Processing AMED records

### Overview

amed\_tools.exe is used to process ETOC records before import to Excel.

Processed records are then imported to the spreadsheet AMED processing.xlsm, where index terms can be added.

The records are exported from Excel, and processing is completed using amed\_tools.exe.

### Processing ETOC records before import to Excel

#### Set up executable files

Copy the following files from [\\ad.bl.uk\bssdata\Utilities\VM tools\amed](file:///\\ad.bl.uk\bssdata\Utilities\VM%20tools\amed) to the folder in which you are working:

* amed\_tools.exe

Ensure that the following files are present in the folder [\\ad.bl.uk\bssdata\AMED\Exports\Lookup](file:///\\ad.bl.uk\bssdata\AMED\Exports\Lookup):

* AMED journal title lookup table.txt
* Amed\_citations.db

#### Processing

From the command line, run:

amed\_pre\_Excel.exe –i <input\_file> --E1

where <input\_file> is the name of the input file. The input file should be a text (.txt) file with one ETOC record per line, as in the example shown in the image below:

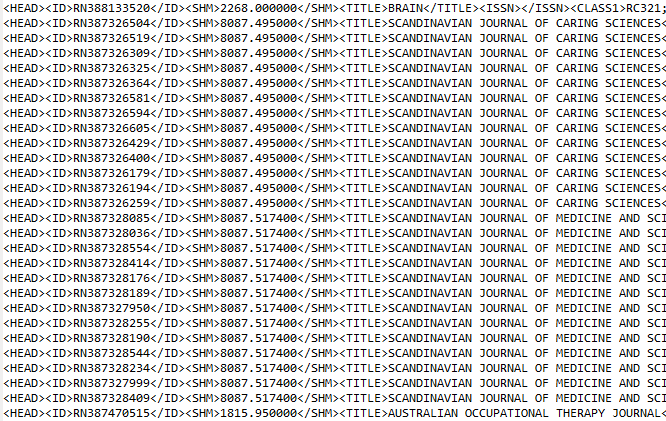
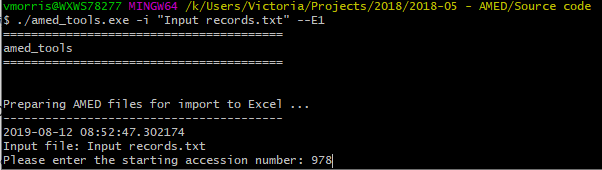
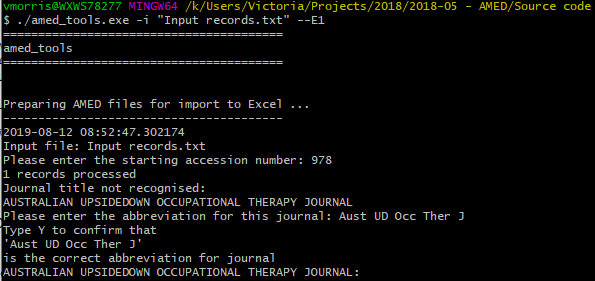


Figure 1: Extract from example file of ETOC records

You will be prompted to enter the accession number for the first record to be processed. Type an integer, then press Enter. The program will automatically pad this number with leading zeros to ensure that the length is correct.



If any journal titles (within ETOC field <TITLE>) are not recognised, you will be prompted to enter abbreviated title(s). These new abbreviated titles will be added to the list of journal titles in the file AMED journal title lookup table.txt once processing is complete.



Citations will be compared against the list stored within the database amed\_citations.db. 20 from the title will be appended to the citation in order to check for duplication.

Records found to be duplicates will be written to the output file amed\_as\_tsv\_duplicates.tsv. None-duplicates will be written to the output file amed\_as\_tsv.tsv.

Once the program has run, you should check the file AMED journal title lookup table.txt to make sure that new journal titles have been added correctly. It is important that the file ends with a blank line.

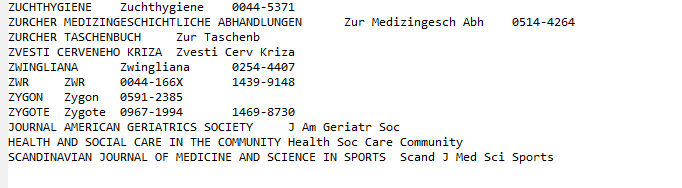


Figure 2: Extract from AMED journal title lookup table.txt showing new journal titles added

The database amed\_citations.db should be checked periodically, using software such as [DB Browser for SQLite](https://sqlitebrowser.org/).

### Processing AMED records in Excel

#### Set up files

Copy the following file from [K:\Utilities\VM tools\amed](file:///\\ad.bl.uk\bssdata\Utilities\VM%20tools\amed) to the folder in which you are working:

* AMED processing.xlsm.

Ensure that the file amed\_as\_tsv.tsv created in the previous step is present in the same folder.

#### Processing

Open the Excel spreadsheet AMED processing.xlsm.

In the Control sheet tab, click Clear index terms to clear any existing data from the workbook.

Next, click Refresh AMED records, to import the records from amed\_as\_tsv.tsv into the workbook.

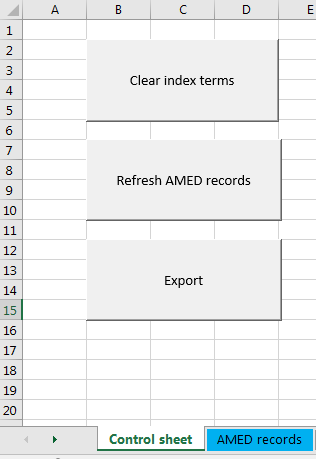


Figure 3: Buttons in the Control sheet

Records will be imported ino the AMED records tab.

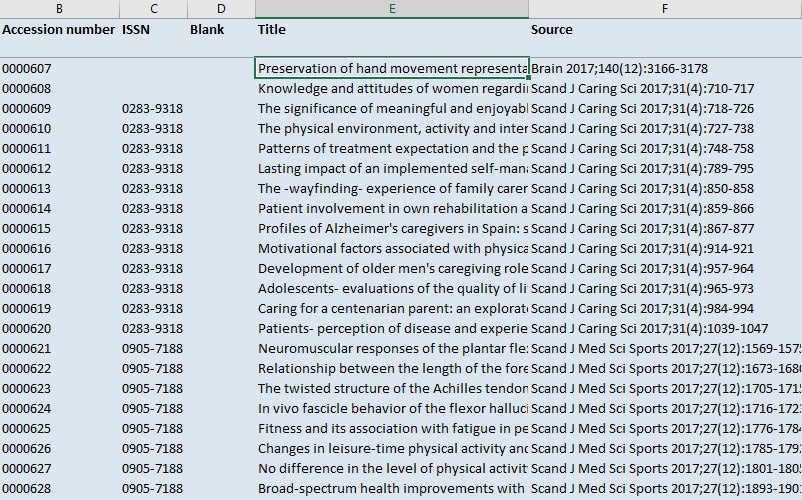


Figure 4: Extract from AMED records

The blue columns contain imported data. Scroll to the right to see green columns, into which index terms can be added.

To aid indexing, if you click on a green cell, the title, citation and abstract from the selected row will display in column A, at the left-hand side of the screen.

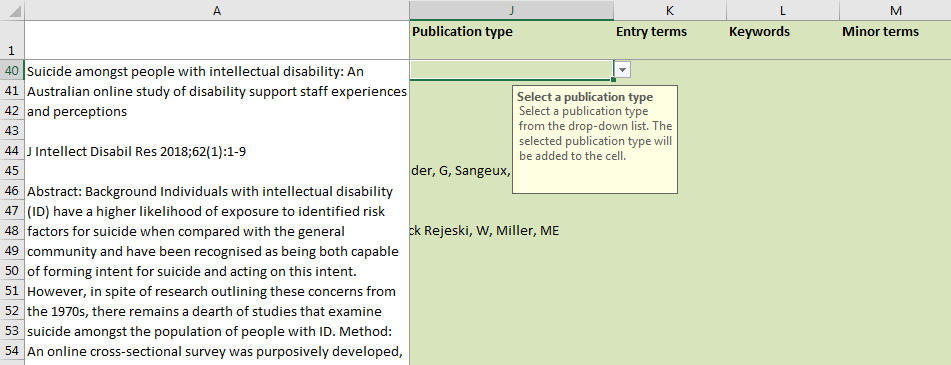
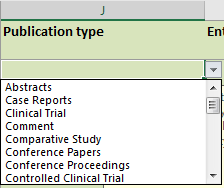


Figure 5: Article title, citation and abstract display in column A

Where index terms are taken from a controlled vocabulary, these terms will can be selected from a drop-down list. Terms selected will be ***added*** to the cell; this enables multiple terms to be selected. To remove terms, edit the cell value in the formula bar.



Controlled vocabularly lists are populated within the tabs AMED thesaurus, Publication types and Languages; these lists can easily be edited as necessary.

Once indexing is complete, return to the Control sheet tab, and click the Export button. The records from the AMED records tab will be exported (in UTF-8 encoding) to a file with a name of the form:

AMED processing output YYYYMMDD.txt

where YYYY MM DD are components of today’s date.

### Processing AMED records exported from Excel

#### Processing

From the command line, run:

amed\_post\_Excel.exe –i <input\_file> --month=<month\_number> --E2

where <input\_file> is the name of the input file, produced in the previous step, and <month\_number> is the number of the month to which the file applies (e.g.--month=04 to process the file for April). The input file should be a text (.txt) file exported from Library Master. If <month\_number> is not specified, the program assumes that the file applies to the current month.

The following output files will be produced:

* amdmonthYY.txt
* amedMMYY for hosts.txt
* AMEDmonth.spl
* F164MMDD.dat
* F164.end
* AMED stats YYYY-MM-DD.txt

where YYYY MM DD are components of the date, and month is a three-letter abbreviation for the name of the month (lower-case).

The file AMED stats YYYY-MM-DD.txt contains a brief summary of the numbers of records processed, and range of accession numbers used.

The fields contained in other files are summarised in the table below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | **Present in … ?** | | | |
| **amdmonthYY.txt** | **amedMMYY for hosts.txt** | **AMEDmonth.spl** | **F164MMDD.dat** |
| ***Output file for*** | | SilverPlatter and Ovid |  | Datastar, Dimdi and EBSCO | Dialog and Dialog2 (also with F164.end) |
| ***Long fields wrapped?*** | | Y | Y | N | Y |
| **Field** | **Description** |  |  |  |  |
| AN | Accession number | Y | Y | Y | Y |
| UD | Date | Y | N | N | N |
| AU | Authors | Y | Y | Y | Y |
| TI | Title | Y | Y | Y | Y |
| SO | Source | Y | Y | Y | Y |
| ET | Entry terms | Y | Y | Y | Y |
| KW | Keywords | Y | Y | Y | Y |
| MT | Minor terms | Y | Y | Y | Y |
| TY | Publication type | Y | Y | Y | Y |
| LA | Language | Y | Y | Y | Y |
| ES | English summary indicator | Y | Y | Y | Y |
| IS | ISSN | Y | Y | Y | Y |
| MD | Abstract indicator | Y | Y | Y | Y |
| AB | Abstract | Y | Y | Y | Y |

### Source code

Code is written in Python, version 3.7, and source code is saved in the folder [\\ad.bl.uk\bssdata\Utilities\VM tools\amed\Source code](file:///\\ad.bl.uk\bssdata\Utilities\VM%20tools\amed\Source%20code).