

The reference source and related information will be captured according to 4.4. One or more reference sources could be cited.

4.7.6.5 Reference source document (Conditional)

The reference source document information should be captured as described in 4.5.

© ISO 2016 – All rights reserved

AVA Documents
Document Sharing Hub

Full document details: PD CEN ISO/TS 19844:2017

4.8 Structure

4.8.1 Structural Representation (Conditional)

Structural information is an essential element for all chemical substances and for other types of substances that have structurally definable elements or modifications. The representation should contain structural information in one or more of the standardised formats as indicated below. A graphical image of the molecular structure should also be provided if available. A complete representation of the structure should be provided and is usually sufficient to define chemical substances. Structural representations can be in a variety of formats or even multiple formats. The commonly used formats include molfile, SMILES, InChI, and cdx. Indication of a preferred structural representation may be defined at regional level. Although most chemical substances may be represented by a single unambiguous structural representation, there are also many substances that can be represented by multiple structural representations or as a mixture of single substances. For the initial implementation and the transmittal of data for substance definitions, an approach that limits submissions to a single or a limited number of unambiguous structural representations is recommended. There are several examples below that illustrate the process in choosing a defining structural representation to submit. A more detailed structural representation guide is provided as part of the Annex B.

For Amine Salts, with the exception of ammonium and quarternium, salts shall be represented as a free base.

EXAMPLE: DIFLOXACIN HYDROCHLORIDE

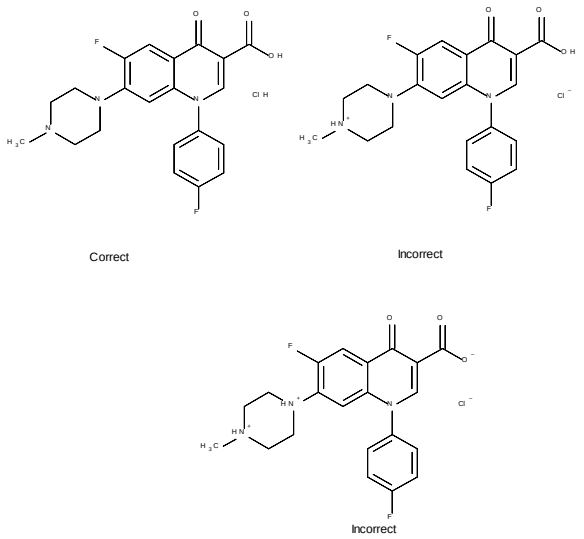


Figure 9 — Correct and incorrect graphical images of the molecular structure of difloxacin hydrochloride

© ISO 2016 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

Metal and ammonium salts: All metal and ammonium salts (NH4+) shall be represented in a charged configuration with all equivalent functional groups ionised. Charge balance will be achieved when necessary by adding hydrogen ions (protons).

EXAMPLE: MONOSODIUM CITRATE

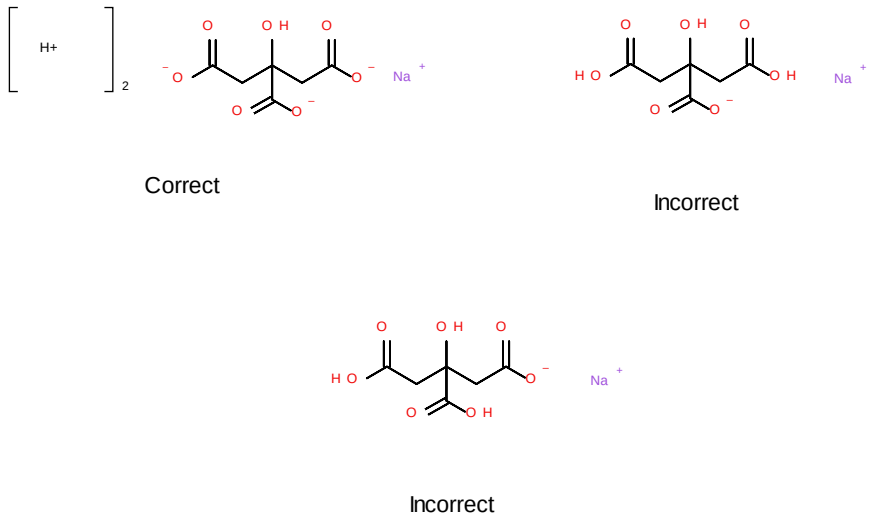


Figure 10 — Correct and incorrect graphical images of the molecular structure of monosodium citrate

Racemic mixtures can be represented as mixtures of enantiomers or by a single structural representation. When using a single structural representation and the substance only contains one chiral centre, a single structure will be drawn that does not contain a chiral bond. If more than one chiral centre exists, relative stereochemistry shall be indicated. The stereochemistry of each moiety of a given substance will be captured separately.

EXAMPLE: OFLOXACIN

A racemic mixture of the R and S isomers. A single structural representation can be used in this case.

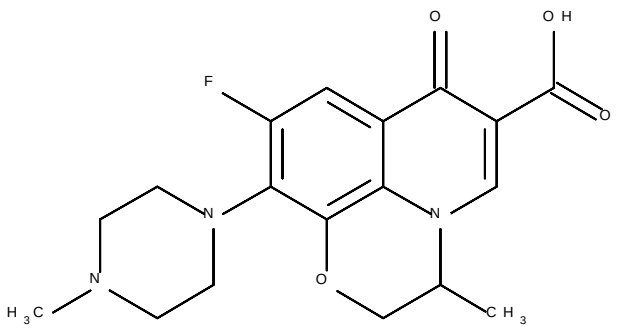


Figure 11 — Graphical image of the molecular structure of difloxacin hydrochloride

© ISO 2016 – All rights reserved

Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017

There are also substances that are mixtures of multiple enantiomers. These substances can also be transmitted in a single structural representation.

Substances can exist in interconverting forms or with a different defined structure depending on the physical state. In cases where the actual structure is a single configuration in a solid physical state, that structure should be captured. When a substance exists in a liquid state and can interconvert to a variety of structures, the least ambiguous single structure should be chosen.

EXAMPLE DEXTROSE (GLUCOSE): Exists as three separate substances in the crystalline state. For a crystalline substance one of the following representations should be chosen.

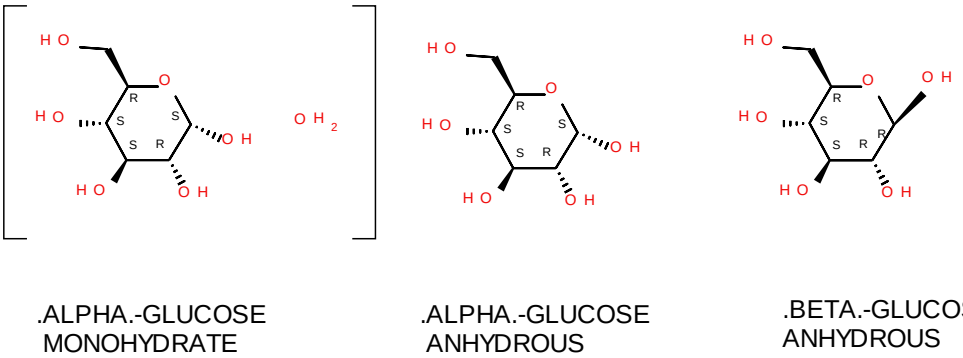


Figure 12 — Graphical images of the molecular structure of the three crystalline states of dextrose

In the liquid state dextrose and many other monosaccharides, disaccharides and higher homologs that end in an aldehyde are often referred to as reducing sugars and exist as mixtures of interconverting substances as illustrated in the structures of glucose below.

In this case the linear single representative structure is chosen to represent dextrose in the liquid or amorphous state. If pyranose or furanose forms were chosen, the representative structure would have one site of stereochemical ambiguity. One general rule is that if there are multiple interconverting structural representations for a substance the preferred representation is the one with the least ambiguity. In Figure 13 the pyranose and furanose structures have one site of stereochemical ambiguity.

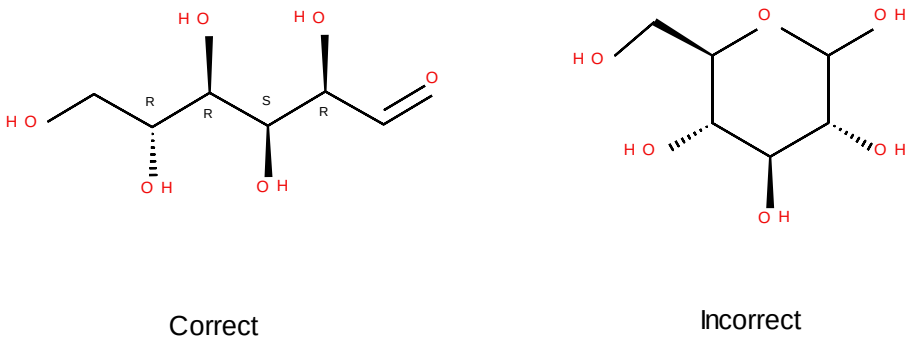


Figure 13 — Correct and incorrect graphical images of the molecular structure of (liquid) dextrose

© ISO 2016 – All rights reserved
Documents
Document Sharing Hub
Full document details: PD CEN ISO/TS 19844:2017