



Public Health  
England

Protecting and improving the nation's health

# **National Drug Treatment Monitoring System (NDTMS)**

## **CSV file format specification**

NDTMS dataset Q

Version 15.0

# Contents

- 1. Introduction .....3
- 2. Treatment provider or prison treatment data export.....4
- 3. Input file .....5
  - 3.1 File naming .....5
  - 3.2 File content .....6
  - 3.3 Extraction criteria .....7
  - 3.4 Client data order.....8
- 4. Implementation of core dataset changes .....9
- 5. NDTMS data model ..... 10
- Appendix A – Revision history..... 11

# 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 the core dataset data items into the National Drug Treatment Monitoring System (NDTMS).

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 dataset and reporting requirements.

Please read this document in association with:

- **NDTMS Technical definitions** – provides the full list of fields that are required in the CSV file and the verification rules for each item
- **NDTMS Reference data** – provides permissible values for each data item
- **NDTMS Geographic information** – provides the locality information eg DAT of residence, Local Authority etc.
- **NDTMS 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 dataset is allocated a letter.

The latest version (commonly referred to as the NDTMS dataset Q) for national data collection will come into effect 1<sup>st</sup> April 2022.

## 2. Treatment provider or prison treatment data export

There are several versions of the NDTMS dataset, the correct one to use for a given provider is dependent upon whether the client cohort is young people or adults and whether the provider is in the community or a secure setting. All headers (as specified in the NDTMS Technical definitions document) should appear in all dataset extracts, but only the fields specified as applicable to that dataset should be populated (see dataset inclusion column in NDTMS Technical definitions document).

The NDTMS system (DAMS) does not require a particular methodology or application to be used to create the CSV input file. As a *technical* minimum the file should contain all new records and all changes to existing records since the previous 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 or secure setting since the NDTMS provider code was established (a full extract)
- all client data from a particular treatment provider or secure setting 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.

#### 3.1 File naming

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

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

Filenames will take one of the following forms:

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

|                   |   |
|-------------------|---|
| AAAAA             | Originating treatment provider code or NDTMS secure setting code. The regional NDTMS team will be responsible for specifying the treatment provider code or NDTMS secure setting code, this will be a |
| YYYYMMDD-yyyymmdd | Dates defining the time span that the file relates to:<br>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  |
| Q                 | The version of the core dataset that was used to produce the extract – see NDTMS Technical definitions document   |
| CSV               | File structure indicator  |

The following is an example filename:

- L0001-20200101-20211001-TREAT-IN-Q.CSV

This file would be a CSV input file from a treatment provider ‘L0001’ containing all client data from 01/01/20 to 01/10/21, working to version ‘Q’ of the dataset.

### 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 dataset 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 intervention

Subsequent rows will contain a row for each client in each of the following instances:

- Treatment Intervention
- Treatment Outcomes Profile (TOP)
- Young Person Outcome Record (YPOR)
- Sub Intervention Review (SIR)
- Client Information Review (CIR)

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

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 Technical definitions document.

|  |
|--|
| FINITIAL,SINITIAL,DOB,SEX,ETHNIC,PC,,,,,,,,,MODAL,MODST,MODEND |
| M,M,1984-10-04,M,,CM14 4,,,,,,,, 4,2004-01-18,2004-01-19       |
| 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 treatment episode including all associated treatment interventions, TOP/YPOR records, SIR details and CIR.

**Full extract:** All relevant activity data on the treatment provider database or secure setting database, regardless of the discharge date

**NDTMS 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 the 1<sup>st</sup> January prior to the previous financial year to the end of the current month. The financial year changes on the 1<sup>st</sup> September, for example:

- for submission of community data on 01/06/2021 the extract should contain all records relating to treatment episodes that are not discharged, or who were discharged after 01/01/2019. For secure settings, this extract should contain all treatment episodes that are not exited, or who were exited on or after 01/01/2019
- for submission of community data on 01/09/2021 the extract should contain all records relating to treatment episodes that are not discharged, or who were discharged after 01/01/2020. For secure settings, this extract should contain all treatment episodes that are not exited, or who were exited on or after 01/01/2020.

The 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 Technical definitions 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 NDTMS 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 code) for each:

- Episode / Treatment intervention
- Episode / TOP / YPOR record
- Episode / SIR record
- Episode / CIR record

Consider the example where client A has had 2 treatment interventions and 2 TOP / YPOR records which are part of one treatment episode and a secondary episode with 4 treatment interventions, one sub intervention, and one CIR. The CSV input file will contain 10 rows of data for that client:

|         |                  |                             |                                     |
|---------|------------------|-----------------------------|-------------------------------------|
| Row N   | Client A details | Treatment episode 1 details | Intervention 1 details              |
| Row N+1 | Client A details | Treatment episode 1 details | Intervention 2 details              |
| Row N+2 | Client A details | Treatment episode 1 details | TOP/ YPOR details 1                 |
| Row N+3 | Client A details | Treatment episode 1 details | TOP/ YPOR details 2                 |
| Row N+4 | Client A details | Treatment episode 2 details | Intervention 1 details              |
| Row N+5 | Client A details | Treatment episode 2 details | Intervention 2 details              |
| Row N+6 | Client A details | Treatment episode 2 details | Intervention 3 details              |
| Row N+7 | Client A details | Treatment episode 2 details | Intervention 4 details              |
| Row N+8 | Client A details | Treatment episode 2 details | Sub Intervention 1 details          |
| Row N+9 | Client A details | Treatment episode 2 details | Client Information Review 1 details |

Note: TOP/ YPOR data, treatment modality data, sub intervention data, and client information review 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.



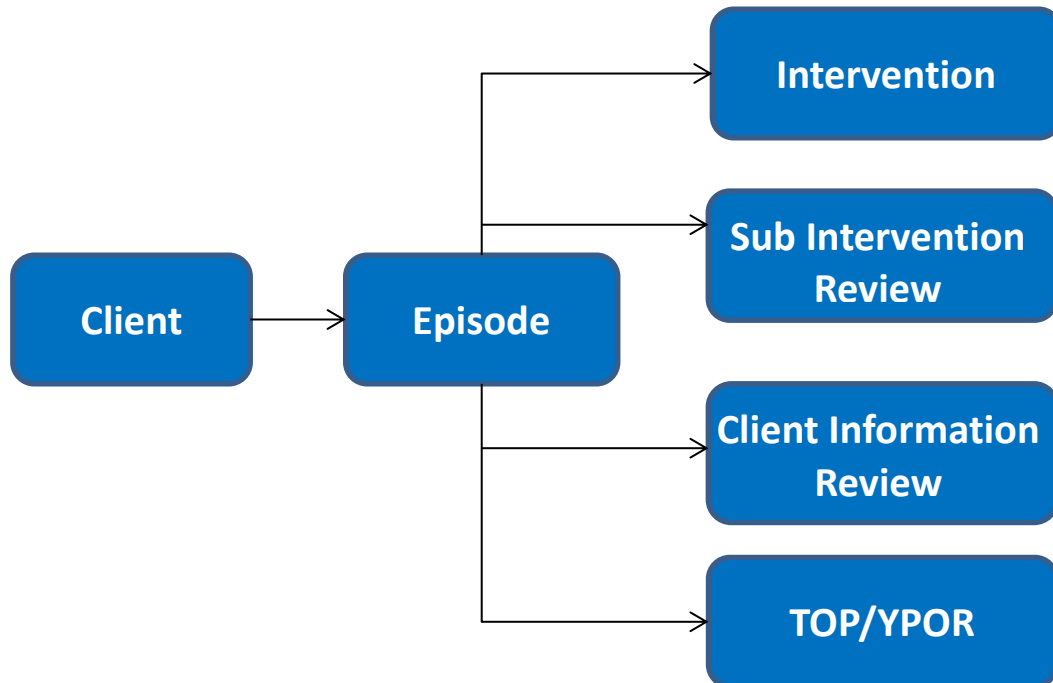
## 4. Implementation of core dataset changes

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

Dataset 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 NDTMS Technical definitions document.

## 5. NDTMS data model



The NDTMS data model for core dataset Q contains a total of 6 entities with the following relationships:

One client can have **many** episodes. Episodes can have **many** interventions, SIRs, TOP/YPOR records and CIRs.

## Appendix A – Revision history

| Version | Author    | Purpose/reason  | Date       |
|---------|-----------|---|------------|
| 10.02   | J Jaswani | Included NDTMS data model in relation to core dataset J   | 01/06/2012 |
| 11.02   | D Mhambi  | Updated to bring in line for NDTMS core dataset L   | 22/04/2013 |
| 11.04   | W Nielsen | Extensive corrections, re-formatting and updating of NDTMS documentation, bringing in line with   | 24/11/2014 |
| 12.0    | LHughes   | Document updated to bring in line for NDTMS core dataset N. Document revised and  | 20/09/2016 |
| 13.0    | HSmith    | Document updated to bring in line for NDTMS core dataset O.<br>Summary of changes:<br>Prison extracts to contain episodes not exited, or exited after the extract start date (changed from discharges).   | 15/10/2017 |
| 14.0    | HSmith    | Document updated to bring in line for NDTMS core dataset P.<br>Key changes: <ul style="list-style-type: none"> <li>• Removal of Time in Treatment fields from the data structure</li> <li>• Removal of Alcohol dataset and therefore removal of the alcohol outcomes record (AOR)</li> <li>• 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</li> </ul> | 30/09/2019 |
| 15.0    | JPalmer   | Document updated to bring in line for NDTMS core dataset Q.<br>Summary: Example filenames/NDTMS Reporting Extract examples updated with dates appropriate to new dataset  | 06/10/2021 |