International Standard

ISO 8472-1

First edition
2024-07

Biotechnology — Data interoperability for stem cell data —

Part 1: Framework

Biotechnologie — Interopérabilité des données associées à des cellules souches —

Partie 1: Cadre

Reference number
ISO 8472-1:2024(en)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

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 document 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).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

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

For an explanation of 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 www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 276, *Biotechnology*.

A list of all parts in the ISO 8472 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Data interoperability addresses the ability of systems and services that create, exchange and consume data to have clear, shared expectations for the contents, context and meaning of that data.

One-off approaches to data handling and data exchange carry hidden costs for persons and the organizations who are impacted by such data. Most people lack agency when it comes to the data that they generate. Many organizations lack access to the data within their own firewalls. The value of the insights gained from such data is limited because the real potential of such data sets is unrealized.

Stem cell research using human embryonic stem cells (hESC), as well as induced pluripotent stem cells (iPSC) or tissue stem cells, is conducted within ethical and regulatory governance frameworks, which can be highly variable between countries. Stem cells offer unique opportunities to develop therapies for a wide variety of currently intractable conditions. Within this field, it has been recognised that it is premature to focus on any one stem cell type, and that research across a broad front is important to moving the entire field towards application and clinical impact. Furthermore, stem cell research has now reached the stage of clinical testing, with hESC and iPSC–based clinical trials commencing. Moreover, projects provide important 'proof of concept' data for the use of pluripotent stem cell- and tissue stem cell - based therapies in regenerative medicine. Finally, stem cells are becoming a key tool for in vitro disease and tissue modelling, drug and toxicity screening for utilization in the pharma-, chemical-, environmental- and other industries.

In the past decades, research, clinical trials, and industrial developments have greatly increased in scope, diversity and breadth. Moreover, in recent years, many stem cell biobanks have been established. Cross-sector collaborations between academic research institutes, enterprises, governments, industries, etc. [such as Chinese Alliance for Stem Cell Resource Centers, European bank for induced pluripotent stem cells (EBiSC), human pluripotent stem cell registry (PSCreg), International alliance for biological standardization (IABS), International stem cell banking initiative (ISCBI).] have been initiated. Presently, due to the increasing amounts of data associated with stem cell biobanking, it is critical that common approaches to working with data (such as data sharing, storage, analysis, etc.) are standardized. This requires a diverse community of data providers, processors and data consumers to work together and exchange data under same structure and framework. Data interoperability standards aim to help stakeholders and the people who create, manage and use the data to address the shared research requirements and industrial/market needs.

Biotechnology — Data interoperability for stem cell data —

Part 1: Framework

1 Scope

This document specifies a framework for data interoperability of data systems, such as databases, data management systems, web interfaces, API, etc. that manage stem cell data.

It is applicable to all human stem cell types. This document does not apply to other animal stem cells or plant stem cells.

This document specifies considerations and requirements of stem cell data for data interoperability, such as cell characteristics, applied technologies, ethical requirements, and data sharing, analysis, and accessibility.

This document describes an interoperable framework for stem cell data, which can be used for existing systems or existing architectures.

The intended audiences for this document are data generators, implementors of IT infrastructure to handle the data, data providers and data consumers.

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 terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
interoperability
ability of systems to provide services to and accept services from other systems and to use the services so exchanged to enable them to operate effectively together

[SOURCE: ISO 21007-1:2005, 2.30]

3.2
data interoperability
interoperability concerning the creation, meaning, computation, use, transfer, and exchange of data

[SOURCE: ISO/IEC 20944-1:2013, 3.21.12.4]

3.3
access
to obtain the use of a resource

[SOURCE: ISO/IEC 2382:2015, 2121274]

3.4
accessibility
ease of reaching and using a service or facility

[SOURCE: ISO 16439:2014, 3.3]

3.5
data access
process that enables users to retrieve or read published data

[SOURCE: ISO 5127:2017, 3.1.11.17]

3.6
operating system
software to control program operation and to provide the services for resource allocation, task scheduling, I/O control, and data management

[SOURCE: ISO 16484-2:2004, 3.140]

3.7
application
software application program
software or a program that is specific to the solution of an application problem

[SOURCE: ISO/IEC 20944-1:2013, 3.6.3.1, modified — "application" added as the preferred term.]

3.8
database
collection of interrelated data stored together with controlled redundancy according to a schema to serve one or more applications

[SOURCE: ISO/IEC 10027:1990, 3.3.11]

3.9
system schema structure
structure behind the scenes that computer systems access for timely and consistent service

3.10
metadata
data that define and describe other data

[SOURCE: ISO/IEC 11179-1:2023, 3.2.26]

3.11
hardware
all or part of the physical components of an information processing system

[SOURCE: ISO/IEC 2382:2015, 2121277]

3.12
software
all or part of the programs, procedures, rules, and associated documentation of an information processing system

[SOURCE: ISO/IEC 2382:2015, 2121278, modified — Notes 1 to 3 to entry have been deleted.]

3.13
development language
computer programming language used to express the sequence of operations to be performed by a computer

[SOURCE: ISO/IEC/IEEE 24765:2017, 3.3034]

3.14
data provenance
data provenance record
record of the ultimate derivation and passage of a piece of data through its various owners or custodians

Note 1 to entry: A data provenance record can include information about creation, update, transcription, abstraction, validation, and transferring ownership of data.

[SOURCE: ISO 8000-2:2022, 3.8.4, modified — "data provenance" added as the preferred term.]

3.15
raw data
data in its originally acquired, direct form from its source before subsequent processing

[SOURCE: ISO 5127:2017, 3.1.10.04]

3.16
data set
dataset
identifiable collection of data

3.17
data sharing
making data (numerical, textual, images, etc.) available to, and findable by, others

Note 1 to entry: Data is not truly shared, if it cannot be found.

3.18
data exchange
storing, accessing, transferring, and archiving of data

[SOURCE: ISO 10303-1:2021, 3.1.31]

3.19
data quality
assessment of the fitness of the data and information to serve its purpose in a given context

[SOURCE: ISO 23952:2020, 3.4.46]

3.20
IT infrastructure
design of an IT system

3.21
application programming interface
API
standard interface and set of function calls between application software and data access libraries of vehicle navigation systems

[SOURCE: ISO 17267:2009, 2.4, modified – “in accordance with this international standard” was deleted from term.]

3.22
stem cell
non-specialized cells with the capacity for self-renewal and differentiation potential, which can differentiate into one or more different types of specialized cells

Note 1 to entry: Most adult stem cells are multipotent stem cells.

[SOURCE: ISO/TS 22859:2022, 3.24]

4 Interoperability requirements and recommendations for stem cell data

4.1 General

4.1.1 The stem cell data interoperability standard framework shall be designed to improve the ability of stem cell researchers and clinical practitioners to find, obtain, share, exchange, analyse, interpret, understand and use stem cell data.

4.1.2 Compatibility with existing ISO standards in biotechnology and information technology should be taken into account in database design.

The framework for data interoperability should fit the purpose and requirements from users including researchers, clinicians and industry actors. This incorporates recommendations to take into account stem cell characteristics and IT-relevant-activities, such as source, process, purpose, database construction, and developments for websites, API, and data formats.

NOTE A list of database examples is given in Table B.1.

4.1.3 Special attention shall be given to best practices for data sharing and interoperability among the international stem cell research community. The data requires the records of technologies for generating the data.

NOTE A list of relevant document(s) and a scheme for data sharing and data exchange is provided in Table A.1.

Relevant local, national and international ethical requirements shall be addressed with regards to data and privacy protection. Applicable local, national and international regulatory requirements for data and privacy protection shall be documented.

4.1.4 Appropriate patterns for different types of stem cells and data sharing should be adopted to facilitate the interoperability of stem cell data.

EXAMPLE The common model can be a data layer, a database layer, website APIs, or websites.

Each model shall be properly designed and developed for all key elements of data interoperability.

4.1.5 The data type, stem cell type, terminology definition, jurisdictional regulations or guidelines, IT-solutions, etc. affect the interoperability.

4.1.6 Stem cell data access policy, data quality, organizational type and culture, available technology and IT infrastructure impact on interoperability. Already existing data may require specific flexibility and agreements in adopting data sharing standards.

The interoperability requirements for stem cell data include considerations about, but not limited to, donor characteristics, ethical provenance, cell type characteristics, technologies used for generation, expansion, cultivation, , security, regulation, risk, data searchability, data accessibility, data usability.

4.2 Characteristics

Through technical and organizational evaluation, the most appropriate level of data interoperability can be selected according to capabilities and needs.

The characteristics include, but are not limited to:

a) stem cell resources and technologies;

b) data type of interest (for data interoperation);

b) stem cell type;

c) stem cell application, e.g. developmental biology research, disease modelling, diagnostics, cell therapy;

d) stem cell characteristics, such as purity, viability, differentiation potential, genomic features, contaminants, key gene expression, etc.;

e) stem cell authentication, i.e. confirmation that the derived stem cell is consistent with the source donor material

f) stem cell self-renewal state, asymmetric versus symmetric division;

g) stem cell identifier, which is a unique reference code used for stem cell data provenance.

4.3 Technologies

The technologies, methods and process shall be documented and ensure the data interoperating process is safe, reliable and traceable. A record of stem cell related data interoperation shall be maintained, includes but not limited to:

a) Stem cell processes, such as isolation, establishment, expansion, passaging, cryopreservation, thawing, sterility etc.

b) Stem cell characterization, such as authentication, cell type determination, differentiation potency, molecular features (RNA, protein, metabolism, methylome, genome stability and features), purity, genetic modification

c) Stem cell quantification, such as cell cycling, viability, passage number

4.4 Ethics

Data interoperation with stem-cell related data shall comply with relevant ethical requirements. The ethical approvals and donor informed consent permissions and requirements shall be integrated within the data in an interoperable format. Existing community guidelines shall be followed, such as International Society for Stem Cell Research (ISSCR) guidelines for stem cell research and clinical translation – fundamental ethical principles.

4.5 Security

In the process of data interoperation, attention to the security of relevant data shall be given and attention to the security and stability of relevant data.

NOTE ISO 27001 can be implemented for this purpose.

4.6 Regulation

The interoperation of stem cell data shall meet the specification requirements of relevant data information, data protection and management of special categories of data such as health and genetic data.

4.7 Risk

Data interoperability should be protected by appropriate role permissions and technical means to ensure that data is not illegally tampered with or falsified, and stable during sharing, exchanging, and interoperation. The risk of ambiguity shall be minimized by applying standard taxonomies, ontologies and vocabulary.

4.8 Data searchability

Data should be searchable through index, services or existing APIs with appropriate identifiers.

4.9 Data accessibility

Data should be accessible through network or certain processes required to protect donors from re-identification, such as a governance process that uses a data access committee (DAC) to control access to special categories of data.

4.10 Data usability

The terms and conditions under which the data may be used shall be clearly stated. These usage conditions shall be associated with the data, so that data consumers are informed about the allowed uses of the data.

5 Framework for interoperable systems based on existing architectures

5.1 Context for interoperability within and between systems

Existing information systems, e.g., operating system, application and database, should be maintained to ensure interoperability between systems.

The system schema structure, and metadata should be documented.

5.2 General description

The stem cell data interoperability framework should be independent of specific hardware, software, networks, databases, and development languages. The data interoperability standards framework should guide the information retrieval, sharing and exchange practices of the research, commercial and clinical application community in the stem cell field. Data provenance should also be recorded and considered.

5.3 Interoperability of entries

The interoperability of entries should include, but is not limited to, the complete life cycle of cells, such as cell line generation, cell culture, cell expansion, cell characterization, cell authentication, genetic modification, cell preparation, cryopreservation, transportation, qualification for clinical use and other typical cell related information.

The system should support different levels for data interoperability, such as raw data layer, meta data, project level layer, national/institutional level layer.

The interoperability of entries should support a single stem cell record and a data set of many stem cell records. Each stem cell line shall be clearly identifiable by a persistent unique identifier through which its descendants are traceable (such as differentiated cells and genetically modified cells).

5.4 Flexibility of interoperability

The level of data interoperability between systems can be evaluated against the framework. Users can choose the level of data interoperability based on their needs and available access. A standard scheme or table can be applied for sharing and exchanging data between systems (see Annex A).

Annex A
(informative)

Document(s) or scheme for data sharing and data exchange

Table A.1 — Document(s) or scheme for data sharing and data exchange

Document(s) or scheme	Title	Description
ISO/IEC 23751	ISO/IEC 23751 Information technology — Cloud computing and distributed platforms — Data sharing agreement (DSA) framework	refer to document scope.
ISO 23903	ISO 23903 Health informatics — Interoperability and integration reference architecture — Model and framework	refer to document scope
ISO/TS 22691	ISO/TS 22691 Health informatics — Token-based health information sharing	refer to document scope
ISO 2146	ISO 2146 Information and documentation — Registry services for libraries and related organizations	refer to document scope
ISO/TS 18876-2	ISO/TS 18876-2 Industrial automation systems and integration — Integration of industrial data for exchange, access and sharing — Part 2: Integration and mapping methodology	refer to document scope
ISO/IEC TR 20943-1	ISO/IEC TR 20943-1 Information technology — Procedures for achieving metadata registry content consistency — Part 1: Data elements	refer to document scope
ISO/TR 3985	ISO/TR 3985 Biotechnology — Data publication — Preliminary considerations and concepts	refer to document scope
ISO/TS 18876-1	ISO/TS 18876-1 Industrial automation systems and integration — Integration of industrial data for exchange, access and sharing — Part 1: Architecture overview and description	refer to document scope

Table A.1 *(continued)*

Document(s) or scheme	Title	Description
ISO 20691	ISO 20691 Biotechnology — Requirements for data formatting and description in the life sciences	refer to document scope
Database/Webservice (hPSCreg)	Human pluripotent stem cell registry	hPSCreg is a global registry for human pluripotent stem cell lines (hPSC-lines). It provides — a central and searchable data hub for available hPSC-lines — a standardized unique persistent identifier and register for hPSC-lines — a data source by verifying the ethical and biological conformity of registered lines based on community standards — interoperable data on hPSC and related data standards