
**Health informatics — Identification of
subjects of health care**

Informatique de santé — Identification des sujets de soins sanitaires



Reference number
ISO/TS 22220:2011(E)

© ISO 2011



COPYRIGHT PROTECTED DOCUMENT

© ISO 2011

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword	v
Introduction.....	vi
0.1 General	vi
0.2 Usage	vii
0.3 Responsibilities	viii
0.4 Training.....	viii
0.5 Business processes.....	viii
1 Scope	1
1.1 General	1
1.2 Objective	1
1.3 Application	1
2 Normative references.....	2
3 Terms and definitions	2
4 Components of data elements	3
4.1 General	3
4.2 Data element structure.....	4
4.3 Summary structure.....	5
5 Subject of care identifiers	6
5.1 General	6
5.2 Subject of care identifier designation	8
5.3 Subject of care identifier geographic area.....	9
5.4 Subject of care identifier issuer.....	10
5.5 Subject of care identifier type	10
6 Subject of care name	11
6.1 General	11
6.2 Family name group.....	17
6.3 Preferred name	22
6.4 Conditional use.....	22
6.5 Name title group	24
6.6 Given name group	26
6.7 Name suffix group	30
6.8 Name usage group	31
6.9 Alternative name representation	34
7 Additional demographic data	35
7.1 General	35
7.2 Date of birth	36
7.3 Date of death.....	40
7.4 Sex	43
7.5 Mother's original family name.....	44
7.6 Country (place) of birth.....	44
7.7 Birth plurality	45
7.8 Birth order	46
7.9 Identification comment	46
8 Subject of care address	47
8.1 General	47
8.2 Address line	48
8.3 Suburb/town/locality	56
8.4 State/territory/province identifier.....	57

8.5	Postal code (ZIP code)	57
8.6	Delivery point identifier	58
8.7	Country identifier	58
8.8	Address type	59
9	Subject of care electronic communications	65
9.1	General.....	65
9.2	Electronic communication medium	66
9.3	Electronic contact preference code.....	66
9.4	Electronic communication details	67
9.5	Electronic contact usage code.....	67
10	Biometric identifiers	68
11	Subject of care linkage.....	69
11.1	General.....	69
11.2	Subject of care linkage identifier	70
11.3	Subject of care linkage relationship	70
Annex A (informative) Collection of data		71
Annex B (informative) Messaging.....		74
Annex C (informative) Data matching		76
Annex D (informative) Guide for implementation of subject of care master indices		80
Annex E (informative) Guidelines for searching for a subject of care.....		83
Annex F (informative) Names condensed guide		85
Bibliography.....		87

.....

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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

In other circumstances, particularly when there is an urgent market requirement for such documents, a technical committee may decide to publish other types of document:

- an ISO Publicly Available Specification (ISO/PAS) represents an agreement between technical experts in an ISO working group and is accepted for publication if it is approved by more than 50 % of the members of the parent committee casting a vote;
- an ISO Technical Specification (ISO/TS) represents an agreement between the members of a technical committee and is accepted for publication if it is approved by 2/3 of the members of the committee casting a vote.

An ISO/PAS or ISO/TS is reviewed after three years in order to decide whether it will be confirmed for a further three years, revised to become an International Standard, or withdrawn. If the ISO/PAS or ISO/TS is confirmed, it is reviewed again after a further three years, at which time it must either be transformed into an International Standard or be withdrawn.

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

ISO/TS 22220 was prepared by Technical Committee ISO/TC 215, *Health informatics*.

This second edition cancels and replaces the first edition (ISO/TS 22220:2009), which has been technically revised.

Introduction

0.1 General

The health care system relies heavily on the ability to uniquely and accurately identify a person when they attend for care. The introduction of computerization into this process requires the clear specification of all elements of information used to support the procedural, as well as the computerized, identification of a subject of care so that the present person is associated with previous health information. Computerization is also important in supporting communication between health care professionals.

Developments in the health care system and the emergence of health networks have amplified the importance of collecting, sharing and exchanging data concerning individual subjects of care between different health care providers and between different information systems.

More effective communication between health care professionals is key to securing closer co-operation, improving the handling of subjects of care in terms of quality and continuity of care and prevention, and promoting health system efficiency.

Reliable identification of the individual has always been a critical part of the health care process. The ability of computerized systems to support and enhance the manual process of identification is vital, as is the ability of these systems to identify individuals when communicating patient information electronically. High quality identification is necessary to ensure that health care professionals have access to patient information, facilitating closer co-ordination and continuity of care and improving service in terms of prevention and follow-up. Modern service delivery networks result in greater flows of subjects of care and services across national, functional, jurisdictional, and professional boundaries. However, high quality identification can be very complex in a more integrated health care environment.

Within health care service delivery environments, the process of positively identifying subjects of care entails matching data supplied manually, electronically or through hard documentation by those subjects of care against data the service provider holds about them. This process occurs both manually, increasingly with computer support, and electronically where systems have to communicate information about individuals securely and accurately. Impediments to high quality identification include variable data quality, inadequately considered manual identification processes, differing data capture requirements and mechanisms, and varying data matching methods.

This Technical Specification identifies the data elements and relevant structure and content of the data used to manually identify individuals in a health care setting. In addition, it provides support to the identification of individuals in a consistent manner between systems that will support the natural changes in usage and application of the various names used by people over time.

This Technical Specification addresses the business requirements of identification as well as the data needed to improve the confidence of health service providers and subjects of care identification. It defines the data used to identify subjects of care and the business processes associated with this activity, whether computerized or manual. It is intended to be used both to support the processes of the identification of subjects of care by individuals and computerized identification in automated matching systems.

0.2 Usage

Within a health care service delivery context, the process of positively identifying individuals entails matching data supplied by those individuals against data the service provider holds about them.

The ability to positively identify individuals and to locate their relevant details is critical to the provision of speedy, safe, high quality, comprehensive and efficient health care. The benefits of positive identification include:

- less time wasted and inconvenience generated in hunting for and/or re-gathering information about the individual, which translates to more efficient health care;
- more complete and accurate information on which to base potentially life-critical clinical decisions;
- fewer duplicate entries for an individual leading to less duplication of testing and prescribing;
- safer treatment from having clinical details for the right individual;
- more complete and accurate information on which to base potential data use and disclosure decisions.

The delivery of health care is undergoing a paradigm change, brought about by changing consumer expectations, technological advances, economic pressures, socio-demographic change and changes in the patterns of health and ill health in communities.

These changes include:

- a) a shift from institution-centred care to subject-centred care, together with greater empowerment of the subjects of care;
- b) greater emphasis on continuity of services supporting quality and safety, health promotion and maintenance;
- c) more integrated health care, in which organizational and administrative barriers are invisible to subjects of care.

These new service directions will necessitate a much greater flow of information on subjects of care and services across functional, jurisdictional, administrative and professional boundaries. In a more integrated health care environment, positive identification is no less critical, but is much more complex. Population mobility and multiple points of access to the health care system lead to the accumulation of subject related data in a variety of fragmented, unrelated repositories. Positive subject of care identification is recognized around the world as a critical success factor for health care reform.

Below are some examples of the many barriers to successfully identifying individuals in health care settings.

- 1) Variable data quality and changes in key identifying information over time.
- 2) The patient's capacity to provide information. In a health care environment, it is important that the identification system can cope with the fact that people's memories and capacity to communicate vary according to their mental and physical capacity and to their willingness to seek and receive care. Information is often provided by third parties (family and friends) who might know the person by a preferred name rather than by the person's formal name.
- 3) Differing data capture requirements and mechanisms and varying data matching methods. This Technical Specification provides a framework for improving the confidence of health service providers and subjects of care alike so that the data being associated with any given individual, and upon which clinical decisions are made, are appropriately associated and suited to the flexibility of the health care setting.

- 4) The need to respect the wishes of the subject of care. The system should be able to accommodate the wishes of an individual who prefers that others not know their full name, or who prefers to be known by a preferred name or nickname. The system should be able to communicate the formal name when required to other systems but also to ensure that the preferred name is used so as not to cause unnecessary stress to the subject of care, or confuse family and friends.

Where permitted by law, data matching can be undertaken in a variety of contexts and settings, including for administrative purposes. However, the specific focus of this Technical Specification is the positive identification of subjects of care for health care service delivery purposes. It is recognized that implementations in different systems and national settings might vary according to local needs.

It is recognized that this Technical Specification can support national client registry projects in health care, but does not represent a registry content or structural specification.

0.3 Responsibilities

The positive and unique identification of subjects of care within and between health care organizations is a critical event in health service delivery, with direct implications for the safety and quality of health care.

It is important that responsibilities for the quality, capture, storage and use of identifying data for subjects of care, including implementation of this Technical Specification are clearly and unambiguously assigned within the organization, and documented in relevant policies, procedures and work instructions.

Users of this Technical Specification should refer to relevant privacy legislation, codes of fair information practice and other guidelines so as not to breach personal privacy in their collection, use, storage and disclosure of subject of care information.

0.4 Training

Relevant staff should receive training that highlights the nature, importance and health benefits of high quality procedures for the capture, storage and use of health identifying data and the safety implications of errors and duplications of subject of care information.

0.5 Business processes

Business processes associated with the capture, storage and use of subject identifying data should be designed and continuously improved to ensure that accurate, consistent and complete data collection, communication and storage practices are used.

Health informatics — Identification of subjects of health care

1 Scope

1.1 General

This Technical Specification indicates the data elements and structure suited to accurate and procedurally appropriate and sensitive identification of individuals in health care in a face-to-face setting supported by computer technology, or through interactions between computer systems. It provides guidelines for improving the positive identification of subjects of care within and between health care organizations.

It defines demographic and other identifying data elements suited to capturing subject of care identification in health care settings, and the wide variety of manual and computer enhanced procedures used for this process.

It provides guidance on the application of these procedures in the manual and the computer environment and makes recommendations about the nature and form of health care identifiers, the management organization to oversee subject of care identification and computer support to be provided for the identification process.

There are additional factors to be considered in providing access to distributed subject of care data, including privacy, security and data transfer mechanisms; these are outside the scope of this Technical Specification.

Application of this Technical Specification will increase the capacity for data access. Authorization of such access is determined by the application of legislation, organizational policies and guidelines, and professional ethics.

It is recognized that specific applications might require additional data to fulfil their purpose. This Technical Specification provides a generic set of identifying information, which is application independent. Implementations in different health care environments and national settings might require the establishment of data sub-sets or priorities.

1.2 Objective

The objective of this Technical Specification is to promote uniform good practice in:

- a) identifying individuals in a face-to-face, or paper-based environment, as well as in and between automated systems;
- b) recording and reporting of subject of care identifying data;
- c) ensuring that data being associated with any given subject of care, and upon which clinical communication and data aggregation are based, are appropriately associated with that individual or organization and no other.

1.3 Application

This Technical Specification is primarily concerned with the use of subject of care identification data to support patient care. It is envisaged that this Technical Specification will be used by health and health-related establishments that create, use or maintain records on subjects of care. It can be used, where appropriate, for collecting data when registering subjects of care or potential subjects of care and when reporting patient information to other systems, clinical and administrative.

Informative guides for business processes associated with capture, storage and use of subject identifying data are included in Annexes A to F.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 2022, *Information technology — Character code structure and extension techniques*

ISO 3166-1:2006, *Codes for the representation of names of countries and their subdivisions — Part 1: Country codes*

AS 4846, *Health care provider identification*

AS 4590-2006, *Interchange of client information*

ASTM E1714-00, *Guide for Properties of a Universal Health Care Identifier (UHID)*

HL7 V2.4, Health Level Seven Standard Version 2.4, *An application Protocol for Electronic Data Exchange in Healthcare Environments*, Health Level Seven Inc., Ann Arbor, Michigan, 2000

HL7 V3, Health Level Seven Standard Version 3, *Core Principles and Properties of Version 3 Models*, Health Level Seven Inc., Ann Arbor, Michigan, 2005

3 Terms and definitions

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

NOTE Individual data elements are defined in subsequent sections.

3.1

capture

deliberate action, which results in the registration of a record into a record keeping system

3.2

subject of care

SOC

subject of health care

any person who uses, or is a potential user of, a health care service

NOTE Subjects of care may also be referred to as patients, health care consumers or subjects of care.

3.3

subject of care identifier

SCI

unique number or code issued for the purpose of identifying a subject of (health) care

3.4

information system

organized collection of hardware, software, supplies, policies, procedures and people that stores, processes and provides access to information

3.5**record**

recorded information, in any form, including data in computer systems, which is created or received and maintained by an organization or person in the transaction of business or the conduct of affairs and kept as evidence of such activity

3.6**registration**

act of giving a record a unique identity in a record keeping system

3.7**storage**

function of storing records for future retrieval and use

4 Components of data elements**4.1 General**

This Technical Specification includes recommendations concerning the data elements most likely to affect the quality of identification of subjects of care. Data elements are expressed in terms of the following interrelated components. Individual organizations should identify the elements of most relevance for identification in their cultural and health setting environment.

Data concepts described in this Technical Specification are listed in Figure 1, which does not show the interrelationships between the sections, nor all the data elements that comprise these concepts, nor data structures.

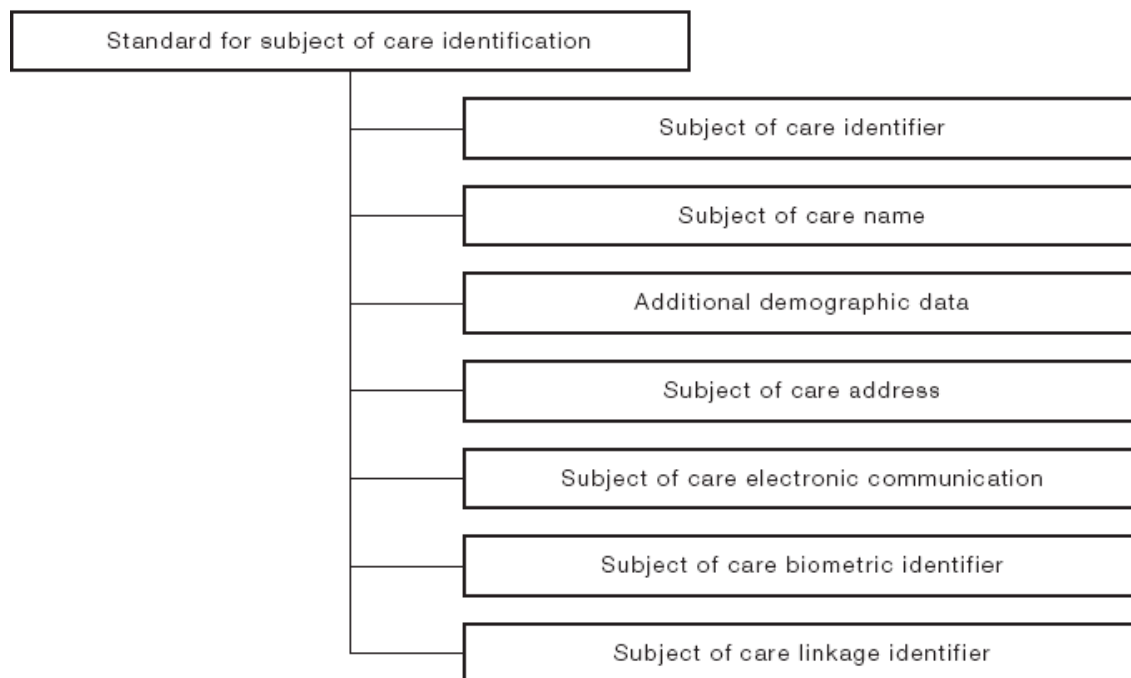


Figure 1 — Data elements and interrelated components

4.2 Data element structure

4.2.1 General

Each data element has been defined according to a set of metadata components based on ISO 11179-3. Most components (definition, data type, representational class, data domain, etc.) describe essential features of the structure of a data element. Some components, such as collection method and comments, describe additional, non-essential features and may be left blank where appropriate.

4.2.2 Synonyms

Synonyms are alternative names for this data element.

4.2.3 Definition

This is a statement that expresses the essential nature of the data element and its differentiation from all other data elements.

4.2.4 Source standards

These are details of established data definitions, or guidelines for data elements, that have been cited in this Technical Specification and are listed in the Bibliography.

4.2.5 Data type

It is recognized that different representations of the values shown in this Technical Specification might be required. Where possible, the data types are described in a manner consistent with HL7 data types. The list below provides examples of data types used in this document.

- Boolean-literal (true/false).
- Number, e.g. ISO/IEC 11404 (only used in this Technical Specification where arithmetic operations are performed).
- Character string.
- Text or unconstrained text.
- Coded text (from an agreed vocabulary or value domain).
- Constrained text (where the text is associated with a formal terminology). The difference between the coded and constrained text is the relationship to a formal, structured terminology, as opposed to a code set, or list of values.
- Unique identifier.
- Dates/times.

Though there are other data types, they are not required within this Technical Specification and thus have not been included.

4.2.6 Data domain

This refers to the values or codes acceptable for representation of the data element. The data elements contained in this Technical Specification are either free text or coded. For each data element that is coded, a code value, a descriptor of the code value and, in some cases, an alternative code (generally an alphabetic code) are provided. The code should be used for communication of this data value, the descriptor is the title of the code value and the alternative code is provided for collection of the data where the use of alphabetic code values is preferred at the point of data collection or for screen viewing. For example, the data domain for the data element "sex" is shown in Table 1 below:

Table 1 — Example of data domain representation

Code	Descriptor	Alternative code
1	Male	M
2	Female	F
3	Indeterminate	I
9	Not stated/inadequately described	N

4.2.7 Guide for use

This is additional guidance to inform the use of the data element.

4.2.8 Verification rules

These are quality control mechanisms that restrict the collection, storage or transferral of non-valid data.

4.2.9 Collection method

This contains comments and advice concerning the actual capture of data for the particular data elements in order to achieve uniformly high quality data.

4.2.10 Comments (optional)

This is any further information relevant to data element collection or storage.

4.3 Summary structure

Table 2 provides a summary guide to the structure of the data elements defined in this Technical Specification. This table indicates that for a given individual there may be multiple identifiers, names, addresses, electronic communications and biometric identifiers, but one set of core demographic data. Each of these data elements is established from a sub-set of data elements.

Table 2 — Summary of data element structure

Section of document	Data elements	Opt. ^a	Data type	Repeat data element ^b
5	Subject of care identifier	R	Text	Y
6	Subject of care name	R	Text	Y
7	Additional demographic data	O	Text	N
8	Subject of care address	O	Text	Y
9	Subject of care electronic communications	O	Text	Y
10	Subject of care biometric identifier	O	Text	Y
11	Subject of care linkage	O	Text	Y
^a Whether the data element is optional (O) or required (R). ^b Whether yes (Y) or no (N).				

5 Subject of care identifiers

5.1 General

This clause includes data elements that jointly comprise a unique identifier for subjects of care. It also outlines subject identifiers.

The combination of the subject identifier and the health care organization, the type of identifier and the name given to the identifier in the organization is one way to indicate unique identification of the subject of care.

The subject of care identifier may also be known as:

- patient ID (HL7);
- person identifier;
- unit record number (URN);
- medical record number (MRN);
- local subject identifier;
- subject identification number;
- enterprise identifier;
- area identifier;
- province/state/territory identifier;
- unique identifier (UID);
- unique health identifier (UHID).

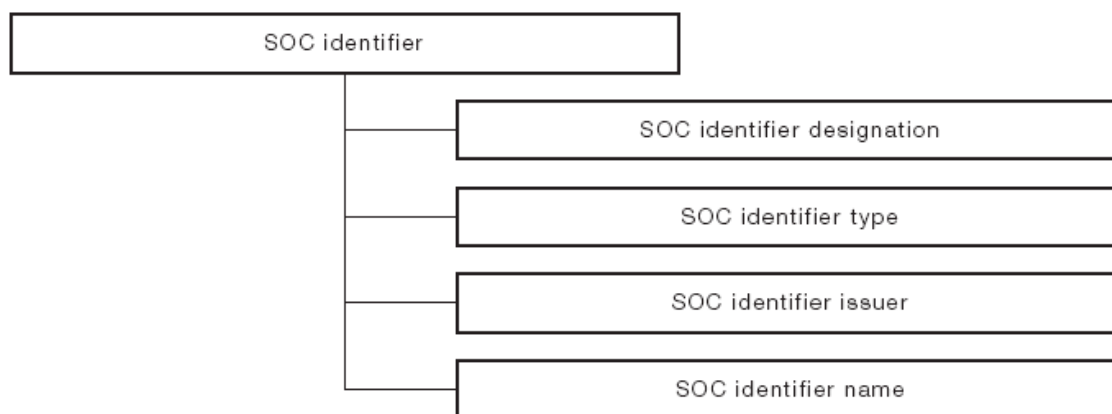


Figure 2 — Data elements for subject of care identifier

Table 3 indicates the data elements used to describe a subject of care identifier. There may be multiple identifiers collected for any one individual. Most subjects of care will have more than one identifier. Each health care organization or health care activity may designate a specific identifier as the one for use in their environment. This identification process would require the specification of the identifier type, identifier issuer and identifier name to be used within that organization/purpose.

Table 3 — Subject of care identifier data elements

Clause/ subclause	Data element name	Opt. ^a	Data type	Repeat data element ^b	Example
5	Subject of care identifier	R	Unique identifier	Y	
5.2	Subject of care identifier designation	R	Unique identifier	Y	12345678
5.3	Subject of care identifier geographic area	R	Coded text	Y	N (National)
5.4	Subject of care identifier issuer	R	Unique identifier	Y	
5.5	Subject of care identifier type	R	Coded text	Y	

^a Whether the data element is optional (O) or required (R).
^b Whether yes (Y) or no (N).

Table 4 provides examples of the identifiers used at a number of organizations for Joe Smith.

Table 4 — Example of subject of care identifiers

SOC identifier designation	SOC identifier geographic area	SOC identifier issuer	SOC identifier type
99876543	1 (local)	AB1345 (The Hill Regional Hospital)	01 (unique identifier for issuer)
NCB 913452	1 (local)	AB1345 (The Hill Regional Hospital)	02 (specialty number — pathology)
XYZ123	2 (area)	ABC4 (Northern Area Health Service)	01 (unique identifier for issuer)
998AAB990	4 (national)	SSA (Social Security Agency)	01 (unique identifier for issuer)
99812341	3 (state/province)	ABC (ABC State Department of Health)	01 (unique identifier for issuer)
3344 2256 2235 3	4 (national)	DOHAU (National Department of Health Australia)	01 (unique identifier for issuer)

The combination of any of the items along one line of Table 4 represents a subject of care identifier. For example, at The Hill Regional Hospital, the medical record number may be identified as the number to be used within that organization as the main identifier.

Some identifiers assigned by government agencies or other regulatory bodies to subjects of care may be for special purposes (billing or claiming benefits). Therefore such identifiers should not generally be used for purposes other than these special purposes. The individual requirements of legislation in individual countries should be applied.

5.2 Subject of care identifier designation

Synonym	<p>Patient identifier number</p> <p>Health care client identifier number</p> <p>UR number</p>
Definition	<p>A number or code assigned to a person by an organization, establishment, agency or domain in order to uniquely identify that person as a subject of health care within that health care organization, establishment, agency or domain.</p>
Source standards	<p>ASTM E1714-00, <i>Guide for Properties of a Universal Health Care Identifier (UHID)</i></p> <p>HL7 V2.4, Health Level Seven Standard Version 2.4 (PID-3 Patient identifier list)</p>
Data type	<p>Unique Identifier</p>
Data domain	<p>Identifier code according to issuer rules for identifier generation.</p>
Guide for use	<p>Individual agencies, establishments or collection authorities may use their own alphabetic, numeric or string coding systems.</p> <p>The combination of the subject of care identifier designation, subject of care identifier type, subject of care identifier issuer and subject of care identifier name uniquely identify the person.</p> <p>ASTM E1714-00 should be used as a guide to the properties of subject of care identifiers.</p>
Validation rules	<p>Field may not be blank.</p>
Collection method (informative)	<p>The following criteria and characteristics of the subject of care identifier (SCI) are adapted from ASTM E1714-95, <i>Guide for Properties of a Universal Healthcare Identifier (UHID)</i>.</p> <p>Atomic: the SCI should be a single data item. It should not contain sub-elements that have meaning outside the context of the entire SCI, nor should it consist of multiple items that are taken together to constitute an identifier.</p> <p>Content free: the SCI should not depend on possibly changing or possibly unknown information pertaining to the subject of care. Including content in the SCI will make it impossible to assign the “correct” identifier if that information is not known. It also leads to invalid situations if the information changes, e.g. what happens to an identifier based on sex if the subject has a sex change procedure.</p> <p>Longevity: the SCI system should be designed to function for the foreseeable future. It should not contain known limitations that will force the system to be restructured or revised radically.</p> <p>Permanent: once assigned, an SCI should remain with the subject of care. It should never be reassigned to another subject, even after the subject's death.</p> <p>Unambiguous: whether represented in automated or handwritten form, an SCI should minimize the risk of misinterpretation. Where using string identifiers, it is important to be aware of possible confusion with the number “0”, with the letter “O” and the number “1” with the letter “l”.</p> <p>Unique: a valid subject of care identifier designation should identify one and only one subject of care.</p>

5.3 Subject of care identifier geographic area

Definition A code representing the geographic area within which this identifier is used.

Source standards Not applicable.

Data type Coded text.

Data domain	Code	Description	Alternative code
	1	Local subject identifier	L
	2	Area, region or district identifier	A
	3	State/province/territory identifier	S
	4	National identifier	N

Guide for use Codes 1–4 are recommended for storage and the alternative codes are suggested for collection of data where the full descriptor cannot be displayed.

The fields can be a multiple occurring field, one occurrence per group of subject of care identifier fields. This is one of the essential elements required for inclusion in a subject of care identifier.

The combination of the subject of care identifier (designation, subject of care identifier geographic location, subject of care identifier issuer and subject of care type) uniquely identifies a client.

- 1) Local subject identifier: the unique identifier assigned by an organization for use within a specific site or small group of sites local to an area or city.
- 2) Area, region or district identifier: the unique identifier assigned by an organization for use within a specific area, such as a health region or district.
- 3) State/province/territory identifier: the unique identifier assigned by an organization for use within a specific political boundary such as a state or territory or province, such as a provincial health care identifier.
- 4) National identifier: the unique identifier assigned by an organization at a national level to identify a health care client, e.g. health insurance number or national health identifier.

Verification rules Not applicable.

Collection method Not applicable.

5.4 Subject of care identifier issuer

Synonym	Patient Identifier Assigning Authority (HL7 PID 3.4)
Definition	The organization, agency or provider that allocates a subject of care identifier designation.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000. (STF-2 ID code <assigning authority>) HL7 V3, Health Level Seven Standard Version 3: Health Level Seven Inc., Ann Arbor, Michigan 2005 (Authority Name)
Data type	Unique identifier.
Data domain	Unique identifier of the issuer of the subject of care identifier.
Guide for use	It is desirable that this field be represented using established, formal identifiers to assist in communication between organizations. As such an identifier does not always exist, implementation might require that the field be free text. Where an established identifier exists for a health care service provider who issued the identifier, the unique identifier of that organization should be used in this field.
Verification rules	Not applicable.
Collection method (informative)	Not applicable.

5.5 Subject of care identifier type

Synonym	Identifier type code.
Definition	The type of the identifier within the organization, e.g. unique patient identifier, health care card, pension card.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-3.5 — Identifier Type Code)
Data type	Coded text.
Data domain	The type of identifier to be established within the health care organization.

Guide for use Each organization that issues health care identifiers should identify the code to be used for each type of identifier they issue.

Subject of care identifier names could include the following examples.

01 — Unique identifier within this organization (to be used where only one number is used for identification within the organization).

02 — Medical record number, patient record number, UR number.

Service specific numbers for numbers issued by a specific service, e.g.:

21 — radiology

22 — pathology

23 — pharmacy

Examples of identification numbers indicating card types:

31 — top level cover card

32 — medium level cover card

33 — lowest level cover card

51 — Insurance/finance/pension number

82 — research purpose number

Verification rules Not applicable.

Collection method (informative) Not applicable.

6 Subject of care name

6.1 General

The subject of care name is a composite data element that is captured through the combination of name title group, given name group, name usage group and name suffix group.

There may be more than one name recorded for each subject of care. At least one name should be captured. There may be multiple titles, given names, suffixes and name usage for any name. Only one name can be the person's preferred name at any given point in time.

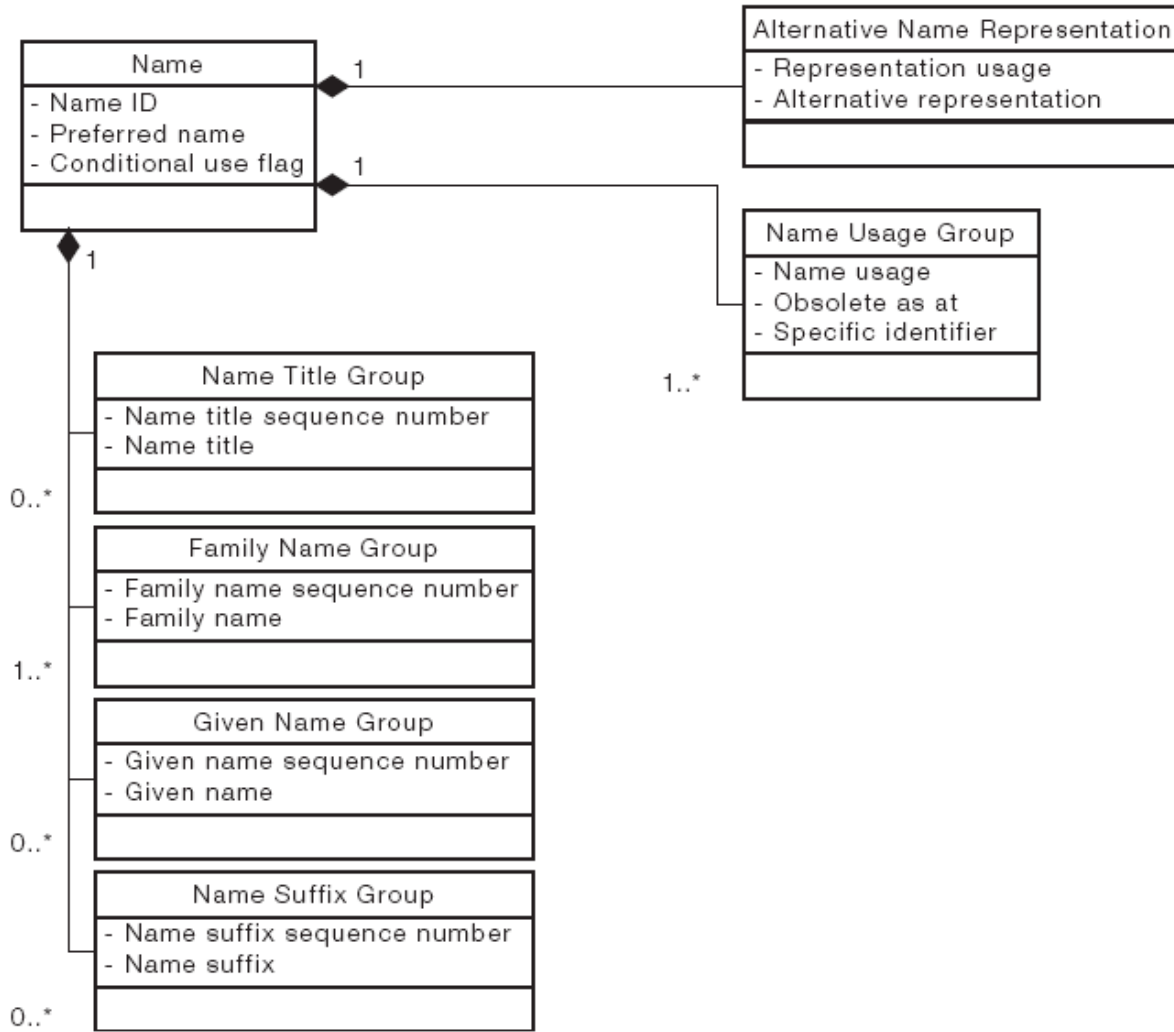


Figure 3 — Relationships of name data elements

Figure 3 indicates the data elements available for the combination data element of name. Table 5 shows a summary example outline of each of these elements. For any given purpose, specific elements of the name might be used, such as for communicating with a patient or for messaging an order. This Technical Specification does not attempt to identify the elements for these purposes but the elements are required to clearly and specifically specify a name.

Figures 4, 5 and 6 provide examples of possible name data elements in Australia.

Table 5 — Subject of care name data elements

Clause	Data element name	Opt. ^a	Data type ^b	Repeat data element ^c	Example ^b
6	Subject of care (SOC) name	R	N/A	Y	N/A
6.2	Family name group	R	Text	Y	N/A
6.2.2	Family name	R	Text	N	Brown
6.2.3	Family name sequence number	R	Number	N	1
6.3	Preferred name	R	Boolean-literal	N	N
6.4	Conditional use	O	Code	N	3 misspelling
6.5	SOC name title group	O	N/A	Y	N/A
6.5.2	SOC name title	R	Text	N	Dr
6.5.3	SOC name title sequence number	R	Number	N	1
6.6	Given name group	O	N/A	Y	N/A
6.6.2	Given name	R	Ext	N	Mary
6.6.3	Given name sequence number	R	Number	N	1
6.7	Name suffix group	O	N/A	Y	N/A
6.7.2	Name suffix	R	Text	N	Jnr (Junior)
6.7.3	Name suffix sequence number	R	Number	N	1
6.8	Name usage group	O	N/A	Y	N/A
6.8.2	Name usage	R	Code	N	1 reporting
6.8.3	Name usage start date	R	Date	N	01042006
6.8.4	Name usage end date	O	Date	N	22062006
6.8.5	Usage identifier	O	Unique identifier	N	113456 Insurance Company unique ID for this person
6.9	Alternative name representation	O	N/A	Y	N/A
6.9.2	Representation usage	R	Code	N	
6.9.3	Alternative representation	R	Text	N	

^a Whether the data element is optional (O) or required (R).
Required (the group may be required or, where the group is optional, the individual data elements within the group may be marked as required. In this case, where the group exists the required elements should be present.

^b N/A = Not applicable.

^c Whether yes (Y) or no (N).

EXAMPLE 1 Event 1 produces name ID number 1: an unnamed newborn girl, registered as her mother's daughter with the name Baby 1 of Jane Jones. This name is a temporary name.

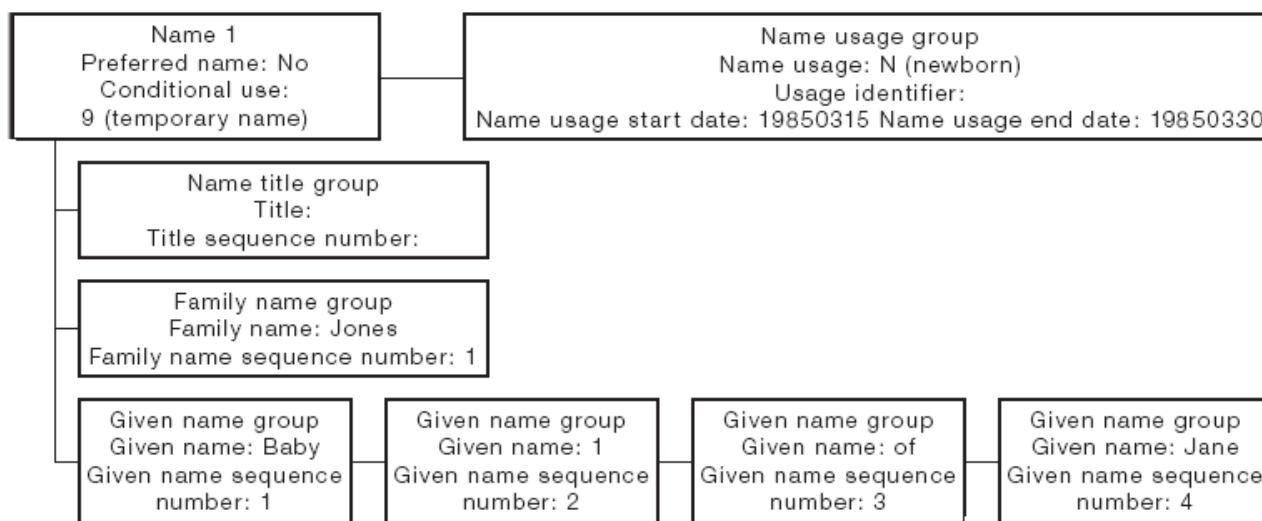


Figure 4 — Example of newborn name

Event 2: The same subject is admitted as a child who now has a name in normal use. This name is also used when claiming benefits from the health insurance company. The example above remains with two changes (the “obsolete as at” date and the preferred name has changed, and a new name set has been added).

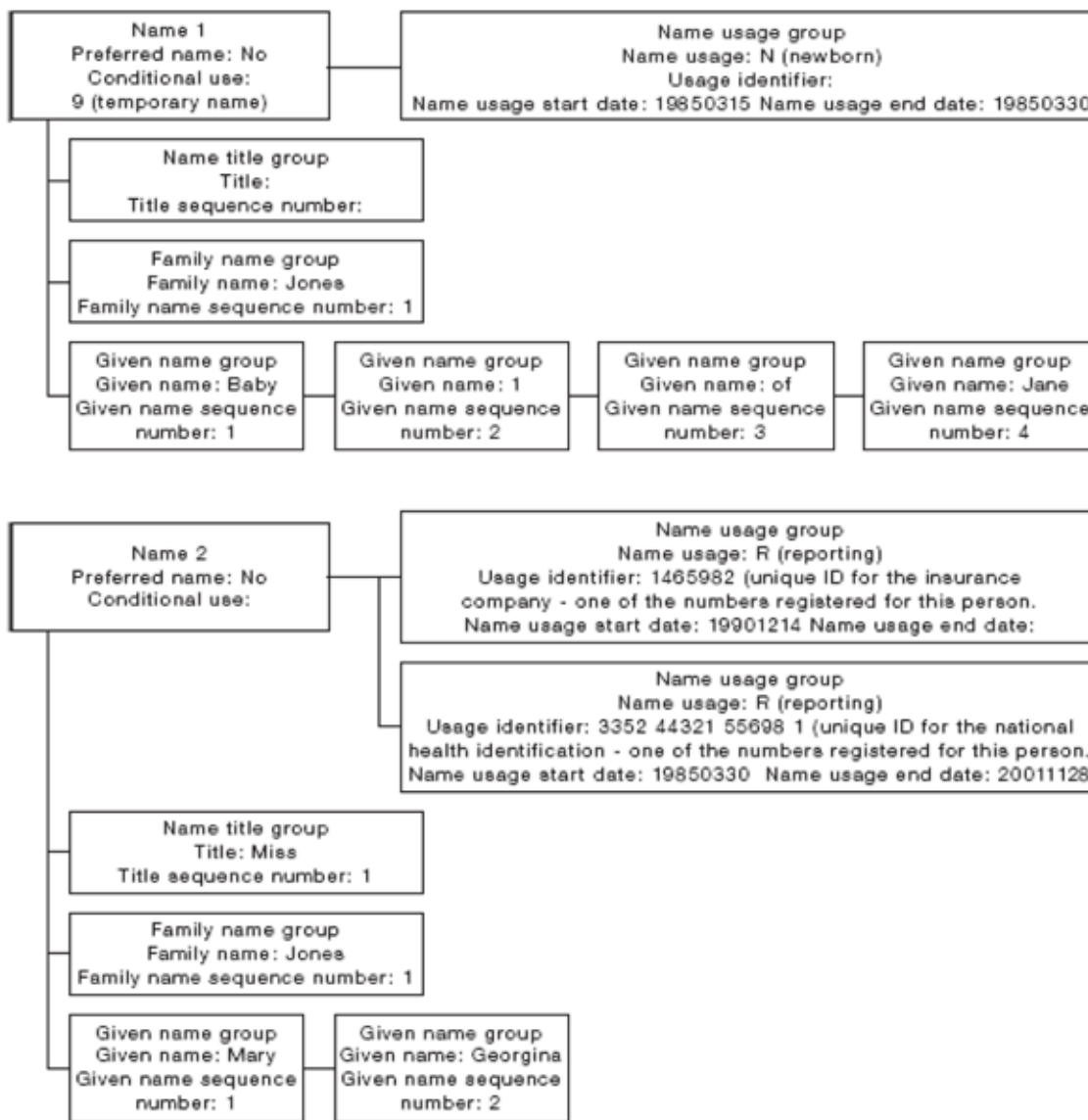


Figure 5 — Name example 2

Event 3: Mary Jones (now a teenager) has decided that she would rather be called Marie but her old name is still the name for reporting for her insurance company so the name usage of her second name is changed and a new preferred name is entered.

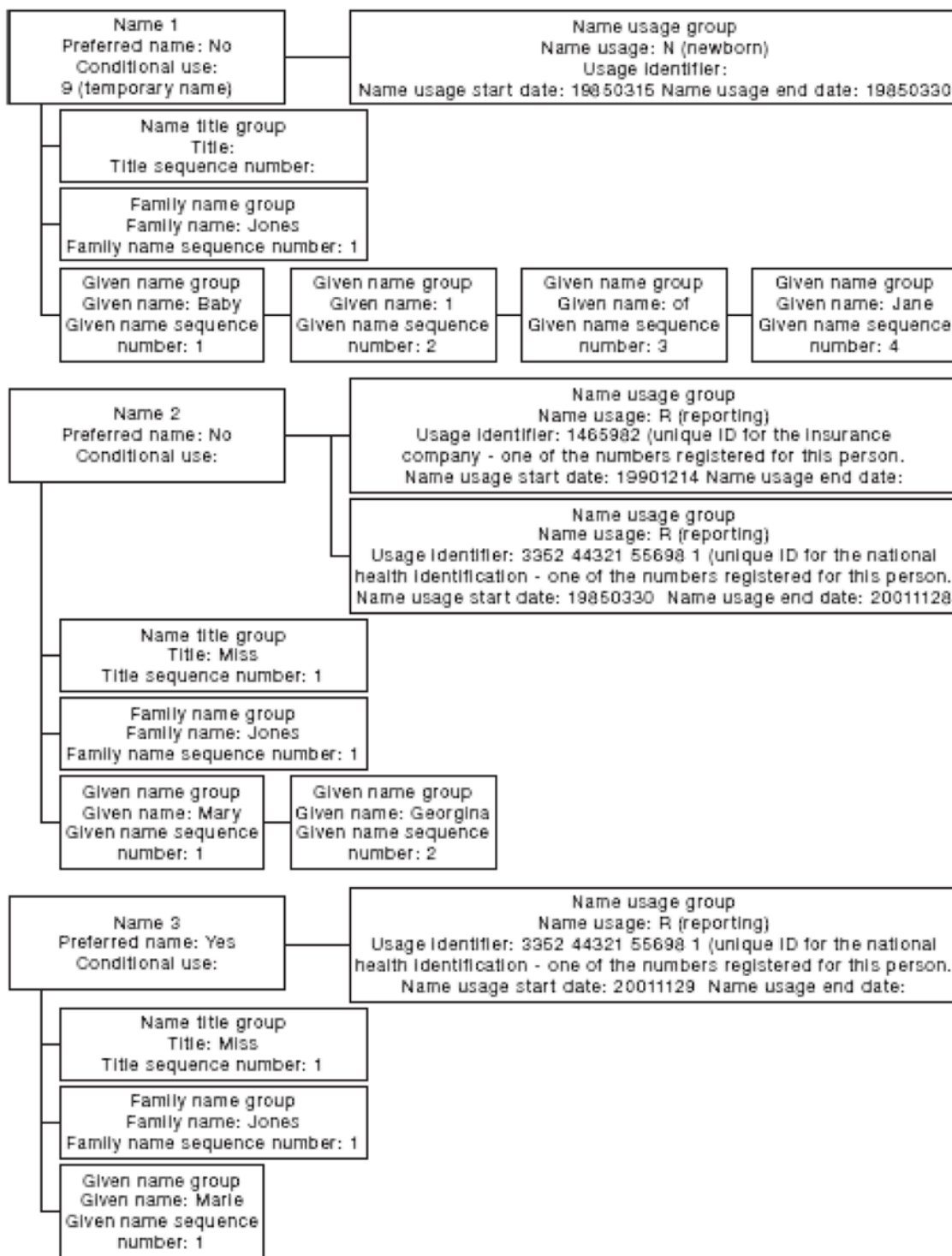


Figure 6 — Name example 3

EXAMPLE 2 Name with alternative representation/s.

.....

6.2 Family name group

6.2.1 General

This group includes each family name element of a specific family name set and indicates the sequence within which the names should be used. The group includes family name and a family name sequence number. Subject of care name is the combination of name title (and sequence number), family name/s (and sequence number/s), given name/s (and sequence number/s), name suffix/s (and sequence number/s) and usage information such as preferred name, name usage and name conditional use.

6.2.2 Family name

Synonyms	Surname Last name
Definition	The part of a name a person usually has in common with some other members of his/her family, as distinguished from his/her given names.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-5 Patient name, in part)
Data type	Text.
Data domain	Not applicable.
Guide for use	Mixed case should be used.

The family name should be recorded in the format preferred by the subject. The format should be the same as that written by the subject on a (pre-) registration form or in the same format as that printed on an identification card, to ensure consistent collection of name data.

Where a person uses two names as a family name each name should be registered as a family name within the group (e.g. for Von Schreider, enter as family name sequence 1 — Von and family name sequence 2 — Schreider).

It is acknowledged that some people use more than one name (e.g. formal name, birth name, married or maiden name, tribal name). Depending on the circumstances, each name should be recorded with the appropriate name usage (see 6.8).

A subject of care should generally be registered using their preferred name as it is more likely to be in common usage and more likely to be used on subsequent visits to the health care establishment and will therefore make identification more accurate and easier.

Use this field to record only the family name. It should not be used to record any other related information, including “cancelled” or “duplicate”.

Verification rules Alphabetic characters, hash (#), punctuation (.',-), special symbols below (though the list is not all inclusive) and spaces.

Eth	ð
Tilde	ñ
Grave	ò
Acute	ó
Circumflex	ô
Diaeresis	ö
Stroke	ø
Macron	ā
Breve	ă
Caron	č
Cedilla	ñ
Double acute	ő
Dot	ı
Ligature	Æ
Ring	Å

Collection method (informative) The following question format might assist with data collection:

- What is your family name?
- Are you known by any other family names that you would like recorded? If so, what are they?
- Please indicate, for each name above, the “usage type” of family name that is to be recorded, e.g. identity card name, other (any other name that you are known by).

NOTE Whenever a subject of care informs the establishment of a change of name, e.g. following marriage or divorce, the former name should be recorded as an other name (see name usage). Previous family names should not be deleted or overwritten.

EXAMPLE 1 Mary Georgina Smith informs the hospital that she has been married and changed her family name to Jones. Record Jones as her preferred family name and record Smith as an other name.

Hyphenated family names

Sometimes subjects of care with hyphenated family names use only one of the two hyphenated names. It is useful to record each of the hyphenated names as an other name. If the subject has a hyphenated family name, e.g. Wilson-Phillips, record Wilson-Phillips in the family name as the preferred name and record Wilson and Phillips separately.

Collection method (informative) **Registering an unidentified subject of care**

The default for unknown family name should be “unknown” in all instances and the name recorded as an other name. A fictitious family name such as Doe shall not be created, as this is an actual family name. When the subject's name becomes known, it shall be recorded as the preferred name and the other name of “unknown” shall not be overwritten.

Registering a pseudonym

This process might be required in order to mask the identity of an individual, e.g. in the case of HIV testing where the subject of care has the right of anonymity. This is the case in many jurisdictions. In this case, a pseudonym (fictitious or partial name) in lieu of their full or actual name is used. It is recommended that the subject be asked to record both the pseudonym (other name) in addition to a legally known name, e.g. identity card name. This requires the local system to be able to identify which name is to be used as the preferred name for the purpose of the test. This might require the temporary change of a name to preferred name, which is changed to an other name after the pseudonym use is over.

It is important to recognise that this is not total anonymity as there is a link to the usual identifiers of the individual. In a jurisdiction where a subject of care is allowed full anonymity, to register a pseudonym against a true identity would be to breach the anonymity. Where anonymity is required or permitted, no link can be made.

Registered unnamed newborn babies

When registering a newborn, the mother's family name should be used as the baby's family name unless instructed otherwise by the mother. Record unnamed babies with a name usage value of N (newborn).

Subjects of care with only one name

Some people do not have a family name and a given name; they have only one name by which they are known. If the subject has only one name, record it in the family name field and leave the given name field blank.

Registering subjects of care from disaster sites

Subjects treated from disaster sites should be recorded with a name usage of “other name”. Local business rules should be developed for consistent recording of disaster site subject details. Care should be taken not to use identical dummy data (family name, given name, date of birth, sex) for two or more subjects from a disaster site. For example, use of the surname “Unknown1”, “Unknown2” in the family name can be used to clearly differentiate between individuals. Some organizations use an injury description in the given name field to assist in identification. The use of a standard start to the name supports the use of search strategies that treat all names that begin in this way as “unreliable”.

If the family name needs to be shortened

If the length of the family name exceeds the length of the field, truncate the family name from the right (that is, dropping the final letters). Also, the last character of the name should be a hash (#) to identify that the name has been truncated.

Use of incomplete names or fictitious names

Some health care facilities permit subjects of care to use a pseudonym (fictitious or partial name) in lieu of their full or actual name. It is recommended that the subject be asked to record both the pseudonym (other name) in addition to a legally known name, e.g. identity card name.

Collection method Baby for adoption

(informative)

(continued)

The word “adoption” should not be used as the family name, given name, or alias for a newborn baby. A newborn baby who is available or scheduled for adoption should be registered in the same way as other newborn babies are registered. However, if a baby born in the hospital is subsequently adopted and is admitted for treatment as a child, the baby is registered under their adopted (current) name, and the previous name should be marked as a usage type of name linkage forbidden by law. Systems should be built to ensure that this link is not used without authorisation.

Punctuation

If special characters form part of the family name they should be included, e.g. hyphenated names should be entered with a hyphen.

Hyphen

There should be no spaces before or after a hyphen, i.e. between the last letter of “Wilson” and the hyphen, or a space between the hyphen and the first letter of “Phillips”.

EXAMPLE 2 Wilson-Phillips.

Apostrophe

There should be no spaces before or after the apostrophe, i.e. between the “O” and the apostrophe, nor a space between the apostrophe and “Brien”.

EXAMPLE 3 O'Brien, D'Agostino.

Full stop

There should be no space before a full stop, i.e. between “St” and the full stop. A space should be left between the full stop and “John”.

EXAMPLE 4 St. John, St. George.

Space

If the subject has recorded their family name as more than one word, displaying spaces in between the words, record their family name in the same way leaving one space between each word.

EXAMPLE 5 Van Der Humm, Le Brun, Mc Donald.

Alternatively, the parts of a family name can be recorded as different family names with a sequence number to indicate the order in which they are to be used. Local practice should dictate the method of use.

Prefixes

Where a family name contains a prefix, such as one to indicate that the subject of care is a widow, this should be entered as part of the family name field. When widowed, some Hungarian women add “Ozvegy” (the abbreviation is “Ozy”) before their married family name, e.g. Mrs. Szabo would become Mrs. Ozy Szabo. That is, “Mrs. Szabo” becomes a name usage of other name and “Mrs. Ozy Szabo” becomes the preferred name.

Misspelled family name

If the subject's family name has been misspelled in error, update the family name with the correct spelling and record the misspelled family name as an other name in the name usage field with conditional use indicating that the name was a misspelling. Recording misspelled names is important for filing and identifying documents that might be issued with previous versions of the subject's name and for future identification of the subject, should they contact the health system again and have the same problem with spelling. Discretion should be used regarding the degree of recording that is maintained.

Alternative name representations.

In some environments, alternative representations of the name are used either as the primary representation (in preference to this field) or in addition to this field. This is a local implementation variation of this Technical Specification (see 6.9).

6.2.3 Family name sequence number

Definition	An indicator of the order of use for family name/s.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan 2000
Data type	Numeric.
Data domain	<ol style="list-style-type: none"> 1 First family name within a name set. 2 Second family name within a name set. 3 Third family name within a name set. 4 Fourth family name within a name set. 5 Fifth family name within a name set. 6 Sixth family name within a name set. 7 Seventh family name within a name set. 8 Eighth family name within a name set. 9 Ninth and subsequent family name within a name set.
Guide for use	<p>To be used in conjunction with family name.</p> <p>Multiple family names may be recorded for a given name. Each family name should have a family name sequence number recorded.</p>
Verification rules	Not applicable.
Collection method	Not applicable.

6.3 Preferred name

Definition Indicates the name by which the subject chooses to be identified.

Source standards Not applicable.

Data type Boolean.

Data domain	Code	Description
	Y	This is the preferred name.
	N	This is not the preferred name.

Guide for use This is the name that will be displayed when this subject of care is referenced. It is to be used on screens, reports, letters and data collections, unless there is a need to use a different name to suit a specific usage (see 6.8, name usage group).

There should only be one preferred name recorded for a subject at any point in time.

Systems should consider the preferred name as the default name, e.g. if there is no identity card name recorded in the system, the preferred name should be used for this purpose.

Verification rules Only one name for any individual subject of care can be allocated as the preferred name at any point in time.

Collection method (informative) Not applicable.

6.4 Conditional use

Definition An indicator of specific conditions or rules that should be applied to a subject of care name.

Source standards HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PV-16 VIP indicator, Table 0099 VIP indicator code)

Data type Coded text.

Data domain	Code	Description
	1	Unreliable information.
	2	Known misspelling.
	3	Name not to be used.
	4	Name linkage forbidden by law.
	6	Special confidentiality requirement.
	9	Temporary name.

Copyright © ISO 2011. All rights reserved.

Guide for use

The following are field value definitions for conditional use.

- Unreliable information. This should be used where it is known that the name recorded is a fictitious or partial name. This name should match any other name in search processes (though other identifying criteria such as date of birth and sex should still be applied).
- Known misspelling. These names might be useful for limited matching purposes as some names are often misspelled in a similar manner, e.g. Graham/Graeme. This indicator allows the user to indicate that this is a misspelling, but that it is one that should be retained for potential future matching.
- Name not to be used. Indicates that this name should not be used when referring to this subject. Certain tribal names may become “not to be used”. Searches on this name will retrieve only the new name (without display of the old name). This value provides a mechanism for management of names that have become taboo. However, there is no legal restriction upon access from this name to the current name, in fact such forced movement to the new name is the objective of the value.
- Name linkage forbidden by law. Indicates that this name and all names prior to it in name sequence are not to be displayed or indicated in any way when searching for or dealing with information and events associated with a name that is subsequent to this one in the name sequence. This value provides a mechanism for management of name change and offers the possibility of maintaining allergy or major health information that would otherwise be lost.

If a search is made that matches the name linkage forbidden by law name, that name is retrieved, along with any other name details with lower name sequence numbers, but without any indication of the names subsequent to this one, or to episodes of care that relate to the newer name/s. This restriction applies to identification systems and to all other information held in any system. If a search is made that retrieves a name with a sequence number subsequent to this one, the display would have no indication of the name linkage forbidden by law name or any names prior to that name.

Country legal requirements for such cases as adoption and witness protection should be considered when developing identification usage systems. Consideration may be given in different cultures and legal situations to permit the barrier between the two names to be raised. This element is not covered here but should be considered when implementing in health information systems. Implementation should also consider the establishment of a new individual as a process for management of adoption and witness protection, in which case this value would not be used and the new individual would have no possibility of linkage to prior health information for personal health care, or for epidemiological or service review.

- Special confidentiality requirements. This may apply to names for which episodes are attached that should only be accessible to specified authorised persons. There should be a specific need to implement this additional security level. Local policy should provide guidance to the use of this code.
- Temporary name. This may be used when the name of the subject has not been determined, e.g. the subject has been given a surrogate name in emergency situations where the subject's name could not be obtained. This flag may be used in circumstances where the subject's name has not yet been registered with the Registry of Births, Deaths and Marriages, e.g. newborn names allocated by registering authorities. This name should match any other name in search processes (though other identifying criteria such as date of birth and sex should still be applied).

Verification rules

Valid codes or blank.

6.5 Name title group

6.5.1 General

This group holds details of each title relevant to a specific family name for this subject of care. The group indicates the actual title and the sequence in which that title should appear before the person's name.

EXAMPLE Dr Rev Brown would have Dr as the first sequenced name title and Rev as the second sequence name title.

6.5.2 Name title

Synonym	Title Honorific Name prefix (HL7)
Definition	An honorific form of address commencing a name, used when addressing a subject of care by name, whether by mail, by phone, or depending upon cultural situation in person.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-5 <i>Patient name</i> , in part)
Data type	Coded text.
Data domain	The following is a list of commonly used abbreviations. This list is not exhaustive.

Name Title	Abbreviation
Admiral	Adm
Bishop	Bish
Brother	Br
Canon	Canon
Captain	Capt
Constable	Con
Corporal	Corp
Dame	Dame
Damen	Dam
Doctor	Dr
Father	Fthr
General	Gen
Herr	Herr
The Honourable	Hon
Madame	Mdm
Mademoiselle	Mmsl

Major	Maj
Master	Mstr
Miss	Miss
Mister	Mr
Mrs	Mrs
Ms	Ms
Pastor	Pst
Private	Prv
Professor	Prof
Reverend	Rev
The Right Honourable	The Rt. Hon
The Right Reverend	The Rt. Rev
Sergeant	Sgt
Sir	Sir
Sister	Sr
The Venerable	The Ven

Guide for use

Subject of care “name” is the combination of name title (and sequence number), family name, given name (and sequence number), name suffix (and sequence number), name usage and name conditional use.

An example of name title is “Mr” for Mister.

Multiple name titles may be recorded. Each name title should have a name title sequence number recorded.

EXAMPLE “Professor Sir John Markham” would have name title sequence numbers assigned as follows:

- 1 Prof
- 2 Sir

That is, “Professor” would have a name title sequence number of 1 and “Sir” would have a name title sequence number of 2.

Name title should not be confused with job title.

Mixed case should be used (rather than upper case only).

Verification rules

Title of Master should only be used for subjects less than 15 years of age.

Title of Doctor and Professor should only be applicable to subjects of greater than 20 years of age.

Titles of Dame, Miss, Mrs, Ms and Sr should only be accepted for females.

Titles of Br, Mstr, Mr and Sir should only be accepted for males.

6.5.3 Name title sequence number

Definition An indicator of the order of use for name titles.

Source standards AS 4846-2006, *Health care provider identification*

Data type Numeric

- Data domain**
- 1 First name title.
 - 2 Second name title.
 - 3 Third name title.
 - 4 Fourth name title.
 - 5 Fifth name title.
 - 6 Sixth name title.
 - 7 Seventh name title.
 - 8 Eighth name title.
 - 9 Ninth and subsequent name title.

Guide for use To be used in conjunction with name title.

Multiple name titles may be recorded for a given name. Each name title should have a name title sequence number recorded.

EXAMPLE “Professor Sir John Markham” would have name title sequence numbers assigned as follows:

- 1 Prof
- 2 Sir

That is, “Professor” would have a name title sequence number of 1 and “Sir” would have a name title sequence number of 2.

Subject of care “name” is the combination of name usage, name title (and sequence number), family name, given name (and sequence number), name suffix (and sequence number) and name for continued use flag.

Verification rules Not applicable.

Collection method (informative) Not applicable.

6.6 Given name group

6.6.1 General

The given name group is associated with a specific family name and set of titles, suffixes and name usage rules. There may be many given name groups, within each of which there is a given name and a given name sequence number.

6.6.2 Given name

Synonym	<p>First name</p> <p>Middle name</p> <p>Forename</p> <p>Second name</p> <p>Other given name</p> <p>Other given name/s</p>
Definition	The subject's identifying name(s) within the family group or by which the subject is uniquely socially identified.
Source standard	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-5 Patient name, in part)
Data type	Text.
Data domain	Not applicable.
Guide for use	<p>Mixed case should be used.</p> <p>Health care establishments may record given names (first and other given names) as a single entry or as individual entries with an associated sequence number. Individual given name entry with a given name sequence number is the preferred method of collection to support more efficient searching of the database. This Technical Specification applies regardless of the format of data recording.</p> <p>Given names should be recorded in the format preferred by the subject. The format should be the same as that written by the subject on a (pre-) registration form or in the same format as that printed on an identification card, to ensure consistent collection of name data.</p> <p>It is acknowledged that some people use more than one given name, e.g. formal name, birth name, nickname or shortened name, or tribal name, depending on the circumstances.</p> <p>A subject of care should generally be registered using their preferred name as it is more likely to be in common usage and more likely to be used on subsequent visits to the health care establishment. The name usage code data element (see 6.8, name usage) can be used to distinguish between the different types of names that might be used by the subject of care.</p> <p>Use this field to record given names. It should not be used to record any other related information, e.g. "cancelled" or "duplicate".</p>
Collection method (informative)	<p>The following question format might assist with data collection:</p> <ul style="list-style-type: none"> — What is your given name? — Are you known by any other given names that you would like recorded? If so, what are they? — Please indicate, for each name above, the name usage that is to be recorded (e.g. identity card name (if different to preferred name); other (any other name that you are known by)).

**Collection method
(informative)**
(continued)

Whenever a subject of care informs the establishment of a change of given name, e.g. they prefer to be known by their middle name, the former name should be recorded according to the appropriate name usage value. Previous given names should not be deleted or overwritten.

EXAMPLE 1 “Mary Georgina Smith” informs the hospital that she prefers to be known as “Georgina”. Record “Georgina” as her preferred given name and record “Mary” as the health identity card given name.

EXAMPLE 2 The establishment is informed that “Baby of Louise Jones” has been named “Mary Jones”. Retain “Baby of Louise” as the newborn name and also record “Mary” as the preferred given name.

Registering an unnamed newborn baby

An unnamed (newborn) baby is to be registered using the mother's given name in conjunction with the prefix “Baby of”. For example, if the baby's mother's given name is Fiona, then record “Baby of Fiona” in the (preferred) given name field for the baby. This name is recoded under the newborn name usage code. If a name is subsequently given, record the new name as the preferred name and retain the newborn name.

Registering unnamed multiple births

An unnamed (newborn) baby from a multiple birth should use their mother's given name plus a reference to the multiple birth. For example, if the baby's mother's given name is “Fiona” and a set of twins is to be registered, then record “Twin 1 of Fiona” in the given name field for the first born baby, and “Twin 2 of Fiona” in the given name field of the second born baby.

In the case of triplets or other multiple births the same logic applies. The following terms should be used for recording multiple births:

Twin	Use Twin	i.e. Twin 1 of Fiona
Triplet	Use Trip	i.e. Trip 1 of Fiona
Quadruplet	Use Quad	i.e. Quad 1 of Fiona
Quintuplet	Use Quin	i.e. Quin 1 of Fiona
Sextuplet	Use Sext	i.e. Sext 1 of Fiona
Septuplet	Use Sept	i.e. Sept 1 of Fiona

These names should be recorded as a temporary name in conditional use and as a newborn name in name usage. When the babies are named, the actual names should be recoded as the preferred name. The newborn name is retained as it might exist on paper work or in other associated systems and there might be a need to be able to link these names together over time.

Shortened or alternate first given name

If the subject uses a shortened or alternate version of their first given name, record this as their preferred name, the actual name as their reporting name associated with a specific reporting agency (e.g. identity card name) and any alternative versions as other names as appropriate.

EXAMPLE 3 The subject's given name is Jennifer but she prefers to be called Jenny. Record “Jenny” as the preferred given name and “Jennifer” as her health identity card name.

EXAMPLE 4 The subject's given name is “Giovanni” but he prefers to be called “John”. Record “John” as the preferred name and “Giovanni” as the health identity card name.

**Collection method
(informative)**
(continued)

Punctuation

If special characters form part of the given names, they should be included. Hyphenated names shall be entered with the hyphen.

Hyphen

There should be no spaces before or after the hyphen, i.e. between the last letter of “Anne” and the hyphen, or a space between the hyphen and the first letter of “Maree”.

EXAMPLE 5 Anne-Maree, Mary-Jane.

Registering an unidentified subject of care

If the subject's given name is not known, record “unknown” in the given name field and use the name usage value of other name, and conditional name use of unreliable name. When the subject's name becomes known, add the actual name as preferred name usage (or other as appropriate). The other name of “unknown” shall not be deleted or overwritten.

Use of first initial

If the subject's given name is not known, but the first letter (initial) of the given name is known, record the first letter in the given name field. A full stop shall not follow the initial.

Subjects of care with only one name

Some people do not have a family name and a given name: they have only one name by which they are known. If the subject has only one name, record it in the family name field and leave the given name field blank.

6.6.3 Given name sequence number

Definition An indicator of the order of use for given names

Source standards AS 4846-2006, *Health care provider identification*

HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000

Data type Numeric

Data domain

- 1 First given name.
- 2 Second given name.
- 3 Third given name.
- 4 Fourth given name.
- 5 Fifth given name.
- 6 Sixth given name.
- 7 Seventh given name.
- 8 Eighth given name.
- 9 Ninth and subsequent given name.

Guide for use To be used in conjunction with given name.

For example: “Mary Georgina Smith” would have “Mary” as a given name sequence number of 1 and “Georgina” with a given name sequence number of 2.

6.7 Name suffix group

6.7.1 General

This group indicates a specific name suffix used with a defined name. The sequence number indicates the sequence in which the suffixes are to be used for display, printing, etc.

6.7.2 Name suffix

Definition	Additional term used following a person's name to identify a subject of care.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-5 Patient name, in part)
Data type	Coded text.
Data domain	Name suffix should be abbreviated. The following is a list of commonly used abbreviations.

Name Suffix	Abbreviation
Junior	Jr
Member of Parliament	MP
Queen's Counsel	QC
Senior	Sr
First	I
Second	II
Third	III
Fourth	IV
Fifth	V
Sixth	VI
Seventh	VII
Eighth	VIII
Ninth	IX
Tenth	X
PhD	PhD
Medical Doctor	MD

Guide for use Mixed case should be used (rather than upper case only).

Subject of care name is the combination of name title (and sequence number), family name, given name (and sequence number), name suffix (and sequence number), name usage and name conditional use.

A name may have more than one name suffix group. Each name suffix group has one suffix and sequence number. The name suffix sequence number field is used to control the sequence in which name suffixes are presented.

EXAMPLE John Markham Jnr MP.

Validation rules Not applicable.

Collection method (informative) Not applicable.

6.7.3 Name suffix sequence number

Definition An indicator of the order of use for name suffix.

Source standards AS 4846-2006, *Health care provider identification*

Data type Numeric.

Data domain

- 1 First name suffix.
- 2 Second name suffix.
- 3 Third name suffix.
- 4 Fourth name suffix.
- 5 Fifth name suffix.
- 6 Sixth name suffix.
- 7 Seventh name suffix.
- 8 Eighth name suffix.
- 9 Ninth and subsequent name suffix.

Guide for use To be used in conjunction with name suffix.

Multiple name suffixes may be recorded. Each name suffix should have a name suffix sequence number recorded.

EXAMPLE John Markham Jnr MP.

In the example above, “Jnr” would have a name suffix sequence number of 1 and “MP” would have a name suffix sequence number of 2

6.8 Name usage group

6.8.1 General

This is a classification that enables differentiation between recorded names for a subject of care. A name may be associated with a specific unique identifier, in which case the usage type should indicate the identifier type and identifier issuer and identifier name in the specific identifier field.

A name may have many name usage groups, but each group should have a name usage indicated and may have associated dates and unique identifier for reporting.

6.8.2 Name usage

Definition A classification that enables differentiation between the usage of names for a subject of care. An individual name may have many name uses.

Source standards HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-5 Patient name, Table 0200 Person Name Usage code)

AS 4590-2006, *Interchange of client information* (Clause 3.6, Person Name Usage code)

Data type Coded text.

Data domain	Code	Description	Alternative code
	1	Reporting	R
	2	Newborn name	N
	3	Professional or business name	B
	4	Maiden name (name at birth) (original name)	M
	5	Registered name (legal name)	L
	8	Other name (alias)	O

Guide for use More than one name can be recorded for a subject of care and each of these names may have more than one usage at any given point in time. Each name should have one or more name usages associated with it. Record as many as required.

Where there is only one name recorded, that name is assumed to be the name for all other purposes, including unique identification and financial reporting. However, where the subject offers more than one name, clarification should be obtained from the subject to ensure accurate recording of the various names. All currently used names, as well as names by which the subject has previously been known, should be recorded if these are known. These names should never be deleted from the system, as there may be existing paper work with the old names, or reference from other agencies to the name used in the past or in error.

Reporting name (R) is the subject's name as it is to be used for reporting, when used with a specific identifier. There should only be one reporting name for any given specific identifier at a time, therefore the combination of usage type, identifier and obsolete, as it should clearly identify the name to be used for reporting.

Newborn name (N) type is reserved for the identification of unnamed newborn babies. It acts as a preferred name until an actual name is available, at which time it is no longer used.

Professional or business name (B) is the name used by the subject of care for business or professional purposes.

Maiden name (M) is the name used by the subject of care prior to marriage.

Other name (O) is any other name that a subject is also known by, or has been known by in the past; that is, **all** other names. This includes misspelled names or name variations that are to be retained as they have been used to identify this subject. More than one other name may be recorded for a subject.

Validation rule Not applicable.

Collection method Not applicable.

6.8.3 Name usage start date

Definition	The date at which this name usage for the name to which the usage is associated starts.
Source standards	Not applicable.
Data type	Date.
Data domain	Valid dates or YYYY0000 (zeros).
Guide for use	<p>This field is linked to a name usage instance to indicate that this usage started for this name group as at the date stored. This allows the computer systems to identify which name to use for a given purpose. It should not exclude a name from searching processes.</p> <p>It is essential that the date be presented in a manner that will be understood in the regional area in which it is used.</p> <p>The presentation may be modified to include “/” or “.” to act as a display delimiter according to local policy.</p>

6.8.4 Name usage end date

Definition	The date at which this name usage for the name to which the usage is associated ceased to be used.
Source standards	Not applicable.
Data type	Date.
Data domain	Valid dates or YYYY0000 (zeros).
Guide for use	<p>This field is linked to a name usage instance to indicate that this usage ends for this name group as at the date stored. This allows the computer systems to identify which name to use for a given purpose. It should not exclude a name from searching processes.</p> <p>It is essential that the date be presented in a manner that will be understood in the regional area in which it is used.</p> <p>The presentation may be modified to include “/” or “.” to act as a display delimiter according to local policy.</p>

6.8.5 Usage identifier

Definition	The combination of identifier type, identifier issuer and identifier name that specifies the link between this name and reporting or other unique identifier usage.
Source standards	Not applicable.
Data type	Text.
Data domain	SOC identifier issuer + SOC identifier issuer + SOC identifier name.
Guide for use	This field is used to provide a link between the name and a unique ID and identification issuer for a purpose, usually reporting.
Validation rule	A unique identifier should exist for this subject of care for this set of identifier information.
Collection method (informative)	It is suggested that a system would provide a set of existing identifiers from which the user can select, rather than expect manual entry of this information.

6.9 Alternative name representation

6.9.1 General

This group of data elements indicate the representation of a name when the alphabetic representation is not the one used within a community. This is sometimes called the domestic name, local representation or local name. Any alternative font or character-based representation of a name set should be included here.

A name may have multiple alternative name representations.

6.9.2 Representation usage

Synonym	Domestic name type Type of local representation of name Alternative character set handling scheme (HL7)
Definition	Name of the representational form used.
Source standards	ISO/IEC 2022:1994, <i>Information technology — Character code structure and extension techniques</i>
Data type	Coded text.
Data domain	Valid language representations.
Guide for use	This field is used to indicate domestic representations, e.g. domestic Russian name, Chinese character representation.
Verification rules	Not applicable.
Collection method (informative)	Not applicable.

6.9.3 Alternative representation

Synonym	Domestic name Character representation Local name representation Domestic name representation
Definition	Alternative representation of the subject of care name using alternative styles of representation such as characters or language character set variations for local display.
Source standards	Not applicable.
Data type	Text.
Data domain	Not applicable.
Guide for use	Name represented using an alternative font or character system. This field is linked to the representation usage element.
Verification rules	Not applicable.
Collection method	In some areas, this representation may be the principal or required representation for the health care system, rather than the alphabetic representation.

7 Additional demographic data

7.1 General

This clause describes nine additional data elements (summarized in Table 6) that may be used, where relevant, to maximize the likelihood of positive identification of a subject of care. These data should only be collected where required for the identification of the subject. Where an item such as date of birth is composed of sub-elements, these are described in full, and the composite item forms a heading in the table. Figure 7 shows the structure of this data group.

Table 6 — Additional data elements

Subclause	Data element name	Opt. ^a	Data type	Repeat data element ^b	Example
7.2	Date of birth	O		N	
7.2.2	Birth date	R	Date	N	19601209
7.2.3	Date of birth accuracy indicator	O	Coded text	N	AAE
7.2.4	Date of birth follow-up indicator	O	Boolean	N	
7.3	Date of death			N	
7.3.2	Death date	R	Date	N	19991208
7.3.3	Estimated date (of death) flag	O	Coded text	N	AAE
7.3.4	Source of death notification	O	Coded text	N	2
7.4	Sex	R	Code	N	1 (male)
7.5	Mother's original family name	O	Text	N	N
7.6	Country (place) of birth	O	Code	N	1301
7.7	Birth plurality	O	Code	N	1 (singleton)
7.8	Birth order	O	Code	N	2 (second of a multiple birth)
7.9	Identification comment	O	Text	N	This is not a duplicate entry with ID 343551234 checked and confirmed as two different people

^a Whether the data element is optional (O) or required (R).
Required (the group may be required or, where the group is optional, the individual data elements within the group may be marked as required. In this case, where the group exists the required elements should be present.

^b Whether yes (Y) or no (N).

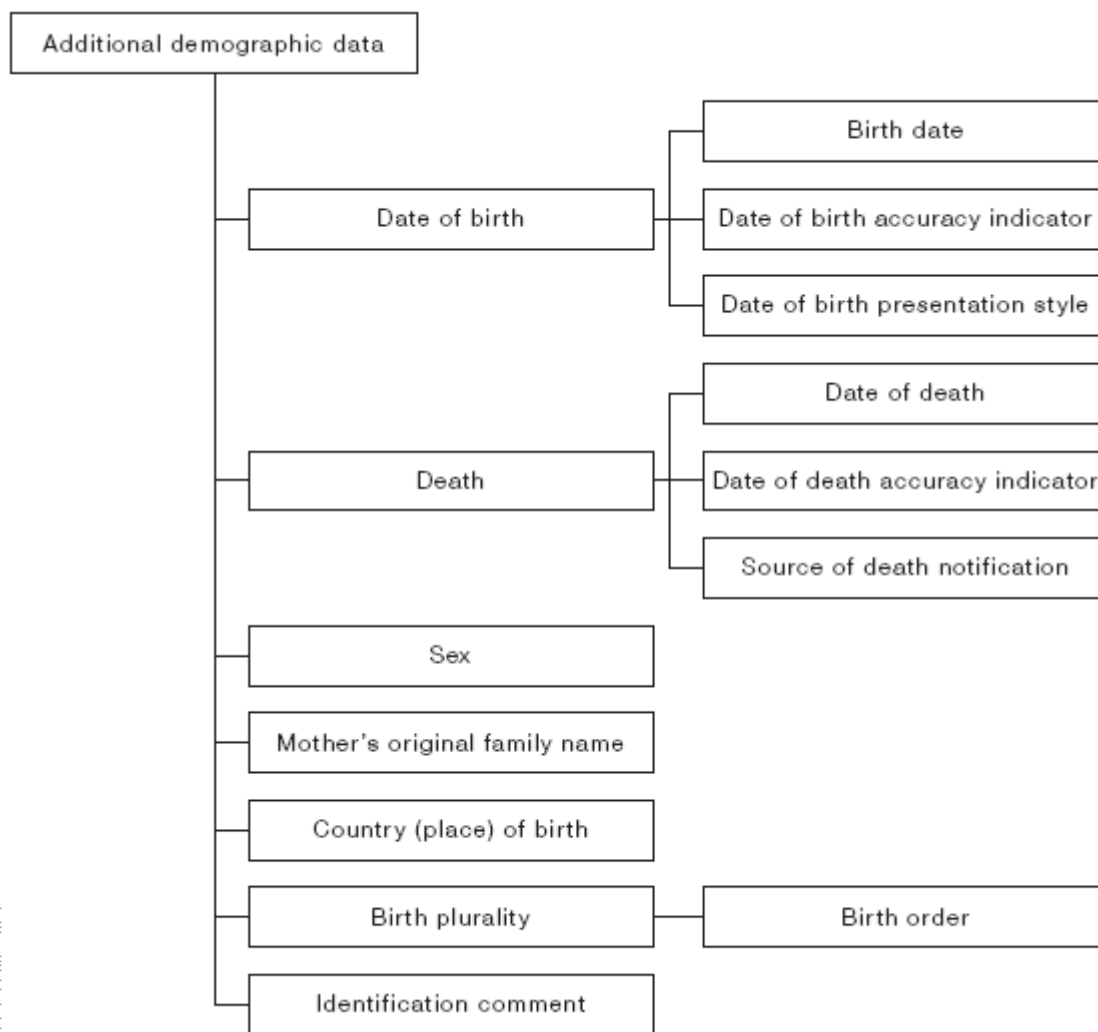


Figure 7 — Structure of additional demographic data

7.2 Date of birth

7.2.1 General

The concept “date of birth” comprises the elements birth date, estimated date (of birth) flag, and date of birth follow-up.

7.2.2 Birth date

Synonym	Birthdate.
Definition	The date of birth of the subject of care.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-7 Date/time of birth)
Data type	Date.
Data domain	Valid dates.
Guide for use	<p>Enter the full date of birth using day, month and year. Use leading zeros if necessary.</p> <p>Date of birth is required for accurate identification. Where it is not accurately known, an approximate date should be used to derive age, for analysis by age at a point in time, for approximate searching and for other applications and reporting within health care.</p> <p>If date of birth is not known or cannot be obtained, provision should be made to collect an estimate age. Collected or estimated age would usually be in years for adults, and to the nearest three months (or less) for children aged less than two years. Additionally, an estimated date flag should be reported in conjunction with all estimated dates of birth.</p>
Verification rules	<p>Birth date should be less than or equal to date of death.</p> <p>This field should:</p> <ul style="list-style-type: none"> — be less than or equal to the date of record creation, otherwise resulting in a rejection error; — not be null; — be a valid date of birth. <p>It might be useful to include a warning edit on the subject registration system or master index to alert staff if they attempt to enter a date of birth more than 120 years previous to the current year.</p>
Collection method (informative)	<p>Punctuation (back slashes or hyphens) or spaces should not be used in data collection. In cases where all components of the date of birth are not known or where an estimate is arrived at from age, use 00 for day and 00 for month and estimate year of birth according to the subject of care's approximate age. As soon as known, or on re-presentation, the date of birth (DOB) field shall be updated.</p> <p>It is recommended that in cases where all components of the date are not known or where an estimate is arrived at from age, a valid date be used together with the date accuracy indicator.</p> <p>This approach is recommended to allow collection of information that is known to be accurate and to identify where further follow-up with the subject of care might be appropriate.</p> <p>It should be noted that countries may determine a different presentation method to meet their national norms, e.g. if the date is presented in European format of DDMMYYYY, this should be consistent within all representations within the given geographic and political region in order to ensure consistent and accurate representation of data.</p>
Comment	This may be used in conjunction with date of birth accuracy indicator.

7.2.3 Date of birth accuracy indicator

Definition An indication of the accuracy of a reported date at the date component level for dates represented in YYYYMMDD format. Where the date is represented in an alternative presentation, the code groups would differ.

Source standards Not applicable.

Data type Coded text.

Data domain Any combination of the values A, E, U representing the corresponding level of accuracy of each date component of the reported date, including:

Code	Description
AAA	Accurate date
EEE	Estimated date
UUU	Unknown date
EAA	Accurate day and month, estimated year
AAU	Unknown day, accurate month and year
UUE	Unknown day and month, estimated year
UUA	Unknown day and month, accurate year

The domain values will be dependent upon the date of birth presentation style value. The examples below are for a presentation style DDMMYYYY.

Data domain	Date component (for format DDMMYYYY)		
	(D)ay	(M)onth	(Y)ear
Accurate	A	A	A
Estimated	E	E	E
Unknown	U	U	U

Guide for use This is used to record the level of certainty or estimation used in recording the subject of care's birth date.

Provision of a date of birth is often a mandatory requirement in data collections. However, at times, the actual date, or part thereof, is not known or is estimated.

This data element is designed to flag the part, or parts, of a date that has been estimated when a date provided is based on an approximation rather than the actual date. This data element can assist with manual searching and record linkage processes, e.g. when the date of birth is a component of the linkage key.

It should be noted that the sequence and content of the data domain is dependent upon the format of date displayed in the country where the information is being used.

EXAMPLE 1 A date has been sourced from a reliable source and it is known to be accurate. The "date accuracy flag" should be recorded as (AAA).

EXAMPLE 2 If only the age of the person is known and there is no certainty of the accuracy of this, then the "date accuracy flag" should be recorded as (EUU). That is the day and month are "unknown" and the year is "estimated".

EXAMPLE 3 If a person was brought in unconscious to an emergency department of a hospital and the only information available was from a relative who only knew the age and the month of the subject's birth, then the "date accuracy flag" should be recorded as (EAU).

Collection constraints: if constraints for the collection of the date are imposed, such as "a valid date should be input in an information system (software) for 'unknown' date components", the accuracy date indicator is recommended to be used along with the date as a way of avoiding the contamination of the valid dates with the same value on the respective date components.

EXAMPLE 4 Some systems use YYYY0107 and some use YYYY0101 when only the year is known. When month and year are known, some use the fifteenth day as the date, i.e. YYYYMM15. Where this occurs in a data collection that is used for reporting or analysis purposes, there will be dates in the collection with the attributes YYYY0107, etc. that are accurate and some that are not accurate. Without a corresponding flag to determine this accuracy, the analysis or report will be contaminated by those estimated dates.

For future users of the data, it might also be essential that they know the accuracy of the date components of a reported date. The date of birth accuracy flag can also be useful for operational purposes to indicate the level of accuracy that a date has been collected at any point in time, indicating whether the stored date needs to be followed up until it reaches its intended minimal required accuracy.

For example, if a person was brought in unconscious to an emergency department of a hospital the level of accuracy of the date collected at that point might not be satisfactory. It is likely that the correct date of birth can be obtained at a later date. The date of birth accuracy flag provides information on the accuracy of the entered dates that might require further action.

Validation rules	Any combination of the codes A, E and or U.
Collection method (informative)	This data element shall always be used in conjunction with a date of birth.
Comment	Most computer systems require a valid date to be recorded in a date field, i.e. the month part should be an integer between 1 and 12, the day part should be an integer between 1 and 31 (with rules about the months with less than 31 days) and the year part should include the century. However, in actual practice, the date or date components are often not known. This means that a date should be included and it should follow the rules for a valid date. It therefore follows that, while such a date will contain valid values according to the rules for a date, the date is in fact an "unknown" or "estimated" date. For future users of the data it is essential they know that a date is accurate, unknown or estimated and which components of the date are accurate, unknown or estimated.

7.2.4 Date of birth follow-up indicator

Definition	Date of birth requires follow-up to obtain a more accurate date.
Synonym	Birthdate follow-up indicator.
Definition	Flag that indicates when the current date of birth requires follow-up to obtain a more accurate date.
Source standards	Not applicable.
Data type	Boolean.

Data domain	Y = the date does need follow-up N = the date does not need follow-up
Guide for use	Where the date is estimated but it is not possible to obtain a more accurate date – as in the case where the subject of care does not know the exact date of birth, this flag allows the system to highlight the need for follow-up.
Verification rules	Where the date of birth accuracy indicator is AAA, date of birth follow-up indicator should be N.
Collection method (informative)	Not applicable.
Comment	May be used in conjunction with date of birth accuracy indicator.

7.3 Date of death

7.3.1 General

This concept “date of death” comprises the elements death date, estimated date (of death) flag.

7.3.2 Death date

Synonym	Date of death.
Definition	The date of death of the subject of care.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-7 date/time of birth)
Data type	Date.
Data domain	Valid dates.
Guide for use	Enter the full date of death using day, month and year. Use leading zeros if necessary.
Validation rule	Where date of birth is collection, date of death should be equal to or greater than the date of birth of the same person.
Collection method (informative)	Punctuation (back slashes or hyphens) or spaces should not be used in data collection. It is recommended that in cases where all components of the date are not known, a valid date be used together with the date of death accuracy indicator.
Comment	It may be used in conjunction with date of death accuracy indicator.

7.3.3 Estimated date (of death) flag

Definition An indication of whether any component of a reported date was estimated.

Source standards Not applicable.

Data type Coded text.

Data domain Any combination of the values A, E, U representing the corresponding level of accuracy of each date component of the reported date including:

Code	Description
AAA	Date is valid, not estimated (default value)
EEE	Date estimated from known information
UUU	The whole date is unknown
EAA	The year is estimated but the day and month are accurate
AAU	The year and month are accurate but the day is unknown
UUE	Year and month are unknown and the day is estimated
UUA	Year and month are unknown, but the day is accurate

The domain values will be dependent upon the date of death presentation style value. The examples below are for a presentation style DDMMYYYY.

Data domain	Date component (for format DDMMYYYY)		
	(D)ay	(M)onth	(Y)ear
Accurate	A	A	A
Estimated	E	E	E
Unknown	U	U	U

Guide for use

This is used to record the level of certainty or estimation used in recording the subject of care's date of death.

This data element is designed to flag the part, or parts, of a date that have been estimated when a date provided is based on an approximation rather than the actual date. This data element may assist with manual searching and record linkage.

It should be noted that the sequence and content of the data domain is dependent upon the format of date display in the country where the information is being used.

EXAMPLE 1 If a date has been obtained from a reliable source and it is known to be accurate, then the date accuracy flag should be recorded as (AAA).

EXAMPLE 2 If a person died unattended and their body was found some time later, it might not be possible to accurately assess the date of death, in which case an estimated day and sometimes month might be necessary (AEE).

Collection constraints

If constraints for the collection of the date are imposed, such as "a valid date should be input in an information system (software) for 'unknown' date components", the accuracy date indicator is recommended to be used along with the date as a way of avoiding the contamination of the valid dates with the same value on the respective date components.

7.3.4 Source of death notification

Definition This indicates the source of information about a subject of care's death. This field provides an indication of the certainty of the information.

Source standards Not applicable.

Data type Coded text.

Data domain	Code	Description
	1	Registry
	2	Health care provider
	3	Relative
	4	Other
	9	Unknown

- Guide for use**
- **Registry.** Notification received from an official registry such as births, deaths or coroner, death certificate. This source is considered to be of the greatest certainty.
 - **Health care provider.** Death is notified directly from a health care provider, other than the person responsible for certification of death. This source is considered to be of very good certainty.
 - **Relative.** Death is highly likely to be certain, but cases of inaccurate reporting of death by relatives has been known and should not be considered equal in certainty to health provider or official register as a source of death information.
 - **Other.** Death is identified through newspapers and other sources. These should be considered a less reliable source of death notification.
 - **Unknown.** Source of information about the subject's death is not known. This is the least reliable source of death notification.

Validation rule Valid codes or blank.

Collection method (informative) This data element should always be used in conjunction with a date of death.

Comment Not applicable.

.....

7.4 Sex

Definition The sex of the subject.

Sex is the biological distinction between male and female. Where there is an inconsistency between anatomical and chromosomal characteristics, sex is based on anatomical characteristics.

Source standards HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-8 Sex)

Data type Coded text.

Data domain	Code	Descriptor	Alternative code
	1	Male	M
	2	Female	F
	3	Intersex or indeterminate	I
	9	Not stated/inadequately described	U

Guide for use This data element indicates the sex of the person for administrative or general communication purposes and may be much less specific than the values used in clinical care.

Code 3, "Intersex or indeterminate", refers to a person who, because of a genetic condition, was born with reproductive organs or sex chromosomes that are not exclusively male or female, or whose sex has not yet been determined for whatever reason.

Code 9, "Not stated/inadequately described", should only be used if the data are not collected at the point of subject contact, or circumstances dictate that the data are not able to be collected.

Verification rules Accept only allowed values.

Field should not be blank.

Code 3, "Intersex or indeterminate" should be confirmed if used for subjects aged over 90 days.

Collection method (informative) The following format should be used for data collection:

What is your (the person's) sex? __ Male __ Female

Codes 1 (Male) and 2 (Female) may be mapped to M and F respectively for collection purposes, but they should be stored within information systems as codes 1 and 2 above.

Operationally, sex is the distinction between male and female, as reported by a subject or as determined by an interviewer. When collecting data on sex by personal interview, asking the sex of the respondent is usually unnecessary and might be inappropriate, or even offensive. It is usually a simple matter to infer the sex of the respondent through observation, or from other cues such as the relationship of the subject(s) accompanying the respondent, or first name. The interviewer may ask whether subjects not present at the interview are male or female.

A person's sex can change during their lifetime as a result of procedures known alternatively as sex change, gender reassignment, transsexual surgery, transgender reassignment or sexual reassignment. Throughout this process, which might be over a considerable period of time, sex could be recorded as either male or female.

Code 3, “Intersex or Indeterminate” is normally used for babies for whom sex has not been determined for whatever reason. It should not generally be used on data collection forms completed by the respondent, and should only be used if the person or respondent volunteers that the person is intersex or where it otherwise becomes clear during the collection process that the individual is neither male nor female.

Code 9, “Not stated/inadequately described”, is not to be used on primary collection forms. It is primarily for use in administrative collections when transferring data from data sets where the item has not been collected.

7.5 Mother's original family name

Synonym	Mother's maiden name Mother's family name Mother's surname
Definition	The original family name of the subject of care's mother.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-6 Mother's maiden name)
Data type	Text.
Data domain	Not applicable.
Guide for use	May be used to confirm the identity of the subject of care.
Verification rules	All letters of the alphabet and additional characters as identified in family name.
Collection method (informative)	See family name.

7.6 Country (place) of birth

Synonym	Birthplace Place of birth
Definition	The country in which the subject of care was born.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-23 Birth place)
Data type	Coded text.
Data domain	Valid country codes.
Guide for use	The four character numeric code should be used for data storage. The full descriptor should be used for data collection or onscreen display where possible. Ensure that staff are aware of synonymous country names: e.g. Netherlands/Holland.

A country, even if it comprises other discrete political entities, such as states, is treated as a single unit for all data domain purposes. Parts of a political entity are not included in different groups. Thus, Hawaii is included in North America (as part of the identified country United States of America), despite being geographically close to, and having similar social and cultural characteristics as, the units classified to Polynesia.



Verification rules Not applicable.

Collection method (informative) Not applicable.

7.7 Birth plurality

Synonym Multiple birth indicator

Definition An indicator of multiple births, showing the total number of births resulting from a single pregnancy.

Source standards HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-24 Multiple birth indicator)

Data type Coded text.

Data domain	Code	Descriptor
	1	Singleton
	2	Twins
	3	Triples
	4	Quadruplets
	5	Quintuplets
	6	Sextuplets
	8	Other
	9	Not stated

Guide for use Plurality of a pregnancy is determined by the number of live births, or by the number of fetuses that remain *in utero* at 20 weeks' gestation and that are subsequently born separately. In multiple pregnancies, or if gestational age is unknown, only live births of any birth weight or gestational age, or fetuses weighing 400 g or more, are taken into account in determining plurality. Fetuses aborted before 20 completed weeks or fetuses compressed in the placenta at 20 or more weeks are excluded.

Verification rules Not applicable.

Collection method (informative) These data should be collected routinely for subjects of care aged 28 days or less. If a subject of care or their parent/guardian/carer/next of kin volunteers this information, it may be recorded.

7.8 Birth order

Synonym	Birth sequence.	
Definition	The sequential order of this subject of care in a multiple birth regardless of live or still birth.	
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-24 Multiple birth order)	
Data type	Coded text.	
Data domain	Code	Descriptor
	1	Singleton or first of a multiple birth
	2	Second of a multiple birth
	3	Third of a multiple birth
	4	Fourth of a multiple birth
	5	Fifth of a multiple birth
	6	Sixth of a multiple birth
	8	Other
	9	Not stated
Guide for use	Newborns registered within 28 days of birth, or later to assist in identification.	
	Stillborns are counted such that, if twins were born, the first stillborn and the second live-born, the second twin would be recorded as code 2 second of a multiple birth (and not 1 singleton or first of a multiple birth).	
Verification rules	Not applicable.	
Collection method (informative)	These data should be collected routinely for subjects of care who are aged 28 days or less. If a subject of care or their parent/guardian/carer/next of kin volunteers this information, it may be recorded.	

7.9 Identification comment

Definition	Comments recorded for a subject of care registration to distinguish between two or more subjects with the same or similar demographic information.	
Source standards	Not applicable.	
Data type	Text.	
Data domain	Not applicable.	
Guide for use	Should only be used to confirm the identity of the subject i.e. this is not a field for comments about a subject's character.	
	EXAMPLE 1	“Two pts with same name. Do not merge with 210465.”
	EXAMPLE 2	“Duplicate registration? Check also 230163.”
	EXAMPLE 3	“Confirm date of birth. If 14/05/75 see also 081035.”
	EXAMPLE 4	“Double-check spelling of family name.”

8 Subject of care address

8.1 General

This clause describes eight data elements used to capture and store address details of subjects of care. The structure outlined in this Technical Specification attempts to simplify data collection while capturing the range of addresses and telephone numbers important to health care establishments. The format of data storage is not as important as the consistent method of recording these data.

Each subject of care address is defined as the combination of data elements set out in Table 7.

Table 7 — Subject of care address data elements

Data element name	Clause	Opt. ^a	Data type	Repeat data element ^b	Example
Subject of care address	8	O	Text	Y	
Address line	8.2	O	Text	Y	Level 7, Room 5
Building/complex sub-unit type abbreviation	8.2.2	O	Coded text	N	APT
Building/complex sub-unit number	8.2.3	O	Text	N	6
Address site name	8.2.4	O	Text	N	Treasury Building
Floor/level number	8.2.5	O	Text	N	L 3
Floor/level type	8.2.6	O	Coded text	N	B (basement)
Street number	8.2.7	O	Text	N	401A
Lot number	8.2.8	O	Text	N	Lot 52A
Street name	8.2.9	O	Text	N	Mortonville
Street type code	8.2.10	O	Coded text	N	Circuit
Street suffix code	8.2.11	O	Coded text	N	N (north)
Suburb/town/locality	8.3	O	Text	N	Upper Conductor West
State/territory/province identifier	8.4	O	Coded text	N	NSW
Postal code (ZIP code)	8.5	O	Coded text	N	25300
Delivery point identifier	8.6	O	Coded text	N	
Country identifier	8.7	O	Coded text	N	2101 (England)
Address type	8.8	O	Text	N	1 (Business)
Address type start date	8.8.2	O	Date	Y	19951012
Address type start date accuracy indicator	8.8.3	O	Coded text	Y	AAE
Address type end date	8.8.4	O	Date	Y	19951012
Address end date accuracy indicator	8.8.5	O	Coded text	Y	EUU
<p>^a Whether the data element is optional (O) or required (R).</p> <p>Required (the group may be required or, where the group is optional, the individual data elements within the group may be marked as required. In this case, where the group exists the required elements should be present.</p> <p>^b Whether yes (Y) or no (N).</p>					

The relationship between subject of care address data elements is illustrated in Figure 8.

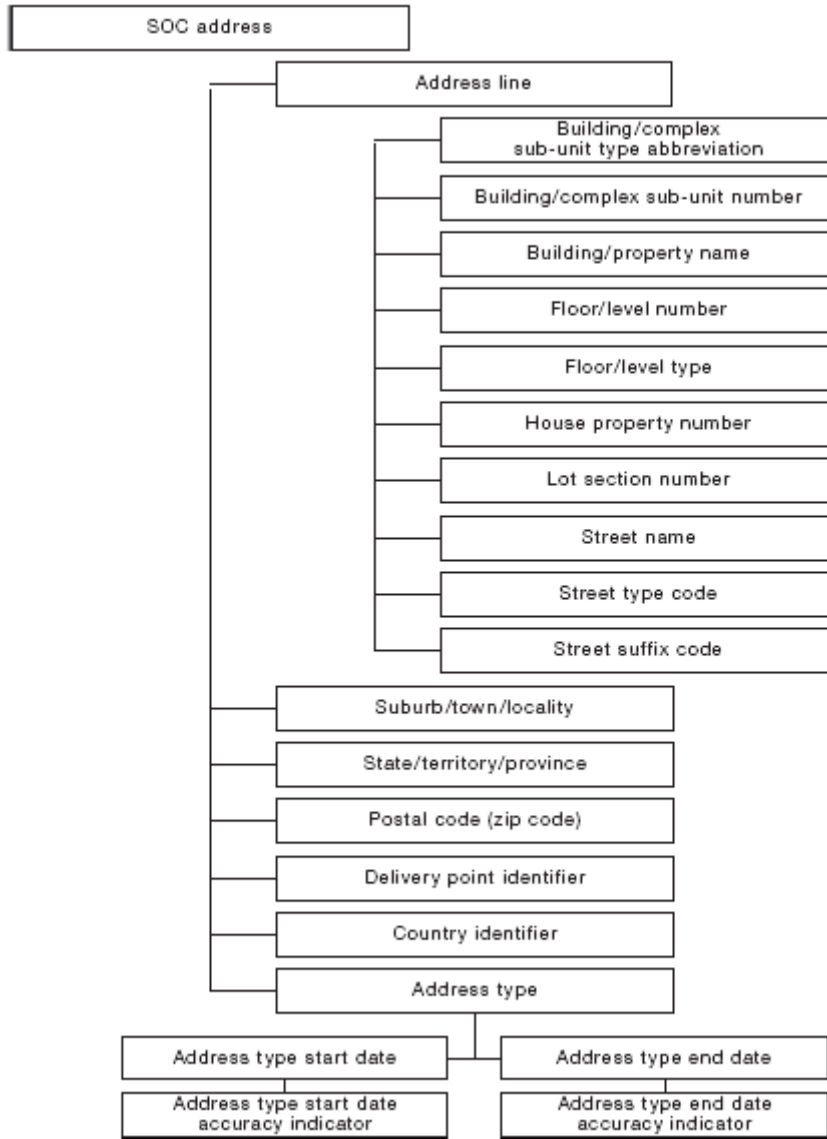


Figure 8 — Data elements for subject of care address

8.2 Address line

8.2.1 General

Definition

A composite of one or more standard address components that describe a low level of geographical/physical description of a location that, used in conjunction with the other high-level address components, i.e. “suburb/town/locality name”, “postal code”, “state/territory/province”, and “country”, forms a complete geographical/physical address.

Source standards

HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address)

Data type

Text.

Data domain

This item is a combination of the following standard address data elements that may be concatenated in the address line in the following sequence:

Building complex sub-unit type abbreviation
 Building/complex sub-unit number
 Address site name
 Floor/level number
 Floor/level type
 House/property number
 Lot/section number
 Street name
 Street type code
 Street suffix code

Guide for use

One complete identification description of a location/site of an address can comprise one, or more than one, instance of address line. Instances of address lines are commonly identified in electronic information systems as address-line 1, address-line 2, etc. The format of data collection is less important than consistent use of conventions in the recording of address data. Hence, address may be collected in an unstructured manner but should ideally be stored in a structured format. Where address line is collected as a stand-alone item, software may be used to parse the address line details to separate the subcomponents.

Multiple address lines may be recorded as required.

Address line can include more than one physical line of text.

All of the relevant “street” details, including building or property name, should be captured in this field. The field is free text, although some commonly used abbreviations are permitted.

The format of data collection is less important than consistent use of conventions in the recording of address data.

No unnecessary punctuation should be added to the address, e.g. no full stop following street type.

**Guide for use
(continued)****Residential facilities**

Enter name of the residential facility (such as the nursing home, caravan park, prison or boarding school) before the street address, if space permits.

Complete street address

The full street address should be recorded. This may be a combination of the above components.

EXAMPLE 1 Level 15 Room 2B 27 James Street

EXAMPLE 2 Level 7 Room 15

 Customs House

 Main Street

EXAMPLE 3 Unit 2A Technology Park

 4 Centre Road

Verification rules Not applicable

Collection method (informative) The format of data collection is less important than consistent use of conventions in the recording of address data. Hence, the address may be collected in an unstructured manner but should ideally be stored in a structured format.

Unknown addresses

Enter “unknown” in the address line field.

No fixed address

Enter “unknown” in lieu of street number and name in the address line field.

8.2.2 Building/complex sub-unit type abbreviation

Definition The specification of the type of a separately identifiable portion within a building or complex, marina, etc. to clearly distinguish it from another.

Source standards HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address)

Data type Coded text.

Data domain	Code	Description	Code	Description
	APT	Apartment	SE	Suite
	CTGE	Cottage	SHED	Shed
	DUP	Duplex	SHOP	Shop
	F	Flat	SITE	Site
	FY	Factory	SL	Stall
	KSK	Kiosk	STU	Studio
	MB	Marine Berth	TNHS	Townhouse
	MSNT	Maisonette	U	Unit
	OFF	Office	VLLA	Villa
	PTHS	Penthouse	WARD	Ward
	RM	Room	WE	Warehouse

Guide for use Addresses may contain multiple instances of building/complex sub-unit type. Record each instance of building/complex sub-unit type with its corresponding building/complex number when appropriate.

EXAMPLE 1 APT 6

EXAMPLE 2 SHOP 3A

EXAMPLE 3 U 6

Verification rule Not applicable.

Collection method (informative) This is a composite part of the address line. It is to be collected in conjunction with the data element building/complex sub-unit number.

8.2.3 Building/complex sub-unit number

Definition	The specification of the number of identifier of a building/complex, marina, etc. to clearly distinguish it from another.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address, part)
Data type	Text.
Data domain	Not applicable.
Guide for use	<p>The building/complex sub-unit number should be recorded with its corresponding building/complex sub-unit type abbreviation. Where applicable, the number may be followed by a string suffix.</p> <p>EXAMPLE 1 APT 6</p> <p>EXAMPLE 2 SHOP 3A</p> <p>EXAMPLE 3 U 6</p>
Verification rule	Not applicable.
Collection method (informative)	<p>This is a composite part of the address line. It is to be collected in conjunction with the data element building/complex sub-unit type abbreviation.</p> <p>Where a building or other type of unit is present in a complex of such buildings or units, the data elements building/unit sub-unit type abbreviation and building/complex sub-unit number should be used in conjunction in that order. An example can be seen in a shop within a shopping complex. Such a shop could have as part of its address line the word “shop” as the type followed by its identifying “number” within the complex, e.g. “209a”. Thus the words “Shop 209a” would form part of the address line.</p>

8.2.4 Address site name

Definition	The full name used to identify the physical building or property as part of its location.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address, part)
Data type	Text.
Data domain	Not applicable.
Guide for use	<p>Usually this information is not abbreviated. It should include any reference to a wing or other components of a building complex, if applicable. A comma is to be used to separate the wing reference from the rest of the building name.</p> <p>Record each building/property name relevant to the address:</p> <ul style="list-style-type: none"> — Building/property name 1 (30 string characters); — Building/property name 2 (30 string characters). <p>EXAMPLE 1 Building — Treasury Building</p> <p>EXAMPLE 2 Property — Brindabella Station</p>
Verification rule	Not applicable.
Collection method (informative)	This is a composite part of the address line.

8.2.5 Floor/level number

Definition	Descriptor used to identify the floor or level of a multi-storey building/complex.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address)
Data type	Text.
Data domain	Not applicable.
Guide for use	The floor/level number should be recorded with its corresponding floor/level type. Some floor/level numbers may be followed by an alphabetic suffix. Examples of floor/level identification include: EXAMPLE FL 1A, L3, LG A
Verification rule	Not applicable.
Collection method (informative)	This is a composite part of the address line and relates to the data element floor/level type. There should be no spaces between the number and the alpha suffix. To be collected in conjunction with floor/level type.

8.2.6 Floor/level type

Definition	Descriptor used to classify the type of floor or level of a multi-storey building/complex.	
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address)	
Data type	Coded text.	
Data domain	Code	Definition
	B	Basement
	FL	Floor
	G	Ground
	L	Level
	LG	Lower Ground
	M	Mezzanine
	UG	Upper Ground
Guide for use	Some floor/level identification may require the floor/level type plus a floor/level number to be recorded. Examples of Floor/level identification: EXAMPLE L 1A, 3, G A	
Verification rule	Not applicable.	
Collection method (informative)	This is a composite part of the address line and relates to the data element floor/level number. To be collected in conjunction with floor/level number where applicable. Some floor/level type entries will often have no corresponding number, e.g. basement, ground, lower ground, mezzanine and upper ground.	

8.2.7 Street number

Definition	The numeric or string reference number of a house or property that is unique within a street name, suburb.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address)
Data type	Text.
Data domain	Not applicable.
Guide for use	<p>Generally, only one street number is used. However, if the house/property number includes a number range, the range of applicable numbers should be included, separated by a hyphen (-), with no spaces between numerals, i.e. 17-19.</p> <ul style="list-style-type: none"> a) House/property number 1 refers to physical house/property number and for ranges is the starting number (5 numeric characters). b) House/property number suffix 1: single character identifying the house/property number suffix (1 string character). c) House/property number 2 refers to a physical house/property number and for ranges is the finishing number (5 numeric characters). d) House/property number suffix 2: single character identifying the house/property number suffix (1 string character) with no space between the numeric and the alpha characters. <p>EXAMPLE: "401A-403B"</p> <p>"401" is the house/property number first in range.</p> <p>"A" is the house/property suffix 1.</p> <p>"403" is house/property number last in range.</p> <p>"B" is house/property suffix 2.</p>
Verification rule	Not applicable.
Collection method (informative)	This is a composite part of the address line.

8.2.8 Lot number

Synonym	Section, allotment number.
Definition	The lot reference allocated to an address in the absence of street numbering.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address)
Data type	Text.
Data domain	Not applicable.
Guide for use	<p>The lot number is suitable for postal purposes as well as the physical identification of addresses.</p> <p>A lot number should be used only when a street number has not been specifically allocated or is not readily identifiable with the property.</p> <p>For identification purposes, the word “lot” should precede the lot number and be separated by a space.</p> <p>EXAMPLE 1 Lot 716</p> <p>EXAMPLE 2 Lot 534A</p> <p>EXAMPLE 3 Lots 716-718</p>
Verification rule	Not applicable.
Collection method (informative)	<p>This is a composite part of the address line.</p> <p>The lot number is positioned before the street name and type, located in the same line containing the street name.</p>

8.2.9 Street name

Definition	The name that identifies a public thoroughfare and differentiates it from others in the same suburb/town/locality.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address)
Data type	Text.
Data domain	Not applicable.
Guide for use	<p>To be used in conjunction with street type code.</p> <p>To be used in conjunction with street suffix code.</p>
Verification rule	Not applicable.
Collection method (informative)	This is a composite part of the address line and relates to the data element house/property number, street suffix code and street type code.

8.2.10 Street type code

Definition	A code that identifies the type of public thoroughfare.			
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address)			
Data type	Coded text.			
Data domain	The following is a list of commonly used street type abbreviations. This list is not exhaustive as different languages and nationalities will have different requirements.			
	Code	Description	Code	Description
	Ally	Alley	Gr	Grove
	Arc	Arcade	Hwy	Highway
	Ave	Avenue	Jnc	Junction
	Bvd	Boulevard	Lane	Lane
	Bypa	Bypass	Ln	Line
	Crc	Circle	Link	Link
	Cct	Circuit	Mews	Mews
	Cl	Close	Pde	Parade
	Crn	Corner	Pl	Place
	Ct	Court	Ridge	Ridge
	Cres	Crescent	Rd	Road
	Cds	Cul-de-sac	Sq	Square
	Dr	Drive	St	Street
	Esp	Esplanade	Tce	Terrace
	Grn	Green		
Guide for use	Street types should be written in full where space permits in order to avoid potential confusion in the case of an emergency and to improve the quality of street type information. Where abbreviations are used, those in this data domain provide a basis.			
Verification rule	Not applicable.			
Collection method (informative)	This is a composite part of the address line. It is to be collected in conjunction with street name and street suffix code.			

8.2.11 Street suffix code

Definition	Term used to qualify street name issued for directional references.			
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address)			
Data type	Coded text.			
Data domain	The list below is indicative and is not exhaustive as different languages and nationalities will have different requirements			
	Code	Description	Code	Description
	CN	Central	NW	North West
	E	East	S	South
	EX	Extension	SE	South East
	LR	Lower	SW	South West
	N	North	UP	Upper
	NE	North East	W	West
Guide for use	Not applicable.			
Verification rule	Not applicable.			
Collection method	This is a composite part of the address line.			
	It shall be used in conjunction with the street name and the street type code.			
	EXAMPLE	Browns Rd W.		

8.3 Suburb/town/locality

Definition	The full name of the general locality containing the specific address of a subject of care.		
Source standards	AS 4590-2006, <i>Interchange of client information</i>		
	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address)		
Data type	Text.		
Data domain	Suburb or town/city or locality.		
Guide for use	Suburb/town/locality may be a town, city, suburb or commonly used location name such as a large agricultural property or Aboriginal community.		
Verification rules	Not applicable.		
Collection method (informative)	Unknown subject address: enter "unknown" in the suburb/town/locality field.		
	No fixed address: enter "no fixed address" in the suburb/town/locality field.		

8.4 State/territory/province identifier

Definition	An identifier of the province, state or territory in which a subject of care resides.
Source standards	Australian Institute of Health and Welfare (AIHW), <i>National Health Data Dictionary (NHDD)</i> : (Knowledgebase ID: 000155 <i>State identifier</i>) HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address)
Data type	Coded text.
Data domain	Codes uniquely identifying the province, state or territory.
Guide for use	Record the state or territory or province as indicated by the subject of care as their address, e.g. Texas. This should be consistent with the ASCII character set and should comply with the official conventions of that country, e.g. Hunan rather than Chinese characters.
Verification rules	Not applicable.
Collection method (informative)	Not applicable.

8.5 Postal code (ZIP code)

Definition	The code for a postal delivery area, aligned with locality, suburb or place for the address of a subject of care, as defined by the postal service.
Synonym	Post code ZIP code
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address)
Data type	Coded text.
Data domain	Valid postal code or blank.
Guide for use	Postcodes should be recorded with punctuation (for data quality and usability purposes).
Verification rules	Not applicable.
Collection method (informative)	Unknown address National standards should be used to represent the concept of unknown address. No fixed address National standards should be used to represent the concept of no fixed address. The use of HL7 V3 Flavors of Null for Address Type was considered. This is one mechanism that could be used to implement these concepts in systems, according to local need, but is not recommended for data collection at the workplace. The current recommendation for collection is to adhere to common practice internationally within clerical practice and to support epidemiological reporting, which uses a specific code in the postal code field to represent these concepts.

8.6 Delivery point identifier

Definition	A unique number assigned to a postal address as designated by the postal service.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000(STF-11 Office/home address <other geographic designation>)
Data type	Coded text.
Data domain	Valid delivery point identifier (DPID) code or blank.
Guide for use	Used to specify postal delivery address groups, such as rural post box groups. The specific domain values are determined by the country of use.
Verification rule	This should be a valid code used by the postal service in the country of the address.
Collection method (informative)	The delivery point identifier (DPID) is assigned electronically from a postal address file system maintained by the postal service.
Comment	The DPID is easily converted to a bar code and can be included on correspondence and address labels. If the bar code is displayed on a standard envelope that passes through a mail-franking machine, a reduced postage cost is offered in some countries.

8.7 Country identifier

Definition	A code representing the country component of a subject of care's address.
Source standards	ISO 3166-1:2006, <i>Codes for the representation of names of countries and their subdivisions — Part 1: Country codes</i> HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 - Address (STF-11 Office/home address <country>)
Data type	Coded text.
Data domain	ISO 3166.
Guide for use	The four character numeric code should be used for data storage. The full descriptor should be used for data collection or onscreen display, where possible.
Verification rules	This should be a valid code/country.
Collection method (informative)	Staff using and collecting this information should be aware of synonymous country names such as Holland/Netherlands.

8.8 Address type

8.8.1 General

Definition	A code representing a type of address.
Source standards	HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-11 Patient address, Table 0190 Address type)

Data type Coded text.

Data domain	Code	Description
	1	Business
	2	Mailing or postal
	3	Temporary accommodation
	4	Residential
	8	No fixed address
	9	Unknown/not stated/inadequately described

Guide for use Multiple addresses may be recorded as required. This field can be a multiple occurring field, and each address should have an address type code.

Code 1 — “Business”: used to indicate an address that is the physical location of a business or office at which a subject of care can be contacted. More than one business address may be recorded.

Code 2 — “Mailing or postal”: used to indicate an address that is only for correspondence and/or billing purposes.

Code 3 — “Temporary accommodation”: temporary accommodation address (such as for a person who usually resides overseas or where the provider is in temporary accommodation due to renovation, or treatment).

Code 4 — “Residential”: used to indicate where a person is living.

NOTE This code is not valid for organizations.

Code 8 — “No fixed address”: used where a person has no fixed address.

Code 9 — “Unknown/not stated/inadequately described”: may also be used where the person has no fixed address or does not wish to have their residential or a correspondence address recorded.

A single address may have multiple address types associated with it. Record as many as required.

Verification rules Not applicable.

Collection method (informative) At least one address should be recorded (this may be an unknown address type).

Health care organizations or establishments should always attempt to collect the residential address of a subject when an occasion of service or admission is provided. When recording the address for a health care provider or organization, the business address should always be collected. In addition, any number of other addresses may also be collected.

Further, addresses may also have effective dates attached to them to indicate which are current and which are past addresses.

Overseas address

For individuals, record the overseas address as the residential address and record a temporary accommodation address as their contact address in the country where they have contacted health services.

8.8.2 Address type start date

Definition The date on which the address type is first applicable to the subject of care.

Source standards Not applicable.

Data type Date.

Data domain Valid date.

Guide for use If the date is estimated in some way, it is recommended that the data element address type start date accuracy indicator also be recorded at the time of record creation to flag the accuracy of the data.

For data integrity, data exchange, future data analysis and/or manipulation of data from diverse sources, the date accuracy indicator should be used in conjunction with the address type start date in all instances to ensure accuracy.

Verification rules This field should:

- a) not be null;
- b) be a valid date;
- c) be less than or equal to the address type end date (where that date is not blank).

Collection method (informative) Punctuation (back slashes or hyphens) or spaces should not be used in data collection.

8.8.3 Address type start date accuracy indicator

Definition	An indication of the accuracy of the address type start date at the component level for the date.
Source standards	Australian Institute of Health and Welfare (AIHW), <i>National Health Data Dictionary (NHDD)</i> (Knowledgebase ID: 000431 Estimated date flag)
Data type	Coded text.
Data domain	Any combination of the values A, E, U representing the corresponding level of accuracy of each date component of the reported date including:

Code	Description
AAA	Accurate date
EEE	Estimated date
UUU	Unknown date
EAA	Accurate day and month, estimated year
AAU	Unknown day, accurate month and year
UUE	Unknown day and month, estimated year
UUA	Unknown day and month, accurate year

The domain values will be dependent upon the date of birth presentation style value. The examples below are for a presentation style DDMMYYYY.

Data domain	Date component (for format DDMMYYYY)		
	(D)ay	(M)onth	(Y)ear
Accurate	A	A	A
Estimated	E	E	E
Unknown	U	U	U

Guide for use This is used to record the level of certainty or estimation used in recording the subject of care's address type start date.

Provision of a subject of care's address type start date might not be exactly known.

This data element is designed to flag the part, or parts, of a date that have been estimated when a date provided is based on an approximation rather than the actual date. This data element may assist with manual searching and record linkage processes, e.g. when the date of birth is a component of the linkage key.

It should be noted that the sequence and content of the data domain are dependent upon the format of date display in the country where the information is being used.

EXAMPLE 1 If a date has been sourced from a reliable source and it is known to be accurate, then the date accuracy flag should be recorded as (AAA).

EXAMPLE 2 If a person is unsure of the exact date they began to use this address, but they can confirm the year, (AAU) would be recorded.

Collection constraints

If constraints for the collection of the date are imposed, such as “a valid date should be input in an information system (software) for 'unknown' date components”, the accuracy date indicator is recommended to be used along with the date as a way of avoiding the contamination of the valid dates with the same value on the respective date components.

EXAMPLE Some systems use YYYY0107 and some use YYYY0101 when only the year is known. When month and year are known, some use the 15th day as the date, i.e. YYYYMM15. Where this occurs in a data collection that is used for reporting or analysis purposes, there will be dates in the collection with the attributes YYYY0107, etc. that are accurate and some that are not accurate. Without a corresponding flag to determine this accuracy, the analysis or report will be contaminated by those estimated dates.

For future users of the data it might also be important that they know the accuracy of the date components of a reported date.

The format of this field and the data domain are dependent upon the format of dates used in the country.

Validation rules	Any combination of the codes A, E and/or U.
Collection method (informative)	This data element should always be used in conjunction with an address type start date.
Comment	Most computer systems require a valid date to be recorded in a date field, i.e. the month part should be an integer between 1 and 12, the day part should be an integer between 1 and 31 with rules about the months with less than 31 days, and the year part should include the century. However, in actual practice, the date or date components are often not known. This means that a date should be included and it should follow the rules for a valid date. It therefore follows that, while such a date will contain valid values according to the rules for a date, the date is in fact an “unknown” or “estimated” date. For future users of the data it is essential they know whether a date is accurate, unknown or estimated, and which components of the date are accurate, unknown or estimated.

8.8.4 Address type end date

Definition	The date on which the address or address type is no longer applicable to the subject of care.
Source standards	Not applicable.
Data type	Date.
Data domain	Valid date.
Guide for use	If the date is estimated in some way, it is recommended that the data element “address type end date accuracy indicator” also be recorded when the fact that this address/type is no longer applicable to the subject of care.

For data integrity, data exchange, future data analysis and/or manipulation of data from diverse sources, the address type end date accuracy indicator should be used in conjunction with the address type end date in all instances to ensure accuracy.

Validation rules	This field should: <ul style="list-style-type: none"> a) not be null; b) be a valid date; c) be greater than or equal to the address type start date.
Collection method (informative)	Punctuation (back slashes or hyphens) or spaces should not be used in collection of these data and display should be governed by local common practice.

8.8.5 Address type end date accuracy indicator

Definition	An indication of the accuracy of the address type end date at the component level for the date.
Source standards	Australian Institute of Health and Welfare (AIHW), <i>National Health Data Dictionary (NHDD)</i> (Knowledgebase ID: 000431 Estimated date flag)
Data type	Coded text.
Data domain	Any combination of the values A, E, U representing the corresponding level of accuracy of each date component of the reported date including:

Code	Description
AAA	Accurate date
EEE	Estimated date
UUU	Unknown date
EAA	Accurate day and month, estimated year
AAU	Unknown day, accurate month and year
UUE	Unknown day and month, estimated year
UUA	Unknown day and month, accurate year

The domain values will be dependent upon the date of birth presentation style value. The examples below are for a presentation style DDMMYYYY.

Data domain	Date component (for format DDMMYYYY)		
	(D)ay	(M)onth	(Y)ear
Accurate	A	A	A
Estimated	E	E	E
Unknown	U	U	U

Guide for use Used to record the level of certainty or estimation used in recording the subject of care's address type end date.

Provision of a subject of care's address type end date might not be exactly known.

This data element is designed to flag the part, or parts, of a date that has been estimated when a date provided is based on an approximation rather than the actual date. This data element may assist with manual searching and record linkage processes, e.g. when the date of birth is a component of the linkage key.

It should be noted that the sequence and content of the data domain is dependent upon the format of date display in the country where the information is being used.

EXAMPLE 1 If a date has been sourced from a reliable source and it is known to be accurate, then the date accuracy flag should be recorded as (AAA).

EXAMPLE 2 If a person is unsure of the date they stopped using this address, but they can confirm the year, it would be recorded as (AUU).

Collection constraints

If constraints for the collection of the date are imposed, such as “a valid date should be input in an information system (software) for 'unknown' date components”, the accuracy date indicator is recommended to be used along with the date as a way of avoiding the contamination of the valid dates with the same value on the respective date components.

For example, some systems use YYYY0107 and some use YYYY0101 when only the year is known. When month and year are known some use the 15th day as the date, i.e. YYYYMM15. Where this occurs in a data collection that is used for reporting or analysis purposes, there will be dates in the collection with the attributes YYYY0107 etc. that are accurate and some that are not accurate. Without a corresponding flag to determine this accuracy, the analysis or report will be contaminated by those estimated dates.

For future users of the data, it might also be important that they know the accuracy of the date components of a reported date.

The format of this field and the data domain are dependent upon the format of dates used in the country.

Validation rules Any combination of the codes A, E and or U.

Collection method (informative) This data element should always be used in conjunction with an address type end date.

Comment Most computer systems require a valid date to be recorded in a date field, i.e. the month part should be an integer between 1 and 12, the day part should be an integer between 1 and 31 with rules about the months with less than 31 days, and the year part should include the century. However, in actual practice, the date or date components are often not known. This means that a date should be included and it should follow the rules for a valid date. It therefore follows that, while such a date will contain valid values according to the rules for a date, the date is in fact an “unknown” or “estimated” date. For future users of the data it is essential they know whether a date is accurate, unknown or estimated, and which components of the date are accurate, unknown or estimated.

9 Subject of care electronic communications

9.1 General

This clause describes data elements used to capture and store the electronic communication contact details of subjects of care shown in Figure 9. Examples of the contact details that may be collected include telephone numbers and email addresses. Each subject of care electronic communication contact detail is defined as the combination of the data elements set out in Table 8. There may be multiple instances of subject of care electronic communication for any individual subject of care.

Table 8 — Subject of care electronic communication data elements

Clause	Data element name	Opt. ^a	Data type	Repeat data element ^b	Example
9	SOC electronic communications	O	String	Y	
9.2	Electronic communication medium	M	Code	N	Phone
9.3	Electronic contact preference code	O	Code	N	Workdays, daytime
9.4	Electronic communication details	M	String	N	
9.5	Electronic contact usage code	O	Code	Y	1 (business)

^a Whether the data element is optional (O) or required (R).

^b Whether yes (Y) or no (N).

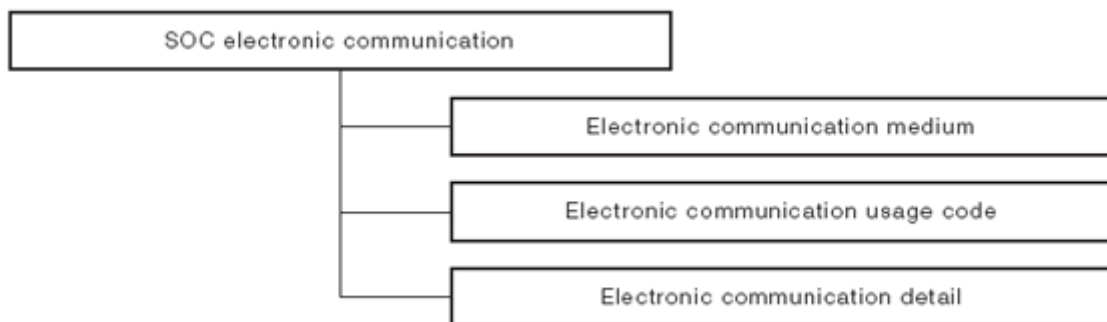


Figure 9 — Electronic communication data element structure

9.2 Electronic communication medium

Definition A code representing a type of communication mechanism used by a subject of care.

Source: AS 4846-2006 *Health care provider identification*

HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (STF-16 Preferred method of contact)

HNBC 98-10, *HealthNet/BC Provider Data Standard Version 1.0*

Data type Coded text.

Data domain	Code	Description	Alternative code
	1	Telephone (excluding mobile)	T
	2	Mobile (cellular) telephone	C
	3	Facsimile machine	F
	4	Pager	B
	5	E-mail	E
	6	URL	U
	8	Other	O

Guide for use Multiple electronic communication addresses, e.g. multiple phone numbers, fax numbers and email, may be recorded as required.

Each instance should have the appropriate electronic communication medium and usage code assigned.

Verification rules Not applicable.

Collection method (informative) This field shall be left blank for unknown electronic communication details.

9.3 Electronic contact preference code

Definition An indication of the preferences for use of this contact type.

Source standards Not applicable.

Data type Coded text

Data domain	Code	Description
	B	Business hours
	D	Day Time hours
	W	Weekend hours
	A	At all times
	E	Evening/night hours

Guide for use This field indicates the times of the day when the associated contact is preferred for use.

Verification rules Not applicable.

9.4 Electronic communication details

Definition	A unique combination of characters used as input to electronic telecommunication equipment for the purpose of contacting a subject of care, i.e. a text string identifying the address for receipt of correspondence.
Source standards	AS 4846-2006, <i>Health care provider identification</i> HL7 V2.4, <i>Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (STF-10 Phone - includes email etc.)</i>
Data type	Text.
Data domain	A text string valid for a specific communication medium including spaces where applicable.
Guide for use	Multiple electronic communication details or addresses, e.g. multiple phone numbers, fax numbers and e-mail, may be recorded as required. Each instance should have an appropriate electronic communication medium and type assigned. Record the full phone number (including any prefixes) with no punctuation (hyphens or brackets).
Verification rules	Not applicable.
Collection method (informative)	<p>Prefix plus telephone number</p> <p>Record the prefix plus telephone number. The default should be the local prefix with an ability to overtype with a different prefix, e.g. 08 8226 6000 or 0417 123456. Systems may record these elements separately or together.</p> <p>Punctuation</p> <p>Punctuation shall not be recorded.</p> <p>EXAMPLE (08) 8226 6000 or 08-8226 6000 would be incorrect.</p> <p>Unknown electronic communication detail</p> <p>This field shall be left blank for unknown electronic communication details.</p>

9.5 Electronic contact usage code

Definition	A code representing the manner of use that a person applies to an electronic communication medium.		
Source standards	AS 4846-2006, <i>Health care provider identification</i> <i>HealthNet/BC Provider ID Standard</i> , B.C. Health Information Standards Council, B.C. Ministry of Health and Ministry Responsible for Seniors, June 1999		
Data type	Coded text.		
Data domain	Code	Description	Alternative code
	1	Business	B
	2	Personal	P
	3	Both business and personal use	A
Guide for use	Multiple communication details, e.g. multiple phone numbers, fax numbers and e-mail, may be recorded as required. Each instance should have an appropriate electronic communication medium and usage code assigned.		

Verification rules Not applicable.

Collection method (informative) This field shall be left blank for unknown electronic communication.

10 Biometric identifiers

Biometric identifiers may be used in addition to conventional identification methods as they can be faster and more reliable. Traditional methods of identification centre around something one possesses, such as a token or driving licence, or something one knows, such as passwords, addresses or names. Unlike these, biometric identifiers are part of the person themselves and therefore cannot be forgotten or stolen.

“Biometric capture devices create electronic digital templates that are encrypted and stored and then compared to encrypted templates derived from “live” images in order to confirm the identity of a person. The templates are generated from complex and proprietary algorithms and are then encrypted using strong cryptographic algorithms to secure and protect them from disclosure. Thus standing alone, biometric templates cannot be reconstructed, decrypted, reverse-engineered, or otherwise manipulated to reveal a person's identity.”

This Technical Specification does not identify the method of recording, nor the structure used within the different forms of biometric identification. This clause serves to identify some of the types of biometric identifiers that can be considered.

Common types of biometric identification include the following examples.

- Fingerprint: records the unique skin pattern of a finger or fingers. These devices capture one or two fingers and create a template for comparison. This process requires identification of the digit/s to which the image relates. There are two types of fingerprint images appropriate to subject of care identification:
 - fingerprint (rolled), which entails rolling each individual finger. “Rolled fingerprints generally have sufficient ridge details to allow classification in almost all cases. Rolled fingerprints provide a great deal of information allowing for highly accurate searches. However, capturing a properly rolled fingerprint is a slow process that requires trained staff, and the operator's manipulation of the subject's fingers often makes the subjects feel 'manhandled'”^[2];
 - fingerprint (flat), also called “plain” fingerprints. These can be captured quickly using inexpensive scanners by individuals with minimal training. They are more difficult to classify than rolled fingerprints and often provide a lower quality image than the rolled fingerprint.
- Facial features identification records the shape of the face, determined by distances between the eyes, ears and nose and other facial characteristics which are stored in a template^[3]. This method can identify an individual from many directions, even following changes such as plastic surgery.
- Voice recognition operates by recording a specific set of words and the method by which an individual says those words. It considers both pitch variations and timing.
- Iris scanning records 247 traits of a person's iris into a template for comparison. It functions with or without glasses or contact lenses.
- Retinal scanning records the structure of a person's retina into a template for comparison.
- Hand geometry records the size and shape of the hand and fingers. Hand geometry evaluates a three dimensional image of the four fingers and part of the hand for comparison. This process requires identification of the hand to which the image relates.
- Signature dynamics record not only the shape and style of a signature but the speed and pressure used in the creation of the image.
- Keystroke dynamics record the rhythmic elements of keystroke entry.

- Lip movement records the movement elements of different parts of the mouth and surrounding structures when specific words and sentences are said.
- Thermal face image records heat patterns in the face.
- Thermal hand image records heat patterns in the hand.
- Gait records a wide range of elements in the body movement when walking or running.
- Blood type records both the blood type and rhesus factor of the blood. This is not a unique identifier but it serves as a suitable additional identifier in health care.
- DNA records the unique pattern of DNA for the individual.

This Technical Specification provides an outline of biometric identifiers. Detailed biometric specifications can be found in ISO/IEC 19785-1.

NOTE Any implementation of biometric identifiers should refer to ISO/IEC 19785-1 as normative.

11 Subject of care linkage

11.1 General

An individual may have multiple linkages established. Linkages can be used to support system activities such as address updates to improve accuracy of this information. These relationships are social and are likely to change over time and should therefore be used with caution and with consideration of the legislative and privacy requirements of the jurisdiction in which it is used. Table 9 indicates data elements for subject of care linkage.

Table 9 — Subject of care linkage data elements

Clause	Data element name	Opt. ^a	Data type	Repeat data element ^b	Example
11	Subject of care linkage	O	Text	Y	1234587ABCM
11.2	Subject of care linkage identifier	O	Unique identifier	Y	1234587ABC
11.3	Subject of care linkage relationship	O	Text	Y	M

^a Whether the data element is optional (O) or required (R).

^b Whether yes (Y) or no (N).

Figure 10 illustrates subject of care linkage data element structure.

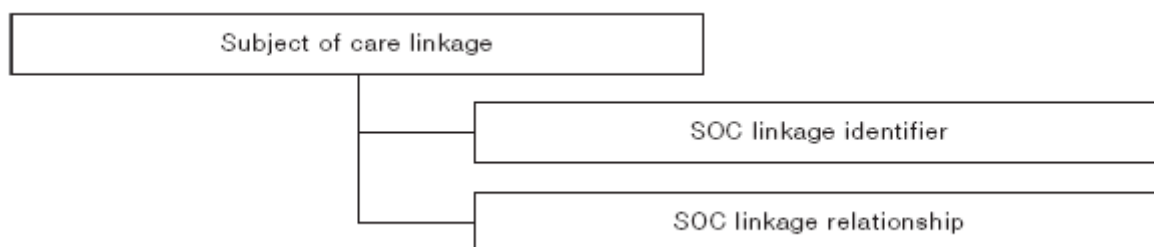


Figure 10 — Subject of care linkage data element structure

11.2 Subject of care linkage identifier

Definition	Person identifier unique within the relevant health care establishment, agency or domain for a person to whom the subject of care is linked.
Source standards	ASTM E1714-00, <i>Guide for Properties of a Universal Health Care Identifier (UHID)</i> Australian Institute of Health and Welfare (AIHW), <i>National Health Data Dictionary (NHDD)</i> HL7 V2.4, Health Level Seven Standard Version 2.4: Health Level Seven Inc., Ann Arbor, Michigan, 2000 (PID-3 Patient identifier list)
Data type	Text.
Data domain	Not applicable.
Guide for use	This is the identifier associated with the individual to whom the link is made, e.g. father's identifier.

11.3 Subject of care linkage relationship

Definition	A code representing the relationship of the linked subject of care to the subject represented by this identifier as defined by the subject of care.	
Source standards	Not applicable.	
Data type	Coded text.	
Data domain	Code	Description
	M	Mother
	F	Father
	S	Sibling
	C	Child
	O	Other relationship
Guide for use	The relationship indicates the relationship of the linked identification to the subject of care.	

EXAMPLE Patient (child) and mother. In the child's identifying information, the linkage to the mother will indicate a type of M, while in the Mother's identifying information the linkage would indicate that the mother has a child. Systems need not have matched "sets" of linkages, though this may often be the case.

It should be noted that the use of linkage relationships are social and should not (with the exception of a newborn in hospital) be assumed to indicate a biological relationship. It should also be noted that there may be significant legal issues to the provision of this information in specific cases, e.g. post-adoption, and that the legislative requirements of individual countries should be taken into account when using these fields.

Annex A (informative)

Collection of data

A.1 General

This Technical Specification addresses the collection and storage of data for the purpose of identifying individual subjects of care. It does not deal with the collection, use and storage of clinical, or other forms of health data. This section identifies some of the existing standards to provide assistance with appropriate management of health information, including privacy, and provides a short summary of the issues.

Users of this Technical Specification should refer to relevant privacy legislation and guidelines so as not to breach personal privacy in their collection, use, storage, transmission and disclosure of identifying information. In the event of any conflict between this Technical Specification and relevant privacy legislation and guidelines, the latter shall prevail.

The governing principle for this Technical Specification is to achieve a balance between personal privacy and the controlled and legitimate use of health information to improve health care.

A.2 Collection of data

Information for subject of care identification should not be collected unlawfully or by unfair means. Where lawful and practicable, subjects of care should have the option of not identifying themselves. Where the subjects choose not to identify themselves, or to withhold specific elements of identifying information, the potential impact of this decision should be explained to them, or their representative.

This Technical Specification does not require that all data elements specified within it will be collected in all circumstances. It is important to consider the impact of withholding or collecting specific elements of data.

When collecting identifying information, the collector should inform the subject of the purpose for collecting the information and how it may be used or disclosed, and by whom, as well as details of retention.

For example, this could be via a pamphlet or verbal explanation. This statement should include, as a minimum:

“The information is needed to positively and uniquely identify you so that your previous, current and future health service records can be associated with the correct person”.

The subject of care should be offered the option of disclosing the information in a private area. The collector of identification information should ensure that each piece of information is correct, up-to-date and changed as necessary.

Collectors should establish quality support programs and audit processes for monitoring the collection and recording of subject of care identification data.

It is recognized that the data held in subject of care identification systems could pose a threat to identity security. Collecting organizations are responsible for ensuring safe and appropriate use of the data collected and that their staff are equally responsible for their actions when accessing, using and collecting all health information.

Information may be collected directly (from the subject of care) or indirectly (from other persons, systems or organizations). For this reason, the SOC has the right to know about and to correct any data about them which might have been indirectly collected.

Consent might be required for data collection, especially in the case where information is collected indirectly, i.e. from other sources.

A.3 Collection difficulties

If a subject of care refuses to provide identifying details, the collector of information has the responsibility to explain the possible consequences of not being able to uniquely identify the subject of care in the health care system.

This Technical Specification does not require that all data elements specified within it will be collected in all circumstances. It is important to consider the impact of withholding or collecting specific elements of data.

Where it is not possible to uniquely identify a subject at the point of care, e.g. if the subject is unconscious, it is the responsibility of the collector to verify information with the subject as soon as practicable afterwards. As with the collection of any identifying information, the collector should inform the subject of the purpose for collecting the information and how it may be used and by whom (see 7.2).

Where it is not possible to uniquely identify a subject of care at the point of care, a temporary identification record should be created. This temporary record should be clearly identifiable as a temporary record. Where lawful and practicable, a subject of care should be able to remain anonymous.

In cases where a language barrier exists, translation services need to be obtained within a reasonable time period to verify information.

A.4 Transmission of data

Data transmitted between systems should be securely transferred. There should be authority to disclose the data transmitted. Encryption should be used where sensitive health information is being transmitted. Email should not be used for communicating personal health information unless additional security protocols are used.

A.5 Storage and data processing

This applies to data at rest (in storage) as well as during any processing and transmission.

It is the responsibility of any person or organization storing or using subject of care identifying information to ensure that processes (computerized and procedural) exist that adequately address the inherent risks of information loss or corruption, and that these processes are reviewed regularly.

Some of these risks, with examples of mitigation strategies, are identified in Table A.1. This list is not intended to be comprehensive.

Table A.1 — Risks and mitigations of information loss or corruption

Risk	Mitigations
Theft	Physical security.
Inappropriate access	Appropriate access controls and prohibitions on inappropriate use in policy, education and technology.
Unauthorized access	Session control (screen locks and time outs).
Unauthorized modification	Authentication (people actually accessing the data are authorized, e.g. through password regimes and practices) including encryption, firewalls, staff training and audit trails.
Loss and/or corruption	Backup and retrieval systems. Virus detection software.

An access model with varying levels of permissions should be provided to enable various restrictions of access to specific personal or sensitive health information.

A.6 Further information

Further information is available from the following legislation, standards, codes and guidelines. An outline of the contents of each is given in Appendix D.

AS/NZS 7799.2:2003, *Information security management — Part 2: Specification for information security management systems* (available, superseded).

AS/NZS ISO/IEC 17799-2006, *Information technology — Code of practice for information security management*.

HB 174-2003, *Information security management — Implementation guide for the health sector*.

A.7 Data access

Multiple access levels should be provided to control multiple access privileges to personally identifiable information. A log of access to these data should be maintained.

Many elements of health data are sensitive. In circumstances of family disruption even the privacy of addresses can be critical. This Technical Specification does not address this issue but methods for the management of this issue, including staff training, should be carefully and appropriately implemented.

Multiple security levels should be provided to enable various restrictions of access to specific personal or sensitive health information.

A.8 Presentation

Effective and accurate identification of subjects of health care requires not only that the data used to identify a subject be standardized, but also that the presentation of the data be standardized to maximize clear and effective communication in both the clinical or operational workflow.

As users of data depend on varying information systems to present even standardized data, varying user interfaces will produce varying data interpretations in certain work environments. User interfaces should highlight and/or prioritise different data elements in a similar manner to support work environments in fast-paced clinical settings and other venues. Therefore, organizations should strive to present standardized data in similar formats where necessary, if the result of standardized data is to be standardized practice.

Annex B **(informative)**

Messaging

B.1 General

This annex describes how health care information systems can communicate information regarding subjects of care in a clinical setting.

Data communication of subject data between health information systems (via interfaces) brings great benefit by automating the subject identification process in several ways. In the simplest case, a system with knowledge of a subject can transmit that information to another system. If the subject is known by both systems, one system can update the information in the other system. Information on a subject can also be linked or merged on a system based on interface transactions received from another system.

Generally speaking, information transmitted on a subject can fall into two categories:

- a) information on the subject that is not likely to change from clinical visit to clinical visit, e.g. subject of care's name, sex, date of birth, address, phone number;
- b) information that is centred on one clinical care event and therefore is likely to change from visit to visit, e.g. subject of care's account number, health care provider, location.

For the purposes of subject of care identification, the information most likely to be helpful is the information that is not likely to change from visit to visit. Each time subject data are presented to a system, that system needs to:

- identify the subject in a foolproof, consistent and efficient manner;
- determine if this subject of care exists already in the database under the same identifiers (matching);
- determine if this subject of care exists already in the database under different identifiers (reconciliation);
- determine whether to link the subject of care records together (keep two distinct records under the existing identifiers, but bring them together within an application so that the entire subject of care history can be viewed), or merge them (all existing subject of care data are brought under one common identifier scheme).

B.2 Health Level Seven

The Health Level Seven (HL7) standard provides a robust template that can be used to facilitate subject of care identification via an interface. HL7 is a series of standards governing the communications and format of data transmission between health care systems. The mission of HL7 is to bring order and consistency to all aspects of health information data transmission.

HL7 is an internationally recognized health care information technology standard, which originated in 1987. It is created and owned by the Health Level Seven organization, which is headquartered in Ann Arbor, Michigan USA. HL7 has international affiliates in Australia, Argentina, Canada, China, Finland, Germany, India, Japan, Korea, the Netherlands, New Zealand, South Africa, Switzerland, Taiwan and the United Kingdom.

The name “Health Level Seven” refers to the top (seventh) level of the ISO Open Systems Interconnection model. This seventh level is also known as the application level. The domain or scope of the HL7 standards includes health administrative and clinical data. Therefore, HL7 is a messaging standard that can be used between health care applications internationally.

B.3 Summary

Electronic messaging is helpful, even essential, in today's health care information world. For electronic messages to be effective, positive identification of the subject that the message (and the data contained within it) pertains to should be achieved in a foolproof, consistent and efficient manner. The general category of data that provides the most value in the process of subject of care identification is data that do not change from visit to visit. In interfaces that utilize the HL7 messaging standard, these data will be found in the patient identification (PID) segment.

Annex C **(informative)**

Data matching

C.1 General

This Technical Specification aims to reduce the incidence and impacts of data matching errors, primarily by improving subject of care identifying information used in data matching. However, this annex also provides an overview of data matching methodologies, as the process of matching is another important determinant of successful identification. Note that if data are to be used for data matching, subjects of care should be informed of this intention, in accordance with privacy principles.

On the basis of matched identifying data, health care records may or may not be linked in order to form a more comprehensive source of clinical information. This Technical Specification aims to improve the capacity to match health care records through consistent and accurate identification of subjects of care. Progression to record linkage is a policy issue and is beyond the scope of this Technical Specification. Health care establishments should consider relevant laws, guidelines and policies regarding record linkage.

Note that if data are to be used for data matching, subjects of care should be informed of this intention, in accordance with principles of privacy and within the requirements of legislation. Health care establishments considering data matching should consider relevant laws, guidelines and policies regarding record linkage.

C.2 Uses of matching data

Data matching might take place for a variety of purposes, including clinical and administrative, as well as research and planning studies. There are significant differences in the generation and use of these two types of matching data, as summarized in Table C.1.

Table C.1 — Data matching usage

	Administrative/clinical	Planning/research
Example	Provision of a subject of care's health records to a doctor.	Mapping of subject of care addresses to optimize locations of new health facilities.
Accuracy	Linkage should be as accurate and as complete as possible because the data might be used for management of individual people's care. There are significant safety issues for errors of identification.	A small (known) proportion of linkage errors can be tolerated as the data are subject to statistical analyses.
	Usually (but not necessarily) relatively conservative (minimizing the number of incorrect links with an associated cost of a high number of missing links).	Usually fairly balanced (minimizing the total number of errors to equalizing the numbers of incorrect links and missing links).
Process	Records located by means of client identifiers.	Records are "de-identified" (client identifiers and other unique information is deleted or disguised). In some communities, it is practically difficult, or impossible, to de-identify data.
Access	Available in "real-time".	Provided in static form as "snapshots" as of a certain date.
Timeliness	Requires "live" connection to all databases.	Uses copies of data. Direct connection to databases is not necessary.
Privacy	Extremely sensitive in respect of privacy and confidentiality as records are individually identified.	De-identified data might still be sensitive in respect of privacy and confidentiality of particular subjects of care or groups of people.
Consent	Individual consent might be required from the subject of care.	Individual consent is not usually necessary.

While research and planning activities will also benefit from improved data matching capability, the primary focus of this Technical Specification is the health care service delivery setting, i.e. clinical and administrative uses.

Clinical and administrative data arising from a subject of care's previous interaction(s) with a health care provider or organization are associated with that specific subject via data matching. This matching can occur in a variety of ways, including:

- a) human inspection of records;
- b) computerized scanning of key variables;
- c) within single repositories, e.g. filing cabinets or information systems;
- d) between repositories, e.g. between separate databases, health service establishments or other organizations.

C.3 Quality issues in data matching

The reality of health care service delivery settings is that newly collected data of variable quality should often be used to match against existing data of variable quality, often leading to considerable imprecision. This variable quality arises from a multitude of sources, including:

- a) ability to capture data well, e.g. due to communication difficulties or trauma;
- b) incorrect recording and transcription;
- c) failure to capture and/or track changes;
- d) failure to search comprehensively for existing data.

In practice, particularly in very large-scale organizations, it is almost inevitable that some matching errors will occur. This risk is exacerbated when attempting to match across services or establishments, in order to provide continuity of care. These errors fall into the two following major categories.

- 1) False non-matches (Type I errors), i.e. failure to match identifying data which in fact is associated with the same subject of care. The health care consequences of these errors include risk of misdiagnosis or inappropriate treatment, based on lack of comprehensive data, and duplication of data collection, diagnostic testing, and/or prescription.
- 2) False matches (Type II errors), i.e. “matching” of records that in fact are not or should not be associated with the same subject of care. Health care consequences associated with these errors include risk of misdiagnosis or inappropriate treatment based on incorrect information and/or breach of confidentiality.

Intelligent matching mechanisms have substantial potential to reduce matching errors.

There are other issues of quality to consider when preparing to develop data matching algorithms. These can vary in different communities. They include the following examples.

- Data variability: data items with a small data set do not offer great variability. For example, sex offers only four options and will therefore not cut the potential matches down to a small number with the use of a single field, whereas family name is a highly variable field and therefore very useful in cutting down the number of potential matches.
- Data accuracy, including time. An item might be highly accurate at the time of collection but it can be highly volatile. For example, in a highly mobile community, an address might not be given as much weight in a matching process as it would in one where the community is stable.
- Data accuracy: collection and recall quality. Collection quality should also be considered. Data items such as family name are often misspelt, especially if your data collection occurs over the phone or where there is a multicultural community where cultural spelling variations might not be understood by all staff registering subjects of care. The problems of matching given names can be even more complex, where names that sound the same can have a wide variety of spellings (Catherine, Katherine, Kathryn) or where nicknames (Bill) have been used and the subject of care's “proper” name (William) hasn't been included in the registration system (which is often the case in legacy systems).

Matching processes involve identification of some data elements that are likely to be highly accurate (such as sex) and some that have high variability (a large number of potential values e.g. surname, street addresses).

C.4 Deterministic and probabilistic data matching

Matching of identifying data is generated by a variety of methods falling into two principal categories: deterministic and probabilistic. These terms actually represent the two ends of a spectrum rather than two completely distinct methods.

Deterministic methods match sets of identifying information on the basis of whether they agree on certain variables. For example, records may be matched on family name, initial of first given name, date of birth and sex. Matches are only made when the details in pairs of records are identical, otherwise they are regarded as being different: “John D'Arcy Langton” and “John D'Arcy Langdon” are not a match in deterministic terms. A match based on an encrypted statistical linkage key (SLK) such as the combination of family name, date of birth and sex is an extreme form of deterministic linkage.

In probabilistic matching, these two names would be regarded as a highly probable match. This method uses probabilities to determine whether a pair of records both refer to the same subject of care. Weights are calculated based on these probabilities in order to quantify the likelihood that a pair of records actually match. For every record pair examined, a weight is assigned based on the information obtained from the comparison of the fields used in the linkage. The value of the total weight is used to assess whether the linked pair is a

true match, a non-match, or a probable match. Depending on the type of comparison, probabilistic weights may be either non-specific or value specific.

General or non-specific weights are based on the agreement/disagreement of a specific field, e.g. using general weights, agreement/disagreement on birth date may be given a static positive/negative weight.

Value-specific weights are based on the agreement of a specific value of the field being compared, e.g. if comparing initials using value specific weights, a match of the value "B" will receive a different weight than a match of "A". In general, rare agreements carry higher weights, e.g. in most parts of Australia two records with the family name "Jones" will generally have a lower weight than two records with the family name "Tshabalala".

Weights are highly dependent on the distribution of values within a field. For example, if a data set has roughly equal proportions of males and females, then an agreement on this field would have only a small positive weight (after all, there is a fifty-fifty chance that any two randomly chosen records will have the same gender). A disagreement on this field would usually have a large negative weight (in most data sets, variation in sex between records belonging to the same subject of care is a rare occurrence).

Note that these processes are not solely the domain of computers. A person undertaking manual inspection of records similarly forms opinions on the probabilities of different records being associated with the same subject of care.

C.5 Prospective and retrospective data matching

Two further variants of matching warrant consideration: prospective (active) and retrospective (passive).

In the prospective (active) case, matching occurs during the current process of registration or clinical care (either on-line or with a manual record at the immediate disposal of the recording personnel). In this case, reconciliation queries can be clarified with the subject "on-the-spot" and data quality can be confirmed or improved. In the retrospective (passive) case, matching occurs subsequently to the episode of registration and/or care, and third-party judgements should be made concerning the likely validity of matches.

C.6 Selection of matching methodologies

The matching methodology used in practice should be tailored to the uses to be made of the results, the nature and quality of the data being matched, and the relative risks associated with Type I and Type II errors. Where the number of records involved is relatively small, or the quality of identifying data is known to be uniformly high, and/or the consequences are such that it might be preferable to miss a possible match than to risk an erroneous match, deterministic methods might be more appropriate. Probabilistic matching is likely to be more useful when the numbers of records involved are high and the quality of identifying data is variable, not known, or problematic.

Where computerized matching is undertaken, it is recommended that:

- a) prospective matching be undertaken wherever feasible, including verification with the subject "on the spot";
- b) underlying matching algorithms be statistically defensible;
- c) matching algorithms have the capacity for user specification of linking variables, adjustment of weightings to local circumstances, reflecting the chosen decision rules and acceptance of spelling variations;
- d) probabilistic matching be accompanied by human intervention to resolve uncertainties arising in the reconciliation processes;
- e) policies and procedures be put in place to articulate decision rules that reflect the characteristics of the population of interest.

Annex D (informative)

Guide for implementation of subject of care master indices

D.1 General

Health care organizations are addressing the need for positive subject of care identification by implementing an enterprise-wide subject of care master index. The term subject of care master index is used here to describe a client registration system that extends across the enterprise to include a range of services and health care establishments or related organizations. For the purposes of this Technical Specification, a patient master index is a registration system limited to one specific organization, service or establishment (site).

A subject of care master index is a permanent record of all clients or potential clients registered by the enterprise (organization of one or more health care establishment sites). It is called an index because it serves as a guide to each client's records by linking the client's name to the unique identification number(s) used on their health records.

The subject of care master index plays a key role in supporting continuity of client care across the organization. As people move within and between services, and health care organizations, the index is used to ascertain whether the person has an existing health record and, if so, gives its reference number. The record can then be retrieved so the health care provider(s) can familiarize themselves with the client's past clinical history and document details of the care and services they provide.

Some organizations use the subject of care master index to support a single subject of care identification number for use across the organization. Other organizations use a subject of care master index to provide a cross-reference to multiple client identification numbers issued to the client by each of the various services, health care organizations or establishments.

Where two or more patient master indices are brought together to create a subject of care master index, the organizations seek to match a subject of care's registration details on one "patient master index" (PMI) with their registration details on another PMI. If the subject of care is registered on both PMIs, that subject of care master index can be used to maintain a cross-reference to both PMIs and the client's records.

The success of data matching through a subject of care master index is largely dependent on the uniformity and quality of subject of care identification data collected, the power reliability and appropriateness of the data matching mechanisms, and the recognition of client registration as a critical health care event.

The following information is provided as a generic guide to the implementation of a subject of care master index, as it relates to issues arising in the ambit of this Technical Specification.

D.2 Sites and systems

Outline the scope of the project by identifying the sites and information systems that will make use of the subject of care master index.

- a) List all sites that will be referencing the subject of care master index at the completion of the project.
- b) For each site, list all enterprise applications that contain subject of care identification data, e.g. pathology, medical imaging.
- c) For each site, list all departmental applications that contain subject of care identification data, e.g. cardiology, emergency, pharmacy.

- d) For each departmental application, list those that are currently on the organization's network. Indicate any that will be networked by the completion of the subject of care master index implementation.
- e) List any of the current systems (enterprise and departmental), which will be phased out by the end of the project. Recognize that cessation of any existing filing identifiers will require conversion of data and re-filing of the information collected. This may be a prohibitive procedure and would be a sound reason for retention of multiple identifiers within an organization.
- f) List any new systems (enterprise and departmental), which will be implemented during the course of the project.

D.3 Subject of care registration

Outline the current subject of care registration practices across the enterprise by identifying the sites and departments that register subjects of care.

- a) For each site, list those that can register subjects of care on the organization's PMI, e.g. health records, admissions, emergency, outpatient clinics.
- b) For each site, list the departments that can register subjects of care on their own systems, e.g. pathology, radiology.
- c) Ascertain whether privately referred subjects of care (who may have also been subjects of care at one or more of the sites) who are registered on a departmental system, e.g. pathology, are to be included on the subject of care master index.

D.4 Interfaces

Outline the current systems interfaces by identifying the systems that can send and receive subject identification information.

- a) Prepare a diagram containing all systems (both enterprise and departmental) that currently have interfaces for the purpose of sending or obtaining subject of care identification data.
- b) Of these systems, indicate which can automatically receive subject of care data broadcast to them.
- c) Indicate which systems send a request (solicit) for subject of care data.
- d) Indicate which systems may receive subject of care data only with client consent.
- e) List those systems whose interfaces currently use messages in the HL7 format (the software vendor should have this information).
- f) If the systems do not currently use HL7, list those that have that potential capability to do so (the software vendor should have this information).
- g) Record details of any other interfaces that exist between systems that send and/or receive client identification information.

D.5 Access

Ensure appropriate access to the subject identification data by establishing and promulgating a comprehensive access policy.

- a) Review the requirement for access by each potential user group.
- b) Review all existing policies across the enterprise regarding access to subject of care identification data with the view to developing one enterprise-wide policy. The policy should take into consideration relevant legislation, regulations, codes and standards related to information privacy. The policy should consider and, where appropriate, cover:
 - 1) collection and entry of data,
 - 2) modifying existing data,
 - 3) read-only access to the data, and
 - 4) clients' access to data about themselves.

D.6 Data validity

Determine which is to be the primary reference system for data validity (for a given site and across all sites), the originating patient master index or the subject of care master index.

D.7 Data integrity

Promote data integrity on the subject of care master index by establishing staff roles and responsibilities.

- a) Determine which staff will have responsibility for ongoing monitoring of data integrity on subject of care master index data.
- b) Establish how the data are to be formatted in the subject of care master index.
- c) Establish which code sets need to be standardized.
- d) Establish the desired standard format(s) in the subject of care master index.
- e) Determine who will have the responsibility for the merging of duplicate records.
- f) Establish criteria to be used to assess records as duplicates.
- g) Establish procedures to assess potential duplicate records.
- h) Establish lines of communication to be used to notify departments, e.g. blood bank, of duplicate registrations.
- i) Establish audit procedures for ongoing monitoring of data quality.

Annex E (informative)

Guidelines for searching for a subject of care

For efficient subject registration and accurate recording of subject history and biographical data, information should be gathered using effective interviewing techniques, i.e. by asking the right questions. The aim of subject registration, or subject identification, is to ensure that any existing subject data, and related health data, will be linked to the relevant subject. Having the correct and complete record of a subject of care will assist in the provision of care to that subject. The importance of finding the correct subject of care identifier for the subject at the time of attendance cannot be overstated. Systems should be searched thoroughly for the subject's identifier. If a previous identifier cannot be found then, and only then, should a new identifier be allocated.

Search methodologies are dependent on the logic of the health information system. The following information should be used as a guide in developing policies appropriate for your organization to maximize the likelihood of appropriately identifying subjects. When selecting an appropriate search methodology, the following should be the guiding principles.

- Staff, no matter how well trained, generally will not search multiple times in order to find a subject of care; they are more likely to assume that they aren't there and issue a new identifier. This is simple human nature; computer systems can assist by doing a quality first cut search built to suit the community of subject of care data.
- The objective of the search is to find any entries in the system that could be the person for whom you are searching, not to exclude those that don't exactly match. For this reason, exact matching systems should be rare as a combination of matching algorithms are proven to be more accurate.
- Search criteria should be defined so that they do not exclude the person for whom you are searching. Research has shown that the most successful criteria for finding a person in your system, even where the details are not exactly the same, are as follows.
 - Surname (with matching algorithms appropriate to the culture of the community such as soundex, dolby or others). It should be noted that sound-a-like systems such as soundex are not effective in Asian or Polynesian communities as the names in these groups have a high incidence of the characters AEIOUWHY. These characters are all ignored in the soundex algorithm. The common soundex algorithm groups S,C,G,K,Q,X,Z together and treats them as if they are the same character.
 - Date of birth (with either age matching, or percentage match algorithms such as where at least six of the eight date characters are the same).
 - Sex (with any entries already recorded as I or U matching searches for males or females).

Though a given name is an excellent identifier, it is a poor element for search criteria as it often causes people to be excluded from the potential matches. A more appropriate process is the display of given names when presenting potential matches, to allow the user to define matches manually.

For people with difficult or long names, the final search should be a search on the first three or four letters of the family name. The system will display all names commencing with the exact letters provided, thus increasing the probability of locating the required record.

If you are using a system that allows given name searching when searching for an Asian name, a search should also be conducted using any known middle names, though it is preferable to have computer matching on all given names, irrespective of sequence. This is necessary as the first name of an Asian name may indicate the title of the client. For example, the Vietnamese name elements Thi = Miss and Van = Mr. When entering these names into any system, ensure that Thi or Van is entered as a name title.

Also note that a sex-assisted search may be unreliable in communities where the sex relationships of given names are not known in the collecting community. For example, in an Asian community, the sex related to the name Roberta might not be clear to the person recording the information (in which case the sex may have been recorded as male). Where it is unclear, a sex of unknown should be recorded. A similar situation applies in an Anglo-Saxon community with Asian names. In situations where it is likely that sex has been poorly or inaccurately recorded the use of sex as a matching criteria should be avoided.

When searching for “common” names, the subject's sex and age should be entered to reduce search time. However, this should not be the only technique used to establish whether the subject is known.

Identification management systems should automatically search “baby of” entries as if their given name matched all other names, thus forcing the user to check the name with newborn baby entries prior to allocation of a new identifier.

In systems where there are likely to be subjects of care returning to the hospital where they were born, the search should have the facility to include the mother's family name to assist in identifying the subject of care.

© ISO 2011

Annex F (informative)

Names condensed guide

F.1 General

There are many naming systems in the world that differ quite markedly from the system adopted in English-speaking countries. For example, some Asian groups such as the Chinese, Vietnamese and Japanese have a “family name first” system in their own environment. However, they might reverse this system in an English-speaking environment and have their family name last. Naming systems in a multicultural country can cause confusion. It is therefore important to be alert and to check, if possible, with the client which part of the name is the family name, given name and middle name (or given names). It is also recommended that the term “Christian name” be replaced in official documents by “given name”. The term “first name” is not suitable to the naming procedures adopted by various cultural and ethnic groups.

F.2 Naming structures

The following examples are a guide only. It is advised that the subject of care be asked which name is their given name at the time of registration if there is any uncertainty about which name is the given name and which is the family name. If the client's name cannot be found on an initial search of the subject of care master index, try performing a search using the names in a different order, e.g. search using the subject's given name as the family name and vice versa.

The following examples were adapted from Reference [21].

Table F.1 — Vietnamese names with family name first and given name last

Name structure			Recorded as	
Family name	Second given name	First given name	Given names	Family name
Dang	Thanh	Thuy	Thuy Thanh	Dang

Table F.2 — Chinese names with family name first, given name last

Name structure			Recorded as	
Family name	Second given name	First given name	Given names	Family name
Zang	Thi	Man	Man Thi	Zang

Table F.3 — Arabic names with prefix

Name structure				Recorded as	
Given name	Middle name	Prefix	Family name	Given names	Family names
Jamal	Ali	El	Haddad	Jamal Ali	El Haddad

Record an alias as Haddad.

Prefixes El and Abu may be stored in title if this is culturally appropriate.

Table F.4 — European names with different gender family names

Ethnic background	Male family name	Female family name
Polish	Kowalski	Kowalska
Macedonian	Karlevski	Karlevska
Russian	Ivanov	Ivanova

NOTE Male family names may end in "i", females in "a". Check both formats if no record is found on initial search of the subject of care master index.

Table F.5 — European names with family names followed by given names

Name structure			Recorded as	
Ethnic background	Family name	Given name	Given name	Family name
Croatian	Bosnjak	Teresa	Teresa	Bosnjak
Serbian	Dragicevic	Zoran	Zoran	Dragicevic
Hungarian	Szabo	Janos	Janos	Szabo
Polish	Kowalski	Piotr	Piotr	Kowalski

Table F.6 — Names with both father's and mother's family names — Spanish

Name structure				Recorded as	
Given name	Middle name	Father's family name	Mother's family name	Given names	Family names
Victor	Hugo	Marillanca	Flores	Victor Hugo	Marillanca Flores

NOTE Mother's family name is usually not used.

Table F.7 — Names with both father's and mother's family names — Portuguese or Filipino

Name structure				Recorded as	
Given name	Middle name	Mother's family name	Father's family name	Given names	Family name
Maria	Manuela	Costa	Ferreira	Maria Manuela Costa	Ferreira

NOTE The mother's family name is usually not used.

Table F.8 — Where mother's family name is after father's family name

Name structure				Recorded as	
Given name	Middle name	Father's family name	Mother's family name	Given names	Family name
Maria Victor	Manuela	Ferriera	Costa	Maria Manuela	Ferreira

NOTE The mother's family name is usually not used.

Bibliography

- [1] ISO/IEC 10646, *Information technology — Universal Coded Character Set (UCS)*
- [2] ISO/IEC 11179-2, *Information technology — Metadata registries (MDR) — Part 2: Classification*
- [3] ISO/IEC 11179-3, *Information technology — Metadata registries (MDR) — Part 3: Registry metamodel and basic attributes*
- [4] ISO/IEC 17799:2005, *Information technology — Security techniques — Code of practice for information security management*
- [5] ISO/IEC 19785-1, *Information technology — Common Biometric Exchange Formats Framework — Part 1: Data element specification*
- [6] ISO/IEC 19785-2, *Information technology — Common Biometric Exchange Formats Framework — Part 2: Procedures for the operation of the Biometric Registration Authority*
- [7] ISO/IEC 19785-3, *Information technology — Common Biometric Exchange Formats Framework — Part 3: Patron format specifications*
- [8] ISO/IEC 11404, *Information Technology — General-Purpose Datatypes*
- [9] ISO/IEC 19784-1, *Information technology — Biometric application programming interface — Part 1: BioAPI specification*
- [10] AS 4400-1995, *Personal privacy protection in health care information systems*
- [11] AS 4700.1-2005, *Implementation of Health Level Seven (HL7) Version 2.4 Patient administration*
- [12] ASTM E1714-95, *Guide for Properties of a Universal Healthcare Identifier (UHID)*
- [13] AS/NZS 7799.2-2003, *Information security management, Part 2: Specification for information security management systems*
- [14] Australian Bureau of Statistics, *Directory of concepts and standards for social, labour and demographic statistics*, 1995
- [15] Australian Commonwealth Department of Health and Ageing, *Home and Community Care Data Dictionary (HACCDD)*, Version 1.0
- [16] Australian Commonwealth Department of Health and Ageing; *Home and Community Care Guidelines*, 2002. Available at: http://www.health.gov.au/internet/wcms/publishing.nsf/content/hacc-pub_pg_npg.htm
- [17] Australian Institute of Health and Welfare; *National Community Services Data Dictionary (NCSDD)*, Available at: <http://www.aihw.gov.au>
- [18] Australian Institute of Health and Welfare; *National Health Data Dictionary (NHDD)*, Available at: <http://www.aihw.gov.au>
- [19] Australian Institute of Health & Welfare; *Metadata Online Registry*, Available at: <http://meteor.aihw.gov.au/content/index.phtml/itemId/181162>, [cited 17 May 2005]
- [20] Australian Bureau of Statistics; *Standard Australian Classification of Countries (SACC)*, ABS Catalogue No 1269.0, 1998

ISO/TS 22220:2011(E)

- [21] Centrelink, *Naming Systems of Ethnic Groups: Ethnic Names Condensed Guide*, AGPS Canberra, 1997
- [22] HB 174-2003, *Information security management — Implementation guide for the health sector*
- [23] HB 222-2006, *Australian Health Care Client and Provider Identification Handbook*
- [24] HNBC 98-10, *HealthNet/BC Provider Data Standard Version 1.0*
- [25] *HealthNet/BC Provider ID Standard*, British Columbia Health Information Standards Council, B.C. Ministry of Health and Ministry Responsible for Seniors, June 1999
- [26] Hicklin, R.A. and C.L. Ready, *Implications of the IDENT/IAFIS Image Quality Study for Visa Fingerprint Processing*. Mitretek Systems, 2002
- [27] International Biometric Industry Association, *Biometrics Overview*, Available at: <http://www.ibia.org/biometrics>
- [28] South Australian Department of Human Services, *Best Practice Guidelines for Patient Master Index Maintenance*, Version 1.0:1998
- [29] South Eastern Sydney Area Health Service, *Patient registration and auditing policy for HOSPAS sites — Unique patient identifier initiative*, 1999/2000
- [30] SPENCE, B., *Biometrics' Role in Physical Access Control*, Loss Prevention and Security Journal, 2003 (14 Aug 2005)

ICS 35.240.80

Price based on 88 pages