hscic | Health & Social Care Information Centre

dm+d XML Transformation Tool

User Guide

Author   Stuart Abbott

Date   17-Sep-2015

Contents

Introduction

The NHS Dictionary of Medicines and Devices (dm+d) XML[1] Transformation Tool is a basic tool to transform the dm+d data from a XML format to both a CSV[2] format and a Microsoft Excel format.

It is provided to demonstrate a method of transforming the dm+d data into alternative formats for research and educational use.

This version of the tool can transform both the main dm+d release (NHSBSA dm+d) and the supplementary or bonus dm+d release (NHSBSA dm+d Supplementary).

It does this through the use of an Extensible Stylesheet Language (XSL) processor and a series of accompanying XSL documents to create the CSV files. A macro enabled Microsoft Excel workbook then loads the CSV files into a series of workbooks and worksheets.

Warning

This tool has not been assured for use in clinical systems or for generating data for use in clinical systems. Any such use would require the user to perform their own assurance that the transformed data is sufficiently accurate for the use case.

The tool has been tested to work on Microsoft Windows 7 and Microsoft Excel 2010/2013.

Contact

If you have any questions, issues, comments or suggestions on this tool please contact the Pharmacy Terminology Service via:

• Email: information.standards@hscic.gov.uk

• Call: 0300 30 34 777

Downloading

The dm+d XML Transformation Tool must be downloaded directly from the TRUD[3] website.

To download the tool go to the TRUD website at https://isd.hscic.gov.uk. If you do not have an account for TRUD you will need to create one.

The dm+d XML Transformation Tool pack can be found in the following areas:

• NHS Dictionary of Medicines and Devices

• Terminology Tools

Once you have located the pack you will need to read and accept the OGL[4] licence for the tool. Subscriptions for the pack should be accepted automatically.

You can then download the tool by clicking on 'Download this release'.

---

[1] Extensible Markup Language, see http://www.w3.org/XML
[2] Comma-separated values. A tabular plain-text file format. See http://en.wikipedia.org/wiki/Comma-separated_values
[3] Technology Reference data Update Distribution (TRUD) website. See https://isd.hscic.gov.uk
[4] Open Government Licence. See http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/

Installing

The tool is downloaded as a compressed .zip file. You will need to move the contents of this compressed file to your c:\ drive. The method for doing this will differ depending on the operating system you are using and whether you have any specific programs for zip files installed.

1. First download to your computer the following from TRUD:
   - A copy of the dm+d XML Transformation Tool
   - A copy of the NHSBSA dm+d release
   - *Optional* – a copy of the NHSBSA dm+d Supplementary release

2. Select the save from the download prompt, you should now have two zip files (or three if you have downloaded the supplementary release as well). The endings of these file names will change with each release.
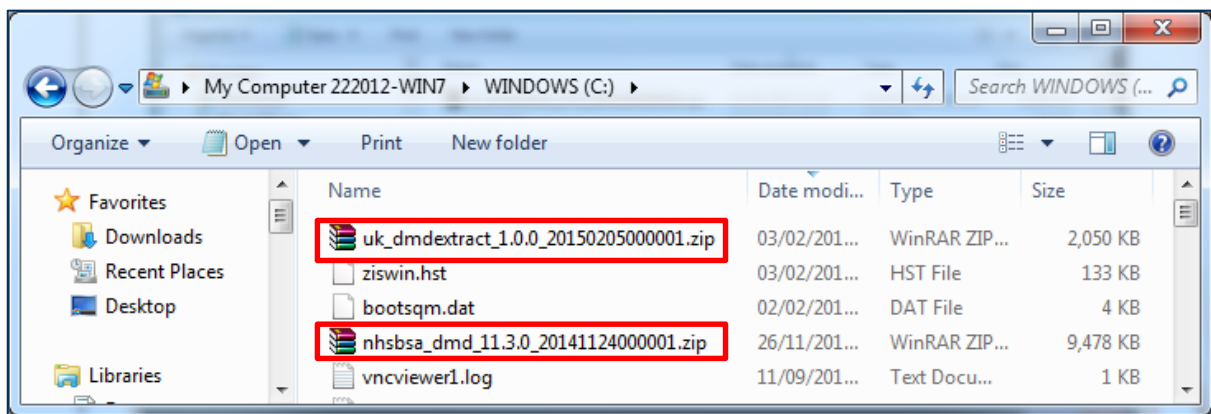
Figure 1 - Example of the two files

3. Select the uk_dmdextract_*n.n.n_yyyymmdd*000001.zip file and extract it (or unzip it) to the C:\ drive.
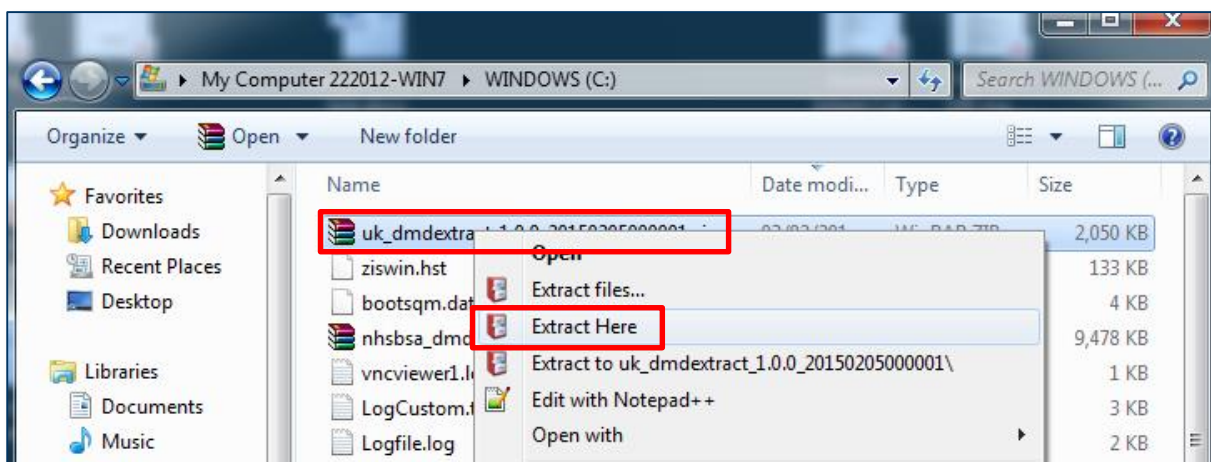

Figure 2 - How to extract the tool using WinRAR

4. This will create the folder structure: the **dmd_extract_tool** in the C:\ drive.
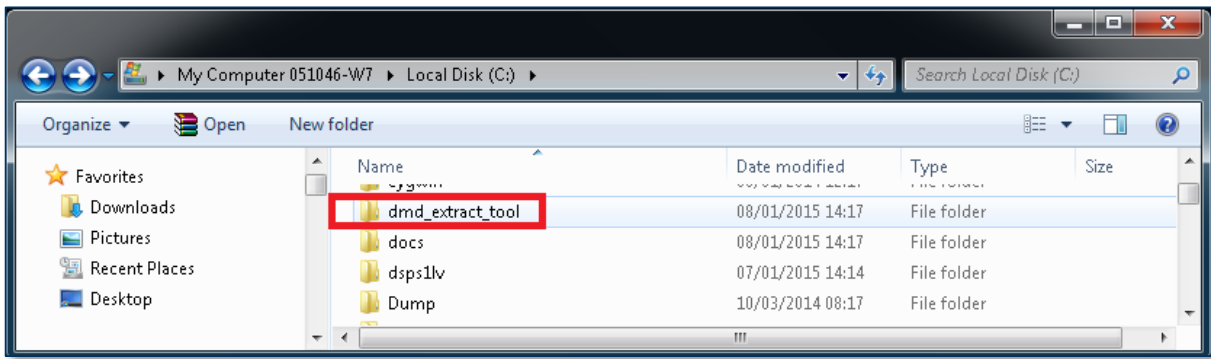

Figure 3 - The unzipped folder in C:

5. The finished path must be C:\dmd_extract_tool\xml-csv.bat
6. The tool is now installed.

Using the tool

Before using the tool, you need to have whichever version of the dm+d files you wish to transform. Releases of dm+d are available from TRUD in the 'NHS Dictionary of Medicines and Devices' pack. Currently only the NHSBSA dm+d subpack can be transformed by the tool.

1. Locate the zip file with the dm+d XML content you downloaded.
   In this example the file is named **nhsbsa_dmd_11.3.0_20141124000001.zip.**
   The dm+d release file should always look like this:
   nhsbsa_dmd_*n.n.n_yyyymmdd*000001.zip but will change to indicate the version of the dm+d data

   Note: the .zip ending may not be visible depending on your settings
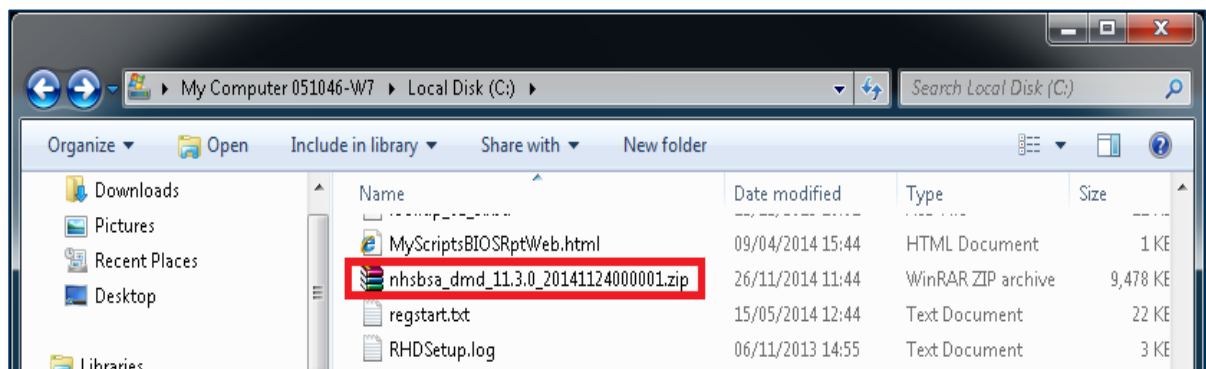
Figure 4 - Example showing the NHSBSA dm+d file

2. Copy the zip file into in to C:\dmd_extract_tool\XMLToUnzipInHere folder. If there are any old zip files in the folder you will need to delete them.
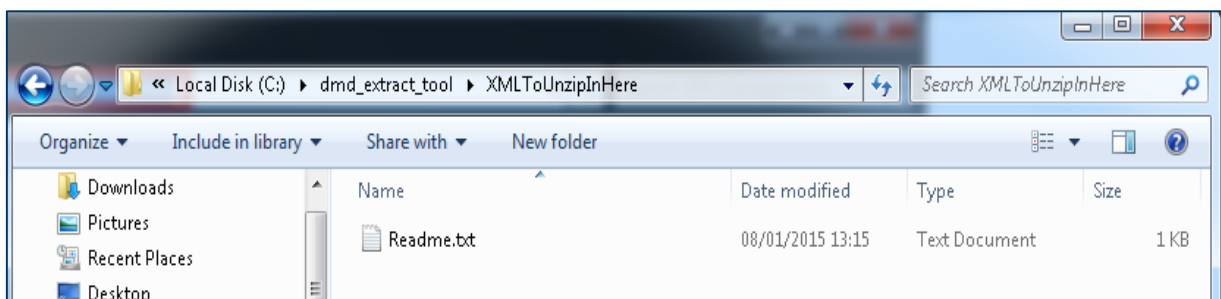

Figure 5 - The empty XML To Unzip In Here folder
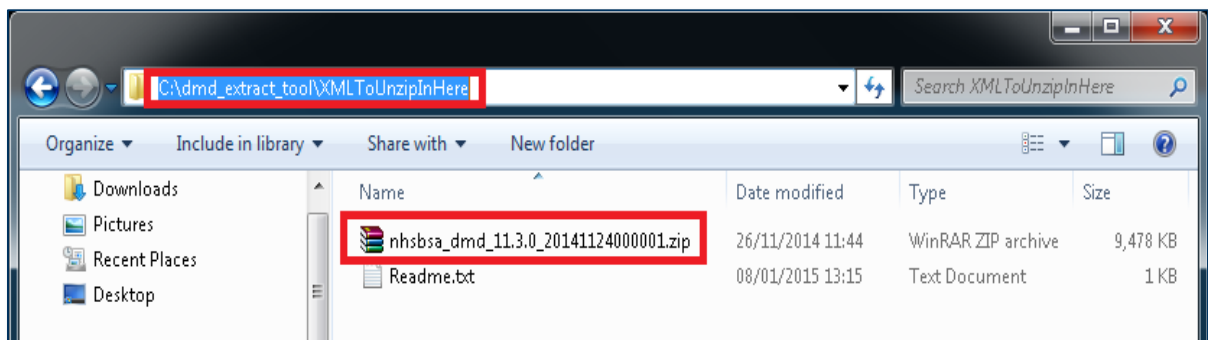
3. Zip file now added to the folder

Figure 6 - The XML to Unzip In Here folder with the NHSBSA dm+d file

    a. Optional – if you want to transform the supplementary files as well, copy the zip file (for example nhsbsa_dmdbonus_9.1.0_20150914000001.zip) into the XMLToUnzipInHere folder as well)

4. Navigate back to C:\dmd_extract_tool and double click the batch file called **xml-csv.bat.**
This will open a command prompt and will create the excel files in C:\dmdDataLoader. This process will take few minutes to run.
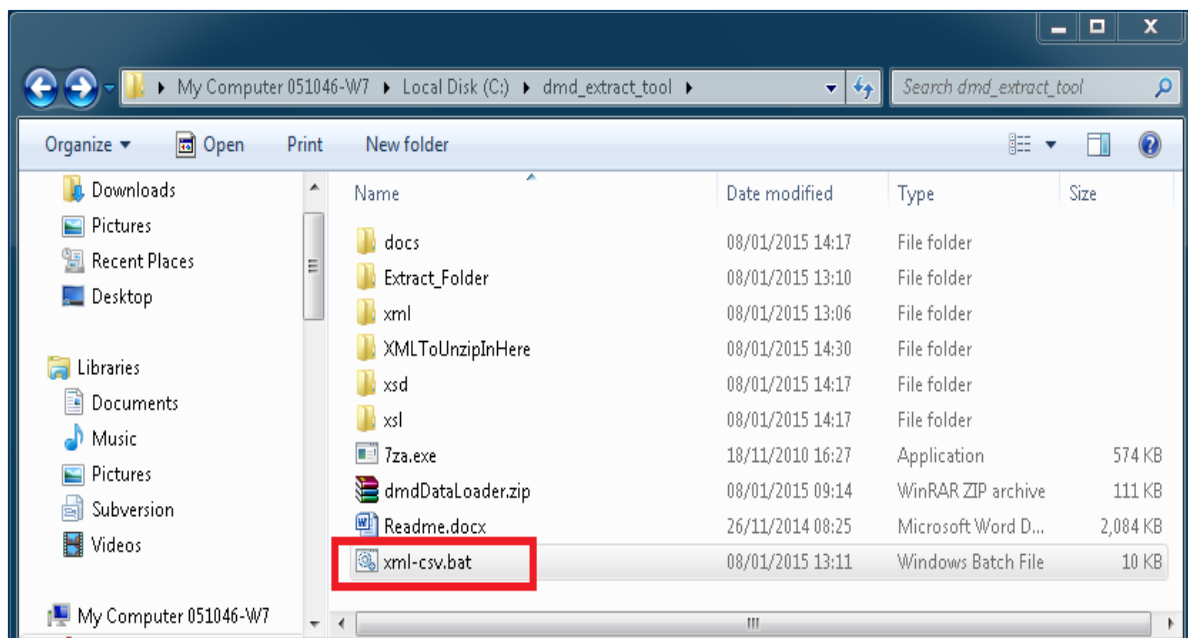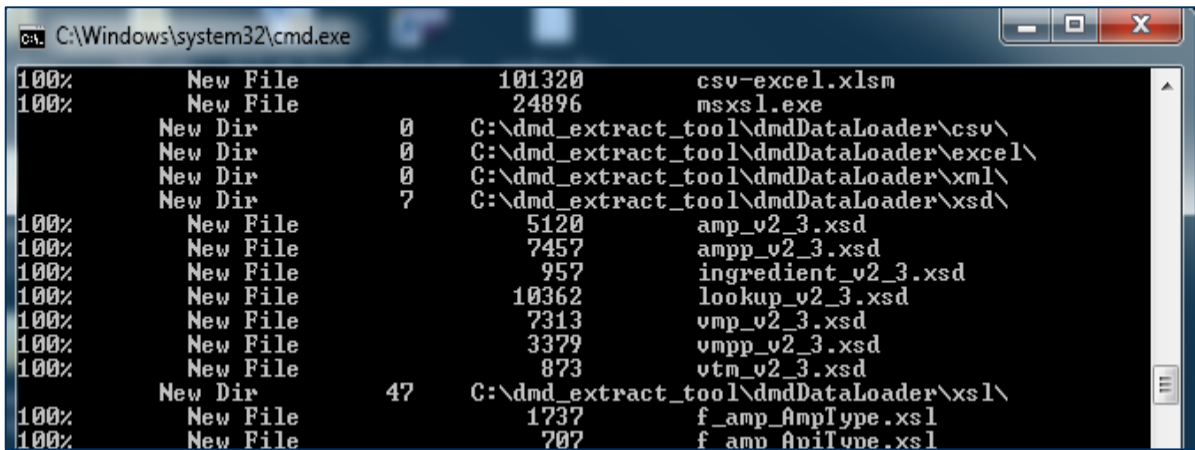

Figure 7 - Location of the xml-csv.bat file

Figure 8 – Command prompt window showing transformation in progress
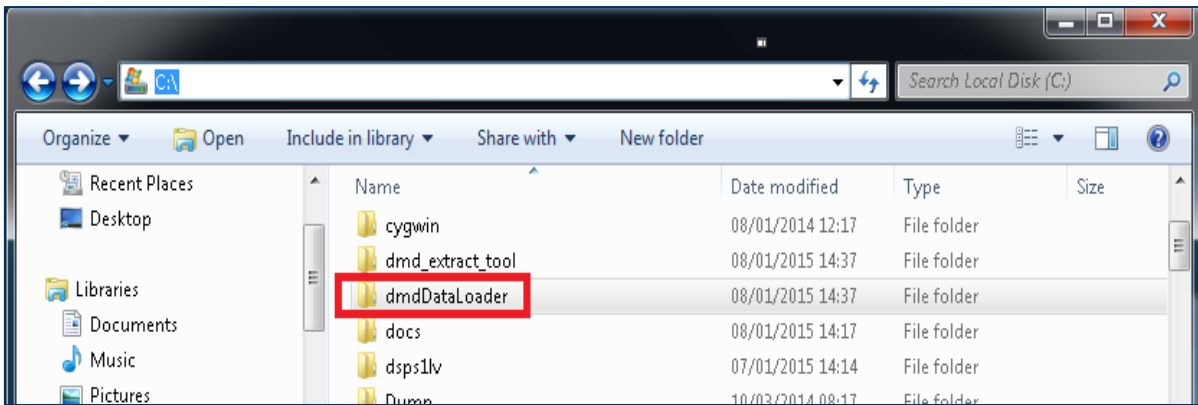

Figure 9 - Location of the dmdDataLoader folder in the C:\ drive

5. During the execution of the batch file **xml-csv.bat** Microsoft Excel will open, select the <mark>**Enable Content**</mark> button so the macro in the excel worksheet can run and load the newly created csv files.
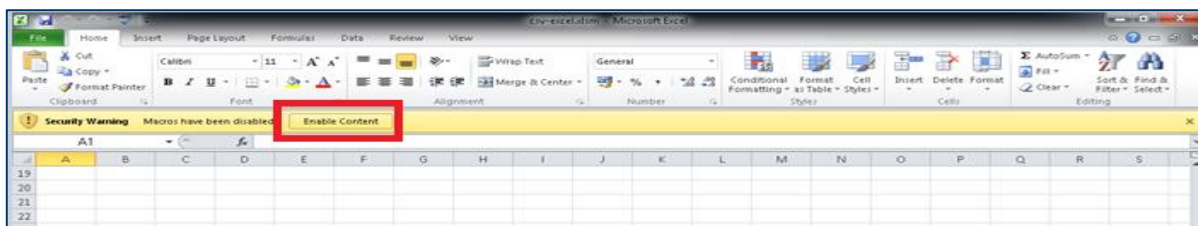

Figure 10 – 'Enable Content' button in Excel

6. You will see Excel opening and closing several times as it creates the workbooks and populates the various worksheets.

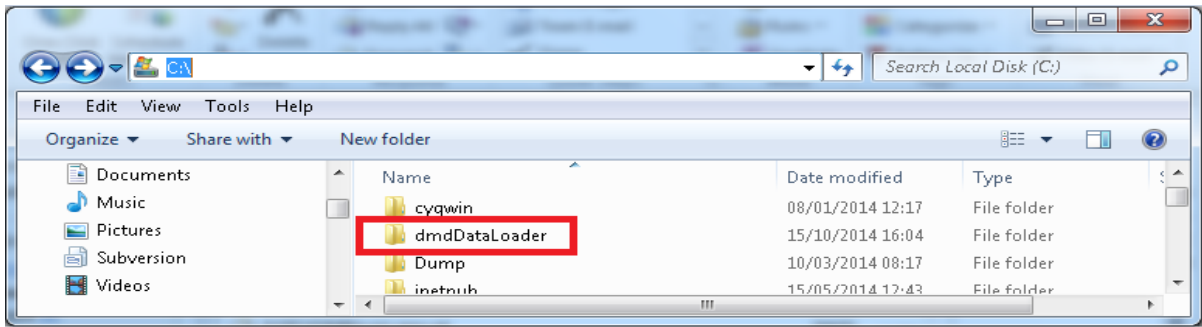7. Once Excel has finished both the Excel workbooks and CSV files can be found in the C:\dmdDataLoader

Figure 11 - The dmdDataLoader folder in the C:\ drive

8. the 7 workbooks can be found in:
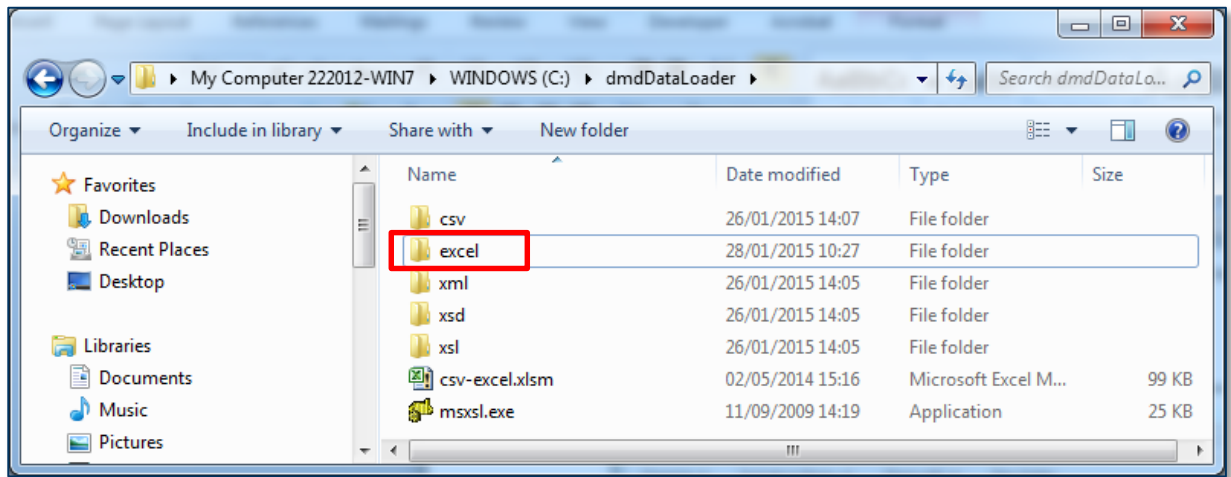
C:\dmdDataLoader\excel

Figure 12 - The Excel folder in C:\dmdDataLoader

a. *Optional* – if you have also transformed the Supplementary files you will see two additional workbooks called f_bnf.xlsx and f_gtin.xlsx
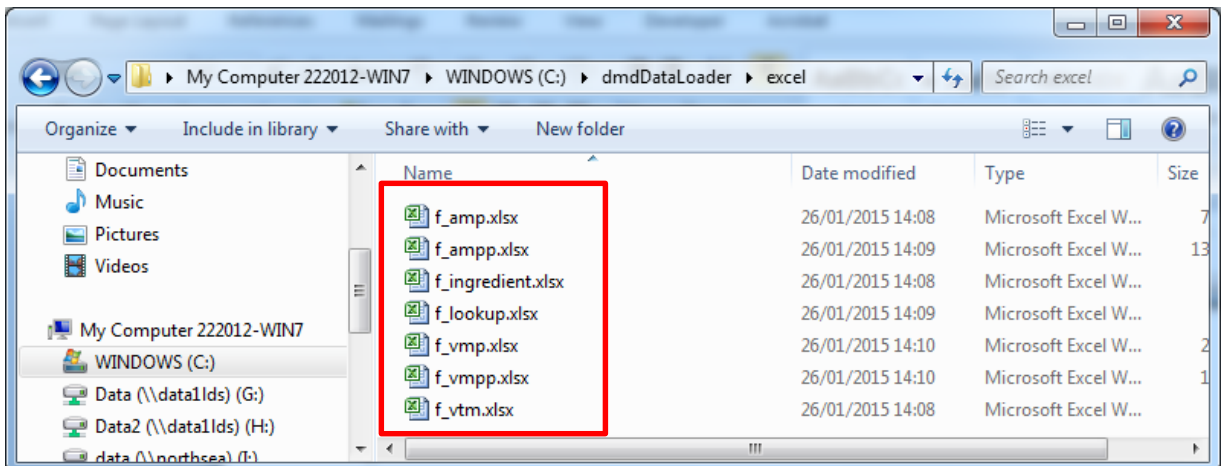

Figure 13 - The Excel folder showing the Excel workbooks

9. The 47 CSV files can be found in

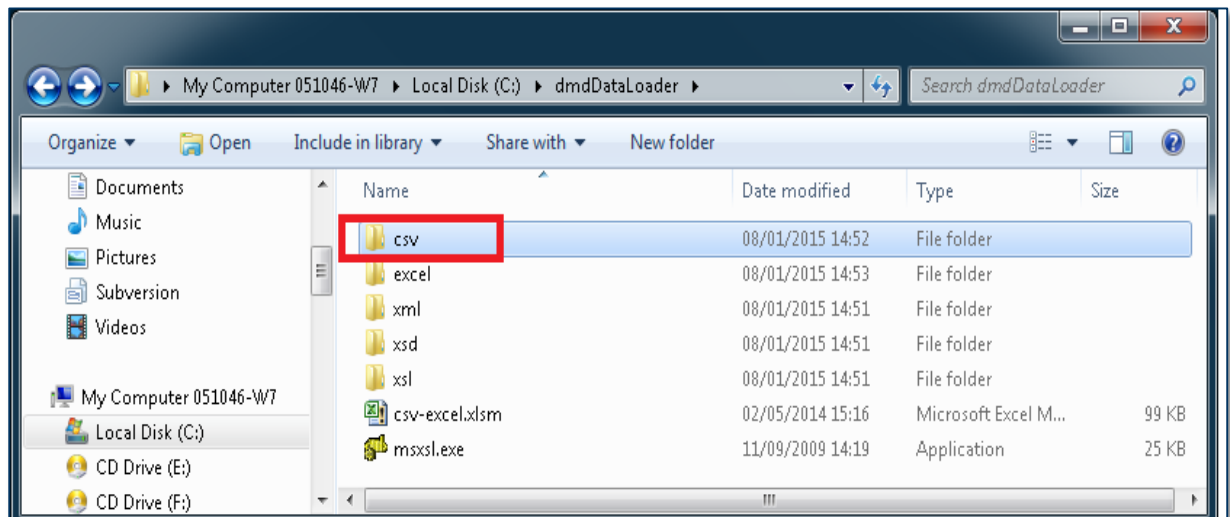   *C:\dmdDataLoader\csv*

Figure 14 - The CSV folder in C:\dmdDataLoader

a. *Optional* - if you have also transformed the Supplementary files you will see 50 CSV files
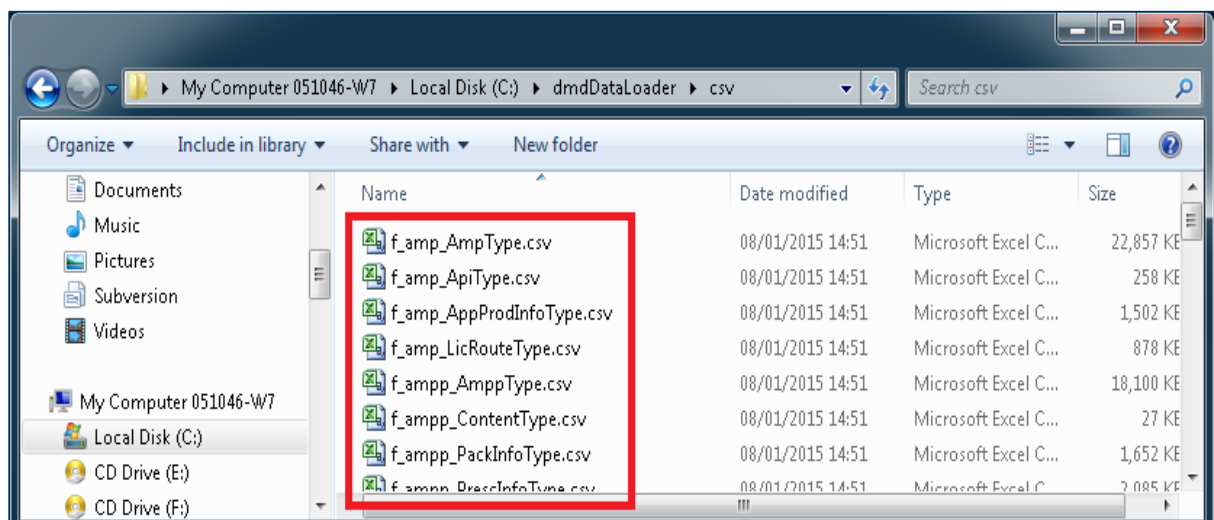

Figure 15 - The CSV folder showing some of the CSV files

If you wish to run the tool again on a different dm+d release you will need to copy any files you wish to keep from the 'excel' and 'csv' folders to a different location as they will be overwritten.