



*National Treatment Agency  
for Substance Misuse*

National Drug Treatment Monitoring System (NDTMS)

**NDTMS DATA SET H**  
**CSV FILE FORMAT SPECIFICATION**

**Author** M. Hinchcliffe  
**Approver** M. Roxburgh  
**Date** 01/03/2011  
**Version** 8.03

**REVISION HISTORY**

<b>Version</b>	<b>Author</b>	<b>Purpose / Reason</b>	<b>Date</b>
3.00	G. Scott	Updated for Core Data Set E	29/05/2007
5.30	G. Scott	Update external references	02/07/2008
6.00	G. Scott	Updated Extract Criteria	11/11/2008
6.01	A. Juna	Updated NDTMS Current Year Extract Criteria	29/04/2009
7.01	M. Roxburgh	Changed extract period	11/11/2009
7.01	J. Jaswani	Added statement around TOP Records being expected in the submission file Expanded on example in Section 3.5 around client data order	12/02/2010
7.02	M. Hinchcliffe	Added section 3.6 to clarify query from supplier	23/04/2010
7.03	M. Hinchcliffe	Updated section 3.5	11/05/2010
7.04	M. Hinchcliffe	Updated section 3.5	01/10/2010
8.00	J. Jaswani	Updated section 3.6	21/10/2010
8.01	M. Hinchcliffe	Updated external links	25/10/2010
8.02	M. Hinchcliffe	Updated section 3.3 – Extraction criteria	24/01/2011
8.03	M. Hinchcliffe	Section 2 'TREATMENT PROVIDER DATA EXPORT' updated to reflect only one treatment agency's data should be present in each discrete file.	01/03/2011

# CONTENTS

<b>1</b>	<b>INTRODUCTION.....</b>	<b>6</b>
<b>2</b>	<b>TREATMENT PROVIDER DATA EXPORT .....</b>	<b>7</b>
<b>3</b>	<b>INPUT FILE .....</b>	<b>8</b>
<b>3.1</b>	<b>FILE NAMING.....</b>	<b>8</b>
<b>3.2</b>	<b>FILE CONTENT .....</b>	<b>8</b>
<b>3.3</b>	<b>EXTRACTION CRITERIA .....</b>	<b>9</b>
<b>3.4</b>	<b>CLIENT DATA ORDER .....</b>	<b>11</b>
<b>3.5</b>	<b>IMPLEMENTATION OF CORE DATA SET CHANGES .....</b>	<b>12</b>
<b>3.6</b>	<b>RISKS .....</b>	<b>12</b>

# 1 INTRODUCTION

This document defines the format of the Comma Separated Variable (CSV) file used as the primary means of inputting the Core Data Set data items into the National Drug Treatment Monitoring System (NDTMS).

## 2 TREATMENT PROVIDER DATA EXPORT

The NDTMS systems (DAMS) do 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, but doing so is not recommended.

The recommended minimum submission period for activity data is the current reporting period (see section 3.4) As National Statistics and performance figures are published monthly from the data, it is assumed 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
- A single client record from a particular Treatment Provider
- Selected client records from a particular Treatment Provider
- A mixture of any of the above

### 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 (CR)

#### 3.1 FILE NAMING

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

- Originating Agency code (where the file contains data from more than one Treatment Provider, the code of the primary Treatment Provider should be used).
- Dates relating to the time span of the extracted client data
- File structure suffix: .CSV
- Filenames will take one of the following forms:
- AAAAA-YYYYMMDD-yyyymmdd-TREAT-IN-X.CSV
- AAAAAA-YYYYMMDD-yyyymmdd-TREAT-IN-X.CSV

The first naming standard is intended for use by third-parties that are providing data, the second naming standard is for use of the NDTMS Core Data Set - Data Collection spreadsheet.

AAAAA	Originating Treatment Provider code (padded with any necessary leading zeros) – The regional NDTMS will be responsible for specifying the Treatment Provider code, typically this will be a 5 character code (only in exceptional cases will it be a 6 character codes).
YYYYMMDD & yyyymmdd	Dates defining the time span that the file relates to: YYYYMMDD = From date, yyyymmdd = to date
TREAT-IN	Literal to denote the type of data contained in the file
X	The version of the Core Data set that was used to produce the extract –see <a href="#">NDTMS Core Data Set - Technical Definition.</a>
CSV	File Structure indicator

The following are example filenames: L0105-20031101-20031130-TREAT-IN-A.CSV

This file would be a CSV input file from Treatment Provider 'L0105' containing all client data relevant to November 2003, working to version "A" of the core data set.

#### 3.2 FILE CONTENT

A CSV Input file will consist of multiple rows/ lines of data.

The first row will be the header record

Subsequent rows will contain a row for each instance of a treatment modality or Treatment Outcomes Profile (TOP) record for a client.

The minimum file size would be a single row (i.e. a header row, with zero client rows).

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. The header row must be constructed as a list of all the field name of the items contains in the file followed by a comma.

The field names are specified in the [NDTMS Core Data Set - Technical Definition.](#) Given that this is a CSV file, the CSV file header values must be used.

The data rows must be in the field order given in the header 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 Core Data Set - Technical Definition](#).

FINITIAL,SINITIAL,DOB,SEX,ETHNIC,PC,,,,,,,,,MODAL,MODST,MODEND
M,M,1984-10-04,M,,CM14 4,,,,,,,, 4,2004-01-18,04,2004-01-19
D,D,1957-03-12,M,B,NW7 3,,,,,,,,,,

The above example also demonstrates the handling of null values. If a field contains a comma, then the field must be wrapped with a single set of quotation marks.

### 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 modality and TOP records.

**Full Extract:** All relevant activity data on the treatment provider database, regardless of the discharge date. Where practical, this should be the default option. The advantage of doing so will be to ensure that the NDTMS and the treatment provder database are fully synchronised, and that performance figures produced from the data are an accurate representation of activity recorded at the provider.

**NDTMS Reporting Extract:** This extract covers the period over which changes to activity data may have an impact on published performance figures. The period required for the reporting extract is the 1<sup>st</sup> January from the previous financial year, with the financial year changing after the July upload. E.g.

- For submission of data on 01/10/2009 the extract should contain all records relating to treatment episodes that are not discharged, or who were discharged after 01/01/08
- For submission of data on 01/07/2010 the extract should contain all relating to treatment episodes that are not discharged, or who were discharged after 01/01/08
- For submission of data on 01/08/2010 the extract should contain all records relating to treatment episodes that are not discharged, or who were discharged after 01/01/09
- For submission of data on 01/07/2011 the extract should contain all records relating to treatment episodes that are not discharged, or who were discharged after 01/01/09
- For submission of data on 01/08/2011 the extract should contain all records relating to treatment episodes that are not discharged, or who were discharged after 01/01/10

This level of extract is the recommended minimum, as changes to data from previous financial years that may affect current performance figures will be updated by the submission,

**NDTMS Current Year Extract:** All records relating to treatment episodes that have a discharge date of on or after 1st January of the start of the financial year or that are currently open. For Healthcare Commission reporting requirements this extract criterion will have to be extended to July of the following year. For example:

For submission of data on August 2009 – July 2010 – start from date would be 1st January 2009.

For submission of data on August 2010 – July 2011 – start from date would be 1st January 2010.

**Note:** This level of extract is technically acceptable, but is not recommended as it cannot be used to amend submitted data from previous financial years data which affects current performance figures.

Note: An open episode means an episode without a Discharge Date

The Extract <from date> is the date that has been used to determine from when the data has been extracted (e.g. 20090101) and the Extract <to date> is the date when the extract was run.

Software suppliers should ensure that their system can support at least all of the above extract criteria.

Treatment Providers should take guidance from their Regional NDTMS database manager as to what extract criteria should be used for their monthly submission.

Records relating to Tier2 treatment modalities in YP services and carers, relatives and concerned others should be excluded from the extract – for further clarification please refer to the NDTMS Core Data Set – Business Definitions for Young People’s Treatment providers.

### 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 episode referral / treatment modality and one row per episode referral /TOP record.

Consider the example where client A has had 2 treatment modalities and two TOP records which are part of one treatment episode and four treatment modalities as part of another treatment episode. The CSV Input file will contain 8 rows of data for that client:

Row N	Client A details	Treatment Episode 1 Details	Modality 1 details
Row N+1	Client A details	Treatment Episode 1 Details	Modality 2 details
Row N+2	Client A details	Treatment Episode 1 Details	TOP Details 1
Row N+3	Client A details	Treatment Episode 1 Details	TOP Details 1
Row N+4	Client A details	Treatment Episode 2 Details	Modality 1 details
Row N+5	Client A details	Treatment Episode 2 Details	Modality 2 details
Row N+6	Client A details	Treatment Episode 2 Details	Modality 3 details
Row N+7	Client A details	Treatment Episode 2 Details	Modality 4 details

**Note: TOP data and treatment modality data CANNOT be in the same row within the CSV Input File.**

The NDTMS does not expect client records to be sorted in any particular order. It may be helpful however for the data to be sorted as described below to aid the manual review / checking of data.

By client (ascending Initials, Date of Birth, Gender)

By Referral Date (in chronological order)

By treatment Modality Start date (in chronological order, with blank treatments last)

By TOP date (in chronological order)

This might give a sort order within the file as follows:

FINAL	SINITIAL	DOB	SEX	REFLD	MODST
G	S	1969-01-15	F	2004-04-12	2004-04-16
G	S	1969-01-15	F	2004-04-12	2004-05-10
J	A	1963-10-19	M	2004-04-04	2004-04-04
J	A	1963-10-19	M	2004-04-10	2004-04-26
M	M	1987-10-04	M	2003-03-31	2004-05-01

### 3.5 IMPLEMENTATION OF CORE DATA SET CHANGES

Core Data Set changes are specified as changes to the schema and as changes to the reference data (the code sets used within the schema).

Data set changes should only be applied to client records that relate to activity during the period the codes are specified. 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 Technical Definition (section 14).

Any changes to the reference data displayed here should be 'triggered' where the associated date is in on or after the implementation date of the current Core Data Set (currently 1 April 2010).

### 3.6 RISKS

Although it is technically acceptable for NDTMS treatment providers to continue to submit the various CSV extract types, the risks associated with submitting an **NDTMS Current Year Extract** are below:

- **For submission of data on August 2010 – July 2011** (*start from date would be 1st January 2010*)

Treatment episodes (and their related treatment interventions and TOP data) discharged between the first of the previous financial year (01/01/2009) and the end of the previous financial year (31/12/2009) will **NOT** be available in this extract type and will therefore invalidate the automatic data mismatch checks on DAMS for these records.