Data Services for Commissioners

NHS England

Consistent Identification Data Items

Supporting documentation

Consistent Identification Data Items to improve data quality, linkage and anonymous identification of patients

This guidance has been produced to supplement the '*National Commissioning Flows resource*'. It provides details of Consistent Identification Data Items for inclusion in collections of patient-level datasets which are submitted for national processing.

Ensuring that patients are identified consistently within commissioning data is crucial to data quality, observing patient preferences for data sharing and for meeting the health needs of all demographic groups. Where patient identifiers are legitimately withheld from provider flows or de-identified in flows to commissioners, it is still important that certain data based on identifiers are available to commissioners to inform their decision-making.

The Consistent Identification Data Items (CIDIs) proposal supports these objectives by:

1. Enabling unique individuals to be consistently de-identified within commissioning datasets
2. Enabling national systems to derive useful data based on date of birth and postcode and flow these derivations to those commissioners or support organisations who do not have a legal basis to receive identifiable data to perform the derivations themselves
3. Enabling the flow of derivations based on postcode and date of birth when providers are required to withhold identifiable data from their outbound data flows for legal and other reasons

Commissioners are consequently better informed and more able to make better commissioning decisions about the services that patients need. Such an approach delivers the benefit of improved health outcomes for all patients.

Consistent Identification Data Items are common fields that exist within the majority of national data flows. There is no requirement to define any new elements, as all items are currently described in the NHS Data Model and Dictionary.

Within local flows, there is no necessity to use fields from the NHS Data Model and Dictionary and the terminology described within it. However, Consistent Identification Data Items are commonly in use in identifiable patient-level data flows, albeit using bespoke terminology. Mapping these items from the local field name to the appropriate specification defined here will be suitable and possible using application of NHS Digital's data landing portal (where the definition is the same).

The Consistent Identification Data Items to be submitted are detailed in Table 1 below.

Table 1: Consistent Identification Data Items

ID	Data Item Name / NHS Data Dictionary Name	Data Item Description	Format	M (Mandatory) or M* (Mandatory with caveats, see Validation for details)	Validation	Rationale
1	**NHS NUMBER**	The NHS NUMBER, the primary identifier of a PERSON, is a unique identifier for a PATIENT within the NHS in England and Wales.	n10	M*	Mandatory unless: 1. Concerns a sensitive record 2. If the number is unavailable 3. Provider has no access to PDS and, thus, should submit name and address instead.	Used as a unique patient identifier and as a potential part of the data linkage within and between different datasets.
2	**NHS NUMBER STATUS INDICATOR CODE**	The NHS NUMBER STATUS INDICATOR of the PATIENT.	an2	M	Mandatory where the NHS Number provided is blank to ensure proper linkage.	Highlights the status of the NHS number.
3	**PATIENT NAME - PERSON NAME STRUCTURED** **or** **PATIENT NAME - PERSON NAME UNSTRUCTURED**	PATIENT NAME is the PERSON NAME where the PERSON NAME CLASSIFICATION is 'Preferred Name' of the PATIENT.	max an70	M*	Mandatory unless NHS NUMBER is populated or the record is sensitive, in line with existing Commissioning Data Set (CDS) guidance.	A potential part of the data linkage within and between different datasets.
4	**PATIENT USUAL ADDRESS - ADDRESS STRUCTURED** **or** **PATIENT USUAL ADDRESS - ADDRESS UNSTRUCTURED**	The usual address nominated by the patient, where the address association type is ‘main permanent residence’ or ‘other permanent residence’.	max an175	M*	Mandatory unless NHS NUMBER is populated or the record is sensitive, in line with existing CDS guidance.	A potential part of the data linkage within and between different datasets.

ID	Data Item Name / NHS Data Dictionary Name	Data Item Description	Format	M (Mandatory) or M* (Mandatory with caveats, see Validation for details)	Validation	Rationale
5	**WITHHELD IDENTITY REASON**	Allows suppliers of records to indicate to recipients of the record that the record has been purposely anonymised for a valid reason.	an2	M	Permitted national code in the national format.	Highlights where a record has been anonymised.
6	**ACTIVITY IDENTIFIER**	A unique number or set of characters that is applicable to only one ACTIVITY for a PATIENT within an ORGANISATION.	an12	M	Must be in the national format.	Used as a unique identifier for activity (by provider/data set) to link between national and local data.
7	**LOCAL PATIENT IDENTIFIER**	This is a number used to identify a PATIENT uniquely within a Health Care Provider. It may be different from the PATIENT's case note number and may be assigned automatically by the computer system.	an10	M	Must be in the correct format.	Used as alternative unique patient identifier, in addition to NHS number.
8	**AGE AT ACTIVITY DATE**	The number of completed years between the person birth date of the patient and the activity date.	n3	M*	Mandatory unless PERSON BIRTH DATE is not blank. Must be in the national format.	Enables age to flow for sensitive records.

ID	Data Item Name / NHS Data Dictionary Name	Data Item Description	Format	M (Mandatory) or M* (Mandatory with caveats, see Validation for details)	Validation	Rationale
9	**PERSON BIRTH DATE**	The date on which a PERSON was born or is officially deemed to have been born.	an10 CCYY-MM-DD	M*	Mandatory unless the record is sensitive. Must be in the national format. NOT ALLOWED Where WITHHELD IDENTITY REASON is populated.	Used in combination with other date fields (for example dates) to derive age. Also a potential part of the data linkage within and between different datasets.
10	**POSTCODE OF USUAL ADDRESS**	The POSTCODE of the ADDRESS nominated by the PATIENT with ADDRESS ASSOCIATION TYPE 'Main Permanent Residence' or 'Other Permanent Residence'.	max an8	M*	Mandatory unless the record is sensitive. NOT ALLOWED Where WITHHELD IDENTITY REASON is populated.	Used to derive location based data items (e.g. LSOA and resident CCG). Also a potential part of the data linkage within and between different datasets.
11	**PERSON STATED GENDER CODE**	PERSON STATED GENDER CODE is self-declared or inferred by observation for those unable to declare their PERSON STATED GENDER.	an1	M	Must be one of the defined national codes in the national format.	Will enable analyses to explore potential health inequalities. Also a potential part of the data linkage within and between different datasets.
12	**GENERAL MEDICAL PRACTICE CODE (PATIENT**	The ORGANISATION CODE of the GP Practice that the	an6	M	Code must be a current live organisation in the	Used to derive the commissioner and GP-based analyses.

ID	Data Item Name / NHS Data Dictionary Name	Data Item Description	Format	M (Mandatory) or M* (Mandatory with caveats, see Validation for details)	Validation	Rationale
	REGISTRATION)	PATIENT is registered with.			national ODS table.	
13	**ORGANISATION CODE (CODE OF PROVIDER)**	The ORGANISATION CODE of the ORGANISATION acting as a Health Care Provider.	max an12	M	Code must be a current live organisation in the national ODS table.	Shows the provider for the activity and will also be a potential part of the data linkage within and between different datasets.
14	**ORGANISATION CODE (CODE OF COMMISSIONER)**	The ORGANISATION CODE of the ORGANISATION commissioning health care.	max an12	M	Code must be a current live organisation in the national ODS table.	Shows the commissioner for the activity.
15	**ORGANISATION CODE (RESIDENCY RESPONSIBILITY)**	The ORGANISATION CODE derived from the PATIENT's POSTCODE OF USUAL ADDRESS.	max an12	M*	Mandatory if record is sensitive and POSTCODE OF USUAL ADDRESS is blank. Code must be a current live organisation in the national ODS table.	Enables CCG/health board etc. to flow where there is no postcode.
16	**ORGANISATION CODE (GP PRACTICE RESPONSIBILITY)**	The ORGANISATION CODE of the ORGANISATION responsible for the GP Practice where the PATIENT is registered, irrespective of whether	max an12	M*	Mandatory if GENERAL MEDICAL PRACTICE CODE (PATIENT REGISTRATION) is blank.	Enables CCG/health board etc. to flow where there is no GP.

ID	Data Item Name / NHS Data Dictionary Name	Data Item Description	Format	M (Mandatory) or M* (Mandatory with caveats, see Validation for details)	Validation	Rationale
		they reside within the boundary of the Clinical Commissioning Group.			Code must be a current live organisation in the national ODS table.	
17	**ORGANISATION CODE (LOCAL PATIENT IDENTIFIER)**	ORGANISATION CODE (LOCAL PATIENT IDENTIFIER) is the ORGANISATION CODE of the ORGANISATION that assigned the LOCAL PATIENT IDENTIFIER.	max an5	M	Must be a current live organisation in the national Organisation Data Service tables.	Used in combination with Organisation Code (Code of Provider) to establish a unique reference for an individual.
18	**ACTIVITY DATE**	The relevant activity date will vary from dataset and the CDS ACTIVITY DATE, or specific dataflow guidance, should be used. In general, the relevant activity date is: A&E/ED = Arrival date OP = Appointment date Finished APC Episode = End date Birth/Delivery = Delivery date	an10 CCYY-MM-DD	M	Must be in the national format.	Enables date-based derivations to be applied and differential date rules to be applied.