

Bring Your Own Data (BYOD) to the Health Data Platform BC (HDPBC)

Process for HDPBC Users

HDPBC Desktop provides projects (organizational or academic research) with the ability to import additional data, application code, documentation, etc. All incoming information will be de-identified, linkable when applicable, and human reviewed. Patient and provider names should be excluded from any imports to the platform for your project.

You can import the information into the secure desktop based on file size and content. The table below outlines the three types of import and qualifying criteria for each of the methods.

For any questions, reach out to MOHAnalytics@gov.bc.ca.

Type of Import	Qualifying Criteria	Transfer Method & Other Info
1. File Import – HDPBC File In Portal	<p>Use this method if your file:</p> <ul style="list-style-type: none"> Is under 24MB AND does not contain personal or sensitive information. The information will not be linked for analysis 	<ul style="list-style-type: none"> HDPBC File In/File Out Portal Imported by requestor
2. Cohort Import – Ministry of Health (MOH) Secure File Transfer Protocol (SFTP)	<p>Use this method if your file:</p> <ul style="list-style-type: none"> Contains personal or sensitive information (must be encrypted, with patient and provider names <i>excluded</i>) AND/OR is over 24MB. <p>The information may be linked to HDPBC data for project analysis if applicable.</p>	<ul style="list-style-type: none"> SFTP Imported by HDPBC Technical Support team Data file should be .csv format Data should be extracted with UTF-8 encoding Metadata information file to be included as .xls file
3. GitHub Repository Imports – Public URL	<p>Use this method to import git Repository.</p> <ul style="list-style-type: none"> The repository must be public. 	<ul style="list-style-type: none"> Public URL import Imported by HDPBC Technical Support team If application code is included in the Git Repository, the HDPBC Security Threat Risk Assessment (STRA) process may be required, which can delay the import of the repository

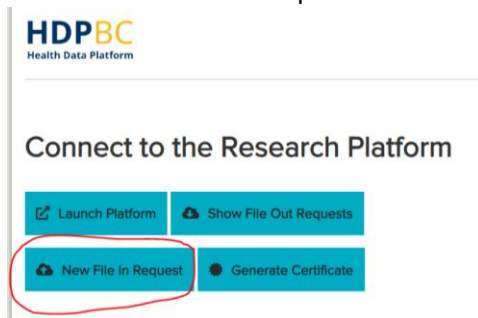
subject to completion of the STRA

1. File Import – HDPBC File In Portal

If your file does not contain sensitive personal information and is under 24MB, you may upload it yourself via the File In portal. For files over 24MB or that contain sensitive personal information, use [Cohort Import and Large Files](#).

Step 1

- **Using incognito browsing**, log in to the HDPBC Desktop: <https://portal-hdpbc.healthbc.org>
- Select “New File In Request”



- Drag and drop your file in. Click Upload Files.

The screenshot shows the 'Request Files In' interface. At the top, it says 'HDPBC Health Data Platform'. In the top right corner, there is a user profile icon and a 'Sign Out' link with the text 'Request new password'. Below this, there is a green notification bar that says 'File Uploads Will Be Automatically Approved' and a yellow warning bar that says 'The file name MUST be less than 200 characters'. The main heading is 'Request Files In'. Underneath, there is a 'Files' section with a large rectangular area containing the text 'Drop files here to upload'. Below this area, it says 'Please select one or more files'. There is a 'Documentation' section with a search box and a 'Select Project' dropdown menu. At the bottom, there is a blue 'Upload Files' button.

Step 2

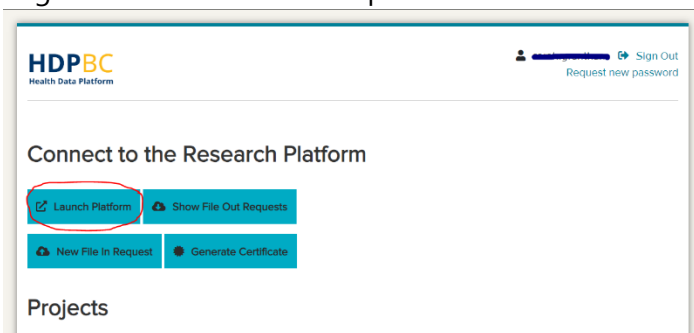
Wait for an email from the HDPBC Support Team.

HDPBC Support will scan the file for viruses and review the data for personal identifiers before releasing them for import to the HDPBC. You will receive an email notification after the files have been reviewed. If they've been approved for release, you can download them from within the HDPBC Desktop.

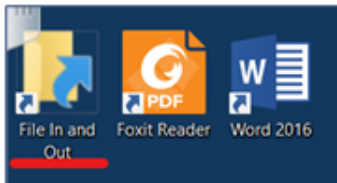
Step 3

Download the files from within the HDPBC Desktop.

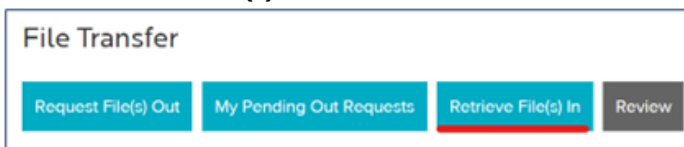
- Log in to the HDPBC Desktop.



- Open the **File In and Out portal** on the Desktop. Log in with your HDPBC credentials.



- Click **Retrieve File(s) In**



- You will see a list of your pending and approved file uploads. Click the view (eye) icon to access the file upload.

HDPBC Health Data Platform PHSABC\nikheel.premasagar Sign Out

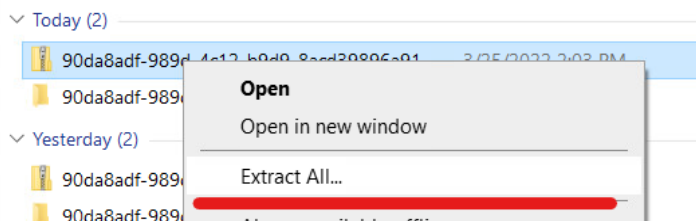
My File Uploads

Pending Requests

Posted	Project	Files	ApplicationStatus	View
02/03/22 15:43		1 (90.8 KB)	Pending	
26/11/21 09:43		1 (14.1 KB)	Pending	

Posted	Project	Files	ApplicationStatus	View
18/03/22 15:40	Testing	6 (868.6 KB)	Approved	
04/02/22 12:40	Testing	1 (95.3 KB)	Approved	
27/01/22 13:06	Testing	1 (27.2 KB)	Approved	

- Download each file individually by clicking the download icon in the request bar or download all the files at once by clicking **Download all as a ZIP**.
 - Your files will download to the "Downloads" folder
- Move the files to your P: drive to save them.
 - ! Files left in the downloads folder will not be saved once you log off from the HDPBC Desktop**
- If you downloaded all your files a one ZIP, right click the folder, select **Extract All**, and follow the prompts.



Step 4

Access the files in your P: drive.

- Open the File Explorer
- Select **This PC**
- Open your P: drive
(Your P: drive has your name in the folder name. Files saved in the P: drive will be available only to you.)

2. Cohort Import – MOH SFTP

IMPORTANT: Applicants must have obtained approval from the provider of the data being imported into the HDPBC to be used for linking and analysis prior to requesting to import it to the HDPBC for their project.

To apply privacy protection rules to the cohort data being imported to your project, the HDPBC Technical Support team needs to know which columns are **direct identifiers**, and **indirect identifiers**.

Direct identifiers, including those with null or proxy values, will be encrypted to facilitate linking of your cohort with your HDPBC project data in the platform. Please include additional descriptive columns to flag these values for use in your analysis, if appropriate.

Ensure that provider and/or patient names have been excluded from your file prior to transferring.

- **Direct Identifiers** are variables that can directly identify individuals, either individually or in combination. e.g., name, email address, IP address, home address, personal health number (PHN), social insurance number (SIN).
- **Indirect Identifiers** are variables that don't directly identify an individual but can be used for indirect re-identification. Whether a variable is an indirect identifier depends on the nature of the data.

Step 1: Prepare a Metadata Information for your cohort data

File: Cohort Metadata information

Format: .xls

Ensure your data is extracted with UTF-8 encoding to prevent the generation or inclusion of special characters that are not accepted by the HDPBC.

Cohort Metadata Information				
Name of File	Attribute Name	Data Type	Type of Information	Identifier Flag (direct or indirect)

Example populated Metadata Information Table:

Cohort Metadata Information				
Name of File	Attribute Name	Data Type	Type of Information	Identifier Flag

test_cohort_rpt_sample.csv	PHN	varchar(50)	Personal Health Number	Direct Identifier
test_cohort_rpt_sample.csv	Name	varchar(50)	First and Last Name	Direct Identifier
test_cohort_rpt_sample.csv	Diag_Code	int	Code representing diagnosis	
test_cohort_rpt_sample.csv	FSA_location	varchar(250)	First three digits of postal code	Indirect Identifier
test_cohort_rpt_sample.csv	home_address	varchar(250)	Home address	Direct Identifier
test_cohort_rpt_sample.csv	attend_time	float	Time spent on patient	
test_cohort_rpt_sample.csv	ICD_10_code	int	ICD 10 Code	
HDPBC_cohort_rpt.csv	PHN	varchar(50)	Personal Health Number	Direct Identifier
HDPBC_cohort_rpt.csv	Mother Name	varchar(50)	First and Last Name of Mother	Direct Identifier
HDPBC_cohort_rpt.csv	Diag_Code	int	Code representing diagnosis	

Step 2: Notify HDPBC Front Counter of Readiness

When your data is ready to transfer, email HDPBC Front Counter (MOHAnalytics@gov.bc.ca). Include your:

<ul style="list-style-type: none"> • Project name • Principal applicant name • Project number • SFTP/technical contact name 	<ul style="list-style-type: none"> • SFTP/technical contact phone number* • SFTP/technical contact email address
---	--

The HDPBC Front Counter will review your request and send it on to the HDPBC Support Team, who will support you in completing your import via MOH SFTP*.

*The MOH SFTP team will contact you via this phone number; please ensure you provide a number you can be reached at.

Step 3: Generate the SSH key

You will generate an SSH key pair with HDPBC so it can be registered on the SFTP server.

In this step, you will create a private SSH key and a public SSH key. The public SSH key will have a **.pub** file name extension. You will share your public SSH key with HDPBC when you submit your cohort data import request in Step 4.

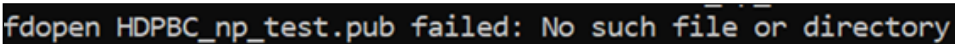
More information on [SSH protocol](#) and [SSH public key authentication](#).

You will start a command prompt from the desktop icon or by the start button, then you will be given specific text to type into the command prompt. This text will create a new SSH key folder, change the directory to the new folder, allow you to enter a file name, and prompt you to enter a passphrase. A passphrase helps protect the keys from unauthorized use.

It is recommended that you name the keys in the following format:
"**HDPBC_firstname_lastname**".

If you get an error:

If you see an error like that shown in Figure 1 and the generated .pub file is empty, you may need to try the operation again. When trying again, do not specify a file name when prompted.



```
fdopen HDPBC_np_test.pub failed: No such file or directory
```

Figure 1. SSH key creation error message.

You can rename the files after they have been created to follow the convention "**HDPBC_FirstName_LastName**".

Step 4: Submit Request to Import Cohort Data

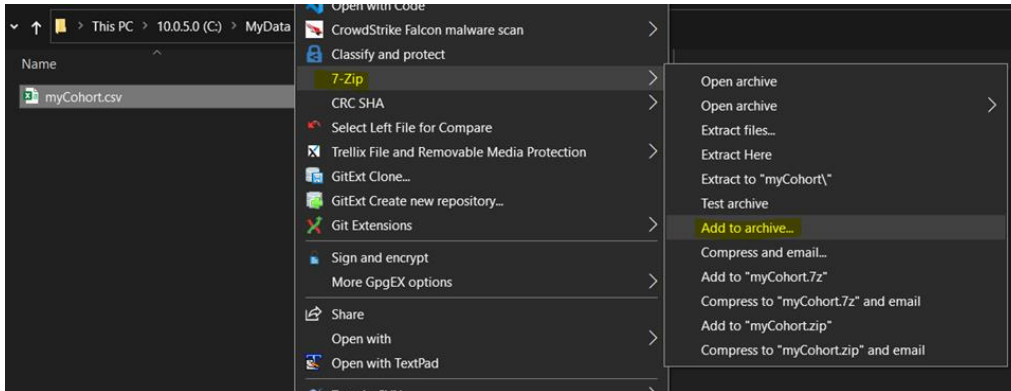
Send an email to MOHAnalytics@gov.bc.ca attaching your completed HDPBC cohort metadata information .xls file, and your **public** SSH key.

Step 5: Wait to receive an HDPBC encryption key

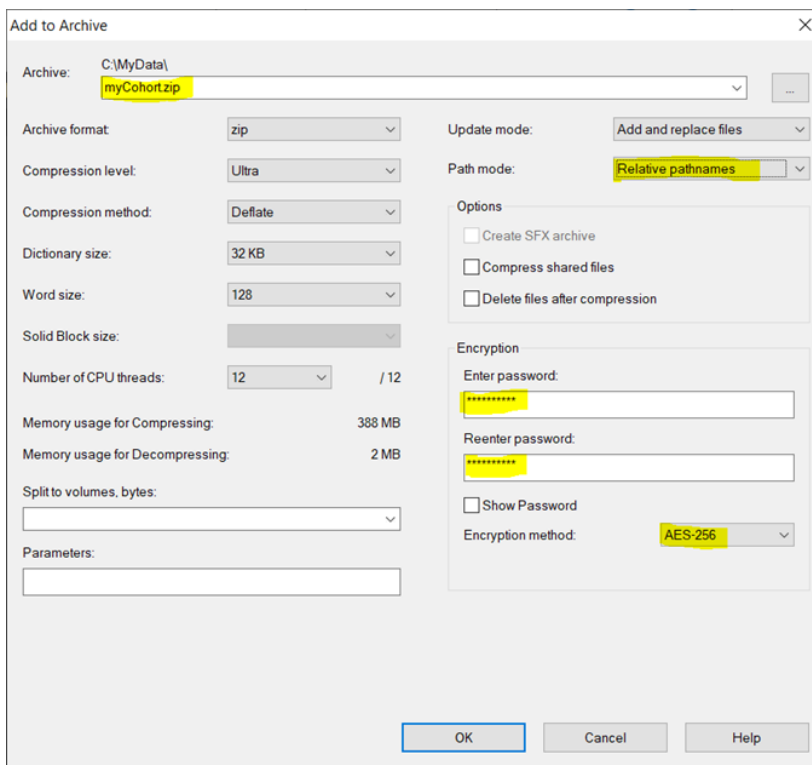
Following the submission of your metadata information and public SSH key, the HDPBC technical team will register your SSH key so that you can send data to HDPBC using the Secure File Transfer protocol (SFTP).

Step 6: Encrypt and compress your file

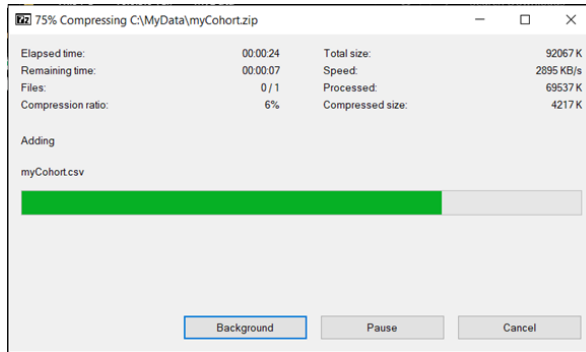
- Download [7-Zip](#) to encrypt and compress your file.
- Right-click on the file and select **7-Zip > Add to archive...** from the 7-zip submenu.



- Enter the file name and ensure **Path mode** is **Relative pathname**. In the **Encryption** section, enter a strong password under **Enter password** and under **Re-enter password**. Ensure **Encryption method** is **AES-256**. Leave the rest of the fields as default.



- Wait for the file to compress.



When complete, this is the file you will transfer via SFTP to HDPBC for your project, in addition to your metadata information table .xls file.

Step 7: SFTP Configuration

You will need to install SFTP Client. To do this, navigate to <https://filezilla-project.org> and click **Download** from the left-hand navigation menu (Figure 2). At this point, you will complete the download and installation steps relevant to your operating system.



Figure 2. FileZilla page. Press the highlighted **Download** button.

Step 8: Use SFTP to Transfer Your Files

In FileZilla, you can click the button at top left (see Figure 3) to start a new connection.

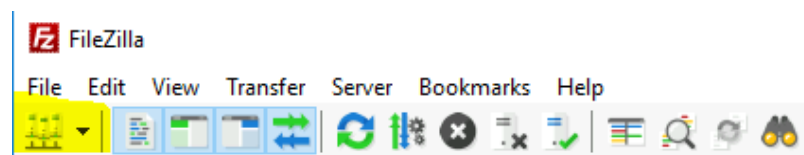


Figure 3. Click the highlighted button to start a new connection.

You will be prompted to fill in the connection information (see Figure 5). It should be completed as specified in Table 1.

Protocol	SFTP
Host	For users on a Health Authority network device: 10.57.149.10 For users outside HA networks (firewall modification): <i>sftpsvcs.healthbc.org</i>
User	<your username> e.g., jane.doe
Key File	1. Click Browse... Navigate to the private key file you created (the file without the .pub extension, e.g., NOT: HDPBC_firstname_lastname.pub) 2. In Logon Type select Key File
Port	Leave blank

Table 1. Connection information for FileZilla.

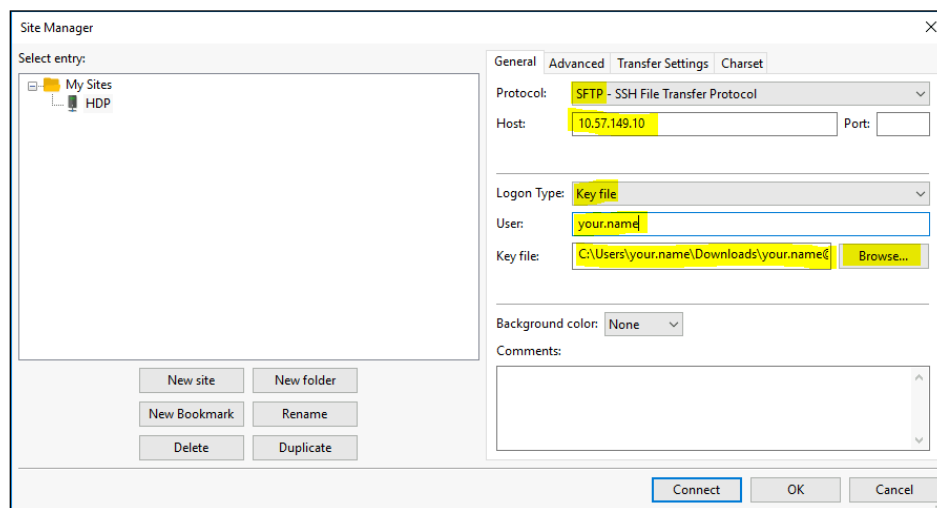


Figure 4. Enter the information as provided in the table above into the highlighted areas.

When you click **Connect**, you may see an “Unknown host key” error. If this happens, you can click **OK** to accept the warning.

If you are successfully connected, the “Remote site” panel to the right will be populated with the folders you have access to.

In the “Local site” pane on the left side of the page, you can navigate to the folder that contains your encrypted (.pgp) file. You need to drag the encrypted file from the **Local site** to the **Remote site** on the right side of the page (Figure 6).

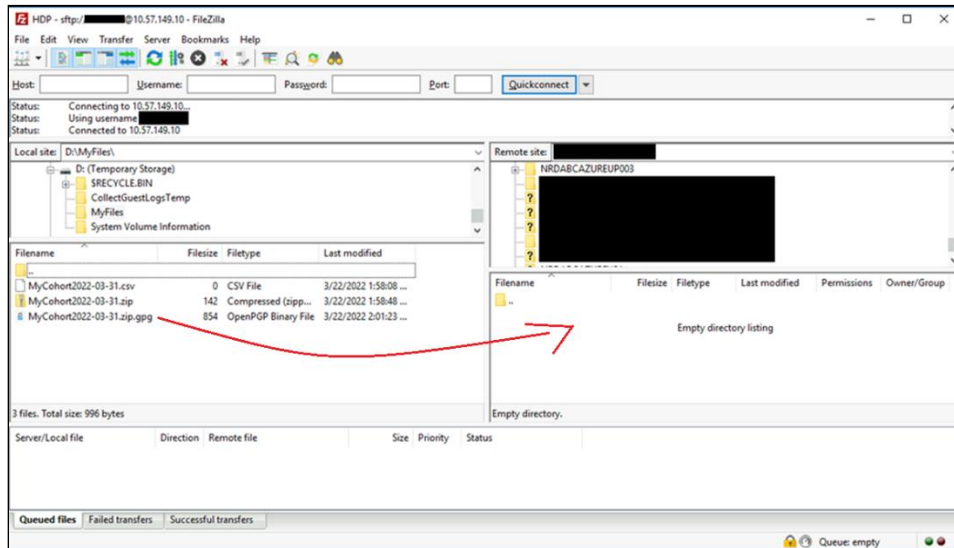


Figure 5. Drag your encrypted folder from the local site to the remote site.

If the file was transferred successfully, you will see the file appear in the 'Remote site' on the right side of the window, and the 'Successful transfers' tab at the bottom of the window will show details about the transfer (Figure 7).

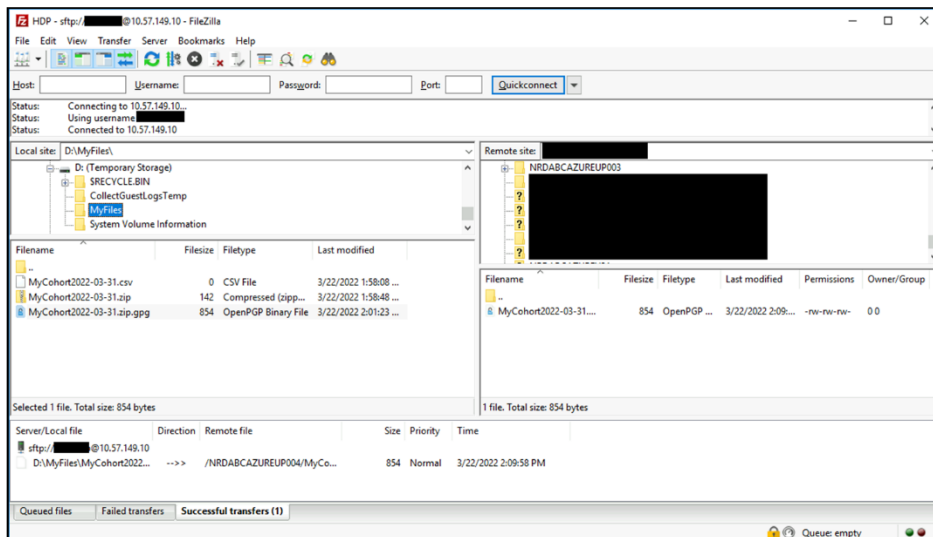


Figure 6. The "Successful Transfers" tab will show details about the transfer.

Step 9: Notify HDPBC that the file transfer is complete

1. Notify the HDPBC Front Counter (MOHAnalytics@gov.bc.ca) that the file has been uploaded for the HDPBC Technical Support team to pick up—***please include your encryption password (refer to Step 4 above) for file retrieval.***

3. GitHub Repository Imports – Public URL

PLEASE NOTE: If application code is included in the Git Repository, the HDPBC Security Threat Risk Assessment (STRA) process may be required, which can delay the import of the repository subject to completion of the STRA.

Step 1

Send the following request details to HDPBC Front Counter (MOHAnalytics@gov.bc.ca):

- Project name
- Principal applicant name
- Project number
- Git Repository information
 - Repository name
 - Location (public URL)

Step 2

Wait for an email confirmation that your Git Repository import has been completed.