



BSI Standards Publication

**Health informatics — Data elements and their metadata for describing structured clinical genomic sequence information in electronic health records**

---

## National foreword

This Published Document is the UK implementation of ISO/TS 20428:2017.

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 2017  
Published by BSI Standards Limited 2017

ISBN 978 0 580 95379 8

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 July 2017.

### Amendments/corrigenda issued since publication

Date	Text affected
------	---------------

---

---

---

**Health informatics — Data elements  
and their metadata for describing  
structured clinical genomic sequence  
information in electronic health  
records**

*Informatique de santé — Éléments de données et leurs métadonnées  
pour décrire l'information structurée de la séquence génomique  
clinique dans les dossiers de santé électroniques*





**COPYRIGHT PROTECTED DOCUMENT**

© ISO 2017. Published in Switzerland

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

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

# Contents

Page

<b>Foreword</b> .....	<b>v</b>
<b>Introduction</b> .....	<b>vi</b>
<b>1 Scope</b> .....	<b>1</b>
<b>2 Normative references</b> .....	<b>1</b>
<b>3 Terms and definitions</b> .....	<b>1</b>
<b>4 Abbreviated terms</b> .....	<b>5</b>
<b>5 Use case scenario</b> .....	<b>6</b>
<b>6 Composition of a clinical sequencing report</b> .....	<b>7</b>
6.1 General.....	7
6.2 Overall interpretation in summary.....	8
6.3 Detailed contents.....	8
<b>7 Fields and their nomenclature of required data</b> .....	<b>9</b>
7.1 General.....	9
7.2 Clinical sequencing orders.....	10
7.2.1 General.....	10
7.2.2 Clinical sequencing order code.....	10
7.2.3 Date and time.....	10
7.2.4 Specimen information.....	11
7.3 Information on subject of care.....	11
7.3.1 General.....	11
7.3.2 Subject of care identifiers.....	11
7.3.3 Subject of care name.....	11
7.3.4 Subject of care birth date.....	11
7.3.5 Subject of care sex.....	11
7.3.6 Subject of care ethnicity.....	11
7.4 Information on legally authorized person ordering clinical sequencing.....	11
7.4.1 General.....	11
7.5 Performing laboratory.....	12
7.5.1 General.....	12
7.5.2 Basic information on performing laboratory.....	12
7.5.3 Information on report generator.....	12
7.5.4 Information of legally confirmed person on sequencing report.....	12
7.6 Associated diseases and phenotypes.....	12
7.7 Biomaterial information.....	12
7.7.1 General.....	12
7.7.2 Types of sample.....	12
7.7.3 Genomic source class in biomaterial.....	12
7.7.4 Conditions of specimen that may limit adequacy of testing.....	12
7.8 Genetic variations.....	13
7.8.1 General.....	13
7.8.2 Gene symbols and names.....	13
7.8.3 Sequence variation information.....	13
7.9 Classification of variants.....	14
7.9.1 General.....	14
7.9.2 Classification of variants based on the pathogeny.....	14
7.9.3 Classification of variants based on clinical relevance.....	15
7.10 Recommended treatment.....	15
7.10.1 General.....	15
7.10.2 Classification of variants based on clinical relevance.....	15
7.10.3 Clinical trial information.....	15
7.10.4 Known protocols related to a variant.....	16
7.10.5 Other recommendation.....	16

7.11	Addendum.....	16
<b>8</b>	<b>Fields and their nomenclature of optional data.....</b>	<b>16</b>
8.1	General.....	16
8.2	Medical history.....	17
8.3	Family history/Pedigree information.....	17
8.4	Reference genome version.....	17
8.5	Racial genomic information.....	18
8.6	Genetic variation.....	18
8.7	Detailed sequencing information.....	18
	8.7.1 Clinical sequencing date.....	18
	8.7.2 Quality control metrics.....	18
	8.7.3 Base calling information.....	18
	8.7.4 Sequencing platform information.....	20
	8.7.5 Analysis platform information.....	20
8.8	References.....	20
	<b>Annex A (informative) Example structure of clinical sequencing report.....</b>	<b>21</b>
	<b>Annex B (informative) Example layout of clinical sequencing report.....</b>	<b>28</b>
	<b>Bibliography.....</b>	<b>32</b>

## 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](http://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](http://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 voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: [www.iso.org/iso/foreword.html](http://www.iso.org/iso/foreword.html).

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

## Introduction

Based on the rapid advancement of sequencing technologies, clinical sequencing has been highlighted as one of methods to realize personalized medicine and precision medicine. There are lots of sequencing data in the public domain with clinical information[1]. In addition, genome-scale clinical sequencing is being adopted broadly in medical practice[2]. Many hospitals have started to sequence patients' whole genome, whole exome, or targeted genes using the next generation sequencing technologies. These genomic data obtained by next generation sequencing technologies can be used for both clinical purposes to diagnose patients and choose the right medications and research purposes. Therefore, the management of genomic and clinical data is increasingly highlighted to precision medicine, clinical trial and translational research[3].

However, until now, there is no international standard for representing clinical sequencing results with a structured format for electronic health records, in consequence, the necessary genomic test results are not efficiently delivered to the clinicians. There are a few related standards for modeling genetic testing results (i.e. ISO 25720 and several HL7 documents from HL7 clinical genomics working group). However, these standards or drafts are mainly focused on the traditional genetic testing results for a single gene test. Based on the rapid development and adoption of next generation sequencing techniques which can detect diverse genetic variants in genome level, there is, therefore, still a need to develop a standard to present clinical sequencing data in such a way they become useful for clinicians[4].

To implement a structured clinical sequencing report in electronic health records, all necessary data fields should be defined and the metadata for each chosen field should be defined. For example, it needs to be determined which vocabulary, in particular gene descriptions and/or disease codes, can be applied in particular fields. In ISO TC 215, GSVML (Genomic Sequence Variation Markup Language) was proposed for interoperability of genomic variants, especially for single nucleotide polymorphism (SNP) data[5]. HL7 is also developing a domain analysis model for genomics using HL7 version 3[6] and fast healthcare interoperability resources (FHIR)[7]. Recently, to facilitate genomic information, SMART on FHIR Genomics has been developed[8],[9]. The Clinical Data Interchange Standard Consortium (CDISC) published a study data tabulation model implementation guide: pharmacogenomics/genetics[10]. Several other international organizations such as Global Alliance for Genomics and Health (GA4GH), Actionable Genome Consortium, and Displaying and Integrating Genetic Information Through the EHR (DIGITize) of Institute of Medicine in US, tried to develop the similar standards. The working group of the American College of Medical Genetics and Genomics Laboratory Quality Assurance Committee published the ACMG clinical laboratory standards for next-generation sequencing[11]. In addition, web-based tools become available that link genotypic information to phenotypic information, and exchanging information and using it in personalized medicine can be very helpful[12].

In this document, to enable the standard use of patient genomic data from clinical sequencing for healthcare purposes as well as for clinical trials and research, the metadata for a clinical sequencing report for electronic health records will be developed. It further explains how and where particular appropriate terminological systems that describe the genomes and/or diseases can be applied in these fields. By defining the necessary fields with structured format based on coded data that adhere themselves to terminological principles such as concept representation and governance, this document can help implement clinical decision support service.

# Health informatics — Data elements and their metadata for describing structured clinical genomic sequence information in electronic health records

## 1 Scope

The document defines the data elements and their necessary metadata to implement a structured clinical genomic sequencing report and their metadata in electronic health records, particularly focusing on the genomic data generated by next generation sequencing technology.

This document

- defines the composition of a structured clinical sequencing report (see [Clause 5](#)),
- defines the required data fields and their metadata for a structured clinical sequencing report (see [Clause 6](#)),
- defines the optional data (see [Clause 7](#)),
- covers the DNA-level variation from human samples using whole genome sequencing, whole exome sequencing, and targeted sequencing (disease-targeted gene panels) by next generation sequencing technologies. Though whole transcriptome sequencing and other technologies are important to provide better patient care and enable precision medicine, this document only deals with DNA-level changes,
- covers mainly clinical applications and clinical research such as clinical trials and translational research which uses clinical data. However, the necessary steps such as de-identification or consent from patient should be applied. The basic research and other scientific areas are outside the scope of this document,
- does not cover the other biological species, i.e. genomes of viruses and microbes, and
- does not cover the Sanger sequencing methods.

## 2 Normative references

There are no normative references in this document.

## 3 Terms and definitions

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

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <http://www.iso.org/obp>
- IEC Electropedia: available at <http://www.electropedia.org/>

### 3.1 allele

one of several alternate forms of a gene which occur at the same locus on homologous chromosomes and which become separated during meiosis and can be recombined following fusion of gametes

[SOURCE: ISO 16577:2016, 3.6]