Table of Contents

About ...................................................................................................................... 3
Consortium Partners ............................................................................................. 3
Accessing OneFlorida i2b2 using UFHealth VPN .................................................. 4
  Without Installing a VPN Client ........................................................................... 4
  Using a VPN Client ............................................................................................... 5
Introduction to i2b2 for OneFlorida ........................................................................ 6
Cohort Identification ............................................................................................... 6
OneFlorida i2b2 Best Practices .............................................................................. 7
  Date Ranges ......................................................................................................... 7
  Threshold Values ................................................................................................ 7
  Age vs. Age at visit ............................................................................................... 7
  Query Constraints and Venn Diagrams ................................................................. 7
OneFlorida i2b2 Interface ...................................................................................... 9
  Login .................................................................................................................... 9
  Main i2b2 Interface ............................................................................................... 9
Navigate Terms ..................................................................................................... 10
  Navigating Terms and Find Tool ......................................................................... 10
Query Tool ............................................................................................................ 11
  Adding New Groups ............................................................................................ 11
  Query Options ..................................................................................................... 12
  Dates ..................................................................................................................... 12
  Occurrences ........................................................................................................ 12
  Exclude ................................................................................................................ 13
  Temporal Constraints ......................................................................................... 13
  Run Query .......................................................................................................... 16
Query Status/Results/Reports ............................................................................. 16
  Show Query Status ............................................................................................. 16
  Graph Results ..................................................................................................... 16
  Query Report ....................................................................................................... 16
Workplace ............................................................................................................ 17
Previous Queries .................................................................................................. 17
  Saving Queries .................................................................................................... 17
  Opening Saved Queries ..................................................................................... 18
  Restoring a Saved Query ................................................................................... 18
  Temporal Dimension Queries .......................................................................... 18
Additional Resources ......................................................................................... 18
About

The OneFlorida Clinical Research Consortium is a collaborative statewide network that seeks to improve health research capacity and opportunities in the State of Florida through the facilitation of clinical and translation research in communities and health care settings. OneFlorida includes 9 unique health systems that provide care for ~15* M or 74% of all Floridians through 4,100 physician providers, 1240 clinic/practice settings and 22 hospitals, with a catchment area covering all 67 Florida counties. The OneFlorida Clinical Research Consortium includes the following partners:

Consortium Partners

*None-de-duplicated between partners and includes Medicaid payer mix
Accessing OneFlorida i2b2 using UFHealth VPN

The OneFlorida i2b2 Instance is contained behind the in the OneFlorida Coordinating Center Fortinet 1200D Firewall within a virtual, FISMA-compliant environment powered by VMWare ESXi 6.5.0. The i2b2 Interface will be available while on a Wired or Secure Wireless Connection inside the confines of Shands, or the Academic Health Center environment. If you are outside this network boundary (Main Campus, Home Network, Wireless Hotspot, etc.) you will be required to use the UF Health VPN Portal (https://vpn.ufhealth.org).

Without Installing a VPN Client

To use the OneFlorida i2b2 on UFHealth VPN:

1. Browse to URL https://vpn.ufhealth.org
2. Authenticate with your Gatorlink username (not UFID) and password. Be sure to select the UFAD domain.
3. If you see a Setup Control Warning, click “Always”

⚠️ If this is your first time connecting to VPN, you may receive a request to restart the computer. After rebooting, please re-visit vpn.ufhealth.org and then start with step 1.

Once logged in, you will see the VPN Portal where you are now connected to the UF Health VPN and to the Shands and Health Science Center (HSC) networks that will allow you to access internal websites, portals, and systems.
4. Go to Web Bookmarks. If you see more than one bookmark, make sure you select the correct i2b2 instance – OneFlorida.

![Web Bookmarks]

**Using a VPN Client**

You can also use the Junos Pulse VPN Client to connect to VPN but you will need to install it on your workstation or laptop first.

1. Log in to the UF Health VPN Portal as described above. [https://vpn.ufhealth.org](https://vpn.ufhealth.org)
2. At the top of the VPN Portal page, find the utility toolbar and choose “Client Apps”
3. Choose the “Start” button to the right of the Junos Pulse Client item
4. Click “Always” on the Setup Control Warning to install Junos Pulse
5. Allow the application to install. The client should open automatically.
6. Click “Connect” next to the listed UF Health connection.
7. When prompted, choose the domain that is associated with your username (UFAD) and check “Save Settings” and click “Connect”
8. Enter your Gatorlink username and password and then click “Connect”
9. Once connected, you will see the listed UF Health connect show as connected. Also, there will be an icon for Junos Pulse in your system tray.
10. You are now successfully connected to UF Health VPN using the Junos Pulse Client. This client will be available in your start menu or applications list.

For UF Health VPN support: [https://ufhealth.org/uf-health-vpn-installation-help](https://ufhealth.org/uf-health-vpn-installation-help)
Introduction to i2b2 for OneFlorida

The OneFlorida Informatics for Integrating Biology and the Bedside (i2b2) allows researchers to query an IRB-approved PCORnet dataset through an Internet browser and virtual private network (VPN) on a local computer and answer count-based questions about OneFlorida patients. The OneFlorida i2b2 uses patient-oriented data stored in an integrated and secure server within the OneFlorida data warehouse. New patients and records are updated quarterly from OneFlorida partner data submissions. While the i2b2 software provides direct access to the consortium data for research use, the aggregation and obfuscation of data in OneFlorida i2b2 simplifies data access and output.

The OneFlorida i2b2 is designed primarily as a cohort identification tool. All queries in i2b2 return a count of unique patients who meet the query criteria. The i2b2 query and analysis tool is approved by the IRB as a de-identified source of data that fully protects patient privacy when returning search results. For all queries, the resulting numbers are obscured and counts are returned with an accuracy of plus or minus three (±3). A researcher can further investigate the distribution of a cohort across demographic characteristics such as age, race and gender.

This document provides a descriptive and visual overview of creating and retrieving queries using OneFlorida i2b2. The data are represented as close to the original PCORnet Common Data Model (http://www.pcornet.org/pcornet-common-data-model/) with minimal deletions, de-duplication, or edits. Therefore, the data are only as good as their source. When interpreting i2b2 query results, we recommend using standard best practices for analyzing secondary data.

Cohort Identification

The OneFlorida i2b2 query and analysis tool is designed primarily as a cohort identification and data exploration tool. Therefore, the answer to an i2b2 query is always the number of unique patients, with a specified degree of uncertainty, that meet the query condition. Researchers can use OneFlorida i2b2 to search for a prospective patient cohort without the need for IRB approval according to the criteria outlined in the OneFlorida i2b2 Data Guide. Once a cohort has been identified, an investigator can use the number of patients in the cohort to submit a study-specific protocol to the IRB, requesting approval to receive the full set of patient data for research or clinical study. Researchers should not expect to see patient cohort data fully exposed in the OneFlorida i2b2.

Examples of cohort discovery could include the following:

- How many men with a diabetes diagnosis were inpatients between January and December 2016?
- How many women ages 45-65 with a breast cancer diagnosis received chemotherapy in the last year?
- How many men over the age of 65 with diabetes mellitus type II have hypertension, pulmonary fibrosis and congestive heart failure?
- How many females age 55 and older have troponin T values greater than 0.09 ng/mL?
OneFlorida i2b2 Best Practices

The following pages outline best practices when using the OneFlorida i2b2 cohort discovery tool. These will help users avoid common pitfalls and problems when making selections and applying query constraints. Guidelines regarding expectations of query performance when utilizing temporal constraints are included. Such queries can affect performance if poorly constructed.

**Date Ranges**
The information in this tool comes from records within a certain date range. This date range is indicated on the top i2b2 title bar, given as 01/2012 to 03/2018 in this document. (Note: the date range will change with the addition of new data).

If a researcher creates a query with a data range beyond that which the data contain, the query will run without errors but return only those patients within the date range of the underlying data.

Data are refreshed and uploaded every 3 months, so any patient encounters after the data refresh will not be reflected in the data. Researchers will need to wait until the next data refresh cycle (i.e. quarterly data) to capture those additional patients. Due to the data characterization and curation process as well as ETL updates, each i2b2 refresh will contain data as current as 2-3 months prior to refresh date.

**Threshold Values**
Pay attention when typing threshold values for facts. For example, a diastolic blood pressure vital measurement of 30 is unreasonable. Most likely, the value should have been 130.

**Age vs. Age at visit**
Oftentimes researchers need to query patient ages at the time of a particular diagnosis or procedure, which is quite different than the current age of patients that have that same diagnosis or procedure. For example, the cohort of patients aged 55-64 years old at the time of a diagnosis of diabetes mellitus type II in 2013 would be different from all patients aged 55-64 years old with diabetes mellitus type II since the beginning of data range.

**Query Constraints and Venn Diagrams**
The query tool uses Venn diagrams to search for multiple disease populations using the OR feature (left diagram). For example, a search for Diabetes mellitus type II OR Essential Hypertension will pull up the population of patients with Diabetes mellitus type II and the population with Essential Hypertension. Using the AND feature (right) will pull up the patient population that has both Diabetes mellitus type II AND Essential Hypertension.
It is possible to search for specific populations using multiple constraints such as Diabetes mellitus type II AND hypertension AND an age constraint. The patient population is the intersection of the Venn circles.
OneFlorida i2b2 Interface

The OneFlorida i2b2 interface is a series of windows where users can navigate or search for query criteria, construct queries, and view query results. The interface is designed as intuitively as possible and the following section highlights basic functionality of constructing and retrieving queries.

Login

After navigating to the OneFlorida i2b2 web address https://i2b2.oneflorida.ahc.ufl.edu/, users will need to enter their UF assigned Gatorlink username and password. The i2b2 OneFlorida Host will be entered by default. Click Login to enter the system.

Main i2b2 Interface

The main i2b2 interface is comprised of five windows where users specify query criteria and query options. The main windows, moving left to right across the interface, include the following:

- **Navigate Terms** – Select data dimensions/concepts for query. Dimensions/Concepts are organized hierarchically
- **Query Tool** – area to drag and drop concepts from Navigate Terms or Search Tool pane to use in a query. Boolean operators (‘And’, ‘Or’, and ‘Not’) are applied to the data elements.
- **Show Query Status/Graph Results/Query Report** – processing status of query and display area for query results and reports
- **Workplace** – “scratch space” holds dimensions/concepts, previous queries and results. Right clicking your username allows you to create a new folder within which to organize saved queries and navigation terms/dimensions.
- **Previous Queries** – log of previous queries and their results
Navigate Terms
Data elements in OneFlorida i2b2 are presented hierarchically, nested by expandable, tree-like icons called data dimensions within eight categories. There are three modifiers used to further constrain queries of the data dimensions.

Please see the OneFlorida i2b2 Data Guide for a full description of query dimensions and modifiers. www.onefloridaconsortium.org/i2b2

Navigating Terms and Find Tool
There are two primary ways to search for query terms: Using **Navigate Terms** and using **Find**. When **Navigate Terms** is highlighted, you can browse through the data dimensions to identify the specific variables you wish to query.

If the items you need are not easily identifiable, select the Find tab and enter your search terms in the blank box. You are able to query on the exact term or a portion of the name. You can also limit the search to one of the main navigating terms by using the Any Category drop down menu. There is also an option to Search by Names or Search by Codes for diagnoses or procedures. When searching by name, you can refine your query using search criteria under the pull down menu labeled Containing.

The Select a Coding System pulldown menu under Search by Codes allows users to search for a specific code, for example an ICD-9 diagnosis or procedure code, or other numeric value, such as Age.

After your query terms have been returned, you can drag and drop them into the Query Tool. Additionally, if you hover over the search results, you are shown where in the Navigating Terms hierarchy that particular query term is located.
Query Tool

The Query Tool is comprised of groups of query criteria users can select for cohort discovery. Data elements from the Navigate Terms and Find Tool or Previous Queries are the building blocks of queries. Boolean logic can be applied within groups (OR) and among groups (AND) and is displayed as graphical boxes at the bottom of the Query Tool.

For example, if we want to search for all patients 65 and older with diabetes mellitus type II, we would drag the age data element from the Demographics dimension to the first group and the diabetes mellitus type II data from the Diagnoses dimension to the second group.

Adding New Groups

Users can add new query groups using the New Group button in the lower right corner of the Query Tool. Once a new window is added, the navigation arrows will highlight yellow allowing users to scroll left and right to view the additional windows.
Query Options

Users can customize the Query Tool by specifying a date range, number of occurrences, and exclusion of terms. These options are at the top of each query group and can be created for each group independently.

Dates

Selecting Dates activates Date Range constraint and users can enter specific start and end dates by selecting the checkboxes to fill the date and verify it using the calendar available through the pull down arrow. This limits your results to only patients who met your criteria during the established date range.

Occurrences

Users can specify the number of occurrences for an event or encounter by selecting Occurs > 0x. Specifying the number of occurrences is useful to uncover frequency of events or encounters, for example, repeat diagnoses of conditions or prescriptions.
Exclude

Exclude removes a query term from the query (Boolean NOT). For example, if you wanted to query only cirrhosis resulting from hepatitis B you can exclude alcoholic cirrhosis of liver by dragging those diagnoses into a group, and select Exclude to remove these terms from the query.

Temporal Constraints

Temporal constraints can include (or exclude) patient(s) according to time-based criteria. The default is *Treat all groups independently*, meaning the query terms could happen at any date within the entire OneFlorida i2b2 temporal range (given on the top of the OneFlorida i2b2 title bar).

Additional options include *groups occur in the same financial encounter* or as a *sequence of events*.

A *single financial encounter* covers one or more episodes of patient care that form a group because they occur close together in time and are defined singularly by payers.

*Sequence of events* allows an investigator to define a specific order that events of interest occur, and in relative time, in which they must have occurred. For example, a patient could be diagnosed with hypertension, then receive a beta-blocker prescription and ultimately be diagnosed with gout.

The population in which events occur contains all the data elements used to define your patient population, or cohort, such as demographic variables and diagnosis and/or procedure codes (similar to a regular query).
Selecting Event 1 and 2 open new group windows where users can drag and drop data elements from the Navigate Terms and Find terms tools. The default number of events is two (2) and users can add new events and remove the last event. (Note: to use Define Sequence of Events, you must have at least two events.)

Group 1 will change to anchoring observation with each new event. The anchoring observation includes the criteria mandatory for the specific event. Essentially, you are creating a unique cohort by event.

In our example, there are three unique events, and we would sequentially construct a query for each event.

Event 1: Diagnosis of hypertension and exclusion of gout diagnosis and anti-inflammatory medications
Event 2: Prescription of beta-blocker medications and exclusion of diuretics

Event 3: Diagnosis of gout and prescription of anti-inflammatory medications for gout attacks

Once the events are defined and data elements added to each event group, you will need to define the order of events. Again, by default, there are two events and if you add additional events, you will need use the Add Temporal Relationship button to include additional event timings to specify that Events 1 and 2 occur before Event 3.
**Run Query**

*Run Query* button executes current query from Query Tool. *Clear Query* removes all query terms and query options from all groups.

After selecting Run Query, users can select multiple options for display of query results. Selecting result type by gender, race or age will return a graphical breakdown of results by category.

**Query Status/Results/Reports**

The window in the lower right corner shows the query status and results and has basic report generating capabilities.

**Show Query Status**

The Show Query Status tab displays the progress, in seconds, of current query execution.

**Graph Results**

The Graph Results tab displays graphical results of query if Gender, Race or Age were selected as result type in the Run Query window.

**Query Report**

The Query Report tab displays a full report of the query including the query definition (e.g. all query constraints and options, and temporal constraints and date ranges) and the number of patients returned. Users can also print this report.
Workplace

The Workplace is useful for saving previous work, retrieving queries, or saving commonly used data elements from the Navigating Terms. Users can drag saved queries from the Previous Queries window, which makes it easy to locate a previous query if you have many saved queries. If you use the same data elements repeatedly in multiple queries, you can drag those data elements into the workplace for easily access instead of looking through the Navigate Terms hierarchy or using the Find tool. Lastly, users can create new sub-folders (by right-clicking their name) under their main user folder to organize saved queries and commonly used data elements.

Right clicking a sub-folder opens a context menu to manage your saved queries and constraints.

- Rename – specify a new name for sub-folder
- Annotate – changes the sub-folder label name (visible when hovering over name)
- Delete – removes sub-folder
- New folder – creates new, nested sub-folder

Essentially, the Workplace acts as a favorites list for commonly used queries and data elements.

Previous Queries

Many queries are complex and investigators will want to rerun or modify similar queries in the future. By default, all queries are automatically saved in the Previous Queries window.

Saving Queries

Queries in the Previous Queries window can be renamed for easy identification or dragged into the Workplace for future access. To rename a query, right click the default name and select Rename. The queries can be more easily identifiable using your naming convention. Then you are able to drag the specific query to your user name folder or project sub-folder in the Workplace window. You can also revise that query, re-run, and rename.
Right clicking a Previous Query will bring up a context menu to rerun, modify, and manage the saved query.

- Display loads a saved query into the Query tool for modification or rerunning.
- Rename allows users to specify a new name for an initial or previously-run query
- Delete removes a saved query.
- Refresh All reloads all saved queries in the Previous Queries window.

Opening Saved Queries

Saved queries can be restored by dragging a query from the Previous Queries or a folder under your Workplace to the Query Name box in the Query Tool, which highlights yellow. This will restore all the query constraints and query options exactly as when it was first run and automatically reruns it. You can then modify the query constraints or options, and add new query groups to build upon the previously run query.

Users can also drag a saved query directly into a group and populates the entire query in that group. The query can then be built upon further adding new query constraints and options.

Restoring a Saved Query

To access a saved query with all of its query constraints and options, drag it from the Previous Query window to the Query Name box at the top of Query Tool to restore each query group. Dragging to the first group will load the saved query but you will not be able to access any of the query constraints or options.

Temporal Dimension Queries

In general, temporal dimension queries can be time consuming to execute and some performance expectations are outlined.

1. Simple 1-5 window multiple dimension queries. We expect to return in <15 seconds.
2. More complex 5-8 window multiple dimension queries with temporal components (this happens before that). We expect to return <180 seconds.
3. Complex queries with temporal functions against observations that are extensively deep, like findings, blood pressures, etc. We expect to return <900 seconds.

Additional Resources

For more detailed information regarding use of OneFlorida i2b2, please visit the i2b2 page of the OneFlorida website for training materials and other resources - http://onefloridaconsortium.org/i2b2