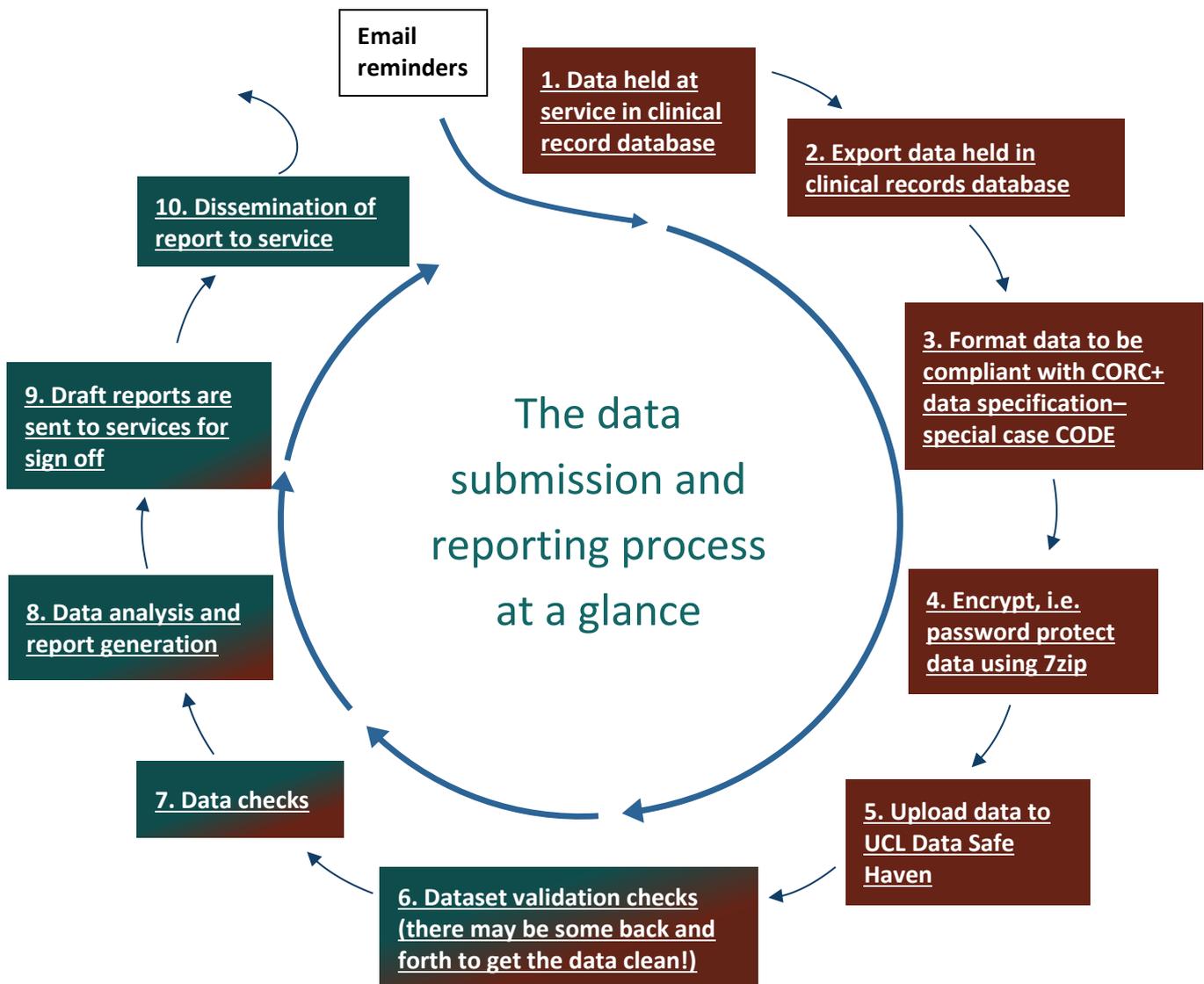


CORC+: an end-to-end guide to data submission and reporting

This end-to-end guide outlines the entire process of exporting and preparing data for submission and the data checks, analysis and report writing that follow. The last few pages of the document include a glossary of terms used throughout the document and appendices.



Items in dark orange are completed by the service; items in turquoise are done by CORC team

You will be allocated a date by which you need to submit your data. We will send you a reminder a month prior to when your submission is due. Once you have submitted the data we will validate it, and confirm the dataset with you. Please note that if data is not submitted in time, CORC reserves the right to reallocate your report and as a result CORC may not be able to deliver a draft report within the stated timeframe.

1. Data held at service in clinical record database

A clinical records database is a place where you store information about the children and young people who engage with your service. These files may be kept in paper format or they can be stored electronically. It may also be that your service utilises a hybrid system where, for example, the routine outcome measures (ROMs) that service users complete are on paper, but their contact details and appointments are in electronic format.

2. Export data held in clinical records database

Exporting data from your clinical records system is typically done by the database or data manager. Some systems allow for data to be transformed and put into the format defined in the [CORC+ data specification](#) at the same time as being exported. Instructions on how to export data from CODE are given in appendix 1. Other systems, do not transform data in the same format as the CORC+ data specification, compliant with the CORC+ variable definitions.

3. Format data to be compliant with CORC+ data specification

After exporting your data, check through the data files to ensure the data is in the format compliant with CORC+ data specification. It might be the case that you will have to format your data once exported.

Examples of formatting that may need to be done are, reorganising data into different files with the right number of columns and in the correct order, putting [dates](#) into the correct format or making sure that [Service Allocated Patient IDs \(SAPIDs\)](#) link up demographic information with period of contact, event and questionnaire data.

Please also ensure that the data files do not contain any personally identifiable information (e.g. NHS numbers, dates of birth, names).

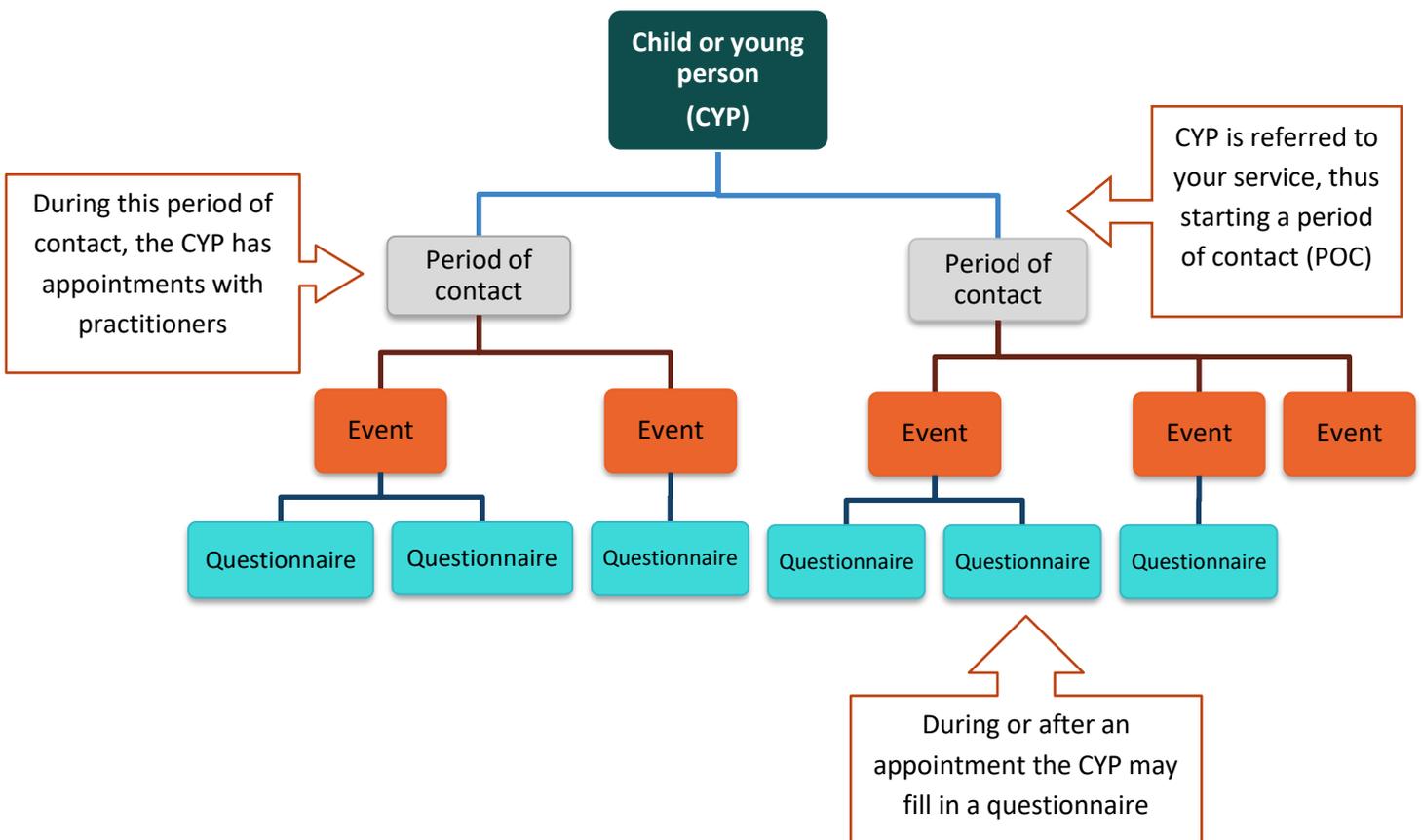
You can use Excel to re-format data, but as Excel auto-formats date and text columns, other free software that might be more appropriate are Notepad, [CSVed](#) or [Notepad++](#). The latter can optionally be used without installing the software. Alternatively, if using Excel to re-format data, please be sure to reformat date columns back to that specified in the CORC+ data specification (yyyy-mm-dd) as when Excel opens a file which contains dates, it will assume that they a) are dates, and b) that they should be stored as dd/mm/yyyy.

The fundamental structure of CORC+ data is to have one csv file containing demographic information (DEM file) of all patients submitted for analysis, one csv file containing information on all periods of contact (POC file) that all the children and young people listed in the DEM file had with the service, and one csv file containing information on all events (EVENT file) that occurred within each period of contact. Outcomes measures are submitted in separate files – i.e. one csv file for each questionnaire

that the children and young people listed in the DEM file completed during a period of contact. Please see the CORC+ data structure diagram below.

The minimum number of csv files you will need to submit is four and details of the number, order and name of columns per file can be found in the [CORC+ data specification](#).

CORC+ Data structure diagram



4. Encrypt (password protect) data using 7zip

Any data you send us needs to be password protected, and because the csv file format which we are asking you to use cannot be password-protected in itself, all your [data files](#) will need to be added to an archive and password-protected.

If you don't have 7zip installed, it is necessary to download and install it.

The software is free and compresses ('zips') as well as password protects your data. You can download the installer from here, <http://www.7-zip.org/>. Instructions on how to use 7zip to password protect your data can be found in [appendix 2](#).

5. Upload data to UCL Data Safe Haven

Once you are ready to submit your data, the CORC team will invite you to register onto the Data Transfer portal, which is hosted by the UCL Data Safe Haven facility.

The UCL Data Safe Haven is a facility hosted by the University College London (UCL) and is compliant with NHS data security requirements. Data uploaded or emailed via the Safe Haven is secure.

The steps in [appendix 3](#) guide you through the process of registering onto the system and uploading data to be shared with the CORC team.

Sometimes it is necessary for us to send some data back to you – we will use the Safe Haven to do that. The process is more involved than regular email in that you will receive a link to the secure email, rather than the email itself, and you can only access the email within a limited amount of time and with a one-off password which you can obtain from one of our team members via a phone call.

Once the data has been uploaded, please notify the CORC team and we aim to be in touch within five working days to let you know of any issues.

6. Dataset validation checks (there may be some back and forth to get the data clean!)

During [data validation](#) and [data cleaning](#) your data is checked for inconsistencies and illegitimate values. This can be quite time-consuming as even issues which seem quite minor (such as a value reading 3, when it should read 03) have a big impact on whether or not the data gets through validation. It pays off in the long run to read the data specification carefully and ask any questions that arise before you make your first submission. Once data has been checked for inconsistencies, illegitimate values and errors, these issues will be outlined to you in an email so that you can iron them out before resubmitting your data.

Typical errors that crop up are [duplicate records](#), [orphan records](#), [date formatting](#) errors, and [file formatting](#) errors. We are able to provide issue guides on request at the time of validation and cleaning to deal with these typical errors that occur.

Once resubmitted, the data is checked again. This cycle of validation checks by the CORC team and cleaning by the service continues until data is of a standard with accepted structure, values and formats.

7. Data checks

Once your data has been validated, the team produces [data checks](#) for you to sign off within two weeks. These are summaries of the number of cases we have for each data item that will appear in your report. You will also be asked to fill out a pro-forma, which provides the CORC team with background information on your service and what you hope to gain from the report.



8. Data analysis and report writing

Once we have confirmed the data checks with you, and we have received the pro-forma, report generation can go ahead.

9. Draft report sent to service for sign off

When you receive the draft report, you then have the opportunity to feedback any comments or minor amendments within one week, before we produce the report in final copy. If no amendments need to be made the final version of the report will be generated.

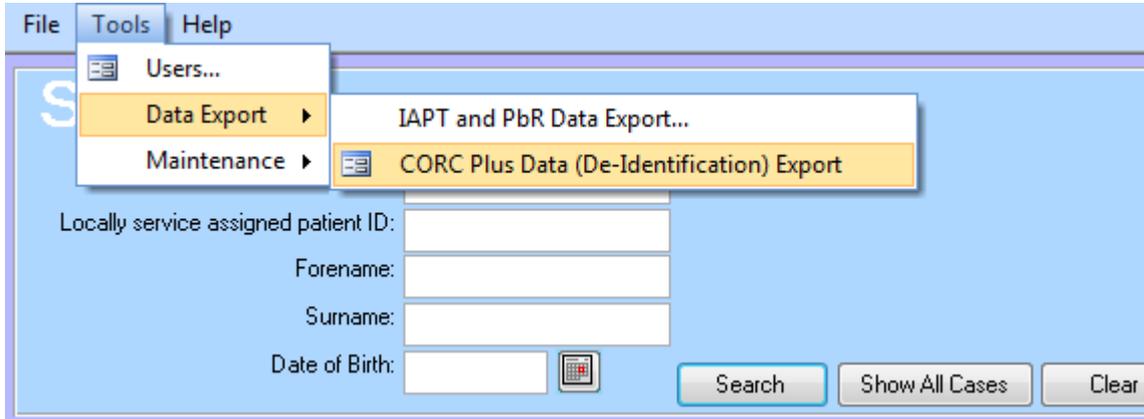
10. Dissemination of report to service

Once you have received the final version of the report, we are unable to make any adjustments.

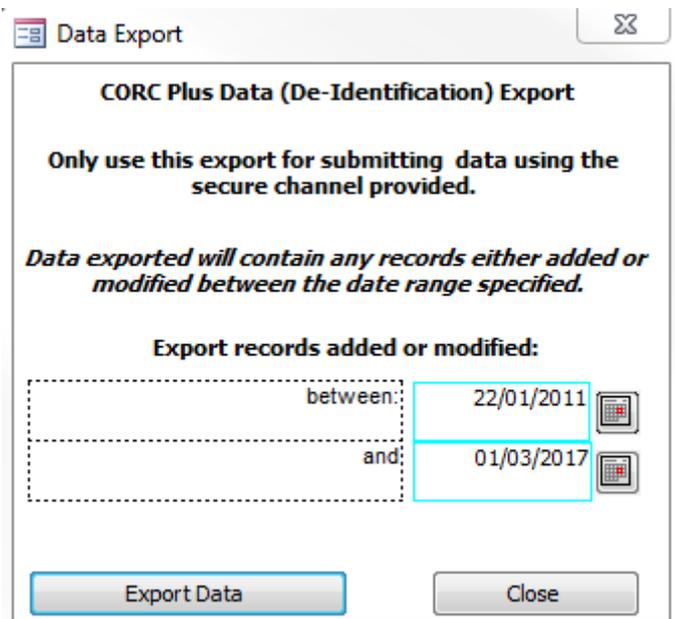
Appendix: 1

Exporting Data from CODE

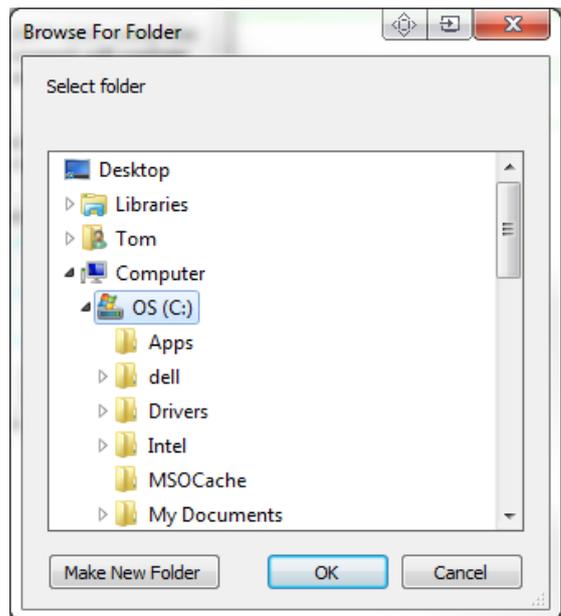
Click on Tools > Data Export > CORC Plus Data (De-Identification) Export



When prompted, please enter the date range you want your data to cover.



Select the folder into which you would like to export the data. Make sure that you are able to save files in that folder.



You can export all data held, or filter by specific dates. For submission to CORC we recommend exporting all data you hold to avoid missing out any data

Appendix: 2

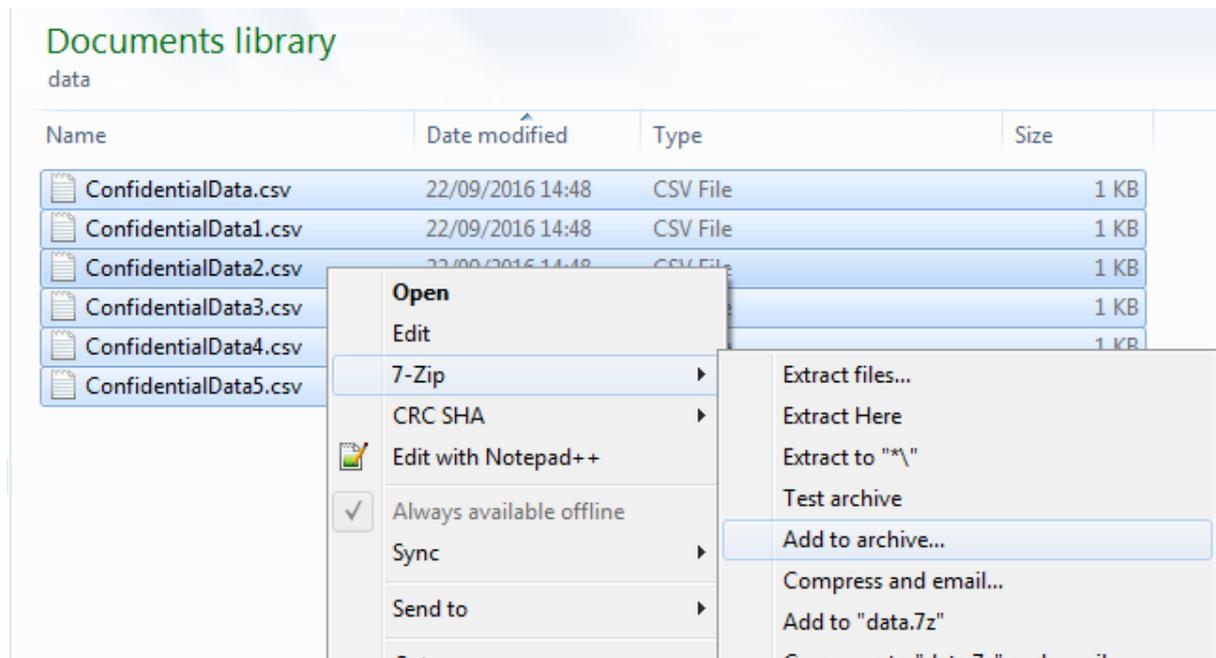
Encrypting Confidential Data

Browse to the file, files or directory that you wish to encrypt and select them.

Documents library
data

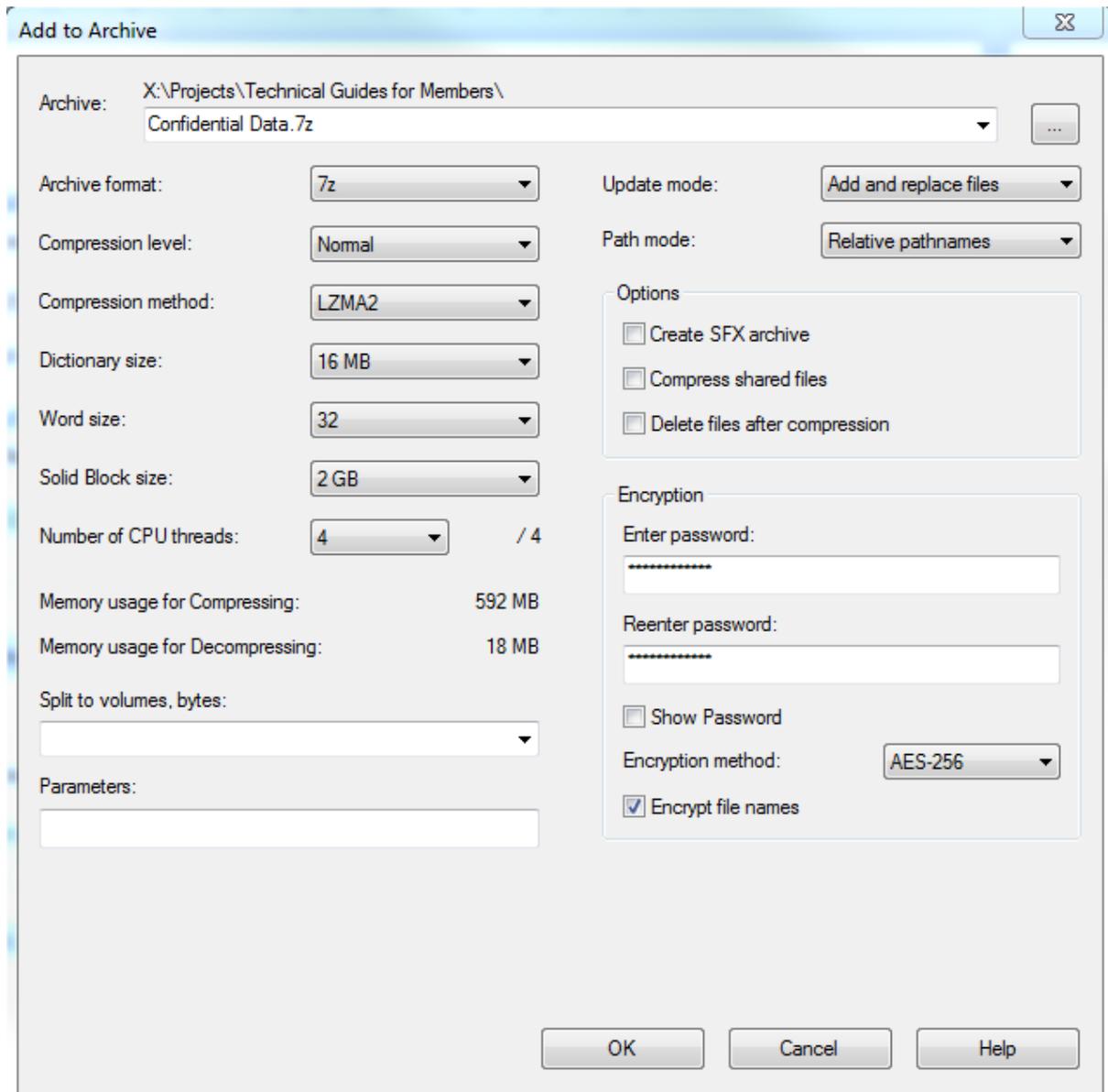
Name	Date modified	Type	Size
ConfidentialData.csv	22/09/2016 14:48	CSV File	1 KB
ConfidentialData1.csv	22/09/2016 14:48	CSV File	1 KB
ConfidentialData2.csv	22/09/2016 14:48	CSV File	1 KB
ConfidentialData3.csv	22/09/2016 14:48	CSV File	1 KB
ConfidentialData4.csv	22/09/2016 14:48	CSV File	1 KB
ConfidentialData5.csv	22/09/2016 14:48	CSV File	1 KB

Right click on the selected items and then select 7-zip -> Add to archive...



It is recommended that the following options are selected.

- Archive format: 7z
- Encryption method: AES-256
- Encrypt file names (not available for zip archives)



You should then have a compressed and encrypted archive of your data.

Documents library
data

Name	Date modified	Type	Size
ConfidentialData.csv	22/09/2016 14:48	CSV File	1 KB
ConfidentialData1.csv	22/09/2016 14:48	CSV File	1 KB
ConfidentialData2.csv	22/09/2016 14:48	CSV File	1 KB
ConfidentialData3.csv	22/09/2016 14:48	CSV File	1 KB
ConfidentialData4.csv	22/09/2016 14:48	CSV File	1 KB
ConfidentialData5.csv	22/09/2016 14:48	CSV File	1 KB
confidentialdata.7z	22/09/2016 15:00	7z Archive	1 KB

Appendix: 3

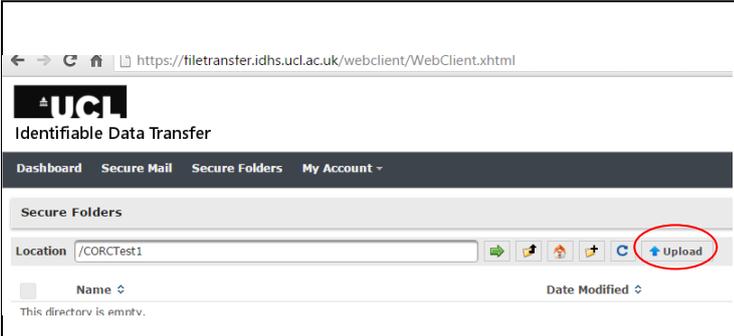
Registering with the UCL Data Safe Haven

1) Register onto the portal

<p>Invitation Request</p> <p>You have been invited to join UCL IDHS File Transfer. Please click the link below to continue with the registration process. Click here to register</p> <p>If the link above does not open, copy and paste the following URL into your browser: https://filetransfer.idhs.ucl.ac.uk/register?token=13905b89-3b27-4d4c-8885-7b69867c5e86&invited=true</p> <p>This link will expire after 24 hours. Please contact your account representative if you have any trouble registering your account.</p>	<ol style="list-style-type: none"> 1. You will receive an automatic email inviting you to register to the Data Transfer portal. Please check your junk mail folder if you don't receive this. <p>Note: This link expires after 24 hours</p>
<p>Please enter the information below and click Submit.</p> <p>Login Credentials</p> <p>User Name * <input type="text" value="exttester13"/></p> <p>Password * <input type="password" value="••••••"/></p> <p>Confirm Password * <input type="password" value="••••••"/></p> <p>Contact Information</p> <p>Email Address <input type="text" value="b.varsani@ucl.ac.uk"/></p> <p>First Name <input type="text" value="Bhavesh"/></p> <p>Last Name <input type="text" value="Varsani ExtTester13"/></p> <p>Organization <input type="text"/></p> <p>Phone <input type="text"/></p> <p><input type="button" value="Submit"/> <input type="button" value="Cancel"/></p>	<ol style="list-style-type: none"> 2. When you follow the link, you will be prompted to create login credentials and register your contact information. 3. Click on Submit
<p style="text-align: center;"></p> <p style="text-align: center;">Your registration is complete and is currently pending approval. A notification will be sent to your email address when you are approved.</p> <div style="text-align: center;"> <p>LOGIN</p> <p>User Name <input type="text"/></p> <p>Password <input type="password"/></p> <p><input type="button" value="Login"/> <input type="button" value="Reset"/></p> </div>	<ol style="list-style-type: none"> 4. You can now log in to the Data Transfer portal on https://filetransfer.idhs.ucl.ac.uk/webclient/Login.xhtml

2) Upload Files

<p>Encrypt the files you would like to send.</p>	<p>See appendix 1</p>
<p>Log in to the Data Transfer portal</p>	<p>https://filetransfer.idhs.ucl.ac.uk/webclient/Login.xhtml</p>

 <p>The screenshot shows a web browser window with the URL https://filetransfer.idhs.ucl.ac.uk/webclient/WebClientL.xhtml. The page header includes the UCL logo and the text 'Identifiable Data Transfer'. Below this is a navigation bar with 'Dashboard', 'Secure Mail', 'Secure Folders', and 'My Account'. The main content area is titled 'Secure Folders' and shows a 'Location' field with the value '/CORCTest1'. To the right of the location field is a toolbar with several icons, including a blue 'Upload' button which is circled in red. Below the toolbar, there are columns for 'Name' and 'Date Modified', and a message that says 'This directory is empty.'</p>	<ol style="list-style-type: none"> 1. Click on Secure Folders 2. Click on Upload. This will allow you to browse your documents and select the encrypted file(s) you want to upload. <p><i>Only encrypted data should be uploaded onto the Data Transfer Portal</i></p>
<p>Once the data has been uploaded, please notify the CORC team on CORC@annafreud.org and give us a call to let us know the password you used on 0207 443 2225.</p>	

Frequently Asked Questions

- 1) I have got as far as registering and creating a folder but the file is not uploading – the system still says ‘file uploading’ after 15 mins.**

It’s not uncommon for NHS workstations to be running very old browser versions - particularly with Internet Explorer, which is tied to older versions for various NHS internal systems. The Data Safe Haven is very browser dependent, with older browser versions providing limited, if not no functionality with DSH services. Please try using a different browser or different computer if possible. Preferred options are Firefox Mozilla, Google Chrome, or Internet Explorer version 11 or higher. If the upload still doesn’t work, please report it to the CORC team.

- 2) I haven’t received the registration email.**

Please check your spam/junk mail folders. If it’s not there, please let the CORC team know.

- 3) I have been trying to login to submit our data, however it is coming up with Invalid user name and/or password.**

If you are sure you are using the correct username and password and it is still not letting you log in, please get in touch with us. We will re-issue the invitation to register, and this will allow you reset the password and login.

Glossary

- data** - Any recorded observations or measurements about the children and young people you see in your service.
- data set** - A collection of observations or measurements. May consist of more than one table of data, which can be linked together by a common field (e.g. PersonID). Note that the term ‘data set’ may also be used to refer to an *output specification* (e.g. on the HSCIC website).
- person-level data set** - Contains separate rows (or ‘records’) of data for each person. This is different to aggregate data, in which the data relating to individual people cannot be distinguished.

- data file** - A collection of observations or measurements that can be accessed using a computer and may comprise all, or part of, a *data set*.

- data set specification** - A collection of definitions of variables/fields and their attributes and relationships to one another. E.g.:

<i>Data item name</i>	<i>Data item description</i>
PersonID	A unique number that identifies a person within the data set
Name	The person’s first name
Height	The person’s height in millimetres

A data set specification defines what a data set should consist of, and aims to make it as complete, consistent and accurate as possible.

- data checks** - Refers to a process by which the researcher working on your report confirms with you the frequencies of cases and data items in your data

- data cleaning** - Refers to a process by which illegitimate values are transformed into accepted values

- data items** - unit of data contained in a record, describing a particular attribute, for example, SAPID

- date format** - The way a date is arranged or set out. There is a large variety of formats for dates in use, which differ in the order of date components , component separators, whether leading zeros are included and whether all four digits of the year are written.

Note: CORC+ files should use a date format of “yyyy-mm-dd”.

- duplicate record** - Records which are identical for all data items.

- data validation** - The process of checking the extent to which data that have been submitted conform to the specification.

- file format** - a standard way that information is encoded for storage in a computer file. Examples are .jpg and .gif files for images, and .mp3 files for sound; but there are also standard file types for text (e.g. .txt), movies (e.g. .mov), and spreadsheet data (e.g. .xml).

Note: CORC+ data submission files should be comma separated value (.csv) files. See CORC+ data set specification for more details.

- orphan record** - records which have no higher-level related record, for example, a questionnaire record which has no related event record

- Service Allocated Patient IDs (SAPID)** - A SAPID is a unique sequence of characters (typically numbers, but could also be letters or a combination of both) that uniquely identifies a patient across your service’s data, in the same way as the NHS number uniquely identifies a patient across the NHS.