Appendix B: Format for data element definitions – ISO/IEC 11179-based standards

All data element definitions included in the *National Health Data Dictionary* are presented in a format based on ISO/IEC Standard 11179 *Specification and Standardization of Data Elements* – the international standard for defining data elements issued by the International Organization for Standardization and the International Electrotechnical Commission. Collectively, the format describes a set of attributes for data definitions. The set of attributes for data definitions used in the *National Health Data Dictionary* is described below.

Administrative status: The operational status (e.g. CURRENT or SUPERSEDED) of the data element or data element concept and the date from which this status is effective. For example, in the Dictionary the latest revision of 'Area of usual residence', effective from 1 July 1997, has a CURRENT status, replacing the previous version of this data element, operational from 1 July 1995 until 30 June 1997, which now has a SUPERSEDED status. No SUPERSEDED data elements are included in this hard copy publication of the Dictionary. However, all data elements, including SUPERSEDED data elements, are included on the Knowledgebase.

Knowledgebase ID: A six-digit number used to identify the data element on the Knowledgebase (previously known as the NHIK). In the Knowledgebase, this number is preceded by an acronym that identifies the registration authority for each data element. The National Health Information Management Group is the registration authority for all data elements included in the Dictionary. The combination of registration authority, Knowledgebase (or NHIK) ID and version number (see below) uniquely identifies each data element in the Knowledgebase.

Version number: A version number for each data element, beginning with 1 for the initial version of the data element, and 2, 3 etc. for each subsequent revision. This meets the ISO/IEC Standard 11179 requirement for 'identification of a data element specification in a series of evolving data element specifications within a registration authority'. A new version number is allocated to a data element or data element concept when changes have been made to one or more of the following attributes of the definition:

- name
- definition
- data domain.

Identifying and definitional attributes

Name: A single or multi-word designation assigned to a data element. This appears in the heading for each unique data definition in the Dictionary.

Data element type: A data element may be:

a. A DATA ELEMENT CONCEPT – a concept that can be represented in the form of a data element, described independently of any particular representation. For example, hospital 'admission' is a process which does not have any particular representation of its own, except through data elements such as 'admission date', 'mode of admission', etc.

b. A DATA ELEMENT – a unit of data for which the definition, identification, representation and permissible values are specified by means of a set of attributes. For example, a hospital 'admission date' is a unit of data for which the definition, identification, representation and permissible values are specified.

c. A DERIVED DATA ELEMENT – a data element for which values are derived by calculation from the values of other data elements. For example, the data element 'length of stay', which is derived by calculating the number of days from 'admission date' to 'separation date' less any 'total leave days'.

d. A COMPOSITE DATA ELEMENT – a data element the values of which represent a grouping of the values of other data elements in a specified order. For example, the data element 'establishment identifier' is a grouping of the data elements 'state identifier', 'establishment type', 'region' and 'establishment number' in that order.

Definition: A statement that expresses the essential nature of a data element and its differentiation from all other data elements.

Context: A designation or description of the application environment or discipline in which a name is applied or from which it originates. For example, the context of 'admission date' is 'admitted patients', while the context of 'capital expenditure – gross' is 'health expenditure'. For the dictionary this attribute may also include the justification for collecting the items and uses of the information.

Relational and representational attributes

Data type: The type of symbol, character or other designation used to represent a data element. Examples include integer, numeric, alphanumeric, etc. For example, the data type for 'intended place of birth' is a numeric drawn from a domain or code set in which numeric characters such as '1=hospital, 4=home' are used to denote a data domain value (see data domain below).

Representational form: Name or description of the form of representation for the data element, such as 'CODE', 'QUANTITATIVE VALUE', and 'DATE'. For example, the representational form for 'country of birth' is 'CODE' because the form of representation is individual numbers that each represent a different country.

Field size (minimum and maximum): The minimum and maximum number, respectively, of storage units (of the corresponding data type) used to represent the data element value. For example, a data element value expressed in dollars may require a minimum field size of one character (1) up to a maximum field size of nine characters (999, 999, 999). Field size does not generally include characters used to mark logical separations of values, e.g. commas, hyphens or slashes.

Representational layout: The layout of characters in data element values expressed by a character string representation. Examples include 'DDMMYYYY' for calendar dates, 'N' for one-digit numeric fields, and '$$$,$$$,$$$' for data elements about expenditure.

Data domain: The set of representations of permissible instances of the data element, according to the representation form, layout, data type and maximum size specified in the corresponding attributes. The set can be specified by name (including an existing classification/code scheme such as ICD-10-AM), by reference to a source (such as the *ABS Directory of concepts and standards for social, labour and demographic statistics*, 1995), or by enumeration of the representation of the instances (for example, for 'compensable status', values are '1=compensable, 2=non-compensable).

Guide for use (optional): Additional comments or advice on the interpretation or application of the attribute 'data domain'. (This attribute has no direct counterpart in the ISO/IEC Standard 11179 but has been included to assist in the clarification of issues relating to the classification of data elements.)

Verification rules (optional): The rules and/or instructions applied for validating and/or verifying data elements occurring in actual communication and/or databases, in addition to the formal screening based on the requirements laid down in the basic attributes.

Collection methods (optional): Comments and advice concerning the actual capture of data for the particular data element, including guidelines on the design of questions for use in collecting information, and the treatment of 'not stated' or non-response data. (This attribute is not specified in the ISO/IEC Standard 11179 but has been added to cover important issues about the actual collection of data.)

Related data (optional): A reference between the data element or data element concept and any related data element or data element concept in the Dictionary, including the type of their relationship. Examples include 'has been superseded by the data element…', 'is calculated using the data element…', and 'supplements the data element…'.

Administrative attributes

Source document (optional): The document from which definitional or representational attributes originate.

Source organisation: The organisation responsible for the source document and/or the development of the data definition (this attribute is not specified in the ISO/IEC Standard 11179 but has been added for completeness). The source organisation is not necessarily the organisation responsible for the ongoing development/maintenance of the data element definition.

National minimum data sets (optional): The name of any national minimum data set established under the auspices of the National Health Information Agreement (NHIA) which includes the particular data element. The date of effect is also included.

Comments (optional): Any additional explanatory remarks on the data element.