Pharmacogenomic Insights (PGXI) Utilities Quick Start Guide

For Pharmacogenomic Insights Utilities version1.0

Contents

Setting Up Your Environment and Configuration

Step 1. Click this this link to download the PGxInsightsUtilities zip file.

Step 2. Unzip the PGxInsightsUtilities zip file to your desired drive or directory. e.g., If you want your working folder to be C:\PGxInsightsUtilities, unzip the file to the C: drive. **Note:** The unzipped PGxInsightsUtilities folder should contain 13 directories and 10 Windows Command Scripts.

Step 3. Open the **Install** directory, double-click on PgxUtilsInstaller.msi, and follow the prompts to install the PGx Insights Utilities. **Note** If you have already installed the Utilities, use the computer search bar to find and remove the installation in the Add or Remove Programs section. After reinstalling, you should be able to run PGxUtils.exe from any command prompt. However, on some 64-bit systems, the executable may not be recognized. If this happens, close all command windows or reboot your system.

Step 4. With your subscription, you should have received login credentials for the PGx Insights server as well as a *Client ID* and *Client Secret*. Use these credentials in the following steps to register your device with a unique API Device Code and acquire demo configuration files, which will be elaborated on in Step 5.

1. Login to the *PGx Insights server* at pgxinsights.com with your Qiagen PGXI user ID and password.
2. Navigate to **</> Get an API Device Code** or pgxinsights.com/getdevicecode to generate a unique API device code:
    a. Enter your **Client ID** and **Client Secret** and click **Next**.
    b. Click the **Copy** button to copy your **User Code** to the clipboard.
    c. Click the link to the *Digital Insights authorization server* to open it in a new browser.
        i. If prompted, re-enter the same login credentials as step 1.
3. Paste the copied **User Code** into the **Register Device** window of the *Digital Insights authorization server* and click **Next** to authorize your API device code.
    a. Follow the prompts to complete the device registration process. **Note** It is recommended that you edit the name of your device in the "Device Name Captured By Us" field.
    b. Confirm your device now appears under the **My Devices** tab of the *Digital Insights authorization server* (QIAGEN Digital Insights OAuth).

Step 5. Return to your *PGx Insights server* browser tab. The **Get an API Device Code** wizard will have advanced to the **Final** page where you will see buttons to download three files:

- JSON configuration file (REQUIRED). Users MUST download and copy this file to the **Config** directory created during Step 2. With that in place, you are ready to use the Utilities.
- Postman configuration file (OPTIONAL). Allows programmers to test various API call types.
- cURL command file (OPTIONAL). Allows users to test your device code configuration. Open this file and paste the contents into a Command Prompt window to confirm that you can retrieve an access token with the configuration information.

Note If you fail to download the file(s) before the browser tab refreshes, you must repeat Steps 4-5.

Reporting Without Demographics

The following steps will enable you to take results data, process it, produce JSON Insights files, and create reports that are only identified by the Sample ID that is represented in results data files. For demo purposes, you will use sample files provided in the **PGxInsightsUtilities** folder.

Step 1. Open the **SampleData** directory and copy the following files to the **InputFolder** directory. **Note** You can include the VCF files in this step, but they may take longer due to their size. See *PreFiltering Large VCF Files* for additional options.

- ThermoFile.csv (Genotypes)
- ThermoCnvFile.csv (Copy Calls)
- Agena Combined.csv

Step 2. Open Command Prompt and Change Directory (cd) to the root folder containing the unzipped software (previously recommended you unzip to root folder C:\PGxInsightsUtilities). Note that all of the command scripts use a relative path to files within the root directory so you need to be at the root for them to work properly without edits.

To process the files that you placed in the **DataFolder** directory (Step 1):

Enter “**ProcessFolder.cmd**” to execute the following command:

```
PgxUtils processfolder --infolder ".\InputFolder" --configfile ".\Config\PGxI_Config.json" --
outtypes "InsightsJSON" --outfolder ".\insightsfiles"
```

This will:

a) Run the PgxUtils executable.

b) Take files from the **InputFolder** directory.

c) Use the **PGxI_Config.json** file from the **Config** directory to authenticate to the server.

d) Call the API with a specification to create JSON file types.

e) Place the JSON output in the **InsightsFiles** directory .

Note: This activity, using the process folder to generate insight files for eventual reports is also capable of processing vcf files. For the purposes of this walkthrough, we opted to utilized ThermoFisher and Agena genetic output. The system is designed to be vendor and format agnostic.

Step 3. To create PDF files from the JSON output:

Enter “**ReportFromJson.cmd**” to execute the following command:

```
pgxutils.exe reportfromjson --infolder ".\InsightsFiles" --templatefolder  ".\Reports" -outfolder
".\SampleReports"
```

This will:

a) Read JSON files from the **InsightFiles** directory

b) Use the default report template from the **Reports** directory

c) Place formatted PDF reports in the **SampleReports** directory

Navigate to the **SampleReports** directory to view all generated PDF files. These reports are considered samples because at this time the data has not been merged with any patient demographic information.

Demographics

Reporting With Demographics

The following steps will enable you to take results data, process it, produce Json Insights files, and create reports that are identified by various demographics represented in a demographics data file. For demo purposes, you will use sample files provided in the **PGxInsightsUtilities** folder.

Step 1. Open the **SampleData** directory and copy the following files to the **InputFolder** directory. **Note** You can include the VCF files in this step, but they may take longer due to their size. See *PreFiltering Large VCF Files* for additional options.

- ThermoFile.csv (Genotypes)
- ThermoCnvFile.csv (Copy Calls) and/or
- Agena Combined.csv

Step 2. Change Directory (cd) to the root folder containing the unzipped software (previously recommended you unzip to root folder C:\PGxInsightsUtilities). Note that all of the command scripts use a relative path to files within the root directory so you need to be at the root for them to work properly without edits.

To process the files that you placed in the **DataFolder** directory (Step 1):

Enter “**ProcessFolder.cmd**” to execute the following command:

```
PgxUtils processfolder --infolder ".\InputFolder" --configfile ".\Config\PGxI_Config.json" -outtypes "InsightsJSON" --outfolder ".\insightsfiles"
```

This will:

a) Run the PgxUtils executable
b) Take files from the **InputFolder** directory
c) Use the **PGxI_Config.json** file from the **Config** directory to authenticate to the server
d) Call the API with a specification to create JSON file types
e) Place the JSON output in the **InsightsFiles** directory

Step 3. To first familiarize you with the data (demographics) that can be mapped into final reports, open the **Sample Demographics** file in the **Demographics** directory. In this file, only **SampleId** is mandatory in order to match the demographics with the test results, all other metadata is optional and individual columns can be added or deleted to suit the needs of your report. Such column changes MUST be mirrored in your report template(s) for the data to be placed properly. Please note the report template and stylesheet can be customized per sample, allowing users to present custom branded information to specific clients. **Note** If you open the Sample Demographics file with Excel, you must close it before step 4 because Excel locks files exclusively so step 4 will fail.

Step 4. To combine demographics data from the Sample Demographics file with the JSON files created in Step 2:

Enter “**ConsolidateFiles.cmd**” to execute the following command:

```
pgxutils.exe consolidatefiles --infolder ".\insightsfiles" --demographicsfile ".\Demographics\Sample Demographics.csv" --outfolder ".\ConsolidatedFiles"
```

This will:

a) Read Insights files from the **InsightFiles** directory.

b) Access the demographics file specified in the **Demographics** directory.

c) Combine the data into new JSON files which are then placed in the **ConsolidatedFiles** directory.

Step 5. To create PDF files from the consolidated JSON output:

Enter “**ReportFromConsolidatedJson.cmd**” to execute the following command:

```
pgxutils reportfromjson --infolder ".\ConsolidatedFiles" --templatefolder  ".\Reports" --
outfolder ".\FinalReports
```

This will:

a) Read JSON files from the **ConsolidatedFiles** directory.

b) If there are reports and stylesheets specified in the demographics files, they will have been captured in the demographics description of the consolidated files and they will be used to format the report. If no specific information is in the demographics, the defaults will be used. All report templates and stylesheets must be in the **Reports** directory based upon the command above.

c) Place formatted PDF reports in the **FinalReports** directory

Navigate to the **FinalReports** directory to view all generated PDF files.

Reporting with HL7 Demographics

The following steps will take you through the process of merging HL7 demographic files with processes patient genetic samples that that have bene converted into Insight.JSON files in the .\insightsfiles folder at the root of the PGxIUtilities command folder.

Step 1: Open the **SampleData** directory and copy the following files to the **InputFolder** directory. **Note** You can include the VCF files in this step, but they may take longer due to their size. See *PreFiltering Large VCF Files* for additional options.

- AdvancedGenotyper.csv
- AdvancedGenotyperCNV.csv

Step 2: Change Directory (cd) to the root folder containing the unzipped software (previously recommended you unzip to root folder C:\PGxInsightsUtilities). Note that all of the command scripts use a relative path to files within the root directory so you need to be at the root for them to work properly without edits.

To process the files that you placed in the **DataFolder** directory (Step 1):

Enter "**ProcessFolder.cmd**" to execute the following command:

```
PgxUtils processfolder --infolder ".\InputFolder" --configfile ".\Config\PGxI_Config.json" -
outtypes "InsightsJSON" --outfolder ".\insightsfiles"
```

This will:

a) Run the PgxUtils executable
b) Take files from the **InputFolder** directory
c) Use the **PGxI_Config.json** file from the **Config** directory to authenticate to the server
d) Call the API with a specification to create JSON file types
e) Place the JSON output in the **InsightsFiles** directory

Step 3: The HL7 demographic file needs to be merged with the patient genetic data that has now been processed and awaits reporting. This is done by matching the Sample ID of the genetic data with the filename of the HL7 file.

> **Ex**: In the AdvancedGenotyper.csv file has a column/header called 'Sample ID'. In this column is typically the name of the patient, or the accession number assigned to the genetic sample. The HL7 file must have the same name as the Sample ID because in the current iteration of the PGxIUtiltiies, the second entry in the PID row of the HL7 is considered optional and not required. Placing the Sample ID there and not ass the filename will result in the data not being merged correctly, if at all.

Place the HL7 demographic file, BetaSample1.hl7, from the 'SampleData' directory into the 'HL7 Orders' folder at the root of the PGxIUtilities command folder.

In a command prompt, enter ‘ConsolidateHL7Files’ to execute the following command:

```
PgxIUtils consolidatefiles --infolder ".\insightsfiles" --hl7orders ".\HL7Orders" --outfolder
".\ConsolidatedFiles" --configfile ".\Config\PGxI_Config.json"
```

This will:

a) Run the PGxiUtils executable.
b) Take data from the ‘insightsfiles’ folder
c) And the HL7 data from the ‘HL7Orders’ folder.
d) Authenticate to the API server
e) Call on the API to run the necessary command to merge the two files based on the common SampleID
f) Place the newly created ‘<sampleID>.Insights.Consolidated.Json’ file(s) in the ‘ConsolidatedFiles’ folder at the root of the PGxIUtilties command directory.

Step 4: Now we need to process the reports from the newly consolidated JSON files. Because the patient demographics have been properly merged with the patient genetic sample results, this will create a new Final Report. This is slightly different from the Sample Reports previously shown as it will include all the necessary patient information: First/Last Name, Provider Name, Provider Location, Sample Date, etc. (All the information present in the demographic file that the report has been configured to pull and generate.)

Note: Depending on the configuration of your reporting templates, not all patient demographic data may be pulled from the HL7 file. This is by design and will likely be worked out with Translational Software ahead of time.

In a command prompt enter ‘Reportfromconsolidatedjson’ to execute the following command:

```
PGxIUtils reportfromjson --infolder ".\ConsolidatedFiles" --templatefolder  ".\Reports"
 --outfolder ".\FinalReports
```

This will:

a) Run the PGxIUtils executable
b) Take the created ‘<sampleID>.Insights.Consolidated.Json’ files from the inights folder.
c) Inject the data from the insights file into a report that is formatted from a designated template in the ‘.\Reports’ folder.
d) Place the newly created report that includes the patient demographic information in the .\FinalReports folder awaiting pickup.

VCF Files

Pre-Filtering Large VCF Files

Some pharmacogenomic systems produce very large files that are unwieldy to work with and use a lot of data transfer time if they are used as-is. A beneficial feature of the PGx Insights Utilities is the ability to filter these files to define the loci that are relevant post-filtering, making the vcf files easier to work with. The filtering process uses the following vcf formatted files in the **Config** directory:

- PgxFilter. Filters to all the loci for which there is reliable clinical evidence.
- MiniFilter. Filters for the minimum variants to demonstrate all features of the final report.

Step 1. Open the **SampleData** directory and copy the LargeVcf.vcf file into the **DataFolder** directory.

Step 2. To filter the vcf file that you placed in the **DataFolder** directory (Step 1):

Enter "**FilterFolder.cmd**" to execute the following command:

```
PgxUtils.exe filterfolder --infolder ".\DataFolder"  --filterfile .\Config\PgxFilter.VCF --outfolder .\inputFolder
```

This will:

a) Read vcf file (input file) from the **DataFolder** directory.
b) Read the filter file from the path provided.
c) Read each line of the input file and:

   - If the line in the inputfile is a comment or header line, it is copied as-is to the output file (filename appended with "filtered" and placed in InputFolder directory).
   - If the line represents a copy number, it is copied as-is to the output folder (**InputFolder** directory).
   - If the build, chromosome, position, any of the alleles matches any of the alleles in a matching line from the filter file, the line is copied as-is to the output file (filename appended with "filtered" and placed in InputFolder directory)..

Navigate to the **InputFolder** directory to find the generated filtered data file(s).

Step 3. Return to Step 2 of Reporting Without Demographics or Reporting With Demographics to complete all remaining steps and produce PDF reports.

If you use the **FilterFolderContinuous.cmd** script, it will continuously check the **DataFolder** for new vcf files.

VCF Scaffolding

Creates a .vcf overlay that defines all of the mutations that are understood and can be useful for PGx purposes. This overlay is created for mutations that actually exist in the genetic results.

VCF Scaffolding is a newly created feature that allows one to process vcf files and then apply a lattice to ensure that where there are expected mutations, they stand out. It does this by creating an overlay that defines all of the mutations that are understood and can be useful for PGx purposes. It is created for mutations that actually exist in the genetic results.

This requires a configured scaffold file to be placed in the in the .\Config and then defined in the command. Default location call is .\Config\NewScaffold.vcf

Step 1. Ensure that the desired and properly configured vcf scaffold file is in placed in the .\Config folder

This will ensure that when the scaffoldfile command is initiated it understands what kind of overlay it needs to build for the vcf files that it is being asked to process.

Step 2. Place the desired vcf file that you wish the system to process into the '.\Datafolder' found at the root of the PGxIUtilites command folder.

In a command prompt, enter 'scaffoldfolder' in the root PGxIUtilities directory to execute the following command:

```
PGxIUtils.exe scaffoldfolder --infolder .\DataFolder  --scaffoldfile .\Config\NewScaffold.vcf
--outfolder .\inputFolder
```

This will:

a) Read and collect the vcf files located in the '.\DataFolder' location at the root of the PGxIUtilities commands folder.
b) Apply an overlay on to it indicating where the expected mutations should be. This overlay is pulled from the .\Config\Newscaffold.vcf file configuration.
c) Place the newly created, scaffolded vcf files in the input folder awaiting the 'processfolder' command to ingest them into the system and return proper insight.json files.

Advanced Options

Automating Processes

Within the **PGxInsightsUtilities** folder, review the provided Windows Command Scripts. Those with “Continuous” in the filename have a parameter called –pollinginterval which determines the time interval in which the input directory will be checked for new data. The pollinginterval value has been set to 10 in all Continuous files but users may adjust that value directly in the script.

Multiple Output Formats

By running **ReportFromJsonContinuous.cmd** in a new window, the software will continuously check the InsightsFiles directory and push reports to the SampleReports folder.

The processfolder command can produce any combination of output types. The output types are specified as InsightsJSON, VariantsCSV, or GenotypesCSV. To modifying the ProcessFolder to produce all three use the following:

```
 PgxUtils processfolder --infolder ".\InputFolder" --configfile ".\Config\PGxI_Config.json" --outtypes
"InsightsJSON" “VariantsCSV” “GenotypesCSV”--outfolder ".\insightsfiles"
```

Will place csv files in output folder for all samples that are run. Csv files will accumulate data for all samples, but json files will be replaced if a new file with the same Sample ID is processed.

Document Revision History

Revision	Description
May 2025	Initial release for Pharmacogenomic Insights Utilities Application version 1.0
September 2025	Updated document contents reflecting added and updated commands and how to use them.