Publication withdrawn

This guidance was withdrawn in April 2024.

For up-to-date information about the National Drug Treatment Monitoring System (NDTMS), see core data set documentation on the NDTMS website.

Public Health England

Protecting and improving the nation's health

National Drug Treatment Monitoring System (NDTMS)

CJIT CSV file format specification

NDTMS data set P

About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing, and reduce health inequalities. We do this through world-leading science, research, knowledge and intelligence, advocacy, partnerships and the delivery of specialist public health services. We are an executive agency of the Department of Health and Social Care, and a distinct delivery organisation with operational autonomy. We provide government, local government, the NHS, Parliament, industry and the public with evidence-based professional, scientific and delivery expertise and support.

Public Health England
Wellington House
133-155 Waterloo Road
London SE1 8UG
Tel: 020 7654 8000
www.gov.uk/phe
Twitter: @PHE_uk
Facebook: www.facebook.com/PublicHealthEngland

Published September 2019
PHE publications
gateway number: GW-749

PHE supports the UN
Sustainable Development Goals

Contents

1. Introduction

The National Drug Treatment Monitoring System (NDTMS) data helps drug treatment demonstrate the outcomes it achieves for the people it treats and in doing so aids accountability for the money invested in it. NDTMS is a national standard and is applicable to young people and adults within community and secure setting based treatment providers.

This document defines the format of the Comma Separated Variable (CSV) file used as the primary means of inputting CJIT Data Set data into the CJIT DAMS system.

The criminal justice dataset is a national standard and is applicable to adults within community based treatment providers.

This document is intended to be a definitive and accessible source for use. It is not intended to be read from end to end, rather as a reference document, which is utilised by a variety of readers including:

- interpreters of data provided from PHE systems
- suppliers of systems to PHE
- suppliers of systems which interface to PHE systems
- PHE/National Drug Treatment Monitoring System (NDTMS) personnel

This document should not be used in isolation, it is part of a package of documents supporting the NDTMS data set and reporting requirements.

Please read this document in association with:

- **NDTMS CJIT technical definition** – defines the complete list of fields and their layout that make up the CJIT dataset.
- **NDTMS CJIT reference data –** provides permissible values for each data item
- **NDTMS geographic information** – provides the locality information eg DAT of residence, Local Authority etc.
- **NDTMS CJIT business definitions** – provides the explicit definition of each of the fields we collect to ensure data is consistent.

To assist with the operational handling of CSV input files, each significant change to the NDTMS data set is allocated a letter.

The latest version (commonly referred to as the NDTMS data set P) for national data collection will come into effect on 1 April 2020.

2. CJIT data export

The CJIT DAMS Portal does not require a particular methodology or application to be used to create the CSV input file. All headers (as specified in the NDTMS CJIT technical definitions document) should appear in all dataset extracts. As a technical minimum, the file should contain all new records and all changes to existing records since the previous months submission (including those records that have been updated but are no longer in the reporting year that have previously been submitted to NDTMS).

The recommended minimum submission period for activity data is the current reporting period (see Section 3.3). As national statistics and performance figures are published monthly from the data, it is expected that a CSV file will be submitted monthly.

Input files can contain any number of client record updates and could represent data from any or all of the following:

- all client data from a particular treatment provider since the NDTMS CJIT provider code was established (a full extract)
- all client data from a particular CJIT treatment provider covering the reporting period (a reporting extract – see Section 3.3)

3. Input file

The CSV input file will be structured according to common CSV file format standards. Each line, including the last, will be terminated by a carriage return. The extract should only contain the code values and other specified values as detailed in the CJIT technical definition and CJIT reference data documents.

3.1 File naming

The filename of the CSV input file must contain the following items of information:

- originating NDTMS CJIT agency code
- dates (YYYYMMDD) relating to the time span of the extracted client data
- literal denoting core data set, for example: TREAT-IN-P
- file structure suffix .CSV

Filenames will take one of the following forms:

AAAAA-YYYYMMDD-yyyymmdd-TREAT-IN-P.CSV

AAAAA	Originating CJIT treatment provider code. The regional NDTMS team will be responsible for specifying the CJIT treatment provider code, this will be a 5 character code
YYYYMMDD-yyyymmdd	Dates defining the time span that the file relates to: YYYYMMDD = from date, yyyymmdd = to date (not case sensitive, only used to show the difference between the 2)
TREAT-IN	Literal to denote the type of data contained in the file
P	The version of the core data set that was used to produce the extract – see NDTMS technical definition
CSV	File structure indicator

The following is an example filename:

CJ001-20180101-20200201-TREAT-IN-P.CSV

This file would be a CSV input file from a treatment provider ‘CJ001’ containing all client data from 1 January 2018 to 1 February 2020, working to version ‘P’ of the data set.

3.2 File content

A CSV input file will consist of multiple rows/lines of data. The minimum file size would be 2 rows:

- the first row will be the header row. The header row is used to define which of the core data set fields are represented within the CSV file, and the order in which they are provided (see below)
- the second row will be the client row which should consist of client demographic information, episode details and the first structured treatment referral or recovery support sub-intervention review

Subsequent rows will contain a data row for each client further instance of the following:

- structured treatment referral
- recovery support sub-intervention review

The field names are specified in the NDTMS CJIT technical definition. Given that this is a CSV file, the CSV file header values must be used in the first row as they appear in the technical definition.

The data rows must be in the order given in the header row and a comma must follow each field. A comma must follow null or empty fields (all rows in the file must contain all fields separated by a comma – even if the field is null).

An example of a CSV input file is given below. Please note that this is an example only, for full details of the required fields and header names, please refer to the NDTMS CJIT technical definition document.

FINITIAL,SINITIAL,DOB,SEX,ETHNIC,PC,……...,SUBMODDT,SUBMID,RECPEER,,
M,M,1984-10-04,M,,CM14 4,……. ,2004-01-18,1111122,Y,,
D,D,1957-03-12,M,B,NW7 3,…….,,,

The above example also demonstrates the handling of null values.

3.3 Extraction criteria

The data to be extracted should conform to one of the following standards and should contain all data relating to a client's treatment episode including all associated structured treatment referrals and recovery support sub-intervention reviews.

Full extract: All relevant activity data on the treatment provider database regardless of the discharge date

Reporting extract: This extract covers the period over which changes to data may have an impact on published performance figures. The period required for the reporting extract is 1 January prior to the previous financial year to the end of the current month. The financial year changes on 1 September, for example:

- for submission of CJIT data on 1 June 2019 the extract should contain all records relating to treatment episodes that are not discharged, or who were discharged after 1 January 2017
- for submission of CJIT data on 1 September 2019 the extract should contain all records relating to treatment episodes that are not discharged, or who were discharged after 1 January 2018

The reporting year extract should also contain any records within the reporting year or prior to the reporting year that have been changed since the last time they were submitted to NDTMS. Where practical, this level of extract should be the default option as changes to data that may affect current performance figures will be updated by the submission.

See the NDTMS CJIT technical definition document for the extraction details of clients that have opted not to share their details with NDTMS.

Extract content

Extracts should only contain the **code values** specified in the reference data document.

3.4 Client data order

Due to the representation of client data in the CSV input file, it is likely that there will be multiple rows of data per client. There must be one row per treatment provider (NDTMS CJIT code) for each:

- episode /Recovery support sub-intervention review
- episode /Referral to structured treatment

Consider the example where client A has had 2 recovery support sub-intervention reviews and 2 structured treatment referrals which are part of one treatment episode and a secondary episode with 4 recovery support sub-intervention reviews and one structured treatment referral. The CSV input file will contain 9 rows of data for that client:

Row N	Client A details	Treatment episode 1 details	RS sub-intervention 1 details
Row N+1	Client A details	Treatment episode 1 details	RS sub-intervention 2 details
Row N+2	Client A details	Treatment episode 1 details	structured treatment referral 1 details
Row N+3	Client A details	Treatment episode 1 details	structured treatment referral 2 details
Row N+4	Client A details	Treatment episode 2 details	RS sub-intervention 1 details
Row N+5	Client A details	Treatment episode 2 details	RS sub-intervention 2 details
Row N+6	Client A details	Treatment episode 2 details	RS sub-intervention 3 details
Row N+7	Client A details	Treatment episode 2 details	RS sub-intervention 4 details
Row N+8	Client A details	Treatment episode 2 details	structured treatment referral 1 details

Note: Structured treatment referral data and recovery support sub-intervention data CANNOT be in the same row within the CSV input file, each dated entity must have their own individual row with individual ID.

NDTMS does not expect client records to be sorted in any particular order.

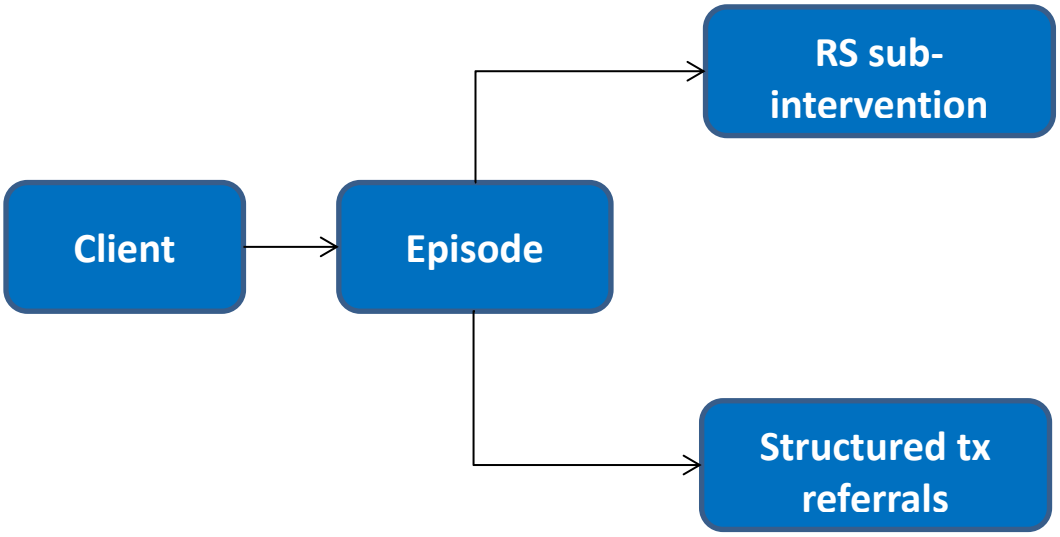
4. Implementation of core data set changes

Core data set changes are specified as changes to the CJIT schema and as changes to the CJIT reference data, which are the code sets used within the schema.

Data set changes should only be applied to client records that are active at the time of the dataset change. These changes **should not** be applied retrospectively to client records that do not meet these criteria. This applies to both schema and reference data code sets.

A full list of reference data items and the associated ‘trigger’ dates for implementation are available in the document NDTMS CJIT technical definition.

5. NDTMS data model



The NDTMS CJIT data model for core data set P contains a total of 4 entities with the following relationships:

One client can have **many** episodes. Episodes can have **many** SIRs and referrals.

Appendix A – Revision history

Version	Author	Purpose/reason	Date
2.0	HSmith	Changes from version 1.01 Key changes: • document updated to bring in line for NDTMS core dataset P • removal of TOP fields from the data structure • addition of CSV needing to contain any previously submitted records (regardless of date) that have been updated, to be included in the agency's reporting extract	12/09/2019