

Document filename:	CDS v6.3 – Specification for Directions and Requests-2021		
Project / Programme	Commissioning Data Sets v6.3		
Document Reference	CDS v6.3/DAPB 0092 IG Spec	Status	Final
Owner	Tom Latham	Version	1.0
Author	Tom Latham	Version issue date	04/05/2021

Information Governance Requirements Specification for Commissioning Data Sets (CDS) v6.3

Document management

Revision History

Version	Date	Summary of Changes
0.1	01/03/2021	Initial version for review by IAO/analytical lead
0.2	03/03/2021	Minor updates following review comments from analytical lead
0.3	26/03/2021	Further updates following wider analytical review. For review by IAO
0.4	06/04/2021	Minor changes in line with IAO recommendations. For IG review
0.5	21/04/2021	Amendments following NHS Digital IG review. For review by DHSC
1.0	04/05/2021	Final version following DHSC approval; for publication alongside direction

Reviewers

This document must be reviewed by the following people:

Reviewer name	Title / Responsibility	Date	Version
Stuart Richardson	Programme Manager – Data Architecture & Secondary Use Transformation; Information Asset Owner for CDS	04/05/2021	1.0
Kieran Baker	Head of Delivery, Business Intelligence & Visual Analytics, Secondary Care	25/03/2021	0.2

Approved by

This document must be approved by the following people:

Name	Signature	Title	Date	Version
Stuart Richardson		Programme Manager – Data Architecture & Secondary Use Transformation; Information Asset Owner for CDS	04/05/2021	1.0

Glossary of Terms

Term / Abbreviation	What it stands for
AHAS	Acute Hospital Activity Statistics
CDS	Commissioning Data Sets
DARS	Data Access Request Service
MHSDS	Mental Health Services Data Set
MSDS	Maternity Services Data Set
SUS	Secondary Uses Service

Document Control:

The controlled copy of this document is maintained in the NHS Digital corporate network. Any copies of this document held outside of that area, in whatever format (e.g. paper, email attachment), are considered to have passed out of control and should be checked for currency and validity.

Purpose of document

This document sets out the requirements for the Commissioning Data Sets (CDS) v6.3 and should be read alongside the Direction issued by the Secretary of State for Health and Social Care ('Spine services (no 2) 2014 Direction').

Introduction / Purpose of data collection

The Commissioning Data Sets (CDS) are patient level, secondary use data sets providing information about NHS provided secondary care activity. CDS is used for a variety of purposes, including acting as the source data for national reporting through Hospital Episode Statistics (HES), the allocation of payments through the National Tariff Payment System and measurement of waiting times through Referral to Treatment (RTT). CDS data is also used to produce curated patient level data extracts to support research and healthcare planning and is provided with appropriate governance to external bodies such as NHS England and NHS Improvement. The data is submitted through the Secondary Uses Service (SUS).

CDS v6.3 is an update to the existing CDS v6.2. The changes are mainly designed to update CDS in line with current clinical and data recording practices, as well as to support recent policy initiatives and enable conformance with other information standards and legislation introduced since CDS v6.2 went live.

The intention is for CDS v6.3 to be a minor, 'tactical' release, in order for changes to be made quickly to allow the flow of SNOMED CT codes (including recently authored codes that relate to COVID-19) as soon as possible.

Data collection

Scope

The scope of CDS includes all secondary care activity undertaken by NHS Hospital Providers in England, including patients receiving private treatment, and NHS patients treated electively in the independent sector (including Any Qualified Providers).

Patient activity in the form of Outpatient Appointments (including the ability to submit Did Not Attends) and Admitted Patient Care (Hospital Admissions) are included.

Source

The data will be collected from NHS Hospital Providers, though in most cases the data submissions themselves are made by accredited middleware providers who submit data on behalf of the NHS Trusts.

Category

Personal data about patients, as well as some data about staff, is included in the CDS v6.3 specification and collected in CDS. This includes the patient's name (only if their identity is unverified, i.e. their NHS Number has not been traced, is unable to be traced or they do not have an NHS Number), address (only if their identity is unverified, as above), postcode of usual address, NHS Number, birth date, stated gender and ethnic category. The patient confidential data items are as defined in the existing '[Submitting CDS Data to SUS](#)' document (see page 31), and are listed below for reference:

- NHS Number
- Local Patient ID
- ALL VGP (very general purpose) fields
- Name
- Patient Usual Address
- Postcode of Usual Address
- Date of Birth
- Hospital Spell Number
- Local CCMDS ID
- Mother's details (on birth CDS):
 - Local Patient ID (Mother)
 - NHS Number (Mother)
 - Birth Date (Mother)
 - Patient Usual Address (Mother)
 - Postcode of Usual Address (Mother)
- Baby's details (for all babies on delivery CDS):
 - Local Patient ID (Baby)
 - NHS Number (Baby)
 - Birth Date (Baby)
- Patient Pathway Identifier
- UBRN Converted

Details of the patient's health conditions and treatments (e.g. diagnoses and procedures) can also be collected through the use of clinical terminology (SNOMED CT) and classifications (ICD-10 and OPCS-4).

Frequency

The data is collected on an ongoing basis via the SUS+ platform¹. CDS is made up of a number of different 'types' relating to different aspects of hospital care, each of which has a specific submission frequency. There are specific national deadlines for data submissions to be received in order for the data to be included in payments, as published annually on [the SUS Payment by Results webpage](#).

NHS England and NHS Improvement have recently specified a requirement in the Standard NHS Contract for weekly data to be submitted, and this took effect from 31 March 2021. The CDS v6.3 ISN will reference this requirement for certain CDS types, whereas others will only need to be submitted on a monthly basis.

¹ SUS+ was implemented in April 2017, replacing the previous legacy version (SUS) as the repository for healthcare data in England. 'SUS' remains in use as a generic term for the Secondary Uses Service as a whole, whereas 'SUS+' is used in this document to refer to the actual landing portal which is used for submissions of CDS data.

Analysis

Internal processing

Healthcare providers collect administrative and clinical information locally to support the care of the patient. The data is submitted to the SUS at NHS Digital, which, as well as making it available to commissioners, also copies the information to a database.

The data is validated at the point of submission by the submission XML schema. All Data Elements are subject to length validation. Some Data Elements are also subject to format and content validation against an enumerated list of permitted values defined in the NHS Data Model and Dictionary. Other specific business rules also apply to selected items, [as set out in the Dictionary](#).

At pre-arranged dates during the year, SUS takes an extract from the SUS+ database and sends it to HES. Data on SUS+ will continue to change, but HES data is fixed as it was when that particular extract was taken.

The HES data quality team then validates and cleans the extract, before deriving new items and making the information available in an Oracle data warehouse. Data quality reports and checks are completed at various stages in the cleaning and processing cycle.

HES is produced using Oracle processing, though this will be decommissioned and replaced by a new asset called Acute Hospital Activity Statistics (AHAS). AHAS will be produced in SUS+ and the Data Processing Service (DPS). AHAS is expected to be introduced between August and October 2021, though the project is still ongoing and these dates are subject to change.

Please see [‘The processing cycle and HES data quality’ pages on the NHS Digital website](#), and in particular [‘The HES Processing Cycle and Data Quality’ document](#), for more information. These pages will be updated to reflect the AHAS changes.

Data linkage

HES data, which is generated from CDS, is linked to other data sets such as Diagnostic Imaging Data (DID) and the Mental Health Services Data Set (MHSDS).

Some of the CDS v6.3 changes will allow additional linkage with specific data sets, for example the introduction of the Care Contact Identifier (Ambulance Service) data item will enable event-based linkage with the forthcoming Ambulance Data Set. HES data is already linked to various other data sources, including Patient Reported Outcome Measures (PROMs), Patient Level Information and Costing System (PLICS), and NHS 111 data.

HES data is also currently used in conjunction with other data – including data from GP extracts, the Personal Demographics Service (PDS) and Maternity Services Data Set (MSDS) – to create the shielded patient list of vulnerable people at “high risk” of complications from COVID-19. In addition, HES data is used in the [COVID-19 Population Risk Assessment](#), which was developed using [QCovid](#), and associated analysis on shielding patient health outcomes. The Population Risk Assessment data also feeds into the shielded patient list.

Consultation

As part of the CDS v6.3 development process and obtaining approval from the Data Alliance Partnership Board (DAPB) for the publication of an updated Information Standards Notice (ISN) for CDS v6.3, a number of consultation activities have been carried out:

- Discussions with national bodies, including Department of Health and Social Care, NHS England and NHS Improvement and NHS X, about the high-level requirements for CDS v6.3. This includes the creation of a CDS Steering Group which is jointly chaired by NHS England and Improvement and NHS Digital
- Regular updates about CDS v6.3 development provided at monthly SUS+ User Group meetings, attended by data submitters, care providers and system suppliers
- A series of webinars held in November 2020 to discuss the changes to CDS as part of the v6.3 uplift and invite feedback. A recording of the webinar was also made available [on the NHS Digital website](#)
- A public consultation in February-March 2021 to ask for feedback on the detailed proposals for CDS v6.3. This was open to any organisation or individual to complete, though data providers and submitters (including NHS Trusts and middleware providers), clinical system suppliers, DARS customers, commissioners and national healthcare bodies were specifically invited to participate
- An Expert Reference Group has been formed from volunteer care providers, middleware providers and system suppliers in order to discuss the changes to CDS v6.3 and proposed technical solutions in more detail through online meetings and via email

Dissemination/Sharing

Regular Dissemination/Sharing

CDS data is used to produce [HES data](#), including monthly and yearly publications. CDS data is processed and used for non-clinical purposes, such as research and planning health services. Because these uses are not to do with direct patient care, they are called 'secondary uses'. This is the HES data set. HES is a database containing details of all admissions, A and E attendances and outpatient appointments at NHS hospitals in England.

HES provides data for the purpose of healthcare analysis to the NHS, government and others including:

- national bodies and regulators, such as the Department of Health, NHS England and NHS Improvement, Public Health England and the CQC
- local Clinical Commissioning Groups (CCGs) and Data Services for Commissioners Regional Offices (DSCROs)
- provider organisations
- government departments
- researchers and commercial healthcare bodies

- National Institute for Clinical Excellence (NICE)
- patients, service users and carers
- the media

The data is used for a variety of purposes, including:

- national policy making
- benchmarking performance against other hospital providers or CCGs
- academic research
- analysing service usage and planning change
- providing advice to ministers and answering a wide range of parliamentary questions
- national and local press articles
- international comparison

The data is provided in accordance with relevant IG principles and requirements. For example, external organisations must apply through the Data Access Request Service (DARS) process in order to receive HES data – see below. Approved users can also access HES data to perform analysis via the [Data Access Environment \(DAE\)](#) and the Monthly Managed Extract Service (MMES), which allows users to receive an extract from HES on a monthly basis.

NHS England and NHS Improvement also receive CDS data as submitted². A pseudonymised copy is sent to the National Commissioning Data Repository (NCDR) web portal, which is a pseudonymised patient-level data repository. The NCDR is used by NHS England and NHS Improvement analytical teams to create their own assets, which are then used to aid in the delivery of a wide range of projects, including NHS Long Term Plan objectives.

In addition, CDS data is accessible to approved researchers from trusted organisations through the NHS Digital Trusted Research Environment (TRE) service for England. TRE contains linked, de-identified health data to enable users to quickly answer COVID-19-related research questions.

Furthermore, CDS data is shared by DSCROs with commissioning organisations – currently CCGs – on at least a monthly basis across the whole of England to support healthcare planning (such as risk stratification) and payment.

Data Access Request Service (DARS)

CDS data has been disseminated through DARS to around 400 customers, and the DARS extracts will be updated to include new data available due to the CDS v6.3 changes. The data are usually pseudonymised before being released to DARS customers. Applicants must have appropriate permission, and adhere to specific information governance requirements, in order to receive data in anything other than an aggregated form.

² NHS England and NHS Improvement receive the CDS data as it is submitted, with the exception of records which are anonymised because they contain legally restricted codes or because they relate to patients who have requested that their records are anonymised.

Where data is disseminated to commissioners for the purpose of Invoice Validation or Risk Stratification, this is covered by an approval under Section 251 of the NHS Act 2006 from the [Confidentiality Advisory Group](#) and is released in an identifiable form.

This is as per NHS [Digital Data Access Request Service \(DARS\)](#) process and all releases are recorded on the [Register of Approved Data Releases](#).

Publication

Data to be published

HES data is published on the NHS Digital website in an aggregated and suppressed form. [Monthly HES publications](#) are produced, along with annual HES publications for Accident and Emergency, Admitted Patient Care, Adult Critical Care, Outpatients and Maternity data. There is the potential for data to be published more frequently in the future (such as in the form of daily or weekly management information reports) as weekly submissions of data become more prevalent and once AHAS is fully established.

We apply a strict statistical disclosure control in accordance with the NHS Digital protocol to all published HES data. This suppresses small numbers to stop people identifying themselves and others, to ensure that patient confidentiality is maintained.

Data prohibited from being published

HES data is only published in an aggregated and suppressed form, and no small numbers are published or routinely released. Various guidelines around the analysis and publication of HES data are available in the [HES Analysis Guide](#).

In addition, data related to legally restricted records will not be published. The SUS+ service uses [a list of legally restricted codes published on the SUS website](#) and submitters must anonymise the record by removing patient identifiable fields before submitting this data. The list of legally restricted codes will be updated to cover relevant SNOMED CT codes before CDS v6.3 submissions start.

System Delivery Function

CDS data is submitted via the existing SUS+ platform, which is covered by a System Level Security Policy (SLSP). No additional IT or communication systems are required in order to facilitate CDS v6.3 submissions.

Change control process

Updates to this specification will be authored when further version updates to CDS are made. Changes will be managed in conjunction with NHS England and NHS Improvement via the CDS Steering Group, and the Department of Health and Social Care will also be

consulted as the directing organisation responsible for the overarching Spine Services (No.2) direction.