



Office for Health
Improvement
& Disparities

National Drug Treatment Monitoring System (NDTMS)

CSV file format specification (Dataset R)

V1.0

Contents

1. Introduction	3
2. Treatment provider or prison treatment data export	5
3. Input file.....	6
3.1 File naming	6
3.2 File content	7
3.3 Extraction criteria	8
3.4 Client data order	9
4. Implementation of dataset changes.....	10
5. NDTMS data model.....	11
Appendix A: Changes implemented in dataset R.....	12

1. Introduction

The National Drug and Alcohol Monitoring System (NDTMS) is an English substance misuse treatment data collection that has been in place since 2004. This is hosted within the Office for Health Improvement and Disparities (OHID) in the Department of Health and Social Care.

Statistics derived from the collection (including those designated as Official and National Statistics) facilitate needs-assessment and targeting of resources, assessment of demand for services, evaluation of intervention/harm-reduction strategy effectiveness and service performance, service quality assurance and analyses of substance misusing populations. NDTMS data currently underpins key indicators that support the government's drug strategy.

NDTMS is a national standard and is applicable to young people and adults within community and secure setting-based treatment providers. The dataset is accredited by NHS Digital and the Information Standard is published under section 250 of the Health and Social Care Act 2012.

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 OHID systems
- Suppliers of systems to OHID
- Suppliers of systems which interface to OHID systems
- OHID/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 conjunction with the following documents, all of which are available on [NDTMS.net](https://www.ndtms.net):

- 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 UTLA of residence
- 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 R) for national data collection will come into effect 1st April 2024.

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 (NDTMSv2) 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 (full extract)
- all client data from a particular treatment provider or secure setting covering the reporting period (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 NDTMS Technical definitions and NDTMS Reference data documents.

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 dataset eg TREAT-IN-R
- file structure suffix .CSV

Filenames will take the following form:

- AAAAA-YYYYMMDD-yyyyymmdd-TREAT-IN-R.CSV

Filename part	Description
AAAAA	Originating treatment provider or secure setting code. The regional NDTMS team will be responsible for specifying the five character treatment provider or secure setting code.
YYYYMMDD-yyyyymmdd	Dates defining the time span that the file relates to: YYYYMMDD = from date, yyyyymmdd = to date (not case sensitive, only used to show the difference between the two)
TREAT-IN	Literal to denote the type of data contained in the file
R	The version of the 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-R.CSV

This file would be a CSV input file from a treatment provider 'L0001' containing all client data from 01/01/22 to 01/10/23, working to version 'R' 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 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's Outcomes 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

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 from 1st January prior to the previous financial year to the end of the current month. The financial year changes on the 1st September. For example:

- for submission of community data on 01/06/2023 the extract should contain all records relating to treatment episodes that are not discharged, or who were discharged after 01/01/2021. For secure settings, this extract should contain all treatment episodes that are not exited, or who were exited on or after 01/01/2021
- for submission of community data on 01/09/2023 the extract should contain all records relating to treatment episodes that are not discharged, or who were discharged after 01/01/2022. For secure settings, this extract should contain all treatment episodes that are not exited, or who were exited on or after 01/01/2022

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 Technical definitions document for the extraction details of clients that have not consented 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 agency 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 dataset changes

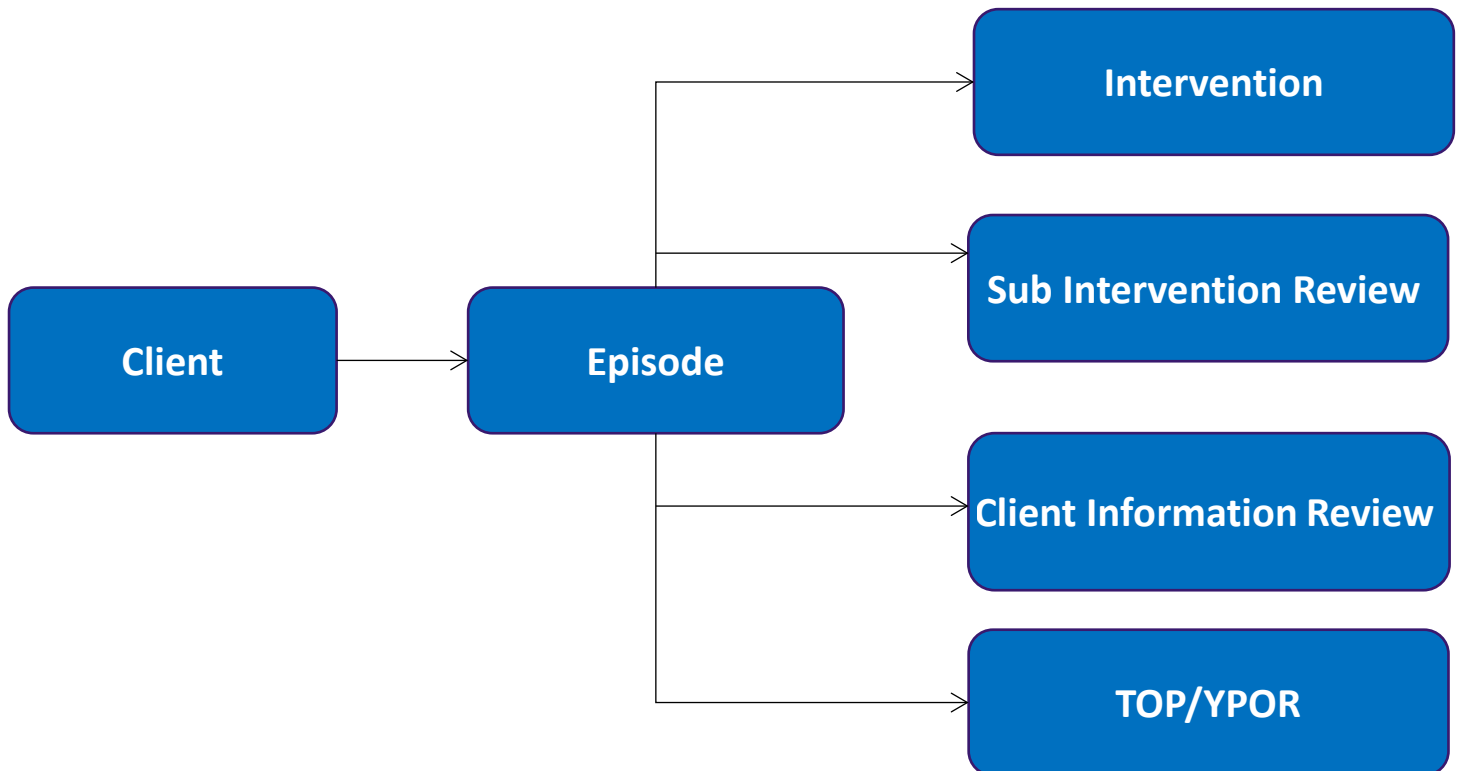
Dataset changes are specified as changes to the NDTMS schema and to the NDTMS reference data, which are the code sets used within the schema.

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 Reference data and NDTMS Technical definitions documents.

5. NDTMS data model

The NDTMS data model for dataset R contains a total of six entities with the following relationships:



One client can have MANY episodes.

Episodes can have MANY interventions, SIRs, TOP/YPOR records and CIRs.

Appendix A: Changes implemented in dataset R

Document updated to bring in line with NDTMS dataset R

Summary

- Example filenames/NDTMS Reporting Extract examples updated with dates appropriate to new dataset