

Investigation of syntaxes for existing interchange formats to be used in healthcare

National foreword

This Published Document which has been prepared under the direction of the Information Systems Technology Policy Committee, by Technical Committee IST/35, is the English language version of the CEN Report entitled *Investigation of syntaxes for existing interchange formats to be used in healthcare*, CR 1350 - September 1993.

Compliance with a British Standard does not of itself confer immunity from legal obligations.

REPORT
RAPPORT
BERICHT

CR 1350:1993

September 1993

English version

**Investigation of syntaxes
for existing interchange formats
to be used in healthcare**

This CEN REPORT has been established by the CEN/TC 251 "Medical informatics" and has been approved by CEN on 1993-07-01.

CEN members are the national standards bodies of Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom.

CEN

European Committee for Standardization
Comité Européen de Normalisation
Europäisches Komitee für Normung

Rue de Stassart 36, B - 1050 Brussels

© CEN 1993 Copyright reserved to all CEN members

Table of Contents

FOREWORD	1
PART I: METHOD, INVESTIGATION, CONCLUSIONS AND RECOMMENDATIONS	3
1. INTRODUCTION	5
1.1. Scope	5
1.1.1. Original Scope	5
1.1.2. Reduced Scope	5
1.2. Definitions and Acronyms	6
1.2.1. Definitions	6
1.2.2. Acronyms	7
2. INTERWORKING BETWEEN HEALTH CARE APPLICATIONS	9
2.1. Message Development Method	10
2.1.1. Scope of the Problem Domain	10
2.1.2. Communication Parties and Communication Roles	10
2.1.3. Scenarios	11
2.1.4. Services	11
2.1.5. Domain Information Model (DIM)	11
2.1.6. General Message Descriptions (GMDs)	12
2.2. Mapping to Interchange Format Dependent Message Descriptions	13
3. FUNCTIONAL REQUIREMENTS FOR INTERCHANGE FORMATS	15
3.1. Support of Information Structures	15
3.2. Support of Datatypes	15
3.3. Encoding	15
3.4. Evolution and Backwards Compatibility	16
3.5. Conformance and Certification	16
3.6. Support and Availability	16
4. INVESTIGATION METHOD	19

5.	EVALUATION CRITERIA	21
5.1.	Scope of Evaluation Criteria	21
5.2.	Domain Dependencies	21
5.3.	Classified Evaluation Results	22
5.4.	Excluded Evaluation Criteria	22
5.5.	Description of the Evaluation Criteria	23
5.5.1.	Supported Information Structures	23
5.5.2.	Support of Datatypes.....	26
5.5.3.	Encoding.....	30
5.5.4.	Evolution and Backwards Compatibility	31
5.5.5.	Conformance	31
5.5.6.	Support and Availability	32
6.	SELECTED INTERCHANGE FORMATS	35
6.1.	Description of Evaluated Interchange Formats	37
6.1.1.	ASN.1	37
6.1.2.	ASTM E1238/HL 7	37
6.1.3.	EDIFACT	38
6.1.4.	EUCLIDES.....	38
6.1.5.	ODA	38
7.	EVALUATION RESULTS AND DISCUSSION	41
7.1.	Supported Information Structures	41
7.2.	Supported Data Types.....	42
7.3.	Encoding.....	43
7.4.	Evolution and Backwards Compatibility	43
7.5.	Conformance	44
7.6.	Support and Availability	44
8.	SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	45
8.1.	Overview of Investigation Process	45
8.2.	Assessment of IFs	45
8.3.	Development of IFDMDs.....	46
8.4.	The Need for Mapping Rules from GMDs to IFDMDs	47
8.5.	Limitations of Evaluated IFs	47

PART II: EXAMPLES AND SUPPLEMENTARY INFORMATION	49
9. EXAMPLE A - LABORATORY COMMUNICATION	51
9.1. Scope of the Problem Domain	51
9.2. Communication Roles.....	51
9.3. Scenarios	51
9.3.1. Scenario I - GP to Laboratory, GP Collects Specimens.....	51
9.3.2. Scenario II - Clinical Department to Laboratory in Same Hospital.....	52
9.4. Services	53
9.5. Domain Information Model (DIM)	54
9.5.1. Specification of Attributes	57
9.5.2. Specification of Operations.....	59
9.6. General Message Descriptions (GMDs)	60
9.6.1. GMD for Lab Request in Scenario I.....	61
9.6.2. GMD for Lab Request in Scenario II	61
9.7. Functional Requirements for Interchange Format	63
10. EXAMPLE B - INTERNAL MEDICINE TRANSFER FOLDER	65
10.1. Scope of Problem Domain	65
10.2. Communication Roles.....	65
10.3. Scenarios	65
10.3.1. Scenario I - Internal Medicine Unit sends Folder to Cardiologist	65
10.4. Services	66
10.5. Domain Information Model	67
10.5.1. Specification of Attributes	68
10.6. General Message Descriptions (GMDs)	71
10.7 Functional Requirements for Interchange Format	72
11. ASN.1	73
11.1. Mapping to Interchange Format Dependent Message Descriptions.....	73
11.2. Evaluation Results	76
11.2.1. Supported Information structures.....	76
11.2.2. Supported Datatypes.....	77
11.2.3. Encoding	77
11.2.4. Evolution and Backwards Compatibility	78
11.2.5. Conformance	78
11.2.6. Support and Availability	79

12.	ASTM.....	81
12.1.	Mapping to Interchange Format Dependent Message Descriptions.....	81
12.1.1.	Mapping scheme.....	81
12.2.	Evaluation Results.....	85
12.2.1.	Supported Information Structures.....	86
12.2.2.	Supported Datatypes.....	86
12.2.3.	Encoding.....	87
12.2.4.	Evolution and Backwards Compatibility.....	87
12.2.5.	Conformance.....	87
12.2.6.	Support and Availability.....	88
13.	EDIFACT.....	89
13.1.	A General Mapping Method for EDIFACT.....	89
13.1.1.	How Mapping Could be Done.....	89
13.2.	Mapping to Interchange Format Dependent Message Descriptions.....	91
13.2.1.	The Objects.....	91
13.2.2.	The Message.....	92
13.3.	Evaluation Results.....	94
13.3.1.	Supported Information Structures.....	94
13.3.2.	Supported Datatypes.....	95
13.3.3.	Encoding.....	95
13.3.4.	Evolution and Backwards Compatibility.....	96
13.3.5.	Conformance.....	96
13.3.6.	Support and Availability.....	96
14.	EUCLIDES.....	97
14.1.	Mapping to Interchange Format Dependent Message Descriptions.....	97
14.1.1.	Mapping from GMD A-1.....	97
14.1.2.	Mapping from GMD A-2.....	103
14.1.3.	Mapping from the Combination of GMD A-1 and GMD A-2.....	104
14.2.	Evaluation Results.....	107
14.2.1.	Supported Information Structures.....	107
14.2.2.	Supported Datatypes.....	108
14.2.3.	Encoding.....	108
14.2.4.	Evolution and Backwards Compatibility.....	108
14.2.5.	Conformance.....	109
14.2.6.	Support and Availability.....	110

15.	ODA.....	111
15.1.	The ODA standard.....	111
15.2.	Mapping to Interchange Format Dependent Message Descriptions.....	111
15.3.	Evaluation Results	116
15.3.1.	Supported Information Structures	116
15.3.2.	Supported Datatypes	117
15.3.3.	Encoding	117
15.3.4.	Evolution and Backwards Compatibility	118
15.3.5.	Conformance	118
15.3.6.	Support and Availability	118
	ANNEXES	119
A.	TERMS OF REFERENCE	119
B.	PROJECT TEAM MEMBERS	121
C.	BIBLIOGRAPHY	123
C.1	Standard Publications	123
C.2	Other Publications	127

Foreword

This technical report was prepared under the direction of the European Committee for Standardization (CEN) and approved by the CEN Technical Committee 251 (TC 251 - Medical Informatics).

A project team, CEN/TC 251/PT004, was established in February 1992 based on mandate BC-IT 211 given by CEN/TC 251 and approved by the CEN/BT. A copy of the terms of reference for the project team is provided in annex A. The project team consisted of 8 members, given annex B.

This technical report has gone through three stages:

First, an Interim Report (INR) was produced and presented for the CEN/TC 251/WG 3 for technical control. Based on the comments received from members of WG 3, the report was updated and a First Working Document (FWD) was produced.

The FWD was submitted to CEN/TC 251 for a three months period of comments. A total of 84 comments were received. Based on these comments, an amendment to the Technical Report was produced and presented to WG3. The amendments, together with the FWD, were accepted by WG3 as a final technical document and forwarded to CEN/TC 251 for approval. CEN/TC 251 approved these documents on 1993-01-26.

This Final Document (FIN) of the Technical Report is an editorial update of the FWD together with the approved amendments. In order to enhance the readability of the report, it has been split into two parts. Part I is the main part and includes the explanation of the method and investigation together with conclusions and recommendations. Part II contains two examples of health care domain analyses and supplementary information about the investigation of each interchange format.

Sigurd From
Leader of CEN/TC 251/PT 004

Oslo, January 1993

PART I: Method, Investigation, Conclusions and Recommendations

1. Introduction

The objective of this technical report is to investigate syntaxes for existing interchange formats (IFs) to be used in health care and to define a strategy for selecting IFs. This requires the establishment of a framework for interworking between health care applications. Today such a framework does not exist and PT 004 has described a preliminary method for development of messages for information exchange within the health care which will be used in the investigation and the definition of a future strategy.

1.1. Scope

1.1.1. Original Scope

In the terms of reference for PT 004 the scope of the investigation was the whole of the health care area. The number of suggested evaluation criteria was large. In the recommended strategy for the project team it is stated "The interchange formats need to be evaluated against a set of properties. The properties will be selected both from health care and technical requirements including efficiency, richness, complexity, ambiguity, flexibility, cost and practicality".

1.1.2. Reduced Scope

Due to resource constraints, PT 004 has reduced the scope of its work. This has been done in three ways:

1. The domain of health care is reduced to two examples from the laboratory communication domain and the internal medicine domain. This selection is explained in chapter 4.
2. The number and type of evaluation criteria were reduced compared to the suggested properties in the terms of reference. Selection of the evaluation criteria was done based on the established method for message development. A more detailed explanation is given in chapter 5.
3. The number of IFs were reduced from a possible 23 down to 5. The selected formats were ASN.1, ASTM E1238, EDIFACT, EUCLIDES and ODA. This selection is explained in chapter 6.

1.2. Definitions and Acronyms

1.2.1. Definitions

Term	Definition
Communication Party	An actor of the problem domain participating in the communication therein.
Communication Role	The role of a party in a communication situation defining the party's behaviour in the communication. Associated with a communication role is a set of services which the party provides to other parties.
Domain Information Model	The model describing common concepts and relationships for a problem domain.
General Message Description	A conceptual description of a message independent of interchange format. A GMD may be regarded as one specific subset of a DIM related to one message type.
Interchange Format	The representation of the data elements and the structure of a message while in transfer between systems. The IF consists of a set of construction elements and a syntax. The representation is technology specific.
Interchange Format Dependent Message Description	The mapping of a conceptual message (GMD) into a message using a specific interchange format.
Problem Domain	The field of health care under consideration in a modelling process. Modified from [Coad-91].
Scenario	Formal description of a class of business activities including the semantics of business agreements, conventions and information content [Open-Edi-91].
Service	A specific behaviour that a communication party in a specific role is responsible for exhibiting. Modified from [Coad-91].
Syntax	The syntax of an interchange format describes the rules for combining the construction elements of the IF.

1.2.2. Acronyms

Acronym	Description
ASCII	American Standard Code for Information Interchange
ASN.1	Abstract Syntax Notation One
ASTM	American Society for Testing and Materials
BER	Basic Encoding Rules
CMIS	Common Management Information Service
CMIP	Common Management Information Protocol
DICOM	Digital Imaging and Communication in Medicine
DIM	Domain Information Model
EDI	Electronic Data Interchange
EDIFACT	Electronic Data Interchange For Administration, Commerce and Transport
GMD	General Message Description
HL7	Health Level 7
IF	Interchange Format
IFDMD	Interchange Format Dependent Message Description
IPI	Image Processing and Interchange
JPEG	Joint Photographic Experts Group (JTC1/SC29/WG10)
MEDIX	Medical Data Interchange Committee
ODA	Open Document Architecture (CCITT)
OSI	Open Systems Interconnection
SGML	Standardized Generalised Mark-up Language
SQL	Structured Query Language

2. Interworking between Health Care Applications

The health service consists of a set of health care parties that collaborate in a network. Each party provides and requests services to and from the other parties. This mutual dependency results in exchange of information between the parties. Today this information is mostly exchanged by phone or mail. The information exchange may be part of any health care activity including clinical, patient management, financial management, logistics, etc.

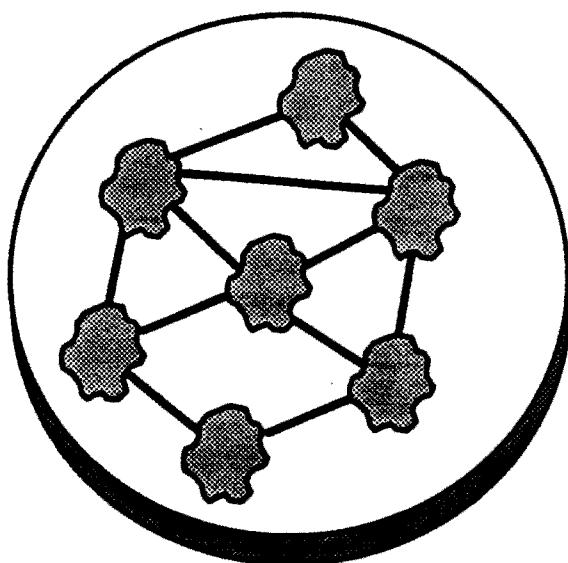


Figure 1: The health care network with health care service requesters and providers

The introduction of electronic communication in the health care should make communication between the parties in the network easier and more efficient and improve the quality of the service provided by the health care parties. In addition new ways of communicating will open the way for new types of services such as remote diagnosis between hospitals.

Each party participating in the network must be supported by computer applications that are able to interwork. Integrating computer applications have two sides: the development of a common description of the health care or application domain and finding good technical solutions to support the requirements in the application domain.

IFs can be viewed as technical solutions to the problem of representing information exchanged between computers. The domain requirements should decide which functions of IFs are important, which again will help selecting an appropriate IF.

PT 004 has used a method for analysing an application domain, finding the domain requirements and creating formal descriptions of the domain. The result of applying this method is a set of message descriptions that are independent of any IF. These descriptions are called General Message Descriptions (GMDs). The GMDs set the functional requirements for the IF.

2.1. Message Development Method

Figure 2 shows the message development method used by PT 004. Each box in the figure illustrates a task in the method and the respective documentation. The method is developed based on results emerged from other standards activities [From-91], [Harrington-93], [KITH-92], [Open-Edi-91].

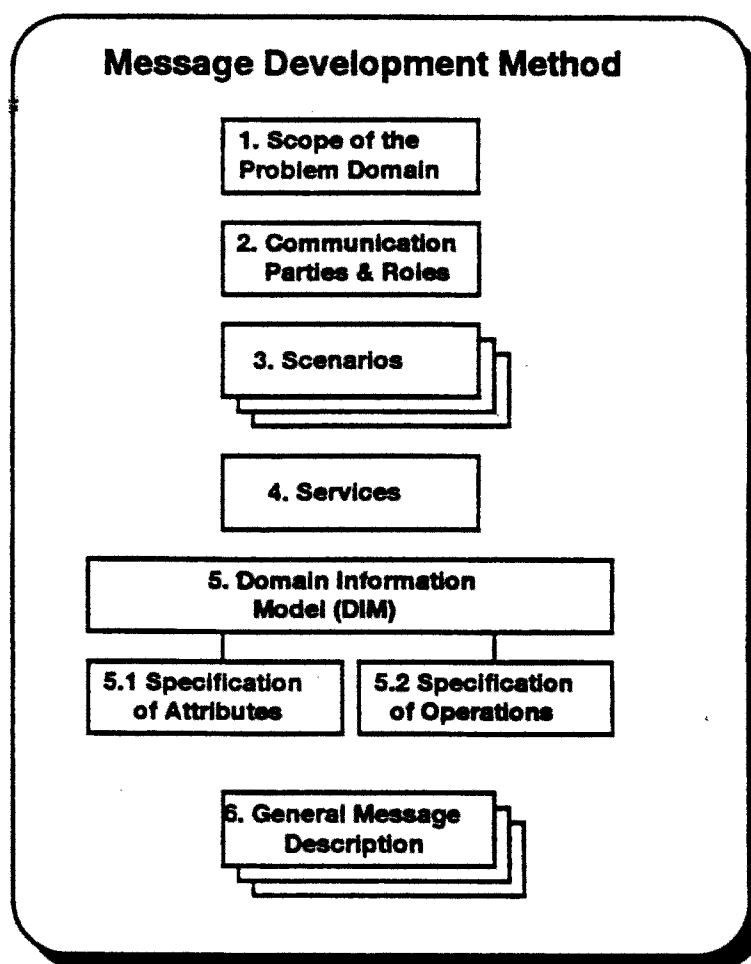


Figure 2: Message Development Method

2.1.1. Scope of the Problem Domain

Health care is a large entity which incorporates many application areas. In analysis, it is important to limit the scope to a manageable size. The field under consideration is called the **problem domain**. Our first task is to define the scope of the problem domain. The scope gives the purpose and the context for the modelling. It is helpful to specify what is both inside and outside the scope.

2.1.2. Communication Parties and Communication Roles

The main actors of the problem domain which participate in the communication are called the **communication parties**. Each communication party has one or more **communication role**

which defines their behaviour in a communication situation. For instance, a doctor requesting a laboratory analysis is a requester. It is the role as requester that is important in a laboratory requesting situation. Similarly the laboratory is the provider. Communication roles may be held by different communication parties. In the example, the doctor as a requester may be substituted by a nurse or a veterinarian who issues a laboratory request.

2.1.3. Scenarios

The activity among the communicating parties in a network may be described by a set of **scenarios**. A scenario is defined as a formal description of a class of business activities [EDI-SWG-91]. As for theatres, scenarios consist of a set of roles played by a number of role-players. In this context the role-players are the communication parties.

Scenarios may be used to describe dynamic aspects of a problem domain. In this report the scenarios are used to show the sequence of messages exchanged between the communication parties in the problem domain.

2.1.4. Services

A set of **services** are associated with each communication role. A service defines a specific behaviour that the party holding the role is responsible for exhibiting when the service is requested.

How detailed a service is described depends on the requested processing in the receiving application. In some cases, for instance when a request is made for an analysis, it may be interesting to specify whether or not the laboratory should interpret the results before returning the answer or whether only the pure measurement results should be returned.

2.1.5. Domain Information Model (DIM)

Like human beings, computer systems need to share a common world of concepts in order to communicate meaningfully. When two computer systems exchange messages only those concepts which are known to both systems can be received and processed by the other system. This is illustrated in figure 3.

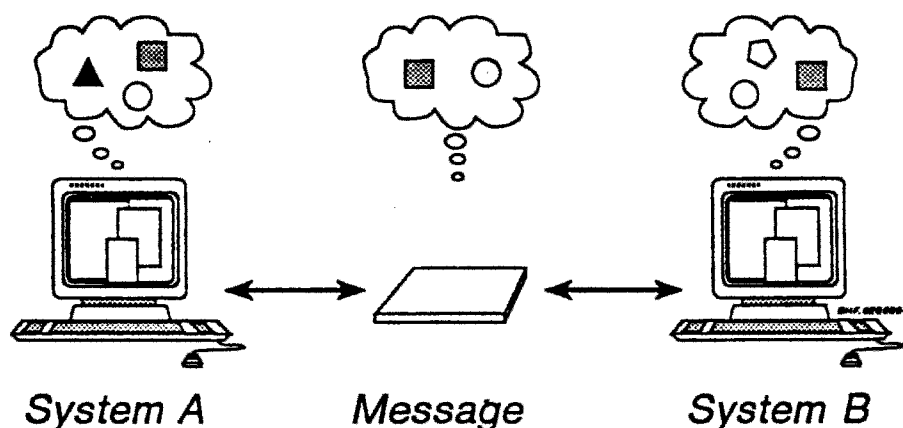


Figure 3: Only shared concepts can be exchanged in messages

A conceptual model for the problem domain should be created. The conceptual model is called the **Domain Information Model (DIM)**. The DIM consists of object types with attributes and operations which relate to things and concepts in the real world. The DIM describes the individual objects and a structure between the objects. The DIM is a static model for the problem domain independent of any particular service or message specification. It serves as a basis for creating GMDs.

When creating a DIM we need to have in mind that the model is primarily to be used for information exchange between applications, not to build specific systems. Only those concepts in the problem domain which need to be shared by the communication parties should be included in the model. Information related to the internal processes in an application program should be omitted. Common concepts in health care may be reused in several problem domains, but the exact attributes and operations may differ from domain to domain.

The DIM must be accompanied by a detailed definition of each object type, its attributes and operations.

In this technical report the OOA technique from Coad and Yourdon [Coad-91] has been chosen. Other object-oriented techniques may as well be used to create DIMs.

2.1.6. General Message Descriptions (GMDs)

A message incorporates a selected amount of information which is exchanged between two systems for a given purpose e.g. a request for a service or a report from a performed service. The actual information elements in a message will vary from time to time depending on the context and the purpose of the message exchange, but all information elements will be found in the DIM. For example, the information needed in a request for a blood test when the specimen is drawn by the requesting doctor, differs from the information needed when the specimen should be taken by the laboratory. The resulting messages will include a different set of information elements.

In the Electronic Data Interchange (EDI) world, messages were initially developed as discrete entities linked only to supporting data directories. For example in the EDIFACT standardisation process, the message types and the support directories are candidates for standardisation rather than the domain information model.

This thinking has its parallel in the paper world where forms have been subject to standardisation. However, focusing only on forms and messages may lead to inconsistent use of concepts between messages which again causes problems when implementing the message descriptions in applications. Increasingly EDIFACT messages are being developed using modelling techniques.

In order to bridge our conceptual model (the DIM) with service specifications and the traditional EDI messages we have introduced an intermediate step called a General Message Description (GMD). GMDs may be seen as a special view of the overall DIM reflecting one message type. GMDs use the same modelling technique as the DIM and are IF independent.

2.2. Mapping to Interchange Format Dependent Message Descriptions

The GMDs cannot themselves be transported between systems. Before information can be exchanged, a description of an electronic message must exist. This description has to be unambiguous and be written in a precise way. PT 004 has mapped the GMDs into IF Dependent Message Descriptions (IFDMDs). The technique used for this mapping varies between IFs. For each object and attribute at the GMD level a construction at the IF level must be found. This is illustrated in figure 4.

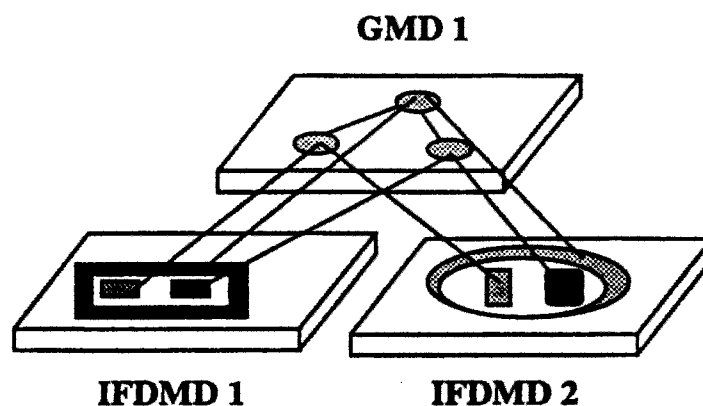


Figure 4: Mapping from a General Message Description (GMD) to two IF Dependent Message Descriptions (IFDMD)

The GMD does not include technical design details. As an example the GMD does not say how an association between object types is implemented, e.g. as a one way or two way pointer. Design details are added during the mapping to a specific IF.

3. Functional Requirements for Interchange Formats

The objective of an IF is to carry the information structure and content of a message across a data network. The IF consists of a set of construction elements and a syntax which describes the rules for combining the construction elements. The syntax and the construction elements form the functionality of the IF.

The goal is to compare the functionality of a number of IFs. Before setting up concrete evaluation criteria those aspects of an IF which we assume influence its functionality will be categorised. The functional requirements for IFs will vary between health care domains.

3.1. Support of Information Structures

In this framework for health care messaging, a GMD describes the information structure and elements of a message. The GMD is not related to any particular IF. Before an instance of a message can be constructed and exchanged electronically, a GMD has to be transformed into an IF dependent description. How easily this mapping can be achieved depends on the ability of the IF to support the information structure of the GMD. A number of common ways to structure information are identified later in chapter 5 where the evaluation criteria are described.

3.2. Support of Datatypes

Another important aspect of the IF is its ability to support different data types. Different health care domains will require support for different data types such as floating numbers, text strings, images and sound. The representation of data types are often related to a specific technology, e.g. the representation of images. The choice of a technology for example to process an image may then influence the choice of IFs.

Some IFs support many data types. Others do not support many data types themselves but are able to encapsulate other IFs.

3.3. Encoding

In the Open System Interconnection (OSI) framework the Presentation Layer is concerned with the representation of information which is communicated between applications. It transforms data represented in the local application syntax into a transfer syntax. The transfer syntax describes the way the data are actually represented in terms of bit patterns while in transit. The transformation to and from a transfer syntax is called encoding and decoding.

The bit patterns of a transfer syntax are determined by an abstract syntax and a set of encoding rules. The abstract syntax defines a formal logical structure of a message without content. The encoding rules specify the transformation from an abstract syntax to a transfer syntax.

The abstract syntax may be viewed as an intermediate step between a logical GMD model and the transfer syntax. In order to be a complete abstract syntax, all details of the message type have to be specified, so that it can be encoded directly into a transfer syntax by applying

encoding rules. One abstract syntax may correspond with multiple transfer syntaxes, as different encoding rules are applied.

An IF may have both an abstract and a transfer syntax. However some IFs do not distinguish clearly between the abstract syntax and the transfer syntax. The encoding rules influence the functionality of an IF. Some transfer syntaxes use character based encoding for human readability under debugging, while others use bit-level encoding for transmission efficiency.

3.4. Evolution and Backwards Compatibility

Integration of health care applications is in an early stage. The application domains in health care are likely to evolve in the short term. A framework for developing standards for integrating health care applications needs to permit both evolution and, as far as possible, provide for backward compatibility as the standards evolve.

In this approach, both the DIM and the service specifications may evolve. An example is the definition of national supersets of GMDs. This restrains the GMDs' and the IFs' ability to support new features. In the DIM and GMDs, evolution and backwards compatibility are solved by inheritance and polymorphism. Inheritance allows a new object to inherit the attributes of its predecessor while adding new features, while polymorphism allows an object to act like one of its predecessors.

There is a requirement that the IFs can support updates in the GMDs both at the object class and attribute level if the evolution of the DIMs and GMDs should not be restricted by the mapping from GMDs to IFDMDs. This requires a mechanism for defining supersets of IFDMDs rather than subsets, which is the method commonly used today.

3.5. Conformance and Certification

When dealing with multiple vendors, each claiming to be conform to a particular standard, the user is faced with the need to verify the vendors' claims. The user's task is considerably simplified if the standard defines what is required to conform to the standard and there is an independent agency which certifies each vendor's claim of conformance to the standard.

It should be possible to conformance test both the syntax of an IF and messages described in an IF. ISO 9646 is a five part standard which defines the requirements for specification of conformance.

3.6. Support and Availability

Before choosing a health care system, the user wants to make sure his investment is sensible. A number of factors should be considered. In order to communicate with other health care parties and with external agencies with whom information needs to be exchanged, systems should use widely accepted standards. International standards are developed by international organisations (principally ISO, IEC and CCITT) based on consensus at an international level. In many countries there are legal requirements for international standards to be used. For IF standards the developing organisation may also provide a mechanism for registration of message types and message components.

Quality software which supports the health care domain and the IF, should be available from a range of vendors before a standard will become in widespread use. There is also a need for people with appropriate skills related to the standard.

4. Investigation Method

The investigation of syntaxes for IFs has been divided into two parts: the establishment of evaluation criteria and the evaluation itself. The method is illustrated in figure 5.

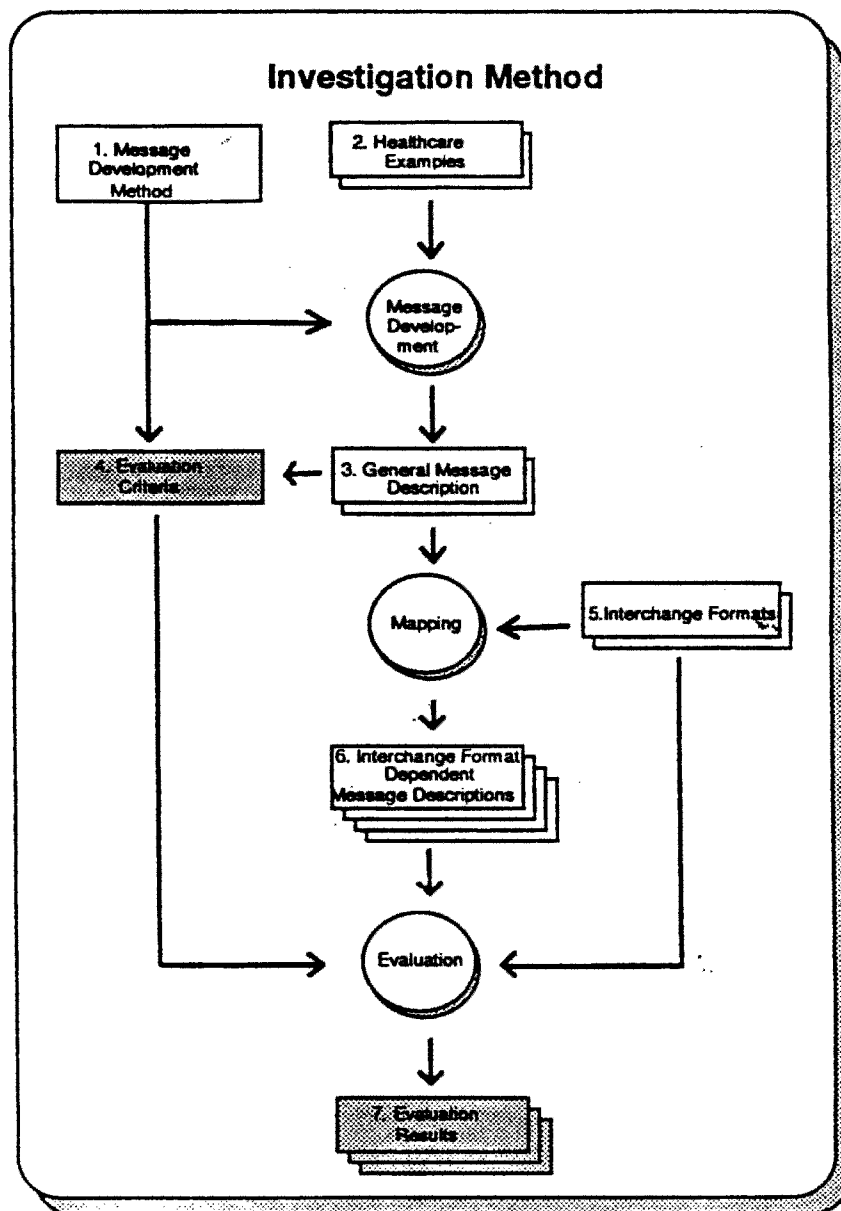


Figure 5: The investigation method defined in this report

In the figure, circles mean activities, boxes mean descriptions. The white boxes are input descriptions to the investigation process. Grey boxes represent important results of the investigation process. There are two results boxes, the evaluation criteria from part one of the investigation and the evaluation results from part two.

In part one our objective was to establish suitable evaluation criteria. The process started with a discussion about what an IF and a syntax is. Further the required functionality of IFs was addressed.

The purpose of this report is to assist in the development of health care messages. Therefore the evaluation criteria were selected with this aim in mind. Similarly a simplified method for message development was established. The simplified message development method and the functional requirements for IFs have been documented in chapter 2 and 3 of this report.

The message development method was used on a limited number of representative examples of health care domains to create General Message Descriptions. Two examples from two different health care domains were selected:

- A. Laboratory communication.
This example should only require messages with simple textual data types.
- B. Internal medicine transfer folder.
This example will require message with images and other non-textual data types.

The result of the message development for the two examples are documented in part II, chapter 9 and 10 of this report. Three GMDs have been established, two related to the laboratory example and one related to the internal medicine example.

Based on the GMDs and the functional requirements for IFs, a set of concrete evaluation criteria were established. These are described in detail in chapter 5.

The next step was to select a few IFs. This report is restricted to five IFs from a list of 23 candidates. Chapter 5 includes a description of the selected IFs.

The GMDs were then mapped into IFDMDs. Only the GMDs of the laboratory example were mapped in this report but additional mappings could be undertaken using the method defined herein. During this mapping process the evaluation criteria were discussed, new criteria were identified, old ones were modified, and some even disappeared. The IFDMDs for each of the IFs are shown in chapter 11 to 15.

The resulting evaluation criteria were used to evaluate each IF. The result of the evaluation is documented with comments in chapter 11 to 15. Chapter 5 summarises the results.

5. Evaluation Criteria

5.1. Scope of Evaluation Criteria

Only evaluation criteria that relate particularly to the syntax of an IF have been considered. Since no formal mapping rules from GMDs to the semantic level of IFDMDs are established today, the mapping rules used in order to create the IFDMDs have not been evaluated.

There has been no evaluation of the semantic aspect i.e. the meaning or content of messages for any of the IFs. Such evaluations are outside the scope of this report.

A number of potential evaluation criteria were discussed. Possible criteria were taken from the terms of reference (see chapter 1.2). Some of these criteria were related to implementation e.g. performance, implementation complexity, efficiency, cost of use, etc. An evaluation based on certain would require prototypes or products for all the IFs tested under equal conditions. This was considered beyond the scope of this report.

Priority has been given to those evaluation criteria which relate directly to the functionality of IFs. This choice was made in order to be able to determine a future strategy for selecting IFs to be used in health care which facilitates the development of further messages.

In chapter 3 the role of the IF was discussed. The functional categories of the IF has been used to categorise the concrete evaluation criteria in this chapter. Six evaluation categories are used:

- A. Supported information structure
- B. Supported data types
- C. Encoding rules
- D. Evolution and Backwards Compatibility
- E. Conformance
- F. Support and availability

In evaluating an IF, users should consider the issues of support and availability described in chapter 3.6. Evaluation criterion of category F are used as a "catch-all" for these issues not examined in detail in this report.

Chapter 5.5 describes the evaluation criteria in detail.

5.2. Domain Dependencies

Not all the functions for an IF are of equal importance for all health care domains. In this report it is indicated for each evaluation criterion in which domain it applies. An evaluation criterion may be applicable to:

- none of the examples in this report
- the laboratory example
- the internal medicine example

- both the laboratory and the internal medicine examples
- to all domains of health care

5.3. Classified Evaluation Results

The IFs offer different functionality. The evaluation criteria relate mainly to various functional aspects of an IF. The objective of this report is to see which IFs offer what functionality. The evaluation results have been classified in terms of:

1. Function supported
2. Function not directly supported.

A comment should be made regarding results of type 2. In some cases, missing functionality on the IF's syntactical level may be compensated by the application itself. Standards like ASTM, EDIFACT etc. comprise more than a syntax. In these standards, functions that can not be implemented at the syntax level are often solved at the semantic level, e.g. the indication of relationships by the use of semantic references.

Solving problems at the 'higher' semantic level makes it more difficult to verify or validate issues at a generic, syntactical level and Sometimes puts a greater burden on the application. It should be stressed here that the evaluation in this report only deals with the syntactical and not the semantic level.

5.4. Excluded Evaluation Criteria

After overall consideration, a number of evaluation criteria were excluded from the evaluation.

One of this was the cost of products and implementations. Making cost an evaluation criterion would require both a detailed analysis of implementation methods related to each IF and market research. Market analysis will give the price of products at a given point in time. Future price development, even in a short time scale, will be difficult to predict. This technical report should stand the test of time, it should have a life time of several years. An evaluation mainly based on today's market, before real standards exist for health care information exchange, will be outdated soon.

A second criterion which was excluded was the issue of efficiency and overhead related to IFs. The overhead a syntax imposes on the data usually depends on the actual content of a message. Some IFs are more efficient when a lot of optional data elements are not present, whereas for others the opposite is true. Some IFs are more efficient for some data types than for others. e.g. efficiency differs considerably when large amounts of native binary data first have to be converted into an allowed character set rather than being transmitted directly.

A third criterion excluded is all matters related to security. Security may be handled on different levels, e.g. information element level or message handling system level. Users will need to assess the approach they wish to follow.

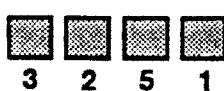
5.5. Description of the Evaluation Criteria

5.5.1. Supported Information Structures

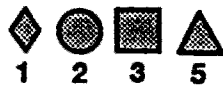
In this report, eleven concepts for structuring information have been selected. They are well known in information modelling and in common use when developing IT systems today, although the terms used for each concept may differ from place to place. In the following description, the term element refers to an instance of an information element that could be either an object instance or an attribute instance. The term type refers to an information element type that could be either an object type or an attribute type.

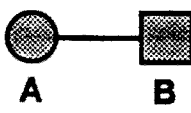
A.1	Optional
	Optional elements are elements of a type that do not have to be presented. Optionality is also used for relationships between object types.
Example	In DIM A, optionality is used in object Patient for the attribute Nationality. This means that a patient object can exist without using the Nationality attribute.
Required in	Laboratory + internal medicine examples

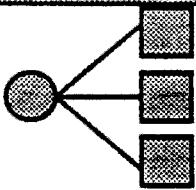
A.2	Choice
	A choice is a selection of one and only one type out of a group of type candidates.
Example	Choice is not used directly in any of the DIMs or GMDs. In DIM A the use of inheritance substitutes the use of choice. Patient Residential Address is modelled as an 'opt list of Address'. Address is an object type which is subtyped into Location Address and Post Address. Each instance of Address could be either one of these subtypes. This is an implicit choice.
Required in	None of the examples

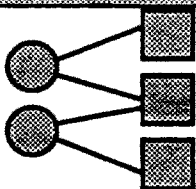
A.3	Repetition (set of)
	Repetition is a collection of elements of the same type whose order is insignificant.
Example	In DIM A the list concept is used to define the attribute Residential Addresses that may include none, one or many Addresses. List is used together with optional.
Required in	Laboratory + internal medicine examples

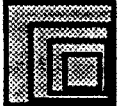
A.4	Unordered sequence (set)
	An unordered sequence is a collection of elements of possible different types whose order is insignificant.
Example	In DIM A the definition of attributes in object types may be seen as an unordered sequence. Each attribute is defined by its local name and not by its position in the object type. An example is the Patient object type.
Required in	Laboratory + internal medicine examples

A.5	Ordered sequence
	An ordered sequence is a collection of elements of the same type or different types whose order is important.
Example	There are no examples in the DIMs or GMDs. If the Name attribute of a Patient had been divided into components, the order of the components could have been used as the sorting criterion for the name. Another example is the order of pictures in a film.
Required in	None of the examples

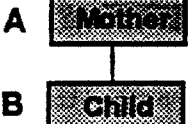
A.6	1-1 relationship
	Cardinality refers to the numbers of elements possible in a relationship between types. The simplest relationship between type A, e.g. a circle and type B, e.g. a square is a one-to-one (1-1) relationship, where one element of type A is related to only one element of type B and vice versa.
Example	In GMD A-1, a 1-1 relationship is used between Request and Patient to illustrate that in one message only one Patient exists per Request.
Required in	Laboratory + internal medicine examples

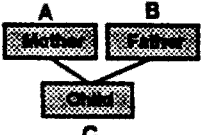
A.7	1-n relationship
 A B	In a one-to-many relationship an element of type A can be related to one or more elements of type B, but every element of type B must only be related to one element of type A.
Example	In DIM A, a 1-n relationship is used between Request and Request Element. This means that one Request may include one or more Request Elements, but one Request Element can only be used in one Request.
Required in	Laboratory + internal medicine examples

A.8	m-n relationship
 A B	In a many-to-many relationship, an element of type A can be related to one or more elements of type B and, in addition, elements of type B can be related to one and more elements of type A.
Example	In DIM A, an m-n relationship is used between Requested Test and Specimen. This illustrates the situation where one test needs several specimens to be performed and one of the specimens (glucose tolerance) is included in another test.
Required in	Laboratory + internal medicine examples

A.9	Recursive structure
	In recursive structures an element of a type contains another element of the same type directly or through another type. The result is a structure with nested elements.
Example	No recursive structure is defined in the DIMs or GMDs. However, in DIM A, the Report Element containing another Report Element is a recursive structure. This means that the composite result in a laboratory report includes several single results.
Required in	None of the examples

Generalisation or specialisation is a very powerful tool for information modelling.
Specialisation can be implemented in a IF by supporting inheritance.

A.10	Single inheritance
	In the case of single inheritance, a subtype B inherits all properties of the supertype A and in addition adds its own properties. Properties of object type may be attributes and/or operations.
Example	In DIM A single inheritance is used to specialise a Request Element into a Requested Test. The Requested Test will cover the attribute Priority defined in Requested Element in addition to its own.
Required in	Laboratory + internal medicine examples

A.11	Multiple inheritance
	In the case of multiple inheritance, a subclass B inherits the properties of two or more superclasses (A,B) and in addition adds its own properties
Example	In DIM A Requested Specimen is a subtype of both Requested Element and Specimen. The new type inherits both the properties of a Request Element and a Specimen.
Required in	Laboratory example

5.5.2. Support of Datatypes

The selection of datatypes as evaluation criteria is based on the examples defined above and a selection of other datatypes that are assumed to be important in health care. Eleven evaluation criteria are identified in this category.

B.1	Encapsulation of other IFs
	Encapsulation of other IFs means that a message based on one IF, can contain data based on another IF.
Remark	Encapsulation should be used with care. In order to process encapsulated data in a message, a computer system has to support all encapsulated IFs. When for instance ASTM, HL7 and EDIFACT are encapsulated in ASN.1 the system has to support ASN.1, ASTM, HL7 and EDIFACT. Encapsulation therefore should be confined to IFs for datatypes that are not covered by the original IF.
Example	Examples of encapsulation are shown in the internal medicine example in chapter 10. The Internal Medicine Transfer Folder (IMTF) contains three objects: a angiography image, a ultrasound image and a ECG which would probably require special IFs. The images require an IF that supports continuous-tone still images. The ECG require an IF that supports wave forms. The resulting IFDMD for the IMTF would contain three different IFs, the continuous-tone IF, the wave form IF and the IF encapsulating these other two IFs.
Required in	Internal Medicine example

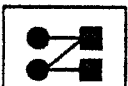
B.2	ASCII text [ISO 646]
	The American Standard Code for Information Interchange (ASCII) as defined in ISO 646 is a 7-bit character set which does not support the special characters of Western European languages, such as the French ç or German ß.
Example	ASCII text is not used in our examples, but ASCII is a subset of the ISO 8 bit character set described in B.3
Required in	Laboratory + internal medicine examples

B.3	Latin alphabet No. 1 [ISO 8859-1]
	The Latin alphabet No. 1 is an 8-bit character set, which is an extension of ASCII. It supports most Western European languages.
Example	The Latin alphabet No. 1 is used for all ISO-STRING elements in the DIMs and GMDs
Required in	Laboratory and internal medicine examples

B.4	Extended character set
	There are a lot of different characters and symbols in use in the world. An 8-bit structure can only encode 256 different characters. Larger structures are proposed, like the ISO 10646 standard, which has a 32-bit structure, which is compatible with ASCII and supports Asiatic languages, or the Unicode standard of a consortium of American corporations which is based on 16 bits. In order to be able to support extended character sets like those mentioned above, the syntax should allow the use of structures larger than 8-bits.
Example	Not used in the DIMs or GMDs in this report.
Required in	None of the examples

B.5	Layout support
	Layout support has to do with the way information is presented to the reader. It can range from simple features like boldface or underlined to the possibility completely to control the way information is presented using different types and sizes of fonts, sizes of margins, footnotes, colour etc. Using an analogy to conventional paper, it can range from a simple typewriter letter to a full coloured glossy magazine. For this evaluation an IF has to support more than the simple features (fonts) in order to be qualified as supporting layout. The support of colour will be indicated in a footnote.
Example	Layout is not used in the DIMs or GMDs in this report.
Required in	None of the examples

B.6	Raster graphics
	Images can be coded in two ways, i.e. as an ordered sequence of pixels (bit mapped or raster graphics) or as a collection of geometrical objects like lines, curves, circles etc. (see next evaluation criterion). A fax is an example of raster graphics.
Example	Raster graphics is not used in the DIMs or GMDs in this report.
Required in	None of the examples

B.7	Geometrical graphics
	Geometrical graphics are built up from geometrical objects like lines, curves, circles etc.
Example	A chart consisting of lines, areas etc. can be coded as geometrical graphics. Geometrical graphics is not used in the DIMs or GMDs in this report.
Required in	None of the examples

B.8	Continuous-tone still image
	A pixel can be coded by one bit per pixel, which means it can be either black or white. Grey scales can be emulated by patterns of black and white pixels (bit mapped). Most raster graphic standards like those used for fax cannot reproduce pictures at a high quality. Continuous tone still images are more sophisticated raster graphics in which grey scales and/or colour can be encoded directly. This requires several bits per pixel and such images take up a lot of space and therefore compression techniques to reduce them are very important.
Example	A typical example of continuous tone still images are X-rays. In DIM B both the ultrasound and angiography images may be implemented as continuous tone still images.
Required in	Internal medicine example

B.9	Moving image
	Moving images consist of (ordered) sequences of still images. Whereas still images already take up a lot of space, moving pictures take even more. Therefore compression techniques which utilise the interdependencies between the individual still pictures which make up the moving image are of the utmost importance.
Example	Moving images are not used in the DIMs or GMDs in this report.
Required in	None of the examples

B.10	Wave form
	A wave form is a sequence of values that represent a curve in time.
Example	In GMD B-1 the ECG may be represented as a wave form.
Required in	Internal medicine example

B.11	Sound
	Sound can be digitised by sampling the analogue signal and encoding it. As is the case for images, data compression is very important to limit the data space needed.
Example	Sound is not used in the DIMs and GMDs in this report.
Required in	None of the examples

B12	Coded datatype
X12345QW	The meaning of codes lies on the semantic level and can therefore never be checked at the syntax level, but a check on datatypes validation of codes is possible, e.g. when a certain attribute can take only the values 1, 2 and 3 as legal codes then it is obvious that when a message contains the value 4 for this attribute there is something wrong. A syntax can have features which make this validation at the syntax level possible.
Example	In DIM A the attribute Sex in object type Patient is enumerated with the values Unknown, Male, Female and Unspecified [ISO 5218].
Required in	Laboratory + internal medicine examples

B13	Undefined binary datatype
101110101	Applications may want to include data in an IF which are unstructured and binary. The IF should be able to carry undefined application specific data.
Example	Binary datatype is not used in the DIMs and GMDs in this report
Required in	None of the examples

5.5.3. Encoding

When evaluating the encoding the only concern is whether the encoding rules may restrict the IF's ability to transfer information types and structures. The encoding rules can impose constraints on the interchange of data. For example binary data and therefore encapsulation is not possible when a syntax uses delimiters for separating data without some length indication. There is always the possibility that the binary data contains bit sequences which are interpreted as being delimiters. In order to be able to exchange these data it first has to be converted so that it contains only allowed bit sequences. Several methods can be used to achieve this (e.g. using release or escape sequences), but they will always make the message longer.

C1	Bit-oriented transfer syntax
	A bit-oriented transfer syntax has a larger field of use than one which is character based and may be required for some domains.
Required in	Internal medicine example

C.2	Variable data length
	The encoding of an IF can also impose constraints on the length of data. Fixed data lengths are generally inefficient. Most IFs therefore support variable data length but sometimes the maximum length is limited.
Required in	Laboratory + internal medicine examples

5.5.4. Evolution and Backwards Compatibility

An IFs' ability to support evolution and backwards compatibility will be evaluated on two levels, by adding new attributes to existing object types and by adding new object types to a GMD.

D.1	Message version handling
	In order to maintain compatibility and to process messages correctly there has to be an efficient mechanism for dealing with different versions of a message type.
Required in	All health care domains

D.2	Adding new attributes to existing object types to a GMD
	A new attribute is added to an old object type in a GMD. The question is whether a new message specification for the IF can easily be generated based on the updated GMD and whether this message is backwards compatible with the previous message specification.
Example	In DIM A a new attribute Title may be added to the object type Patient.
Required in	All health care domains

D.3	Adding new object types to a GMD
	A new object type may be added to a GMD. The question is whether a new message specification for the IF can easily be generated based on the updated GMD and whether this message is backwards compatible with the previous message specification.
Example	In DIM A a new object type Medication may be added as a part object to Clinical Information.
Required in	All health care domains

5.5.5. Conformance

The ability to conform to an IF can be separated into two levels. First the implementation must process the syntax of the IF correctly. Secondly the message described in the IF must be handled correctly. In order to be able to test a message specification, the syntax needs to use a formal method for describing messages.

E.1	General syntax conformance
	Is it possible to verify the IF's transfer syntax ?
Required in	All health care domains

E.2	Message syntax conformance
	Does the IF have the possibility to define unambiguously a message through a formal specification ?
Required in	All health care domains

5.5.6. Support and Availability

The following seven criteria have been selected to evaluate an IF's support and availability.

F.1	International standard
	An important aspect of a IF is whether or not it is an international standard. International standards are likely to have more support than, for instance, local ones.
Required in	All health care domains

F.2	Used in health care
	This aspect deals with the support the IF has in the health care sector. If for example everybody already uses a specific syntax which is not an international standard, then it is likely to be very difficult to introduce a new one, even if it is a standard.
Required in	All health care domains

F.3	Used in other sectors
	Economies of scale are important with respect to price/performance ratios for applications. When a syntax is used in other sectors it is most likely that the user base is larger then when it is confined only to the health care sector. It is also an advantage using an IF which is used in other sectors with which health care has to exchange information.
Required in	All health care domains

F.4	Available off the shelf software tools
	Developing applications is of course much easier when good tools are easily available, than when they have to be developed from scratch.
Required in	All health care domains

F.5	Available skill on the IF
	In order to be able to develop software based on a specific IF, persons with knowledge and skill in it have to be available.
Required in	All health care domains

F.6	Registration mechanism for message type
	Does there exist an authority and a procedure where registration of health care messages will be possible.
Required in	All health care domains

6. Selected Interchange Formats

During the last twenty years or so a large number of IFs have been developed and adopted for use, to meet the diverse needs of the health care community. Many standards have been developed to meet the wider needs of telecommunications and the computer industry, for example ASN.1, ODA and FTAM. EDIFACT has its origins in international trading. Some, ACR/NEMA (now DICOM) have arisen from specialised applications such as radiology. Others have evolved to meet the needs of groups of users and suppliers, such as HL7. More recently specific proposals have been developed by collaborative research programmes, such as the AIM project EUCLIDES.

A comprehensive list of all the known IFs irrespective of the origin or status, was compiled. In an attempt to classify the IFs two broad categories were identified, depending on whether the need and development primarily came from health care or from the information technology industry. The IFs were then given a priority for further investigation. The list is shown in table 1.

Table 1: List of considered IFs

Interchange Format	Description	Origin	Priority
AAPM	The AAPM standard was established in 1982 by the American Association of Physicists in Medicine for the interchange of images by means of magnetic tape.	HC	4
ACR-NEMA/DICOM [ACR-NEMA-88]	The American College of Radiology and the National Electrical Manufacturers' Association compiled a standard for the interchange of images and the corresponding data between devices of different suppliers. The original standard was for point-to-point communications. Version 3 of the standard, called 'Digital Imaging and Communication in Medicine (DICOM) standard, specifies OSI compatibility.	HC	2
ANS X12	The X12 standard, developed by the American National Standards Institute (ANSI) is an American equivalent to the EDIFACT standard.	IT	2
ASN.1/BER [ISO 8824], [ISO 8825]	Abstract Syntax Notation One is the notation for abstract syntax definition used within the OSI framework. The abstract syntaxes defined in ASN.1 may be encoded in different transfer syntaxes using different encoding rules. The Basic Encoding Rules (BER) defines the most commonly used encoding rules used with ASN.1. See chapter 6.1.1 for more information.	IT	1
ASTM E 1238 [ASTM-91]	The ASTM E 1238 specification covers the two-way digital transmission of requests for, and results of, tests, diagnostic studies and care provider observations between requesters (e.g. hospital information systems) and producers (e.g. clinical laboratory systems). See chapter 6.1.2 for more information.	HC	1

ASTM E1467 [ASTM-92]	The ASTM E1467 specification of the American Society for Testing and Materials (ASTM) covers the transmission of digitally recorded electrophysiological wave forms data (EEG, MEG, ERP, etc.) and related textual annotations between computer systems.	HC	4
CMIS/CMIP [ISO 9595] [ISO 9596]	Common Management Information Service (CMIS) and Protocol (CMIP) are the names of the OSI Management service and protocol standards. These standards are used to transport and manage data networks. CMIS may also be used for ordinary application-to-application communication. CMIS is currently the only OSI standard which support an object-oriented description of information. IEEE P1157 MEDIX has described a framework where CMIS is used for communication between tightly coupled health care applications.	IT	2
COST B2/INTERFILE	In the framework of the COST B2 project, established at the initiative of a number of European associations for nuclear medicine, a standard protocol has been developed for the interchange of images in nuclear medicine. The standard is an extension of the AAPM standard.	HC	3
EDIFACT [ISO 9735]	EDIFACT stands for Electronic Data Interchange For Administration, Commerce and Transport and was defined by the United Nations Economic Commission for Europe for the exchange of structured, character based data. See chapter 6.1.3 for more information.	IT	1
EUCLIDES [EUCLIDES-92]	EUCLIDES was developed under an AIM project for the exchange of clinical laboratory data between independent computer systems. See chapter 6.1.4 for more information.	HC	1
FTAM [ISO 8571]	FTAM stands for File Transfer, Access and Management and has been developed within the OSI framework. FTAM permits the transfer and manipulation of files.	IT	2
HL7 [HL7-90]	HL7 stands for Health Level 7, referring to the highest level of the OSI reference model. However HL7 is not compliant with the OSI framework. The standard was developed by an ad hoc group of system vendors and users for the interchange of data in the health care sector. The group works closely together with ASTM and the intention is that both standards will merge.	HC	2 covered by ASTM E1238
IPI [ISO DIS 12087]	A project has been started by the Computer graphics Commission of the ISO to arrive at a standard for Image Processing and Interchange (IPI).	IT	4
JPEG [ISO DIS 10918]	The Joint Photographic Experts Group of JTC1/SC29/WG10 have defined a generally applicable compression standard and the encoding scheme for almost all continuous-tone still images.	IT	4
MEDIX [Harrington-90]	IEEE P1157 Medical Data Interchange (MEDIX) Committee has described a framework for information exchange in health care. The group is supported by IEEE and has in its scope the whole communication in health care. MEDIX does not define new IFs but bases itself on available international standards. The MEDIX framework uses EDIFACT, ODA, SGML and CMIP as vehicles for exchanging information.	IT	2 covered by EDIFACT, ODA, SGML and CMIS
MIB/MDDL [IEEE-90]	IEEE P1073 Medical Information Bus is a standard under development for communication with medical devices.	HC	3

MPEG [ISO DIS 11172]	MPEG is a ISO standard for moving images being developed by JTC1/SC29/WG11	IT	3
ODA/ODIF [ISO 8613]	Open Document Architecture (ODA) is an OSI-standard for the interchange of electronic documents. ODIF is its IF. See chapter 6.1.5 for more information	IT	1
PAPYRUS [Ratib-90]	Papyrus, which was developed by the University Hospital of Geneva, is based on the ACR/NEMA standard and describes a standard file format for the storage and interchange of radiology images.	HC	3
SCP-ECG [Willems-90]	Within the AIM SCP-ECG project a standard communication protocol has been developed for computerised electrocardiography.	HC	2
SGML [ISO 8879]	Standardized Generalised Mark-up Language (SGML) is an ISO standard for marking up documents.	IT	2
SPI [Siemens-87]	SPI is a proprietary standard based on ACR/NEMA developed by Philips and Siemens and is applied in commercial PACS-systems.	HC	4
SQL [ISO 9075]	The Structured Query Language (SQL) is a syntax for a querying relational databases.	IT	2

Consideration was given to whether the IF was relevant to health care. This was based upon its status - an international, national or industry standard and its present or potential use in health care. The IFs were then given four different priorities. Those IFs with highest priority where given a number 1, those with the lowest priority where given the number 4. The IFs with the highest priority (EDIFACT, ASN.1, ASTM 1238, EUCLIDES, ODA) were selected for a detailed study.

6.1. Description of Evaluated Interchange Formats

6.1.1. ASN.1

ASN.1 is a "meta-syntax" which is a general mechanism for defining syntaxes. It was first developed in order to define the X.400 standard (and was formalised as CCITT Recommendation X.409(1984)). This definition, which also contained the "Basic Encoding Rules" for encoding ASN.1 syntaxes for transmission over a communications link, was adopted by ISO as the pair of standards ISO 8824 and 8825, the former containing the specification of ASN.1 with the latter containing the encoding rules. More recently, additional encoding rules (packed, distinguished and lightweight) have been developed and the two standards are in the process of revision into multipart standards.

The ASN.1 notation is in widespread use in the OSI community, and virtually all OSI standards are defined in it.

6.1.2. ASTM E1238/HL 7

The ASTM E1238 Standard Specifications for transferring clinical observations between independent computer systems cover the exchange of requests for, and results of, tests, diagnostic study and care provider observations. Recently a new version (ASTM E1238-91) has been developed in close co-operation with HL7.