National Drug Treatment Monitoring System (NDTMS)

CJIT CSV FILE FORMAT SPECIFICATION
# REVISION HISTORY

<table>
<thead>
<tr>
<th>Version</th>
<th>Author</th>
<th>Purpose / Reason</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CONTENTS

1 INTRODUCTION ................................................................................................................................. 6
2 EXPECTED CSV HEADER RESPONSES ............................................................................................ 8
3 INPUT FILE .......................................................................................................................................... 9
3.1 Extract Content ................................................................................................................................. 9
3.2 FILE NAMING ................................................................................................................................... 9
3.3 FILE CONTENT ................................................................................................................................. 10
3.4 EXTRACTION CRITERIA .................................................................................................................... 11
3.5 CLIENT DATA ORDER ....................................................................................................................... 12
4 FILES CONTAINING SUB-MODALITY, TREATMENT OUTCOMES PROFILE (TOP)/
   ALCOHOL OUTCOME RECORDS (AOR) AND REFERRED TO STRUCTURED TREATMENT DATA ................................................................. 13
4.1 IMPLEMENTATION OF CORE DATA SET CHANGES ................................................................. 14
4.2 CJIT DATA MODEL ......................................................................................................................... 15
1 INTRODUCTION

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 upload Portal.
Treatment Provider OR PRISON TREATMENT Data Export

The CJIT Upload Portal 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 month's submission. Files will be loaded on the CJIT Upload Portal on a delete and replace basis.

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 or Prison Establishment
- A single client record from a particular Treatment Provider or Prison Establishment
- Selected client records from a particular Treatment Provider or Prison Establishment
- A mixture of any of the above
2 EXPECTED CSV HEADER RESPONSES

The permissible responses contained under each header should strictly be code values as specified in the reference data document.

Text values or Special characters such as exclamation marks, full stops, commas, question marks, open and closed inverted commas and the like are not permitted as valid responses.

Please see the reference data document 1.03 for a full list of permitted code values and the Technical definition 1.04 for the rules and layout pertaining to each header.
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 Extract Content

The Full National Extract should only contain the code values and other specified values contained in the the reference data document.

3.2 FILE NAMING

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

- Originating Agency code/NDTMS Prison Code (where the file contains data from more than one Treatment Provider, the code of the primary Treatment Provider should be used) or NDTMS Prison code.
- Dates relating to the time span of the extracted client data
- File structure suffix: .CSV
- Filenames will take one of the following forms:
  - AAAA-YYYYMMDD-yyyyymmd-TREAT-IN-L.CSV
  - AAAAAA-YYYYMMDD-yyyyymmd-TREAT-IN-L.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 CJIT Core Data Set.

<table>
<thead>
<tr>
<th>AAAAA</th>
<th>Originating Treatment Provider code or NDTMS Prison Code (padded with any necessary leading zeros) – The regional NDTMS or Prison Support Team will be responsible for specifying the Treatment Provider code or NDTMS Prison code, typically this will be a 5 character code (only in exceptional cases will it be a 6 character codes).</th>
</tr>
</thead>
<tbody>
<tr>
<td>YYYYMMDD &amp; yyyyymmd</td>
<td>Dates defining the time span that the file relates to: YYYYMMDD = From date, yyyyymmd = to date</td>
</tr>
<tr>
<td>TREAT-IN</td>
<td>Literal to denote the type of data Contained in the file</td>
</tr>
<tr>
<td>L</td>
<td>The version of the Core Data set that was used to produce the extract –see CJIT Technical Definition version 1.04</td>
</tr>
<tr>
<td>CSV</td>
<td>File Structure indicator</td>
</tr>
</tbody>
</table>

The following are example filenames:

A0001-20031101-20031130-TREAT-IN-L.CSV

This file would be a CSV input file from Prison Establishment ‘A0001’ containing all client data relevant to November 2003, working to version “L” of the core data set.

L0001-20031101-20031130-TREAT-IN-L.CSV

This file would be a CSV input file from Treatment Provider ‘L0001’ containing all client data relevant the submission month, working to version “L” of the care data set.
3.3 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 Sub-Modality, Treatment Outcomes Profile (TOP), Alcohol Outcome Record, and Refered to structured treatment.

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

The header row is used to define which of the CJIT 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 names of the items contained in the file followed by a comma.

The field names are specified in the CJIT Technical Definition 1.04 and associated CJIT Mappings document version 1.00. 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 CJIT Core Data Set - Technical Definition version 1.04.

<table>
<thead>
<tr>
<th>FINITAL, SINITIAL, DOB, SEX, ETHNIC, PC, .........., MODAL, MODST, MODEND</th>
</tr>
</thead>
<tbody>
<tr>
<td>D,D,1957-03-12,M,B,NW7 3,..........,</td>
</tr>
</tbody>
</table>

The above example also demonstrates the handling of null values..
3.4 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/AOR details.

Full Extract: All relevant activity data on the treatment provider database or Prison establishment 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 provider and Prison establishment database are fully synchronised, and that performance figures produced from the data are an accurate representation of activity recorded at the provider or prison.

CJIT 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 1st January from the previous financial year, with the financial year changing after the July upload. E.g.

- For submission of data on 01/01/2014 the extract should contain all records relating to treatment episodes that are not closed, or who were discharged after 01/12/2013

- For submission of data on 01/02/2014 the extract should contain all records relating to treatment episodes that are not closed, or who were discharged after 01/03/2014
3.5 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/one row per episode referral TOP/AOR record, one row per episode referral/sub modality record and one row per episode referral/Referal to structured treatment.

Consider the example where client A has three Sub-Modality, one Refered to Structured and two TOP records which are part of one treatment episode The CSV Input file will contain 6 rows of data for that client:

**Client A Example**

<table>
<thead>
<tr>
<th>Row N</th>
<th>Client A details</th>
<th>Treatment Details</th>
<th>Episode</th>
<th>Sub-Modality details</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Client A details</td>
<td>Treatment Details</td>
<td>Episode 1</td>
<td>Sub-Modality details1</td>
</tr>
<tr>
<td></td>
<td>Client A details</td>
<td>Treatment Details</td>
<td>Episode 1</td>
<td>Sub-Modality details2</td>
</tr>
<tr>
<td></td>
<td>Client A details</td>
<td>Treatment Details</td>
<td>Episode 1</td>
<td>Sub-Modality details3</td>
</tr>
<tr>
<td></td>
<td>Client A details</td>
<td>Treatment Details</td>
<td>Episode 1</td>
<td>Refered to Structured treatment 1</td>
</tr>
<tr>
<td></td>
<td>Client A details</td>
<td>Treatment Details</td>
<td>Episode 1</td>
<td>Refered to Structured treatment 2</td>
</tr>
<tr>
<td></td>
<td>Client A details</td>
<td>Treatment Details</td>
<td>Episode 1</td>
<td>TOP/AOR Details 1</td>
</tr>
<tr>
<td></td>
<td>Client A details</td>
<td>Treatment Details</td>
<td>Episode 2</td>
<td>TOP/AOR Details 2</td>
</tr>
</tbody>
</table>
4 FILES CONTAINING SUB-MODALITY, TREATMENT OUTCOMES PROFILE (TOP)/ ALCOHOL OUTCOME RECORDS (AOR) AND REFERRED TO STRUCTURED TREATMENT DATA

When providing Sub Modality data, NO TOP/AOR or Referred to Structured Treatment data may be, provide on the same row.

When providing TOP/AOR data, NO Sub Modality or Referred to Structured Treatment data may be provided on the same row.

When providing Referred to Structured Treatment data, No Sub Modality data or TOP/AOR data may be provided on the same row.

In all cases, both client and episode data must be included in all rows.

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/Triage (in chronological order)
- By treatment Sub-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:

<table>
<thead>
<tr>
<th>FINITAL</th>
<th>SINITIAL</th>
<th>DOB</th>
<th>SEX</th>
<th>REFLD</th>
<th>SUBMODDT</th>
</tr>
</thead>
</table>
4.1 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.
4.2 CJIT DATA MODEL

The CJIT data model for Core Data Set L contains a total of 5 entities. The model elements contain the following relationship:

a. ONE CLIENT can have MANY EPISODES
b. AN EPISODE can have MANY SUB INTERVENTIONS
c. AN EPISODE can have MANY TOP RECORDS
d. AN EPISODE can have MANY AOR Records
e. AN EPISODE can have MANY REFERRED TO STRUCTURED TREATMENT

Key Fields

1. Sub modality = Sub Modality element
   - Key Field = Sub Intervention Assessment Date (this is a KEY field which when populated will create a new sub modality record)

2. Referred to structured treatment = Referred to structured treatment / Intensity element
   - Key Field = Referral to structured treatment Date (this is a KEY field which when populated will create a new Referred to structured treatment)

3. TOP/AOR = TOP/AOR element
   - Key Field = TOP Date (this is a KEY field which when populated will create a new TOP)