INTERNATIONAL STANDARD

ISO 29585

First edition
2023-06

Health informatics — Framework for healthcare and related data reporting

Reference number
ISO 29585:2023(E)

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 215, *Health informatics*.

This first edition of ISO 29585 cancels and replaces ISO/TR 22221:2006 and ISO/TS 29585:2010, which have been technically revised.

The main changes are as follows:

— consideration of the impact of developments such as the availability of big-data and federation of services;

— each requirement has an identified actor responsible for its delivery and each requirement is intended to be clear and unambiguous.

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

0.1 Background

A considerable amount of data is collected during the provision of care and treatment, some of it specific to the patient being treated, and some of it not. The primary purpose of this information is to support and improve individual patient care and much of it is held under professional and legal obligations of confidentiality. However, this information, often in conjunction with other records, is of value for many other purposes to support healthcare for groups of patients or for populations.

Healthcare data reporting provides many benefits. The health and well-being of the population are improved by activities such as disease surveillance, screening, needs assessment and preventative activities such as identifying the relationship between infected water and cholera resulting in better sewers. Research has led to major benefits in health practice such as the cure of duodenal ulcers, prevention of spina bifida, effective treatment of breast cancer and the carrying out of hip replacements. Research has also reduced risks through a greater understanding of HIV prevention, the relationship between smoking and lung cancer and the ill effects of the use of aspirin for children. The regulation of new medicines and other treatments relies on evidence of safety and efficacy from clinical trials.

Providing appropriate conditions are met, these data can legitimately be used to support these other purposes. In practice, such healthcare data reporting covers a wide spectrum including:

— Protecting the health of the public through surveillance and immediate response to infectious disease and other environmental threats to health, monitoring adverse effects of therapeutic interventions and informing and evaluating screening.

— Providing better information to the general public about healthy lifestyles.

— Improving the quality and safety of care or reducing the impact of new risks to population.

— Improving the management of the health system, for example by supporting the more efficient commissioning of services and value-based care.

— Improving the quality of clinical care within an institution, for example through the audit of clinical practice.

— Identifying patients who interact with multiple parts of the health system in order to monitor equity of access and provision:

  — ensuring consistent care for people who interact with multiple parts of the system,

  — monitoring equity of access and provision.

— Ensuring that health policy is evidence-based through carrying out empirical research.

0.2 Healthcare data reporting

Where the term "clinical data warehouse" implied a specific, bounded, repository of data, with specific functions, recent developments have greatly increased the ways of addressing potential applications. For instance:

— The era of "big data" offering new sources and modes of data, with a massive increase in data capture and use, including structured, unstructured, text, images, near real-time, combination of data sources, e.g. personal device data, also social determinant of health data to inform population health and a wide range of presentation and visualization tools.

— The establishment of federated services that can link data sources which previously could not be combined and, hence, supporting distributed queries. These federated approaches can support moving from hierarchical views of data to multi-layered and multi-dimensional approaches, the separation of data sources and data consumers, distributed queries and moving from data warehouses / data marts to data lakes and data labs.

— The potential for analysing data on a much wider scale, particularly for areas such as rare diseases where federated big data enables studies requiring this population size.

— The push for transparency of data has further reinforced the opportunities and responsibilities of sharing the value of such analysis with a wider public.

In view of these developments, this document provides a framework for healthcare and data reporting, addressing both the opportunities and the responsibilities of the handling of the data. Figure 1 summarizes the stages, products and actors through the lifecycle.

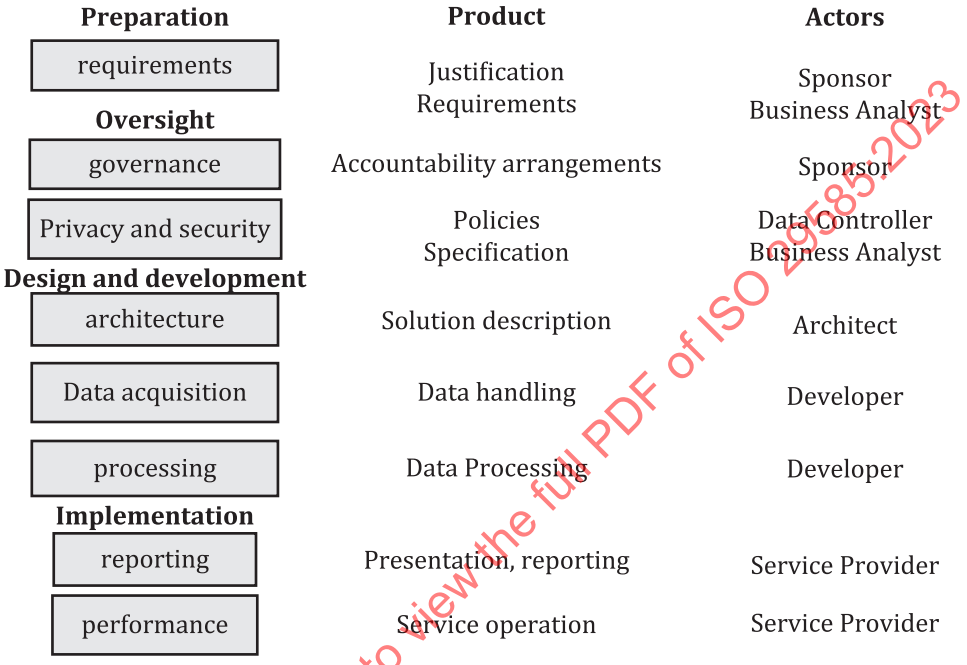

Figure 1 — Lifecycle for a healthcare data reporting service

Clauses 5 to 12 specify requirements, each of which is allocated to one actor. Requirements are individually referenced by actor (e.g. SPnnn for sponsor, DCnnn for data controller, ANnnn for business analyst, ARnnn for architect, DVnnn for developer and PRnnn for service provider).

Health informatics — Framework for healthcare and related data reporting

1 Scope

This document deals with the reporting of data to support improved public health, more effective health care and better health outcomes.

This document provides guidance and requirements for those developing or deploying a healthcare data reporting service, addressing data capture, processing, aggregation and data modelling and architecture and technology approaches.

The role of a healthcare data reporting service is to enable data analyses in support of effective policies and decision making, to improve quality of care, to improve health services organizations and to influence learning and research. This document has relevance to both developing and more established health systems. It enables meaningful comparison of programs and outcomes.

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.

IEC 62304, *Medical device software — Software life cycle processes*

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
analyst
member of the technical community who is skilled and trained to define problems and to analyze, develop, and express algorithms

EXAMPLE Systems engineer, business analyst.

3.2
architect
person, team, or organization responsible for the process of defining a collection of hardware and software components and their interfaces to establish the framework for the development of a computer system

[SOURCE: ISO/IEC/IEEE 24765:2017, modified — Combined definitions of "architect" (3.209) and "architectural design" (3.211).]

3.3
business analyst
person who bridges the gap of understanding between business and technology to accurately define software requirements and carefully control scope

3.4
clinical data warehouse
CDW
grouping of data accessible by a single data management system, possibly of diverse sources, pertaining to a health system or sub-system and enabling secondary data analysis for questions relevant to understanding the functioning of that health system, and hence supporting proper maintenance and improvement of that health system, e.g. public health services

Note 1 to entry: A CDW tends not to be used in real time. However, depending on the rapidity of transfer of data to the data warehouse, and data integrity, near real-time applications are not excluded.

3.5
dashboard
user interface based on predetermined reports, indicators and data fields, upon which the end user can apply filters and graphical display methods to answer predetermined business questions and which is suited to regular use with minimal training

3.6
data controller
organization that determines what information will be processed and why

Note 1 to entry: The data processor is the one that does the actual processing. Controllers are responsible for creating privacy notices, implementing mechanisms to ensure that individuals can exercise their data subject rights and adopting measures to ensure the data processing meets the GDPR's (general data protection regulation) principle of privacy by design and by default.

3.7
data custodian
role within the processing entity (IT department) that handles the data daily

3.8
data dictionary
database used for data that refer to the use and structure of other data, i.e. a database for the storage of metadata

3.9
data element
unit of data that is considered in context to be indivisible

3.10
data mart
subject area of interest within or standalone from the data warehouse dimension

EXAMPLE An inpatient data mart.

Note 1 to entry: Data marts can also exist as a standalone database tuned for query and analysis, independent of a data warehouse.

Note 2 to entry: Data marts are typically suitable to adhere to localized requirements such as GDPR (general data protection regulation) in the European Union, via clear specification of purpose for analysis, permissions of data subjects, and data minimalization procedures.

3.11
data warehouse dimension
subject-oriented, often hierarchical business relevant grouping of data

3.12
developer
individual or organization that performs development activities (including requirements analysis, design, testing through acceptance) during the system or software life-cycle process

[SOURCE: ISO/IEC 25000:2014, 4.6]

3.13
drill down
exploration of multidimensional data which makes it possible to move down from one level of detail to a more detailed level depending on the granularity of data

EXAMPLE Number of patients by departments and/or by services.

3.14
episode of care
identifiable grouping of healthcare-related activities characterized by the entity relationship between the subject of care and a healthcare provider, such grouping determined by the healthcare provider

3.15
health indicator
single summary measure, most often expressed in quantitative terms, that represents a key dimension of health status, the healthcare system, or related factors

Note 1 to entry: A health indicator is informative and also sensitive to variations over time and across jurisdictions.

[SOURCE: ISO 21667:2010, 2.2]

3.16
healthcare data reporting service
managed service to provide reporting of data to support improved public health, more effective health care and better health outcomes

3.17
metadata
information stored in the data dictionary that describes the content of a document

3.18
master data management
enablement of a program that provides for an organization's data definitions, source locations, ownership and maintenance rules

3.19
organization
unique framework of authority within which a person or persons act, or are designated to act towards some purpose

[SOURCE: ISO/IEC 6523-1:1998, 3.1, modified — Removed note to entry.]

3.20
performance indicator
measure that supports the evaluation of an aspect of performance and its change over time

3.21
service provider
organization or part of an organization that manages and delivers a service or services to the customer

Note 1 to entry: A customer can be internal or external to the service provider's organization.

3.22
sponsor
person or group who provides resources and support for the project, program, or portfolio and is accountable for enabling success

[SOURCE: ISO/IEC TR 24587:2021, 3.15]

3.23
star schema
dimensional modelling concept that refers to a collection of fact and dimension tables

4 Abbreviated terms

AES	Advanced Encryption Standard
API	Application Programming Interface
DPO	Data Protection Officer
EHR	Electronic Health Record
ELT	Extract, Load, Transform
ETL	Extract, Transform, Load
GDPR	General Data Protection Regulation
HL7®[a)]	Health Level 7
ICD®[b)]	International Classification of Diseases
LOINC®[c)]	Logical Observation Identifiers, Names and Codes
MBUN	Meaningless But Unique Number
NLP	Natural Language Processing
OCR	Optical Character Recognition
PIA	Privacy Impact Assessment [020 – amended]
RBAC	Role-based Access Control
SLA	Service Level Agreement
SNOMED CT®[d)]	Systematized Nomenclature of Medicine — Clinical Terms
TRE	Trusted Research Environment

[a] HL7 is the registered trademark of Health Level Seven International. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named.

[b] ICD is the trademark of the WHO. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named.

[c] LOINC is the registered trademark of Regenstrief Institute. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named.

[d] SNOMED CT is the registered trademark of the International Health Terminology Standards Development Organisation (IHTSDO). This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named.

5 Preparing: Requirements and planning

5.1 Overview

Clause 5 describes steps to be taken when planning the development of healthcare data reporting service or the extension of existing services. Potential benefits and uses are described in Annex A.

The sponsor and the business analyst are responsible for specifying requirements.

A healthcare data reporting service typically becomes more valued than originally anticipated and grow in size, complexity and rate of access.

SP001 The sponsor should ensure that the healthcare data reporting service be viewed as an on-going development and not as a fixed project with an endpoint.

SP002 The sponsor should provide an “extensibility” plan can include import and export to other systems and communications with other systems, which retain the integrity of the data.

5.2 Prioritization of requirements

There are many factors relevant to the prioritization of requirements.

SP003 A sponsor wishing to develop, extend or make use of the healthcare data reporting service should justify the purposes of use prior to commencing implementation.

SP004 The sponsor shall have a clear value proposition for the foreseen applications.

SP005 The sponsor should, when developing new services, include engagement with initial information providers, users, service providers and other relevant systems with which the healthcare data reporting service is expected to exchange information/services.

SP006 The sponsor shall ensure that proposals are designed to achieve a clear outcome for users or the system. The sponsor shall understand how outputs will result in better provision and/or outcomes for people and the health and care system.

SP007 The sponsor shall document the justification for the intended purpose(s) of the healthcare data reporting service.

SP008 The sponsor should ensure budgets balance costs and performance needs.

SP009 The sponsor can enter into contractual requirements for the healthcare data reporting service with current information systems and service suppliers.

SP010 The sponsor of the healthcare data reporting service shall ensure that the service is scalable.

5.3 Users

Consideration should be given to multiple levels of reporting, such as national, regional, local and international.

Commercial users, government entities, regulators, professional bodies and educational establishments can exist in some form at all levels.

Level	Example
International	agencies such as the World Health Organization (WHO), research bodies such as the Commonwealth Fund or groupings such as the European Union
	governments, government agencies (e.g. analysis and reporting centres), regulators, professional bodies, universities, medical research
Regional	depending on the country, can be state, province or regional government, or health organizations
Local	local care organizations (e.g. health care providers or hospitals), local government for environment, education, housing, other commercial users, e.g. pharmaceutical companies

It is often appropriate to have reporting at each of these levels, each attuned to the analysis and reporting requirements of the sponsoring organization.

SP011 The sponsor should ensure the healthcare data reporting service provides policy and strategic reporting to meet the needs of the stakeholders.

SP012 The sponsor should ensure the healthcare data reporting service has the ability to support day-to-day requirements for the intended stakeholders.

5.4 Data requirements

SP013 The sponsor of the healthcare data reporting service shall ensure that the service has availability of data and corresponding metadata from source systems.

SP014 The sponsor of the healthcare data reporting service shall ensure that the service includes data quality measures that reflect fitness for purpose.

SP015 The sponsor shall document a structured process that reviews current and planned arrangements for handling of personal data.

SP016 The sponsor shall identify, establish and use standards for handling health data.

SP017 The sponsor shall demonstrate that the product collects, stores and processes users' information in a safe and fair way, the handling of personal information.

5.5 Services and non-functional requirements

Provisioning through cloud-based services places more emphasis on supplier and consumer relationships. All the following features are important for the effectiveness of any reporting, based on agreed measures and metrics

SP018 The sponsor of the healthcare data reporting service shall ensure that the service is provided with appropriate Service Level Agreements, or equivalent, regarding ongoing technical support with suitable availability from a helpdesk or similar.

SP019 The sponsor of the healthcare data reporting service shall ensure that the service perform periodic backups and test restores as specified by Service Level Agreements (SLA).

SP020 The sponsor of the healthcare data reporting service shall ensure that the service has a plan for disaster recovery.

SP021 The sponsor shall ensure services are reviewed at least annually to identify and improve processes, which have caused breaches or near misses, or which force staff to use workarounds which compromise data security.

SP022 The sponsor of the healthcare data reporting service shall ensure that the service accommodate the highly dimensional and complex nature of healthcare data and associated analysis

SP023 The sponsor shall ensure that the data requirements take account of the types of output through which the data will be reported.

SP024 The sponsor should ensure that the development of outputs for clinical use involves both technical and clinical expertise in the form of a clinical product owner.

6 Governance

6.1 Principles

Clause 6 considers the governance issues of responsible data organization, management and use.

The primary actors in this clause are the sponsor and, in the context of guarding data access, the data custodian.

SP025 The sponsor shall define a governing structure for establishing policies and decision-making process regarding scope, access, further development, etc.

SP026 The sponsor shall base governance on data protection principles appropriate to country(ies) of operation.

SP027 The sponsor shall ensure there is a risk assessment and control system in place.

SP028 The sponsor shall ensure that governance arrangements include conformity with mechanisms for assuring that all plans have been completed and actions undertaken satisfactorily. Relevant International Standards for data governance include ISO/IEC 38505-1.

SP029 The sponsor shall ensure proposals are reviewed by an appropriate independent body (e.g. ethics committee).

SP030 The sponsor shall identify who is responsible for creating and enforcing policies that specify how data should be managed, used and maintained.

SP031 The sponsor of the healthcare data reporting service shall ensure that the service has audit policies, based on information governance principles (e.g. to ensure no identifiable personal data is revealed to the service provider except where unavoidable, and then all such access should be recorded and processes in place to detect misuse).

SP032 The sponsor shall ensure that, as the healthcare data reporting service is considered a key system, it is included within an overall business continuity plan.

7 Privacy and security of the data

7.1 Overview

Clause 7 describes general considerations regarding privacy and security. It is based on the premise that, prior to consideration of the architecture, there needs to be detailed assessment and planning for addressing confidentiality of personal data, to enable and support privacy by design.

The primary actors responsible in this clause are the sponsor, the business analyst responsible for specifying requirements and the data controller responsible for safeguarding data access.

This document is not a security framework, but it is intended that, within this healthcare data reporting framework, there is a corresponding security framework. Examples include ISO/IEC 27000, NIST SP 800-53, NIST SP 800-171, NIST Cybersecurity Framework, CIS Controls, HITRUST CSF[13] and COBIT[14].

SP033 The sponsor of the healthcare data reporting service shall ensure that the service addresses privacy and security aspects.

7.2 Principles

The following principles underpin the privacy measures for the healthcare reporting service:

DC001 The data controller shall ensure that data are collected for specified, explicit and legitimate purposes.

DC002 The data controller shall ensure that data are not further processed in a manner that is incompatible with those purposes for which it was collected.

DC003 The data controller shall ensure that collected data are adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ("data minimization").

DC004 The data controller should ensure that data are kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed ("storage limitation").

DC005 The data controller shall ensure data are processed in a manner that ensures appropriate security of the personal data ("availability, integrity and confidentiality").

DC006 The data controller shall be responsible for, and be able to demonstrate compliance with, these principles ("accountability").

DC007 The data controller shall ensure that, where data is obtained from other sources, there are data sharing agreements in place.

DC008 The data controller shall ensure that any conditions in the data sharing agreements are met.

DC009 The data controller shall ensure that, in a distributed environment such as reporting with multiple stakeholders and widely distributed users, lines of accountability are clear and adequate.

DC010 The data controller shall ensure that the responsibility for the service is unambiguous, e.g. a recognized entity that accountable for data management, use, retention and destruction.

The custodian of the healthcare data reporting service is often the organization funding its development or the one on whose premises it is located. However, this is not always the case.

For example, in the European Union, the GDPR emphasizes the need for transparency over how personal data are used. The provision of privacy information to individuals (typically through a privacy notice) describes how their personal data will be processed, with whom their personal data will be shared and what their rights are. Such information helps individuals to be enabled to make informed decisions in relation to their personal data.

The application of the process and the standards implies consideration of individual systems and data flows. Documents underpinning this are data flow maps, privacy impact assessments (PIA) and privacy notices. See Annex B for further details and information about the GDPR's Data Protection Impact Assess (DPIA) requirement.

SP034 The sponsor shall ensure that privacy notices are made available to individuals where requested and whenever else it is possible.

SP035 The sponsor shall ensure that privacy notices provide contact details of the data controller and data protection officer.

SP036 The sponsor shall respond to objections to the handling of confidential information.

SP037 The sponsor of the healthcare data reporting service shall ensure that the service communicates to individuals what personal data are being collected and processed and why.

7.3 Policies

DC011 The data controller shall ensure that there are data sharing agreements between organizations contributing data to a healthcare data reporting service and the organization that manages the healthcare reporting service.

DC012 The data controller shall ensure that organizations seeking to implement external data linkage develop policies addressing technical aspects like input data quality standards, formats, specification of encryption and access keys and key control.

DC013 The data controller shall ensure that policies for pseudonymization consider how to protect pseudonymized data from being linked with other (e.g. older) personally identifiable data.

DC014 The data controller shall ensure that policies for pseudonymization require that all data are accompanied by metadata describing its permitted use, disclosure and retention, whether or not it is identifiable, pseudonymized, anonymized or aggregate.

DC015 The data controller shall ensure that the healthcare data reporting service honours patients' right to know who accessed their personal identifiable information

DC016 The data controller shall ensure that by having effective policies and procedures in place, organizations can demonstrate good practice, maintaining records of processing, appointing a Data Protection Officer (DPO) and carrying out PIAs and / or DPIAs, as required.

DC017 The data controller should ensure that, at the minimum, an audit policy be created that triggers an event for administrator review when a query is made which might identify a small group of patients or members or a single patient or member.

DC018 The data controller should ensure that policies are not so stringent so as to be impractical to adopt (e.g. where required audit trails greatly exceed the data being reported).

DC019 The data controller shall ensure that policies are regularly reviewed (e.g. annually) to ensure that they are appropriate to the current state of the healthcare data reporting service.

DC020 The data controller shall ensure that policies are regularly reviewed to ensure that they address relevant risks.

DC021 The data controller shall ensure that policies are reviewed regularly to ensure that they keep current with applicable requirements.

7.4 Processes - Security

For detailed guidance on security as it relates to healthcare data, including the technical safeguards below, see ISO 27799.

National bodies (e.g. NIST and FISMA in the US) can provide further information and guidance.

AN001 The business analyst shall ensure that user access controls address the specific sets of data and the business function requirements associated with the user.

AN002 The business analyst shall ensure that user access controls specify access rights by user organization if patient-level data is involved.

AN003 The business analyst shall ensure that user access controls specify access for data extraction.

AN004 The business analyst shall ensure that user access controls specify access for on-line reports.

AN005 The business analyst shall provide a mechanism to set the level of data confidentiality.

AN006 The business analyst shall ensure confidential data are only accessible to users with the need to access (see DC006 and 9.3).

AN007 The business analyst shall ensure that access to confidential data is removed from users when it is no longer needed.

AN008 The business analyst shall ensure that the service has strong security and privacy safeguards.

AN009 The business analyst shall ensure that the service provides deidentification services such as pseudonymization.

DC022 The data controller shall ensure that the data do not have linkage to individuals except where specific permission is given.

DC023 The data controller shall ensure that, where there is linkage of records, access to personally identifiable data (for both subjects and providers of care) is tightly restricted to only individuals who have the need to know and permission to access as defined by organization policies.

DC024 The data controller shall ensure personal identifiers are removed from patient-level data whenever possible to reduce the risk of re-identification, e.g. by combining enough attributes of an individual in a small population cohort, identities can be inferred.

DC025 The data controller may be required to take additional steps to prevent re-identification, for example, in the case of small population cohorts in which the combination of attributes from individual records can be used to identify a subject of care.

DC026 The data controller shall ensure that databases and infrastructure for the healthcare data reporting service are secured.

DC027 The data controller should classify healthcare data reporting service data elements into categories of privacy sensitivity [e.g. HL7 Data Segmentation for Privacy (DS4P)][16] to establish appropriate data warehouse security.

DC028 The data controller shall ensure that personal data are processed in a manner that ensures appropriate security of the personal data, including protection against unauthorized or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organizational measures (integrity and confidentiality).

PR001 The service provider shall implement safeguards for handling small numbers, where a particular query generates a very small result set from which an individual's identity might be inferred.

AN010 The business analyst shall ensure that the service is able to demonstrate that privacy measures are in place.

AN011 The business analyst shall ensure that there is an analysis of all data flows to ensure security and confidentiality are maintained.

AN012 The business analyst shall establish and document the purpose of arrangements to handle confidential information.

AN013 The business analyst shall consider the impact of anonymization vs identifiable data to support routine business processes.

AN014 The business analyst shall ensure that, whilst there is an ethical responsibility to use data to manage the healthcare system, the confidentiality of the data is protected.

7.5 Processes: Pseudonymization and anonymization

Pseudonymization is an important technique in healthcare data reporting service environments where aggregate data are not sufficiently granular for approved data use.

DC029 The data controller shall ensure that the healthcare data reporting service audits include reviews of general patterns of data access and use, specific re-identification of pseudonymized data, security management processes and practices and processes related to data quality and integrity.

AN015 The business analyst shall ensure that user access controls specify level of data identifiability (i.e. aggregate, anonymized, pseudonymized or patient identifiable).

DC030 The data controller shall ensure that the healthcare data reporting service security includes access control (such as role-based access control (RBAC)), definition and assignment of accessible functions, pseudonymization and audit capabilities.

NOTE The impact of anonymization vs identifiable data can vary by the extent to which local, regional or national users need access to patient-level identifiable data.

7.6 Process: Auditing

Audit is an essential and on-going part of maintaining good practices in the same manner that clinical audit is known to help maintain quality of clinical care.

DC031 The data controller for the healthcare data reporting service shall develop an audit plan.

SP039 The sponsor should develop a process including audit and escalation, with evidence that policies are in effect, are monitored and are reviewed.

DC032 The data controller for the healthcare data reporting service shall review the audit plan's enforcement and results regularly.

8 Data

8.1 Overview

Since 2010, there has been massive expansion of volume and characteristics of data. Examples are provided in Annex C with reference to datatypes as specified in ISO 21090.

AN016 The business analyst should ensure that the strategy for developing and populating the healthcare data reporting service accounts for data items with multiple sources.

AN017 The business analyst should ensure data is collected as close to the source as technically feasible.

AN018 The business analyst should ensure systems that inform international, national or regional databases submit data in a standardized format, and to a specified timeframe.

AN019 The business analyst should ensure there is an extensibility plan defining data integrity and interoperability with other systems.

AN020 The business analyst should ensure use of standards for data representation whenever possible, e.g. ISO 13972.

AN021 The business analyst should specify why data can be accessed/used.

AN022 The business analyst shall specify who is permitted to see the data.

AN023 The business analyst shall specify what classes of data can be accessed.

AN024 The business analyst shall specify how the data is protected and accessed

AN025 The business analyst should ensure that source systems have comprehensive strategies to ensure data quality, e.g. via term selections, or radio buttons in the user interface, offering predetermined data specifications, or via mappings from source data to the required exchange format.

AN026 The business analyst should make business decisions during development regarding the quality of existing historical information and the historical data to be made available.

AN027 The business analyst should ensure that original primary data ("atomic data", or "single data element") is the preferred source for the healthcare reporting service.

8.2 Data definitions

A short-term imperative might require re-use of current datasets, but if this approach does not meet current business requirements a more radical approach could be needed to develop a new source for the data, e.g. with a revised extract specification.

AN028 The business analyst can provide reference to agreed data collection structure or format such as OHDSI OMOP[2] or ISO 13972.

AN029 The business analyst can specify the level of granularity within existing fields so they can be abstracted and summarized for use at regional, national or international levels.

AN030 The business analyst shall ensure that the service has capabilities to ensure integrity of the data for its intended purpose.

AN031 The business analyst shall ensure that the service provides data validation.

AN032 The business analyst should specify when the data is accessible, when it was accessed, and what should happen to the data afterwards.

AN033 The business analyst shall ensure that data acquired include provenance of the data origin, with information such as date/ time, source, episode of care.

AN034 The business analyst shall ensure that the data requirements include the purpose for which the data are used.

AN035 The business analyst shall ensure that the data requirements include ways in which the data are collected.

AN036 The business analyst should ensure that the data requirements include the forms and/or characteristics of data required.

AN037 The business analyst should ensure that the data requirements include potential sources of data.

AN038 The business analyst should ensure that the data requirements include data characteristics such as validation criteria.

AN039 The business analyst should ensure that the data definitions are maintained in accordance with ISO 13972.

AN040 The business analyst shall ensure that the data definitional work includes stakeholders and their detailed business requirements, addressing the scope and the purposes for which data are collected (e.g. patient care, shared decision making, performance management, healthcare purchasing, outcomes management, planning, reimbursement, etc.).

AN041 The business analyst shall ensure that the data definitional work assesses the availability of the data required to meet the business requirements to understand the implications of proposed data collection.

AN042 The business analyst shall ensure that the data definitions identify coding scheme(s) to be used.

8.3 Data models

The development of a conceptual model will assist in data conformance and/or master data management (MDM) through an organization-wide ratification of common healthcare-related concepts.

At the physical level, most healthcare data reporting service design and deployment efforts will involve dimensional modelling techniques combined with traditional entity/relationship (E/R) normalized models.

AN043 The business analyst should carry out comprehensive data quality assessments on source data as part of determining the work effort for a new healthcare data reporting service initiative.

AN044 The business analyst should give business direction for handling data quality situations related to business practice; for example, if a null value is acceptable in certain business scenarios and unacceptable in others.

AN045 The business analyst should provide business rules to detect quality issues in the data staging area.

AN046 The business analyst should require that data found to be erroneous is amended in the source system and not the healthcare data reporting service, requiring the removal of erroneous records and loading updated data.

AN047 The business analyst should allow for complementary enrichment of data supported providing the original data are not altered (e.g. derivations based on postcode or ZIP code, or standard calculations on source data, such as a Body Mass Index, which is derived from weight and height).

AN048 The business analyst shall ensure that the service is designed to support reconciliation of data from source to presentation.

8.4 Dimensions

Data dimensions often have hierarchical classification. For example, providers could be physicians who could be further classified according to their speciality and/or to their unit of service, and/or to their seniority, etc. These differences within a dimension become "attributes" of that dimension.

Further detail is provided in Annex D.

AN049 The business analyst should ensure that a data element acquire metadata and business rules at each level of the reporting hierarchy.

AN050 The business analyst shall ensure that the service knows if changes occur in the nature of the dataset during the history of the service. For example, a laboratory analytical method test change of reference ranges.

AN051 The business analyst should consider how the analyses that come out of the healthcare data reporting service and the decision making at the point of care are closely related.

AN052 The business analyst shall ensure that the service enable linkage of data (e.g. combining separate data elements relating to the same patient).

AN053 The business analyst shall ensure that the service enable linkage of data with a temporal nature (e.g. combining separate data elements in a time series relating to the same patient).

AN054 The business analyst shall ensure that the service enable linkage of data with a logical similarity (e.g. combining separate data elements stored in different source systems but relating to the same patient). Further information can be found in ISO 18308.

AN055 Similar to mastering data, the business analyst shall ensure that the service is not the source for applying new business logic to datasets, e.g. hierarchical groupings.

AR001 The architect shall define healthcare data reporting service dimensions in accordance with business need and relation to process.

AR002 The architect shall ensure that the data architecture works to a set of common, conformed dimensions and measures.

9 Architecture

9.1 Components

Clause 9 addresses the approach to architecture for healthcare data reporting. This acknowledges there is a wide spectrum of approaches including:

— Contained, consented data managed and accessed by user.

— Submission of, and access to, data curated in a trusted research environment (TRE) with potential for access to anonymized or aggregate results.

— Data lake as persistent store of personal information.

— Data lab as temporary subset of data defined or created for a specific purpose.

The architecture in Figure 2 shows those components that are typical in a data reporting environment.

The main components are the source systems, extract, transform and load (ETL)/extract, load and transform (ELT) and data architecture layer, data pipeline, data storage, presentation layer and security layer. These are complemented by supporting components, including the development and test environments, application and infrastructure management, information governance and scheduling.

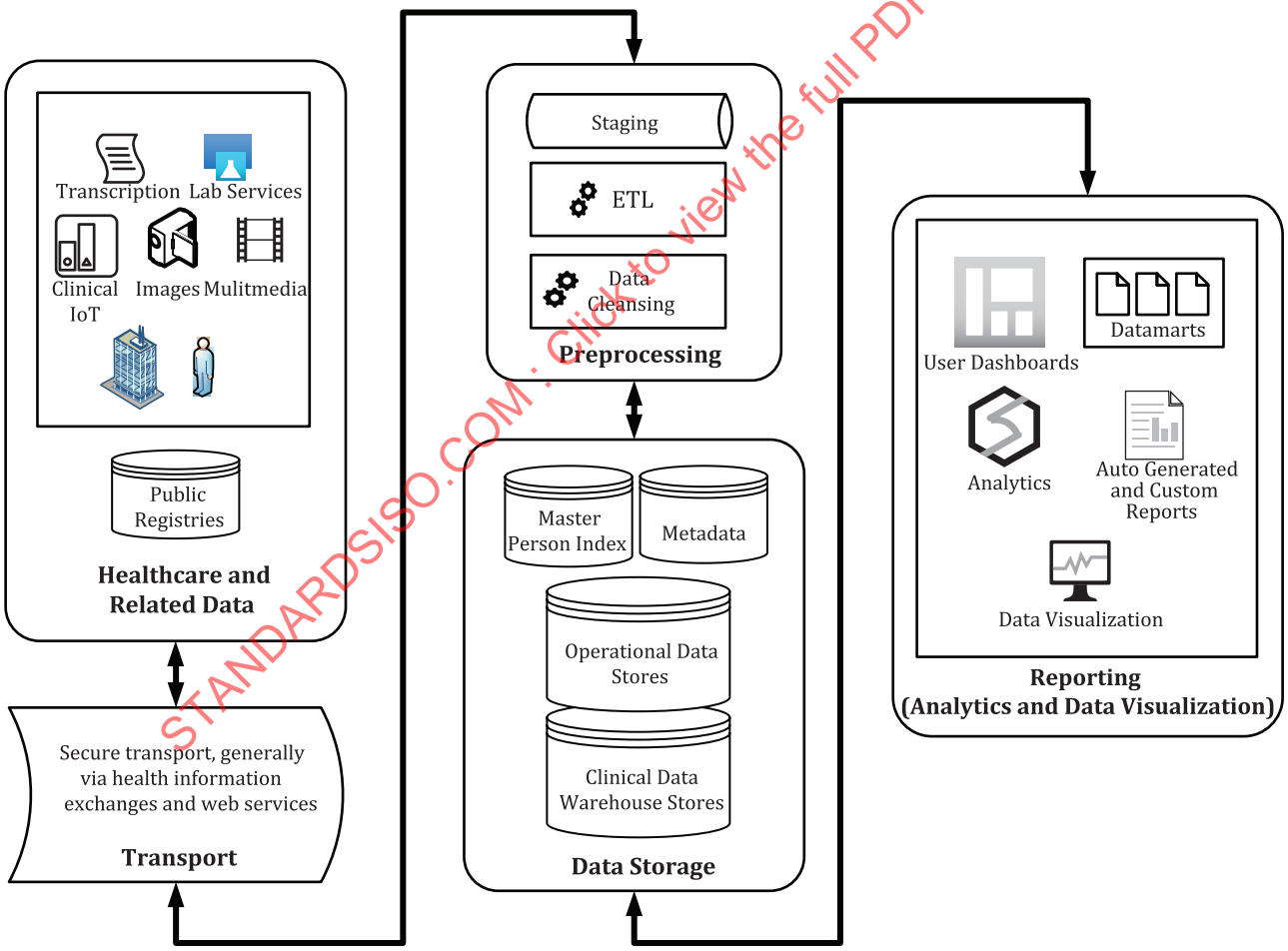

Figure 2 — Healthcare data reporting components

AR003 The architect should ensure that the transport, pre-processing and data storage layer are both source-system and application-technology independent.

AR004 The architect should ensure that the healthcare data reporting service is protected from both short-term and long-term source system changes and technology landscape changes.

AR005 The architect of the healthcare data reporting service shall ensure that the service be implemented as a layered, component-based architecture to improve reusability and to maximize deployment options.

AR006 The architect shall ensure that all healthcare data reporting service data shall be identified with provenance data such as a time-period and/or episode of care.

AR007 The architect should ensure that the healthcare data reporting service data are static or non-volatile in most instances.

AR008 The architect shall ensure that the healthcare data reporting service data are subject-oriented.

AR009 The architect shall ensure that the healthcare data reporting service data are atomic/granular – although most queries will involve aggregation, the complex analytical demands of healthcare require that data be stored in granular manner.

AR010 The architect shall ensure that new healthcare data reporting service data is added but never removed (logical delete).

AR011 The architect shall ensure that the design reflects requirements regarding volumes of data and speed of reporting.

AR012 The architect of the healthcare data reporting service shall ensure that the service provides direct access to an operational data store (ODS).

AR013 The architect should ensure that the ODS handle both query and transaction processing.

AR014 The architect should ensure that data access is optimized to reduce input/output and support analysis (e.g. through the use of indices, star schema, parallelism).

AR015 The architect should ensure that dimensions reflect different structures and functions of organization support.

AR016 The architect should ensure that dimensions allow for complex measurement questions across continuum of care without legacy systems impact.

AR017 The architect shall ensure that the data architecture underpinning healthcare data reporting service provides clarity in order to reduce end-user misunderstanding and silos of information.

AR018 The architect shall define with precision how data will be arranged within healthcare data reporting service and used to communicate requirements to stakeholders.

AR019 The architect can base the data model or conceptual model on clinical information models, archetypes or FHIR resources and /or compositions of these.

AR020 In addition to physical models, the designer should ensure the healthcare data reporting service is developed from logical models.

AR021 The architect should ensure that data models are used to fully describe source and target healthcare data reporting service data constructs.

AR022 The architect should ensure that entity definitions are clearly defined and endorsed by the organization's data governance.

AR023 The architect should ensure that data acquisition develop and maintain comprehensive information models.

AR024 The architect should ensure that data acquisition develop and maintain comprehensive data dictionaries.

AR025 The architect should ensure data acquisition has a formalized approach for the agreement and appraisal of datasets.

AR026 The architect should ensure data acquisition has agreements on accepted data standards.

AR027 The architect of the healthcare data reporting service shall ensure that the service are scalable at each data loading level. A scalable healthcare data reporting service ensures that increases in demand and throughput can be accommodated.

PR002 The service provider should ensure the platform for the healthcare data reporting service accommodate sandboxes for test analytics.

9.2 Data management

AR028 The architect should adopt a strategic approach to data management which uses a consistent set of standards for data, for data sharing and (by implication) pseudonymization and anonymization.

AR029 The architect should ensure the healthcare data reporting service is standardized. Gathered data from a variety of sources are merged into a coherent whole with multidimensional components. This process is primarily achieved through standard definitions of key common dimensions, notably subject of care (patient), provider, and service delivery location.

AR030 The architect should ensure that data are recorded in a standardized way as this can allow useful information to be gathered from multiple sources, to join up care and provide a basis for comparative analysis.

AR031 The architect should develop a conceptual model or reference model for the enterprise of concern.

AR032 The architect should use a conceptual model using standard notation, such as Unified Modelling Language or UML as in ISO 13972, Archetype Definition Language as in ISO 13606-1 and ISO 13606-2, or XML/JSON as in HL7 FHIR®[23].

AR033 The architect should use relevant International Standard for usability such as ISO 9241-210 and IEC 62366-1.

9.3 Metadata

For effective delivery of business intelligence (BI), the Designer shall ensure the healthcare data reporting service maintains/creates business metadata.

NOTE 1 ISO 19115-1 is an International Standard that defines what information should exist in a metadata document.

NOTE 2 See the ISO/IEC 11179 series.

An understanding of the technical and business meaning of a data element allows for more effective utilization of the analysis of information related to the element.

AR033 The architect shall specify the metadata that describes the transition from the operational systems to healthcare data reporting service and from healthcare data reporting service to data storage.

AR034 The architect shall specify business metadata to describe the data to the business user.

AR035 The architect shall specify technical and business metadata independently of the architecture of the data warehouse and of the application.

AR036 The architect shall specify metadata that ensures the same information produces the same results across data storage.

AR037 The architect shall specify business metadata for elements in the healthcare data reporting service to ensure that the service is consistent for the same elements within the data storage.

AR038 The architect shall specify metadata required for each field covering definition, values, alignment with business process, length, association to other fields, units of measure, special considerations, unique identifier of clinical information models used, etc.

AR039 The architect shall adopt standard naming conventions for metadata.

10 Data loading

10.1 Principles

Data loading gathers and populates healthcare data reporting service with data.

Traditionally "Extract, transform and load" (ETL) involves extracting data from data sources, transforming those data to meet operational needs, and then loading those data into the data warehouse. Tools-based ETL is typically simpler, faster and more cost effective in large and/or complex projects.

More recently, some have used "Extract, Load and Transform" (ELT) as an alternative process although consideration should be given to the consequential risks to data quality.

Poor data quality can render a healthcare data reporting service unusable. Data quality issues generally originate in the source systems, their applications or operational processes.

Challenges to extraction and transformation processing:

— If the source system is complex and/or poorly documented, then determining which data to extract can be difficult.

— Source systems cannot typically be modified, nor can performance or availability be adjusted.

AR040 The architect of the healthcare data reporting service shall ensure that the service provides appropriate tools for extract, transform and load (ETL) or extract, load, and transform (ELT) tools for input and output.

AR041 The architect should ensure that data cleansing take place within the source system, allowing healthcare reporting service data processing (ETL/ELT) to pick up the cleansed records.

AR042 The architect should provide the requirement specification including logic for the loading, mapping, processing, managing privacy/security and reporting of received data.

AR043 The architect shall involve the patients and the public in consultations on the use of data unless there is specific, justified reason.

AR044 The architect should consider repository reporting capabilities when evaluating loading tools.

AR045 The architect can adopt a strategic approach to data management that supports various data warehouses at different reporting or organizational levels.

Too often, healthcare data reporting service source systems are legacy applications that have been in production for many years, poorly documented, and have been built and maintained by many different people.

Healthcare data reporting service performs two main but very different operations: The first is loading and updating, the second is searching; protecting the first from the second promotes availability through the reduction of resource conflict.

10.2 Data acquisition

In pursuit of relevant, timely and fit-for-purpose data, the following provisions apply.

DV001 The developer shall ensure that data cleansing is part of the data acquisition process.

DV002 The developer shall ensure that each extract specification works with pioneer sites to test data collection feasibility.

DV003 The developer should ensure that all business rules that have been used to trap erroneous records are fully documented.

DV004 The developer may ensure cleansing takes place at the reporting level, if procedures to carry out this work are made explicit and transparent and meet safety requirements.

DV005 The developer should ensure that data aggregation occurs at the time of data transfer so that the aggregate is a new data field in the healthcare data reporting service or occur during or following data query including the creation of graphs or tables or the application of statistics.

DV006 The developer should conform to good practices if opting to hand-code data loading, building metadata repositories and associated browsers. Although tools for data profiling are commonly available, some working involving relationships between data elements and rules-based analysis will most likely require custom coding.

DV007 The developer should ensure that processing is automated by default, and only manual when all other options have been explored.

AR046 The architect should ensure a single source system be identified for each data item that will reside in the data warehouse.

PR003 The service provider should ensure that the healthcare data reporting service retain information regarding the data source and acquisition (provenance). For examples of guidelines surrounding the management of data provenance, see Reference [3].

PR004 The service provider should ensure that the healthcare data reporting service provide application program interfaces (API) for authenticated partner access.

PR005 The service provider support quality checking/processing of ingested data.

PR006 The service provider should ensure that the healthcare data reporting service support quality checking/processing of data staged for export.

10.3 Data requirements

AR047 The architect shall ensure that each extract specification include dataset definition, including required data dictionary cross-referencing and documentation of the metadata.

PR007 The service provider should ensure that the healthcare data reporting service receive real or near-real-time data feeds from any of the source systems, providing the infrastructure supports it.

AR048 The architect can include within each extract specification reference to, and assessment of, current datasets.

AR049 The architect should ensure that each extract specification include a published structure for the transfer of the datasets from local systems.

AR050 The architect should ensure that each extract specification identify standards requirements, either using existing International Standards such as ISO 13606-1 and ISO 13606-2, ISO 13972 or identifying the need for specific standards to meet a particular purpose.

AR051 The architect should develop a plan to achieve a single source if one is not readily available.

AR052 The architect shall define clear rules to resolve conflicts from various sources in the interim.

10.4 Data quality

Effective secondary use of data depends on knowing and defining the quality of primary data.

PR008 The service provider should ensure that data acquired is validated upon receipt.

PR009 The service provider should ensure that electronic feeds are leveraged whenever possible.

PR010 The service provider should ensure that data acquired is loaded as soon as possible.

PR011 The service provider should ensure that data acquisition meets local requirements (e.g. for regulatory purposes or for government reporting) providing a basis for the submission of timely and accurate data, with the potential for sanctions to be applied in the event of non-delivery.

PR012 The service provider should ensure that a clear audit trail of data from source to presentation shall be available to provide traceability and build confidence in the organization's management of the data.

PR013 The service provider should ensure that data originate from a robust, independent source system.

PR014 The service provider should ensure that data acquired are of high quality (e.g. accurate, valid, complete, timely, etc.).

PR015 The service provider can include within data acquisition further work to improve data quality, with validation at source, and regular reporting on performance.

PR017 The service provider should ensure that the healthcare data reporting service demonstrates its accuracy and completeness to build and maintain end user trust. The easiest way to do this is to show that the data entering the system match the data being presented, with exceptions fully explained.

Normal practice is to use the healthcare reporting service processing to trap erroneous records and feed this information back to the source business for action.

10.5 Data loading

Understanding of source systems should be the first step in data loading planning. Once there is a thorough understanding of the source data, an appropriate data loading strategy can be designed.

AR054 The architect should avoid manual steps in the data warehouse processing as they can introduce delays and performance bottlenecks.

AR055 The architect shall consider an appropriate approach (e.g. ELT vs. ETL) depending on latency and throughput requirements.

AR056 The architect should develop a source to target map during the data loading phase of the healthcare reporting service's development.

AR057 The architect should ensure that such changes are recorded in healthcare data reporting service metadata.

AR058 The architect can specify steps from source data to target data for analysis that consist of either syntactical mappings (the field or tag names of the (type) of data), and/or semantical mappings (the content data are mapped from one semantic description or code to another).

AR059 The architect should ensure transformation of relationships between source and target data should are configurable via a graphical user interface (GUI).

PR019 The service provider should ensure that data derivation is consistent (e.g. through the use of maintained reference tables).

PR020 The service provider should ensure that the healthcare data reporting service enable linkage of data stemming from different clinical records (e.g. combining separate data elements relating to the same patient or the same paediatrician for birth records analysis).

DV008 The developer should ensure that business rules are applied programmatically with an appropriate level of automatic reporting and notification of anomalies.

PR021 The service provider shall record performance benchmark data for both core data processing and end-user experience.

PR022 The service provider should ensure that the healthcare data reporting service is designed to require no manual intervention for the normal loading and processing of source data.

DV009 The developer of the healthcare data reporting service should, whenever feasible, employ proven commercial off-the-shelf (COTS) packages or mature open-source solutions rather than in-house ad-hoc coding.

DV010 The developer shall ensure that access to healthcare data reporting service conforms to security and privacy policy.

DV011 As healthcare data reporting service will contain sensitive data, the developer shall ensure access is controlled.

AR060 The architect shall ensure that there are assurance plans in line with IEC 62304 that address validation, verification, load and performance, regression testing, security, unit, system and integration testing.

PR023 The service provider should ensure that the healthcare data reporting service accommodate extraction to big data analytics such as HADOOP, SPARK, and Delta Lake.

Data loading performance tuning is highly dependent on the technology being used.

Repository reporting capabilities should be a key consideration in data loading tool evaluations.

10.6 Data management

AR061 The architect should plan a Master Data Management programme to provide the organization with a clear set of data definitions, source locations, owners and maintenance rules.

AR062 The architect shall ensure that data archiving is in line with data protection requirements.

AR063 The architect shall ensure that, although archiving ensures that the live system only has relevant data readily available in near real-time, the archive is accessible for retrieval, i.e. irrespective of changes in software, hardware, and data models.

11 Reporting

11.1 Principles

Clause 11 includes provisions on reporting. Further detail is provided in Annex E.

The primary actors for the provisions in this clause are the developer and the service provider.

Key principles for the environment and supporting requirements are as follows.

DV012 The developer should provide presentation facilities that allow end-users a choice of access types. Web-enabled deployments and their associated technologies are inherently easier to implement and deploy to a variety of devices.

PR024 The service provider should ensure that the healthcare data reporting service is robust and resilient (e.g. fail-over capability).

PR025 The service provider should ensure that the healthcare data reporting service provides timely reporting data/feedback; for local organizations this might be close to real time to support day-to-day decision-making, whereas for regional or national organizations, reporting could be weekly or monthly.

PR026 The service provider should ensure that the healthcare data reporting service is flexible enough to address the analytic needs of diverse stakeholders.

AR064 The architect shall ensure various forms of presentation are supported.

PR027 The service provider can support advanced analytics and machine learning as part of the healthcare data reporting service.

11.2 Policies

PR028 The service provider should ensure that the healthcare data reporting service supports web-based end-user access for both consumer and analyst profiles.

PR029 The service provider should ensure that the healthcare data reporting service handles some unacceptable values automatically, according to business direction.

PR030 The service provider should ensure that the healthcare data reporting service support the presentation of data at all levels: national, regional, local, etc.

PR031 The data custodian should ensure that data are accurate and, where necessary, kept up to date ('currency').

PR032 The service provider should ensure that the healthcare data reporting service facilitates metadata access to assist end users in interpretation of the data.

PR033 The service provider shall ensure people are properly informed about how and when data about them is shared so that they feel reassured that their data is being used fairly and equitably.

PR035 The service provider should ensure that the healthcare data reporting service is a basis from which information service providers can access, use and add value to nationally produced data.

PR036 The service provider shall equip all staff to handle, store and transmit personal confidential data securely, whether in electronic or paper form.

PR037 The service provider should ensure that the healthcare data reporting service have sufficient power to meet initial performance requirements.

PR038 The service provider shall train all staff to ensure that personal confidential data is only shared for appropriate purposes.

PR039 The service provider shall ensure that all staff understand their responsibilities including their obligation to handle information responsibly and their personal accountability for deliberate or avoidable breaches.

PR040 The service provider shall ensure that all staff complete appropriate annual data security training and pass a mandatory test.

PR041 The service provider shall implement measures to prevent data security breaches (e.g. through controls over user access).

PR042 The service provider shall ensure the processing of personal data is transparent in relation to the data subject

PR043 The service provider shall ensure that any information supplied relating to the processing of personal data is easily accessible, easy to understand and written in clear and plain language.

PR044 The service provider shall ensure that, where the case is made for access to data relating to identifiable individuals, the informed consent of these individuals is obtained wherever feasible.

PR045 The service provider shall ensure that, where use of identifiable data is required, and where patient consent cannot be obtained, a full justification is provided.

PR046 The service provider shall ensure that the process of determining and granting access to data is transparent and follows principles of good communication with all parties to achieve the appropriate balance between individual privacy and public benefit.

PR047 The service provider should ensure that the healthcare data reporting service provides user educational material.

PR048 The service provider should ensure that the healthcare data reporting service supports educational functions.

Reports from healthcare data reporting service inform decision-making and policy development. However, similar reports in different contexts do not necessarily mean that the processes in those contexts are identical.

As systems are increasingly networked, then certain data, useful at a local level, should be abstracted for use at higher levels.

The more often that there is a direct relationship between local and higher levels, the more informative the information at a higher level, with the added possibility of returning to a lower level to better inform on that context.

AR065 The architect should adopt a process for coordinating and rationalizing the definitions for reporting at different levels. Ideally, this would enable the federation or linkage of data warehouses to enable “summarization up” and “drill down” facilities.

AR066 The architect should include within process planning whether the intention is to provide transaction processing or data warehouse reporting and intelligence facilities.

AR067 The architect should include within process planning details of any agreement on standardization around algorithms, geographic and other identifiers, data derivations, and the construction of indicators.

AR068 The architect should ensure that dynamic querying and reports are subject to the same criteria as for static reports, clearly describing the units of the displayed data, the composition of any aggregate used, the date and time of the report, and including interpretative notes as metadata for the end user.

11.3 Data marts

Data marts are patterns used to retrieve information and are typically subject/context/area specific. In addition to basic marts of discharge data, medication data, etc. defining standards for these common dimensions expands the scope of questions addressable in the clinical reporting service.

DV012 The developer should ensure that data marts support the persistence of conformed patterns for later use.

DV013 The developer should use conformed dimensions to support the building of derived or consolidated marts for specific areas of analytic focus (e.g. chronic disease patient groups like diabetics).

DV014 The developer should ensure that data marts only maintain data for the length of time needed for analysis.

DV015 The developer should ensure that a data mart is deleted if no further analysis is done with its contents.

DV016 The developer should ensure that it is not possible for users to access stale data unless via special requests to the system provider to invoke processes to build new data marts.

DV017 The developer should ensure that tools are available for extraction of data into data cubes for time series analysis.

DV018 The developer should ensure that the means used to secure data are evaluated against impact on query performance.

DV019 The developer should allow for the choices of data for an indicator to be determined locally.

DV020 In some circumstances, the developer should use aggregated data to define both the indicator and the methodology for collecting the underlying source data themselves.

DV022 The developer should ensure that all frequently searched columns are indexed.

PR049 The service provider shall ensure that data for reporting is provided in unidentifiable (aggregate or anonymized) form

PR050 The service provider shall ensure that all users of data for purposes not defined in applicable patient consent are subject to enforceable standards regarding privacy and security of data.

AR068 The architect shall ensure that data marts have defined purpose, i.e. the type of analysis for which they were created.

AR069 The architect shall ensure that data marts are able to identify and document the permissions of data subjects.

AR070 The architect shall ensure that data marts are able to define the data minimization policy that is the minimum data set to carry out the defined analysis. This minimum data set should be specified in detail.

PR051 The service provider shall specify life cycle requirements for the development of healthcare reporting, showing it is suitable for its stated purpose and provides a robust and stable service.

AR071 The architect shall consider privacy implications in the selection and application of spatial data analysis functions for a data warehouse.

11.4 Indicators

Health system indicators represent a means of measuring and comparing performance, e.g. length of stay, outcomes, etc.

Indicators depend on data aggregation as well as predefined special measures.

Healthcare data reporting service is a natural source of health indicator information as the interpretation of indicators is facilitated through the presentation, "drill down", metadata and other features of the healthcare reporting service.

The principles of good health indicators are essentially like the principles of good use of the healthcare reporting service.

DV023 The developer should ensure that proposed indicators for a given need be subject to extensive validation prior to use, and evaluation after use.

DV024 The developer should ensure that the role of an indicator in enabling performance evaluation is the subject of consensus by peers.

PR052 The service provider should ensure the healthcare data reporting service track the change in indicators through normal processing.

PR053 The service provider should ensure that the product references the most appropriate, up-to-date open standards to ensure data quality and interoperability.

AR072 The architect should aim to use the same indicators to serve a range of different purposes and users to support both clinical policy and organizational decision making. The display and interpretation of the changes in these indicators over time is important.

AR073 The architect should take into account known indicators when determining which primary data should be collected and when

Table 1 describes a conceptual framework for health indicators. Table 1 is a high-level view of the framework. It serves to illustrate how different sections of the framework depend on different sources and dimensions of primary data. The use of the word dimension in Table 1 is pertinent to the dimensions of that framework but not necessarily directly equivalent to a data dimension in the healthcare reporting service. Indicator development and instruments for measurement of outcomes have a similar lifecycle.

Table 1 — Conceptual framework for health indicators [SOURCE: ISO 21667:2010, Table 1]

Dimensions		Sub-dimensions					
1	Health status	Wellbeing	Health conditions		Human function		Deaths
2	Non-medical determinants of health	Health behaviours	Socioeconomic factors	Social and community factors	Environmental factors		Genetic factors
3	Health system performance	Acceptability Continuity	Accessibility Effectiveness		Appropriateness Efficiency		Competence Safety
4	Community and health system characteristics	Resources		Population		Health system	

Equity

The interRAI organization and International Consortium for Health Outcomes Measurement (ICHOM) are good examples of international collaboration in developing performance indicators.

11.5 Performance

The simplest definition of acceptable performance is that the healthcare data reporting service shall be able to meet all the business requirements in the required time scales.

DV025 The developer should ensure a healthcare data reporting service is scalable, e.g. data volume capacity.

PR054 The business analyst shall document performance requirements for the healthcare data reporting service (e.g. for data load, processing, visualization, volumes).

PR055 The service provider shall be prepared to respond appropriately to processing or security incidents or near misses, logging and reporting such incidents to relevant authorities.

PR056 The service provider shall ensure that a Disaster Recovery and Business Continuity Plan is in place; this will be explicit in covering how any risk to patient data and patient health is limited and mitigated. For instance, ISO/IEC 27031 describes the concepts and principles of information and communication technology (ICT) readiness for business continuity and provides a framework of methods.

PR057 The service provider can support processing on data ingestion, e.g. optical character recognition (OCR) and/or natural language processing (NLP).

PR058 The service provider shall ensure that processing is balanced, where a balanced system ensures the available processing and storage resources are used optimally.

PR059 The service provider shall ensure that processing is resilient. A resilient system ensures that failures are handled well, e.g. are detected and (auto)recovery measures are in place.

12 Operation and service delivery

12.1 Service specification

This clause includes provisions on performance. Further detail is provided in Annex E.

The primary actor in this clause is the service provider.

SP040 The sponsor can include staff training, internal audits, pseudonymization, encryption, a process for regularly testing, assessing, and evaluating measures and breach management.

SP041 The sponsor should ensure that policies for people are accompanied by clearly defined sanctions that will be put into effect for deliberate breach or carelessness.

SP042 The sponsor of the healthcare data reporting service shall ensure that the service has transparent performance regimes to demonstrate their effectiveness and reliability.

SP043 For healthcare data reporting service initiatives to be successful, the sponsor shall ensure that query and reporting performance are acceptable contingent on service level agreements (SLA).

AN056 The business analyst shall ensure that the service can demonstrate and ensure continuity of service, particularly where current arrangements are being replaced.

AN057 The business analyst shall ensure that the service are reliable, based on agreed measures such as failures, actions to restart.

Healthcare data reporting service disaster recovery plans require periodic testing in compliance with SLA.

PR060 The service provider shall ensure that the product is clinically safe to use, e.g. see ISO 13485.

PR061 The service provider shall ensure that outputs and user interfaces are easy to use and accessible to all users.

PR062 The service provider shall ensure that network traffic is encrypted employing industry standard algorithms, e.g. currently recommended DES (Data Encryption Standard), Triple-DES or AES (Advanced Encryption Standard) algorithms.

PR063 The service provider shall ensure that data are encrypted in transit.

PR064 The service provider should ensure that data are encrypted at rest.

PR065 The service provider shall ensure that encryption is never used as a substitute for effective access control. Encrypting stored data provides assurance of data security, but it can be complex, adding system and query performance overhead.

PR066 The service provider shall ensure that encryption keys are stored securely. Many commercial solutions exist for this difficult task.

PR067 The service provider should strongly discourage the creation of proprietary processes for this purpose.

PR068 The service provider should ensure the auditing functions within healthcare data reporting service enable records of users and usage to be generated to enable the audit to answer the question, “Who has accessed which healthcare reporting service data and in what manner?” (for example, as for a forensic audit).

PR069 The service provider shall be able to audit user activity.

PR070 The service provider shall ensure that the audit functionality is itself subject to the requirements for healthcare reporting systems defined in this document.

PR071 The service provider shall ensure that modern reporting tools can be used to apply AI on audit logs to detect problems.

PR072 The service provider should ensure that systems monitor query statements and data access requests.

PR073 The service provider shall ensure that auditing allows checking of unsuccessful query select statements to find users testing their access boundaries and snooping.

PR074 The service provider should ensure that, within healthcare data reporting service, functionality records access and transaction content in terms of recording the queries and datasets accessed by users to answer the question, "What did they do?".

PR075 The service provider should ensure that only results of queries (not standard extracts) or online accesses involving patient identifiable data are stored.

PR076 The service provider should establish a user housekeeping function to ensure only bona fide users of patient identifiable data are allowed continued access, and to remove expired users from the licensed users list of the healthcare reporting service.

PR077 The service provider shall ensure that Healthcare reporting service-derived data, distributed through healthcare reporting service-based mailboxes, are deleted after a defined period.

PR078 The service provider should ensure that external users are subject to the same conditions in their contracts for access and receipt of data.

PR079 The service provider shall provide mechanisms for assuring that the healthcare data reporting service requirements have been completed and actions undertaken in a manner consistent with trust, identity, privacy, protection, safety, and security.

PR080 The service provider should ensure that the Disaster Recovery and Business Continuity Plan is explicit in covering how any risk to patient data and patient health will be minimized and mitigated. ISO/IEC 27031 describes the concepts and principles of information and communication technology (ICT) readiness for business continuity and provides a framework of methods.

PR081 The service provider should ensure that the platform for the healthcare data reporting service supports failover capabilities with regionally distant data centre to mitigate downtime in the event of catastrophic emergencies.

12.2 Service management

PR082 The service provider should ensure that the platform for healthcare data reporting service supports clustering with failover configurations to ensure zero downtime.

PR083 The service provider should ensure that the platform for healthcare data reporting service system has backup power systems.

PR084 The service provider should ensure that the healthcare data reporting service maximizes the value provided back to the point-of-care provider who is providing the source data (e.g. meets their needs and motivations in wanting to improve the care they provide to their patients).

PR085 The service provider should ensure that the healthcare data reporting service supports aggregation of data from multiple sources.

PR086 The service provider should ensure that the healthcare data reporting service supports normalization of data from multiple sources.

PR087 The service provider should ensure that the healthcare data reporting service supports data anonymization or segmentation to protect the identity and health information of individuals while retaining the usability of the data for its secondary uses, e.g. research, population health reporting.

PR088 The service provider should ensure that the healthcare data reporting service provides an abstract/summary reporting information systems and business planning.

PR089 The service provider should ensure that the healthcare data reporting service supports metadata for proper management of the data serving multiple purposes, e.g. quick retrieval, redundancy mitigation, etc.

PR090 The service provider should ensure that the healthcare data reporting service provides access to analytic tools (descriptive and inferential).

PR091 The service provider should ensure that the healthcare data reporting service provides business intelligence tools for graphical depictions.

PR092 The service provider should ensure that the healthcare data reporting service provides interfaces (e.g. ad hoc query builders, interactive reports, dashboards, portals).

PR093 The service provider should document supplier development(s).

PR094 The service provider should document supplier testing and user assurance.

PR095 The service provider should manage the issue of change control note to suppliers.

PR096 The service provider should document upgrade of local solutions.

PR097 The service provider shall test data capture and submission with healthcare data reporting service suppliers.

PR098 The service provider shall perform integration testing to ensure national and local solutions work together.

PR099 The service provider should document local deployment(s) in each community of service.

PR100 The service provider should monitor, test, and migrate results as part of the local migration to submission of the updated clinical datasets.

PR101 The service provider should ensure that data reporting begins once data flows.

PR102 The service provider shall ensure that healthcare data reporting service programs contain comprehensive training for users.

PR103 The service provider shall ensure that healthcare data reporting service programs contain comprehensive training for maintainers.

PR104 The service provider can ensure healthcare data reporting service programs contain comprehensive training for data suppliers.

PR105 The service provider shall ensure that the healthcare data reporting service programs contain training on data quality with supporting documentation.

PR106 The service provider should document data quality best practices.

PR107 The service provider should document approved uses of the data.

PR108 The service provider should document data limitations and documentation.

PR109 The service provider should document healthcare data reporting service processes and policies.

PR110 The service provider should document access to reports and associated metadata.

PR111 The service provider should document approved analytical techniques.

PR112 The service provider should document analysis and use of reports.

PR113 The service provider should document privacy and security of the data.

Annex A
(informative)

Potential benefits, uses and services

A.1 Benefits

There are many benefits to be achieved by having a coordinated approach to the development of healthcare reporting to support local, regional, or national-level analysis and reporting:

— Healthcare reporting service can include consistency of data collection and analysis across a jurisdiction.

— Healthcare reporting service can include comprehensive coverage of data collection.

— Healthcare reporting service can include cohesion of information collection enabling, for instance, linkage of patient data across primary, community and acute settings for those receiving long-term care.

— Healthcare reporting service can include timeliness of data that are collated directly from local sources on a regular schedule.

— Healthcare reporting service can include a secure environment that enables patient privacy to be maintained.

— Healthcare reporting service can include increased ability for sharing (particularly of aggregated data) for comparative purposes.

— Healthcare reporting service can include a common approach to derivation of data.

There are also some potential "disbenefits" aside from the obvious issues of security and privacy.

It is possible that performance and monitoring information could be perceived to be commercially confidential or even a threat to local managers.

Additionally, in the case of data being used to inform therapeutic activity, poor quality data might lead to patient harm.

A.2 Uses and services

There is a wide range of possible uses and services:

— Healthcare reporting service can include managing the ongoing intake and quality of the data and metadata.

— Healthcare reporting service can include pseudonymization.

— Healthcare reporting service can include management of access rights.

— Healthcare reporting service can include processing (e.g. derivations, aggregations, transformation, etc.).

— Healthcare reporting service can include development of data dimensions, data views, and reporting.

— Healthcare reporting service can include analysis and interpretation of the data and its derivations.

— Healthcare reporting service can include tools (e.g. for data presentation).

— Healthcare reporting service can include end-user education and support.

— Healthcare reporting service can service wider range of users irrespective of their desire to abstract hierarchically.

— Healthcare reporting service can have international views or peer collaboration and comparisons.

— Healthcare reporting service can support greater ranges of data types to enable hypothesis testing, simulations (digital twins), outcome assessments and broader population health assessments. Further detail regarding data types is provided in Annex C.

— Healthcare reporting service capabilities can support the monitoring of trends in (near) real-time from passive, retrospective analysis, to data driven action as part of a learning health system.

Annex B
(informative)

Privacy impact assessment

A Privacy Impact Assessment (PIA) assists organizations in identifying and minimizing the privacy risks of new projects or policies. A PIA is likely to begin early in the life of a project and run alongside the development process. The specific approach to a PIA can vary from country to country, dependent on local legislation.

A PIA should describe the information flows of the project, explain what information is used, what it is used for, from whom it is obtained, to whom it is disclosed, who will have access and any other necessary information.

A PIA should identify privacy and related risks. Some will be risks to individuals, e.g. damage caused by inaccurate data, a security breach or upset caused by an unnecessary intrusion on privacy.

A PIA should identify and evaluate privacy-focused solutions, explaining how to address each risk.

A PIA should include sign off of responsible parties in the organization(s) and record the PIA's outcomes, ensuring that the privacy risks have been signed-off at an appropriate level.

A PIA should integrate the outcomes with delivery in the project plan. It might be necessary to return to the PIA at various stages of the project's development and implementation in order to revise the PIA as the details of the project change.

The organization should make security integral to the design and ensure that the product meets industry best practice security standards.

At-scale adoption and uptake should ensure that systems provide assurance that they are implementing good data security and that personal information is handled appropriately (e.g. by testing for conformance to ISO/IEC 27001 and NIST SP 800-53).

In addition to the privacy impact assessment, a data protection impact assessment might be useful or even required. A Data Protection Impact Assessment (DPIA) is required under the GDPR any time a new project starts that is likely to involve "a high risk" to the rights and freedoms of natural persons (see Reference [4]).