Welcome to i2b2 Training

Open CHROME or FIREFOX and navigate to

https://i2b2-uab.hs.uab.edu

Dale Johnson
Matt Wyatt
Ayme Miles

i2b2support@uabmc.edu
● Decide what criteria patients must meet to be included in your research.
● Using those criteria, build an i2b2 query.
● i2b2 will always deliver a count of DISTINCT patients meeting your criteria.
● If the count looks good, re-run the query and create a PATIENT SET.
● Once you have a patient set you can view a wide variety of de-identified patient data.
● If you want to see PHI, MRNs, FFNs, specific locations and physicians, patient contact information, etc., get your IRB approved.
● Send the i2b2 team your IRB documents and let us know what PHI you want AND the name of your i2b2 query.
● Within a week or so (depending on complexity), we’ll send you your data.
● If you want to further refine your query with data not yet in i2b2, let us know.
● If you think your counts are low, let us know and we’ll investigate.
● If you are very knowledgeable about the data in your area of research and you can’t find what you need, let us know. We need advice
Training Agenda

- What is i2b2 and helpful links
- Screen layout, panel features, basic navigation
- Ontologies and Concepts
- Query building with ontology concepts; grouping concepts
- Query execution, query time-out
- Editing old queries, executing queries again
- Run Query – all the options, race, gender graphs/counts, etc.
- Printing Queries
- Advanced Query features – temporal constraints, occurrence counts, date ranges
- Concept popups – numeric and text constraints on concepts
- What makes a query take more time?
- I2b2 row counts
- UAB custom feature for limited data set downloading

What is i2b2?

I2b2 stands for “Informatics for Integrating Biology and the Bedside”. It is an NIH-funded National Center for Biomedical Computing based at Partners HealthCare System. I2b2 was designed primarily for cohort identification, allowing users to perform an enterprise-wide search on a de-identified database of health information to determine the existence of a set of patients meeting certain inclusion or exclusion criteria. I2b2 allows you to determine the feasibility of a study by showing whether or not a viable number of patients exist who match your study-specific criteria.

A lot of universities use i2b2. A quick Google search for “universities using i2b2” will bring up thousands of hits. I2b2 is fairly customizable, so the look and content of UAB’s ontologies/concepts might be very different from other examples you see online at other institutions. At UAB, I2b2 data is de-identified. An i2b2 user cannot retrieve a patient identifier or encounter identifier from i2b2.

<table>
<thead>
<tr>
<th>UAB’s i2b2 login page</th>
<th><a href="https://i2b2-uab.hs.uab.edu/webclient/uab_i2b2_Helper.php">https://i2b2-uab.hs.uab.edu/webclient/uab_i2b2_Helper.php</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>UAB CCTS i2b2 information web site</td>
<td><a href="https://www.uab.edu/ccts/research-commons/informatics/i2b2">https://www.uab.edu/ccts/research-commons/informatics/i2b2</a></td>
</tr>
<tr>
<td>I2b2’s official web site</td>
<td><a href="https://www.i2b2.org/">https://www.i2b2.org/</a></td>
</tr>
<tr>
<td>I2b2 community wiki</td>
<td><a href="http://community.i2b2.org/wiki/display/dash/i2b2+Dashboard+Home">http://community.i2b2.org/wiki/display/dash/i2b2+Dashboard+Home</a></td>
</tr>
<tr>
<td>I2b2 HELP email</td>
<td><a href="mailto:i2b2support@uabmc.edu">i2b2support@uabmc.edu</a></td>
</tr>
</tbody>
</table>

ICD10 Browser: http://apps.who.int/classifications/icd10/browse/2016/en#/Q04.6
I2b2 Overview

1. Navigate Terms - This is where you will locate search concepts using a hierarchical folder structure.
2. Find Terms - In this pane, you can find search concepts using a representative word, phrase or specific code to describe the concept.
3. Workplace - This is where you will create personal folders for storing your queries.
4. Previous Queries - This pane shows a log of previous queries for the current user.
5. Query Tool - This is your main work area where you will define and refine the criteria for your query.
6. Query Status - You will be able to monitor progress and see results in this pane.
Overview of the Navigate Terms, Workplace, and Previous Queries Panels
The ontology or concept hierarchy

With a drag from previous queries, you can save queries you might need again - create a query library. Rename queries here with a right-click.

You can also drag favorite concepts into your folder.

You can drag patient sets into your folder.

All the queries you've run lately. Can change settings to show many or few.

Can rename queries with a right-click.

Find Concepts without using the Ontology
Can search by concept name or by specific concept code

Search previous queries by name
Overview of the Query Tool and Query Results Panels

Each query group can contain one or more concepts. There is no limit to the number of query groups you can add. You can only see 3 at a time. You can see how many groups you have and get more groups or scroll through the groups.

You can right-click any concept in a group if you need to delete it.

Within a query group, a patient only has to meet one condition to be included in the answer.

You can reverse a group to mean "none of these" using the EXCLUDE button atop the panel.

This is ALWAYS DISTINCT PATIENTS. No duplicates are in this count, ever.

<table>
<thead>
<tr>
<th>Query name</th>
<th>Can drag previous queries here to check their results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal Constraint:</td>
<td>Group 1</td>
</tr>
<tr>
<td>Order</td>
<td>Concept X</td>
</tr>
<tr>
<td>1</td>
<td>* Hypertension and heart disease with heart failure (ICD10:11.0)</td>
</tr>
<tr>
<td>2</td>
<td>* Hypertension and chronic kidney disease with heart failure and stage 1 through stage 4 chronic kidney disease, or unspecified chronic kidney disease (ICD10:11.3)</td>
</tr>
<tr>
<td>3</td>
<td>* Heart failure (ICD10:60)</td>
</tr>
</tbody>
</table>

Number of patients

822

For Query "Three-Panel Query"
i2b2 Ontologies

UAB’s i2b2 has a growing set of ontology folders. As you open folders, you’ll see more folders. Eventually, at the bottom of the folder tree, you’ll find ‘leaf’ concepts that cannot be further expanded. We are constantly adding new types of facts to i2b2 and the ontology is always your guide to what we’ve loaded.

Every institution using i2b2 makes its own ontology choices. If you see i2b2 documentation from other institutions, you’ll notice that their ontologies vary widely from each other and from UAB’s.

The ontology usually displays various medical concepts as a set of coding schemes. You’ve probably heard of some of these coding schemes; for instance, the category of terms for diagnosis can be described using ICD-9, ICD-10, or SNOMED codes. Many ontology folders at UAB, such as Procedures, Diagnoses, DRGs, and Lab Tests, use coding schemes that you might find all over the country:

- Medications: Cerner MULTUM medications codes
- Diagnoses: ICD9, ICD10, and SNOMED codes
- Procedures: CPT, ICD10 PCS, ICD9 codes
- Lab Tests: LOINC codes

Other folders do not have a true hierarchy or a recognizable scheme, such as Immunizations or Microbiology Tests. There are relatively few distinct concepts in these folders, so it is easier to simply list them in the top folder.

Some folders have an ontology developed here at UAB. The Blood Pressure readings ontology is an example.

Those numbers beside the concepts and folders are the DISTINCT patient count for that concept/folder.

After every data refresh, we recompute these counts. If you ever run a query on a single concept and get a different count, the query is right and we’re probably in the middle of updating all counts (which can take several hours).
i2b2 Ontologies - continued

The ontology simply organizes concepts so that you can build a query that finds all UAB patients who have an association with those concepts. But the ontology itself does not help i2b2 count patients in any way.

You’ll notice that many concepts will appear in multiple places in the ontology trees. For example, if you search for “acetaminophen-oxycodone”, you’ll see these results:

The tooltips on those results will show you that acetaminophen-oxycodone is classified as (1) an analgesic, (2) a central nervous system agent, and (3) a narcotic analgesic combination. Plus, it has two different description displays.

A different type of example: We put the top 50 lab tests into a separate folder for the convenience of users. But all of those lab tests also appear in other LOINC folders at normal points in the ontology. It does not matter how you find a lab test – use the Top 50 folder or the traditional lab folders - the query will get the same count because it is using the code, not the ontology folder structure.

ALT (Blood) (UABLBS:315870) appears in two places but note that the distinct patient count is the same for both.

So which one should you search with? Fortunately, it does not matter.

When you drag a concept into a query, it is the concept code (MULTUM:d03431 in the example above) that is used in the query, not where you found it in the ontology. So no matter which result you drag over, you will still get all patients with a relationship to that concept code.
What Kinds of Patient Information is in i2b2?

Our process: We do our best to load all encounters from the EDW. Once we have all encounters, we load all patients associated with those encounters, then we load a select subset of medical information for those encounters/patients.

UAB Encounters - Most of the time, an encounter is a patient visit. But...

- UAB has a lot of different visit types, including lab-only visits and radiology visits. If you ask how many times a patient has been to UAB, the number can be much larger than expected because of all the visit types we have.
- UAB has both regular financial encounters and something called the ‘documentation’ encounter. Generally, a patient’s long-term problems (diagnoses) are attached to the documentation encounter while more episodic diagnoses (like a bout of pneumonia) are attached to financial encounters. Financial encounters have specific dates but documentation encounters live practically forever. This means that restricting some events/facts to the same financial encounter might not get the answer you expect.
- Allergies and some invasive procedures are tied to the documentation encounters.
- We load past and some invasive procedures are tied to the documentation encounters.

Clinical diagnoses – we only load those that are marked as Final, Billing, or Discharge diagnoses.

We have limited procedure codes (ones used for billing purposes) and limited DRG data. These data sets do not go very far back in time. (See below.) If you need to build queries with specific clinical data that we have not loaded into i2b2, we can still help you. Just send us an email at i2b2support@uabmc.edu

The date ranges and counts for all our fact families (as of 11/21/2016) are below:

<table>
<thead>
<tr>
<th>Fact Family</th>
<th>Total Fact Counts</th>
<th>Distinct Patients</th>
<th>Distinct Encounters</th>
<th>Earliest Date</th>
<th>Latest Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allergies</td>
<td>1646780</td>
<td>818638</td>
<td>818638</td>
<td>07/30/2003 16:32:00</td>
<td>11/04/2016 06:58:00</td>
</tr>
<tr>
<td>BMI and Waist Measurements</td>
<td>3782881</td>
<td>562852</td>
<td>3265081</td>
<td>05/12/2008 21:44:00</td>
<td>11/01/2016 04:41:42</td>
</tr>
<tr>
<td>Clinical Diagnoses</td>
<td>4805169</td>
<td>383643</td>
<td>3055777</td>
<td>05/19/1995 05:00:00</td>
<td>12/22/2016 06:00:00</td>
</tr>
<tr>
<td>Lab Orders</td>
<td>141875875</td>
<td>529950</td>
<td>2603974</td>
<td>06/25/2009 17:54:00</td>
<td>11/04/2016 05:25:59</td>
</tr>
<tr>
<td>Tumor/Cancer Data</td>
<td>8008157</td>
<td>123917</td>
<td>168448</td>
<td>12/31/1965 00:00:00</td>
<td>09/26/2016 00:00:00</td>
</tr>
<tr>
<td>Long-Term Problems</td>
<td>2976735</td>
<td>432530</td>
<td>432530</td>
<td>11/22/1965 00:00:00</td>
<td>08/15/4002 00:00:00</td>
</tr>
<tr>
<td>Procedures</td>
<td>526184</td>
<td>136227</td>
<td>193244</td>
<td>02/17/2014 00:00:00</td>
<td>11/01/2016 00:00:00</td>
</tr>
<tr>
<td>DRGs</td>
<td>134461</td>
<td>48752</td>
<td>67457</td>
<td>01/14/2014 14:06:00</td>
<td>10/28/2016 16:00:00</td>
</tr>
<tr>
<td>Blood Pressure Measurements</td>
<td>48425976</td>
<td>574164</td>
<td>3796029</td>
<td>08/05/2003 17:00:00</td>
<td>11/01/2016 05:25:08</td>
</tr>
<tr>
<td>Inpatient Medications</td>
<td>36100284</td>
<td>341748</td>
<td>869742</td>
<td>05/18/2008 14:44:00</td>
<td>11/05/2016 05:18:02</td>
</tr>
<tr>
<td>Height/Weight Measurements</td>
<td>17570643</td>
<td>593767</td>
<td>3784187</td>
<td>08/28/2003 16:43:00</td>
<td>11/01/2016 05:20:49</td>
</tr>
<tr>
<td>Lab Panels</td>
<td>9677955</td>
<td>453094</td>
<td>2035572</td>
<td>10/26/2009 17:04:00</td>
<td>12/15/2016 19:41:00</td>
</tr>
<tr>
<td>Service</td>
<td>ID</td>
<td>Code 1</td>
<td>Code 2</td>
<td>Start Date 1</td>
<td>Start Time 1</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td>---------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Encounter Insurance</td>
<td>11369918</td>
<td>695140</td>
<td>8260259</td>
<td>12/01/2003</td>
<td>01:19:00</td>
</tr>
<tr>
<td>Immunizations</td>
<td>203480</td>
<td>101336</td>
<td>168758</td>
<td>07/26/2008</td>
<td>15:10:00</td>
</tr>
<tr>
<td>Microbiology Tests</td>
<td>992190</td>
<td>171723</td>
<td>381345</td>
<td>01/26/2008</td>
<td>15:23:00</td>
</tr>
<tr>
<td>Less Common Clinical Events</td>
<td>296</td>
<td>75</td>
<td>77</td>
<td>06/29/2011</td>
<td>19:25:00</td>
</tr>
<tr>
<td>Lab Powerplans</td>
<td>1880573</td>
<td>297420</td>
<td>760454</td>
<td>05/18/2008</td>
<td>00:00:00</td>
</tr>
</tbody>
</table>
What Kinds of Patient Information is in i2b2? - continued

When you forget about fact families and just look at facts per year, you can see that most of the data in our EDW (and in i2b2) is fairly young:

<table>
<thead>
<tr>
<th># Facts</th>
<th>Calendar Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>47345765</td>
<td>2016</td>
</tr>
<tr>
<td>51534937</td>
<td>2015</td>
</tr>
<tr>
<td>49700108</td>
<td>2014</td>
</tr>
<tr>
<td>47958801</td>
<td>2013</td>
</tr>
<tr>
<td>44108483</td>
<td>2012</td>
</tr>
<tr>
<td>37569899</td>
<td>2011</td>
</tr>
<tr>
<td>15149276</td>
<td>2010</td>
</tr>
<tr>
<td>856823</td>
<td>2009</td>
</tr>
<tr>
<td>772430</td>
<td>2008</td>
</tr>
<tr>
<td>384785</td>
<td>2007</td>
</tr>
<tr>
<td>390158</td>
<td>2006</td>
</tr>
<tr>
<td>403529</td>
<td>2005</td>
</tr>
<tr>
<td>406917</td>
<td>2004</td>
</tr>
<tr>
<td>349931</td>
<td>2003</td>
</tr>
</tbody>
</table>

So, if you need to get data from before 2010, just be very aware that the data coverage will be uneven.
**Build Queries by Dragging Concepts to Query Groups**

To build a query, you drag a folder of concepts or single concepts into query groups.

But you have to be pretty careful about putting concepts into groups. THESE are important signposts.

So the query above is almost certainly NOT correct. The user probably wanted everyone who has had an imaging procedure, with a BMI under 30, and male. BUT, the query above gets everyone who is male OR who has had an imaging procedure OR has a BMI under 30. The resulting count will be HUGE. So, if you create a query and have only one query group and you have a lot of concepts in that group, ask yourself “is this correct?” and “is my result count too large?”

**EXERCISE Set 1**

1. You want everybody with sitting diastolic BP > 100, taking aspirin (now or previously), with toenail fungus. How many query groups do you need?
2. Same as #1 but you only want MALES. How many query groups?
3. You want everybody with either lupus or gout (or both) who has had a flu shot some time in their medical history and who is pregnant. How many query groups?
Finding Concepts in the Ontologies

If you know the code for the concept, type the code and limit the results to the concept group you want. (Code searches need to be exact – you will not get all ICD10 codes that start with M if you simply enter M – or you’ll get ‘No Records Found’.)

If you do not know the code and only have a description, search by name within the concept group you want (diagnoses below). You can search by partial words.

With Name searches, you might enter a search word that has too many matches. Then, you’ll see this warning:

The number of terms that were returned exceeded the maximum number currently set as 200. Please try again with a more specific search or increase the maximum number of terms that can be returned as defined in the options screen.

You can either enter a more specific search string or change the number of results that you’ll accept by clicking the ‘Show Options’ button and changing the maximum number of items to display. Then, click FIND again.
Finding Concepts in the Ontologies – continued

When searching, tool tips tell you where the concept is found in the hierarchy. For example, search for “troponin I” in Lab Tests and you’ll get several hits. Which one do you want?

<table>
<thead>
<tr>
<th>Search by Names</th>
<th>Search by Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>troponin I</td>
<td>Lab Tests (LOINC Ontology)</td>
</tr>
</tbody>
</table>

The first one is a Lab Panel, the others are plain labs. Find this out by hovering over each concept.

The thing to notice is that a concept can live in many different branches of an ontology. But when you use that concept in a query, it will count all patients with the concept.

Look at the end of the tooltips – those concept codes tell you what concepts are identical and which are different.
Executing Queries

Once you have your query built, you execute it by clicking the ‘Run Query’ button under the query groups.

This brings up the ‘Run Query’ dialog box. 99% of the time, you just want to see one checkmark – by the ‘Number of Patients’ choice. (More on the other choices later in this document.) If you click OK, the query starts running.

When it is running, you’ll see ‘Running Query:’ in the Query Status Panel. You’ll also see the ‘Run Query’ button has become a ‘Cancel Query’ button. (You can cancel the query at any time while it is running.) You’ll also see the query in the ‘Previous Queries’ panel as ‘INCOMPLETE/PROCESSING/QUEUED’ to indicate that the query is still getting an answer for you.
4. Hand On: Search for ‘influenza’ but limit the search to immunizations only. How many concepts did i2b2 find?

5. Hands On: How many people 80 years old or older have had some type of Flu vaccine?
6. Hands On: Regarding the query we just ran, how many of those are daily alcohol users?
7. Hands On: Rename the query you just ran to something more meaningful.

8. Out loud: How many seconds did the query for #6 take to get an answer? (HINT: Show Query Status tab)
9. Hands On: How could you remove the ‘daily alcohol user’ concept from the last query? (two ways)
11. Out loud: How is the result below ever possible? What are those numbers to the side of each concept?
Query Time-Out

If your query is complex, it might take more than 90 seconds. In that case, you’ll see a popup with this message: *Your query has timed out and has been rescheduled to run in the background. The results will appear in “Previous Queries”*.

When this happens, in the Previous Queries window, you’ll still see ‘INCOMPLETE’ and ‘PROCESSING’.

You can log out and return to i2b2 later to check the results. OR, you can wait. You can run other queries as you wait or you can refresh the ‘Previous Queries’ panel until the query has an answer – you’ll see FINISHED when the answer is ready.

But when your query is FINISHED, how do you see the query’s result? (This also applies if you executed a query yesterday and want to see its results again today.) Simply find your query in the ‘Previous Queries’ panel and drag the query folder over to the Query Name box. When you have it in the right place, everything suddenly turns yellow. Let go and you’ll see your results down in the results window at the bottom right.

If you ever want to give your query more time before time-out, click the ‘Show Options’ button in the query panel and increase the number of seconds you’re willing to wait. Any changes you make will be remembered until you log out.
Editing/Altering Queries

When you drag a previous query into the ‘Query Tool’ panel, you can change it in any way you want – remove individual concepts (with a right-click delete), remove entire query groups (by clicking the X in the upper right corner of the group), add more groups, etc.

When you execute the query again, i2b2 creates a new version of the query, without in any way changing the previous query, and that new version goes into your ‘Previous Queries’ list. i2b2 remembers everything you do – no queries are deleted/updated when you edit old queries.

Run Query Options

Sometimes you want a little more than just a distinct patient count. This list can change but currently it includes...

- **Patient Set:** Tells i2b2 to remember all the patients that match your query conditions. This patient set can be used to get partial or full patient details (if you have an IRB that allows this) for you and your research partners.
- **Encounter Set:** Same as patient set but i2b2 will remember all the financial encounters for all the patients that match your query conditions.
- **Number of Patients:** what we normally want.
- **Gender Patient Breakdown:** i2b2 will produce a simple graph and counts by gender
- **Vital Status Patient Breakdown:** i2b2 will produce a simple graph and counts for patients by living/deceased count. (This information is not 100% accurate since patients may have passed away elsewhere and UAB has no way to know this.)
- **Race Patient Breakdown:** i2b2 will produce a simple graph and counts by race
- **Breakdowns normally appear in the Graph Results tab and in the Show Query Status tab, as shown here:**
Race patient breakdown for "Acquired hemo" at 06:34:33:
- *American Indian or Alaska Native: 0*
- *Asian: 0*
- *Black or African American: 116*
- *Declined/Refuse: 5*
- *Hispanic or Latino: 4*
- *Multiple: 29*
12. Hands On: CLEAR your query building window. Now, say you can’t remember the answer for the ‘daily alcohol/flu/80+yo’ query above. How can you see your previous answer without building and executing that query again? (HINT: Previous Queries -> Query Tool)
   a. With that query back, do a search through MICROBIOLOGY concepts (find by name) for ‘bal/bronch wash colony’. How many concepts were found?
   b. Now you want to modify the query to only count patients who have also had a bronch wash (in addition to all the other conditions in that query). How can you add that new condition? (HINT: Is there a button to add a new query group panel?)

13. Hands On: CLEAR your query building window. Do a search in Diagnoses only for the disease ‘acquired hemolytic anemia’. (Containing is fine for this search.)
   a. Drag the first folder in the search results to the query window.
   b. How many patients have this? Do you trust the number beside the folder? How could you validate it?
   c. What is the breakdown by gender and by race for this problem? (Look at both Show Query Status and Graph Results)
Printing Queries

When you execute a query, you can ask for a printable, English-language version. To do this, click the Printer icon in the upper right of the query results panel.

This will bring up a window with all the query details – every group and concept is described. At the bottom of the window, you’ll see how long the query took to execute and the answers you received. Use your browser’s print features to produce hard copies of your query.

This print window has a couple of quirks:

It will not display the breakdown graphs, only the counts. This window is completely independent of i2b2, so if you leave it open, i2b2 does not know. If you click the print button several times for different queries, the window does not clear its previous contents - it just adds more content at the bottom! So, best
Advanced Query Options

Temporal Constraints

- When set to ‘Treat all groups independently’, the patients will have facts that occur at any time in their history. The particular encounters on which facts hang is not considered or matched up. This is the default.
- When set to ‘Selected groups occur in the same financial encounter’, all the patient facts across all groups must be on the same financial encounter. This is more restrictive and will result in a lower result count. An example might be that you only want patients who took a particular medication and who had a specific lab result during the same hospital stay.
- Define sequence of Events: This option is complex and allows you to set up a sequence of facts and order them as you please, even letting you specify the number of days between various facts.

“Occurs N times” option

You can specify that a patient should be counted only if they have X distinct occurrences of something. Here, ‘distinct occurrences’ means i2b2 will count the distinct patient/encounter/provider/start date combinations for concepts you have put into the query group.
Advanced Query Options – continued

Limiting a Query Group to a specific date range

Every query group has a Dates popup in the upper left corner. Click that to get a date range dialog box. To set a FROM date, check the box and supply that date. The same applies to the TO date.

You do not have to supply both dates. For example, to get any facts we have before 01/01/2012, just enter that as the TO date and leave the FROM date unchecked. Or to get all data from 06/01/2015 forward to today, just enter that as the FROM date and leave the TO date unchecked.

The date you set for the group applies to all facts within that group. If you need different date ranges for different facts, you have to put those facts into their own groups, not the same group.

Dates you enter are treated as MIDNIGHT on that date. So if you want all the vaccines given on a single date, you won’t get the right answer if you put the same date in both the FROM and TO boxes. Instead, put something like 06/01/2015 in the FROM box and 06/02/2015 in the TO box to get a true count for a 24-hour period.

EXERCISE Set 4

14. Hands On: How many patients have had a flu shot and a bronch wash in the same encounter? (Can you use a previous query to do this quickly?)

15. Hands On: How many patients have had a flu shot and a bronch wash in the same encounter in all of 2015? That is, 1/1/2015 through 12/31/2015? (Dates on one panel vs. all panels?)

16. Out loud: How could you email someone all the details of this query and its results? (HINT: Print Query button)

17. Out loud: What can you say about this query panel?

Note: for some reason, the default FROM date in i2b2 is 12/01/1979. This is just a random date and does not indicate the age of our oldest data. UAB does not have any significant data that goes back that far.
Concept Pop-Ups

Many concepts have either numeric or text values. For example, lab results will often have a numeric result; a vital sign like BMI will have an associated value; medications have dosage, form, frequency, etc.

I2b2 allows you to query for specific values like these when you first drag a concept to a query group. You’ll see a ‘Choose value’ dialog box when the drag is completed. And, you are much more likely to see a popup when you drag a leaf concept instead of a folder.

In any case, for text or numeric results, when the popup appears, you can

- Click the ‘No Value’ radio button to say that you don’t care what the value is – you only care about the existence of the concept in association with a set of patients.
- Select a single value by clicking that value and clicking OK.
- Select multiple values by holding down the SHIFT key as you click each value you want to include.

Once you make a selection, you’ll see your choices in the query group. If you made a mistake, right-click the concept in the query group and choose ‘Set Value...’
18. Hands On: Clear the query builder. Alkaline Phosphatase is one of UAB’s top 50 most-ordered labs. How many people have a result of < 25 for this test?

19. Hands On: Now change the query to get everybody with that lab, regardless of result value. (Right-click and set value to what?)

20. Hands On: How many people have a Troponin I lab test – any result for the lab test along with Angina NOS (ICD10 I20.9), treating each concept independently? (NO need to execute yet.)

21. Hands On: Set the Troponin level to get those with a lab result of 5 or greater and run the query. (HINT: right-click)

22. Hands On: Get the number of patients taking medication ‘isoniazid’ and medication ‘pyrazinamide’ on the same encounter. You don’t care about the medication details like format, frequency, dosage, etc.
   
   a. Of the three isoniazid medications, does it matter which one you choose?

23. Hands On: How many patients are taking isoniazid or pyrazinamide or both over the last six months?

24. Hands On: Get a count of all patients who are taking isoniazid BUT NOT pyrazinamide on the same encounter (last six months).
What can make a query take more than 90 seconds to execute?

- Queries that use concept folders instead of individual concepts
- The more groups your query has, the longer it will take
- Date ranges on groups increase query complexity
- Asking for gender, race, and other breakdowns in your query results, especially if your query’s patient count is large
- Limiting concepts by value – numeric or text
- Temporal constraints aside from ‘treat all groups independently’
- Asking for a patient set and/or encounter set adds more time, especially if the query’s patient count is large
I2b2 allows research institutions to build custom features called plugins. The list of plugins available to you can be found by clicking ‘Analysis Tools’ in the top right menu.

Now the right half of the i2b2 window will clear. The UAB plugin should be at the top of the list of plugins in the bottom right corner. Click that plugin button to see the UAB Self-Service Limited Data Set Download plugin.

This UAB plugin lets you download specific types of data for patient sets that you have created.
Only patient sets containing 500 or fewer distinct patients are available for download. If your patient set is larger than 500 distinct patients, you should contact the i2b2 team for assistance. There is an email link on the top of this page for this situation (and for any other questions you might have about getting data). **IMPORTANT:** With each download, you acknowledge that you have read and understood the terms and conditions of receiving access to Limited Data Set (LDS) from the UAB i2b2 system. This agreement applies for as long as you retain the LDS.

To create a patient set when you run a query, you should click 'Patient Set' in the 'Run Query' dialog box, before you click OK. This tells i2b2 to remember all the patients that match your query conditions. This patient set can be used to get partial or full patient details (if you have an IRB that allows this) for you and your research partners.

You patient sets appear in this dropdown, with the most recently created first. You’ll see the query name, the patient count, and the date the set was created. You won’t see patient sets with more than 500 distinct patients.

You can download one set of data at a time. We do this because the amount of data for even a small number of patients is quite large (and some researchers want medications but not vitals, for example). You can choose to get aggregate counts instead of all patient details. You will get ALL patient concepts, not just the ones you put into your query.

You must supply your research project’s title and description. You must supply the Blazer IDs of all people with whom you will share this data. If you will not be sharing this data with anyone else, this list can be left blank.
When you click Submit, your data download begins. If the amount of data is large it might take 2-3 minutes but most downloads will take less time.

When the download completes and you extract the zipped data, you’ll notice that the data is pipe delimited and in a form that can be loaded into Excel for further analysis. In these data sets, you’ll see a special patient identifier and/or a special encounter identifier which is NOT any UAB identifier you can make use of anywhere else (that is, these are not MRNs, FFNs, or any other key). If you download several sets of data for a patient set, the patient numbers in each set WILL correlate, so you can match patient diagnoses to patient labs using these special patient/encounter IDs.

Below you see a detailed data download of patient problems on the left and aggregate counts for the same data on the right.

---

**EXERCISE Set 6**

25. Out loud: For the result of any of the last few queries, how could you get a list of all the medications (with counts by distinct patient) being taken by the patients in that query group, without all sorts of patient details that you do not want.

26. Hands On: For the result of any query above, how could you get all their diagnoses? Do it!

27. Hands On: Take the result of any query above and get all their immunizations.
Sequence of Events Queries in i2b2

Example: From the set of all patients with a LDL level over 100, who are subsequently put on certain statins, how many go on to have heart attacks? You want those with an LDL lab result who are put on the statin drug at least 90 days later or more and then have a heart attack over a year after the start of taking the statin.

This kind of query involves

1. The ‘Population in which events occur’ - A starting pool of patients that meet a set of i2b2 conditions
   - The starting pool is optional – if you want the starting pool to be all patients in the entire i2b2 database, you don’t have to do anything further. (This is probably what you want to do.)
   - The starting pool can be as complex as any imaginable i2b2 query – you are not limited to something simple.
2. One or more discrete events that occur within that starting pool of patients
   - You can have as many events as you want. I2b2 will initially give you slots for two events but you can get more than two with a button click.
   - You don’t have to create the events in any particular order but it might help you to define event 1 as the first thing that happens and event 2 as the second thing, etc.
3. The sequence in which all your events should occur and the amount of time between each event in the sequence.

For the example above,

1. Set the temporal constraint to ‘Define sequence of events’

   In our example:

   Event 1: High LDL result

   Event 2: Starts taking statin drug

   Event 3: Heart attack
2. In the event dropdown, select Event 1

3. Find the LDL lab test you want, drag it to group 1, and set the value
4. Now, select 'Event 2' from the event dropdown

5. Find the statin drugs you want and drag that to group 1

6. At this point, you need a third event but the event dropdown only shows events 1 and 2. Click the 'New Event' button to get a slot for Event 3.

7. After clicking the button, you can select Event 3 from the dropdown:
8. Now you can drag the heart attack diagnosis code group over

9. At this point, review what you’ve put into each group by selecting Event 1 from the dropdown, then Event 2, then Event 3. In each case, you see the concepts you dragged into those groups.

10. If everything looks good, now you can define the sequence for your three events. Select ‘Define Order of Events’ from the dropdown

11. You’ll see this:
In our example:

Event 1: High LDL result
Event 2: Starts on statin drug at least 90 days later or more
Event 3: Heart attack over one year after starting statin drug

12. We want everybody who had an LDL lab result > 100 who went on a statin drug at least 90 days later, so ...

13. Now we need another set of boxes so we can put a year between events 2 and 3. Do this by clicking ‘Add Temporal Relationship’ at the bottom of the query panel.
14. Now you can set the time between events 2 and 3 as over one year

15. Finally, you can click ‘Run Query’. You have to wait a bit longer for your answer due to the complexity of the query.
Please note:

- These complicated queries will take longer to run than normal queries.
- Like all other queries, you drag them into the query tool to see their results later.
- Like all other queries, you drag them into the query tool if you want to create a new version of the query. For example, with our query above, we could drag it over and set the starting population to be everybody with a BMI over 30. The rest of the query doesn’t need changing, perhaps, so all the other events are in place and the new version is immediately runnable.
- The print query button is very useful to show others how you defined these queries BUT the display of the number of days between a larger set of events is not as detailed as you might want. (The query will run correctly but the print display is not as detailed.)
- I2b2’s default query name has a ‘(t)’ at the beginning to help you remember this is a temporal query, but you can always change the name to anything you want:

  - (t) LDL (Blood) (UA* HMG-CoA reduc* cardiac infar@07:48:12 [1-2]
  - (t) High LDL, Statin, then heart attack [1-26-2017] [rdale]
  - Chol (Blood) (U@06:46:32 [1-26-2017] [rdale]
  - (t) * Mood [affect* antiparkinson* Edema, protein@06:34:27 [1-26-2
  - (t) * Pedal cycle r* Bacterial inf* analgesics@06:29:02 [1-26-2017]