

University of Minnesota
Clinical and Translational Science Institute

i2b2 User Guide

i2b2 is a tool for discovering research cohorts
using existing, de-identified, clinical data

Table of Contents

I.	Introduction.....	3
	How can i2b2 help you with your research?	3
	Observational Studies and Retrospective Cohort Studies:	3
	Future recruitment:	3
	Before you start.....	3
	Get Started.....	3
	Your ethical and legal responsibilities.....	4
II.	Navigating the Workbench.....	5
	Helpful Hints in this Guide.....	5
	Workbench Screen.....	5
	Resize Workspace	6
III.	Constructing and Running your Query.....	7
	Selecting Query Criteria	7
	Building Your Query (with a worked example)	7
	Modifiers: Principal vs. Secondary diagnosis.....	8
	Using Find Terms to Help Develop Your Query.....	9
	Editing Your Query.....	11
	Refining Your Query.....	12
	Constraining Dates of Encounter.....	12
	Adding Exclusion Criteria	12
	Multiple Occurance	13
	Running Your Query.....	14
	Query Status Tab	14
IV.	Advanced Query Tools.....	15
	Previous Queries.....	15
	Temporal Constraint	15
	Tips for Optimizing your Query.....	16
V.	Additional Help	16

I. Introduction

Informatics for Integrating Biology and the Bedside (i2b2) is an informatics framework designed to simplify the process of using existing, de-identified, clinical data for preliminary cohort discovery. The i2b2 web client provides "drag-and-drop" selection of query criteria and outputs the number of patients in the database based on the submitted criteria.

How can i2b2 help you with your research?

Future recruitment

If you are designing a future study and recruiting patients, i2b2 will help you determine the feasibility of your study. With i2b2 you can discover how many patients with your demographic and clinical criteria have been seen in a given time period at our partnering clinical sites. Of course, past patient visits are no guarantee of future visits, but they provide an excellent estimate of the number of patients you can expect to visit our hospitals and clinics. It is often very helpful to include this sort of feasibility information in your grant applications.

Observational Studies and Retrospective Cohort Studies

If you are planning a study that will involve analyzing data collected during routine clinical care – an observational study or retrospective cohort study - i2b2 can be extremely helpful. For such a study, you may want to explore the size of the available patient cohort before you request a full dataset (requesting data is done through the CTSI's Data Access service). Exploring the impact of various search criteria – such as demographics, diagnostics, therapies, and procedures – on the size of the patient cohort will help you refine your study and optimize your data request.

If you find other ways that i2b2 helps you with your research, [please let us know](#).

Before you start

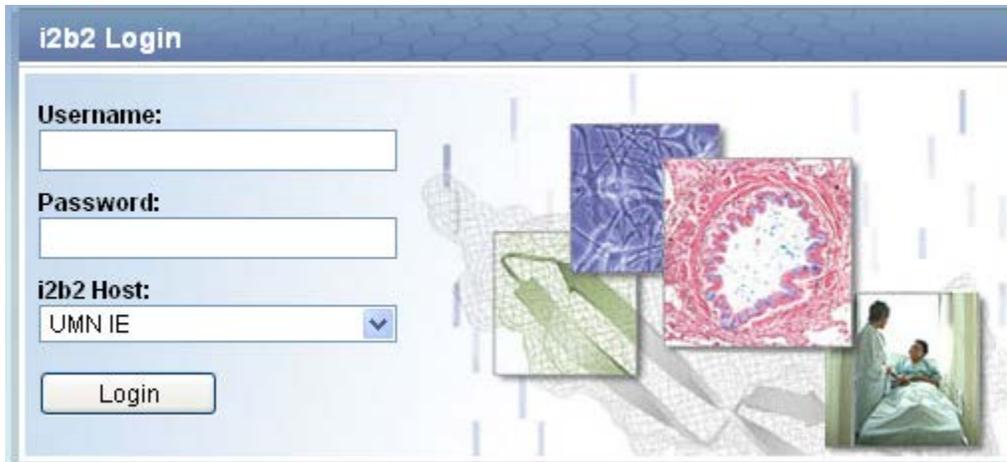
Before you can begin using i2b2 you have to be added to the list of active users. Please send an email to ics@umn.edu with the subject "i2b2 access request" and your x.500 in the body of the email. We will do our best to activate your access within 48 hours of receiving your request (and usually much sooner than that).

Once you get an email confirming your access, make sure you have the right browser available and an appropriate network connection. The best browsers for running the i2b2 web service are Internet Explorer, Safari, and Firefox. If a pop-up blocker is installed on your computer, please configure it to allow the i2b2 web address.

If your computer is not plugged into a University network jack or connected to the University's secure wireless connection, you will need to use the University's Virtual Private Network (VPN) connection (<http://www.oit.umn.edu/vpn/>) to login. If you have trouble installing or using VPN, please contact the Office of Information Technology for support (see <http://www.oit.umn.edu/vpn/technical-support/index.htm>).

Get Started

The URL for i2b2 is <https://i2b2.ahc.umn.edu/webclient/>

The image shows a web-based login form titled "i2b2 Login". It features three input fields: "Username:" with an empty text box, "Password:" with an empty text box, and "i2b2 Host:" with a dropdown menu currently showing "UMN IE". Below these fields is a "Login" button. The background of the form is a collage of medical and scientific images, including a DNA double helix, a microscope, a brain scan, and a patient in a hospital bed.

- Enter your x.500 and password.
- The default i2b2 Host, UMN IE, should remain the same.
- Click the Login button to submit your login credentials.

Your ethical and legal responsibilities

You will be prompted with the following Terms-of-Data-Access Agreement:

i2b2 query tool is a powerful research tool available to authorized users for the purpose of querying clinical data for authorized research purposes.

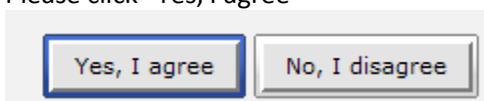
Authorization and access to this query tool is managed by the University of Minnesota Clinical and Translational Science Institute (CTSI). The data that will be queried are the assets of Fairview Health Services. You are accountable for following all applicable laws and Fairview and University policies. Violation of this agreement or applicable laws and policies will result in immediate termination of your access to this tool and may subject you to disciplinary action up to and including termination.

Clicking the "Yes, I agree" button below is our record of your agreement to the terms of this data access agreement and your attestation that you understand and agree to fulfill your responsibilities outlined below.

You are responsible for:

- *Safeguarding your access credentials and query results;*
- *Exercising exemplary ethical conduct when using this tool;*
- *Restricting queries to approved research topics;*
- *Ensuring query data is not used to attempt to identify any patients;*
- *Obtaining a signed confidentiality agreement from authorized collaborators;*
- *Reporting any inappropriate use of the tool or query data to the University Privacy and Security Officer at privacy@umn.edu*

Please click "Yes, I agree"

The image shows two buttons side-by-side. The left button is labeled "Yes, I agree" and is highlighted with a blue border. The right button is labeled "No, I disagree".

II. Navigating the Workbench

Helpful Hints in this Guide



For those who prefer to learn a software tool by exploring independently rather than following a User Guide, a light bulb icon is used within this guide to highlight helpful and nonobvious tips.

Workbench Screen

Following the validation of your login credentials, the **i2b2 Query & Analysis Tool** screen is displayed and available for use. The basic **i2b2** workbench screen provides six primary tabbed panes for identifying, building and running cohort discovery queries. These panes are indicated in the figure below and described in the accompanying text.

The screenshot shows the i2b2 Query & Analysis Tool interface. The top navigation bar includes the tool name, project name (AHC IE i2b2), user name (Zohara Cohen), and various utility links (Find Patients, Analysis Tools, Message Log, Help, Logout). The main workspace is divided into six numbered panes:

- 1. Navigate Terms:** A hierarchical folder structure containing categories like Demographics, Diagnoses ICD9, Outpatient Vitals, Problems, Procedures CPT, and Visits.
- 2. Find Terms:** A search input field for finding specific terms.
- 3. Workplace:** An empty area for building queries, with a yellow box indicating "drop a term on here".
- 4. Previous Queries:** A log of previous queries for the current user.
- 5. Query Tool:** The main area for defining query criteria, including a Query Name field, a Temporal Constraint dropdown, and three query groups (Group 1, Group 2, Group 3) with sub-headers for Dates, Occurs > 0x, and Exclude.
- 6. Query Status:** A panel for monitoring query progress and results.

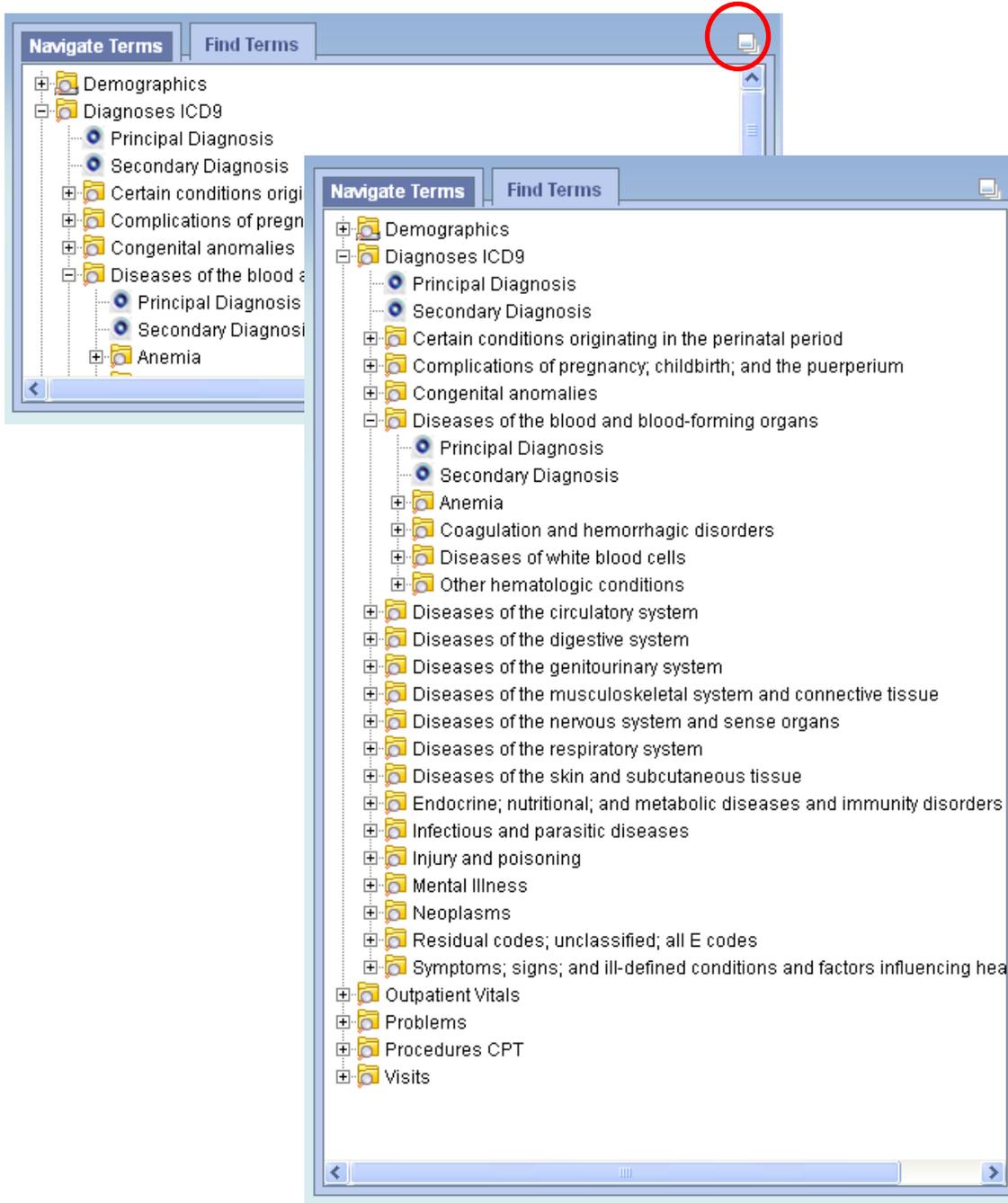
At the bottom of the Query Tool pane, there are buttons for Run Query, Clear, Print Query, and a status indicator showing 0 Groups. Navigation arrows and a New Group button are also present.

- 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 pane is inactive in the current UMN version of i2b2.
- 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.

Resize Workspace



The **Resize Workspace** icon will toggle between the default tabbed pane size and a vertically expanded view of the pane. This can help minimize the amount of vertical scrolling. Any tabbed pane that includes this **Resize Workspace** icon will allow you to expand the tab vertically to the full available height of the browser window. This is a very useful option!



III. Constructing and Running your Query

Selecting Query Criteria

The **Navigate Terms** pane provides a hierarchical tree of search concepts that can be used to find and specify query criteria. You can search through demographics, diagnoses, vitals, problems, procedures, and features of the visit.

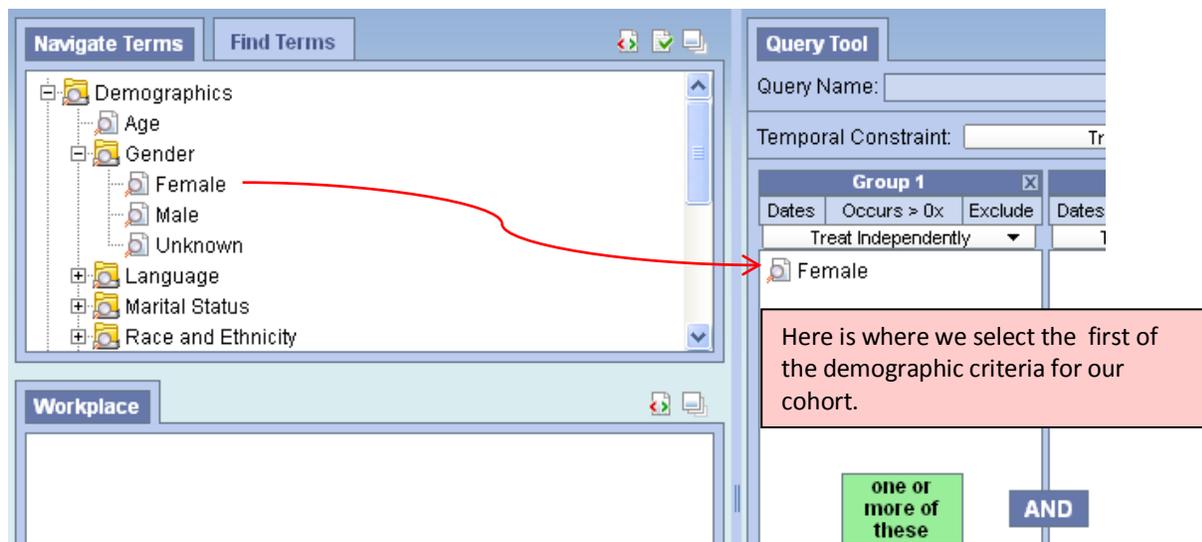
Building Your Query (with a worked example)

Imagine that you want to conduct a trial of a new migraine drug and are designing your study and recruitment plan. You may want to know: how many patients in our database are women, 18 years or older, get migraine headaches with auras (ICD9 codes: 346.0, 346.00, 346.01, 346.02, 346.03), and who are not currently pregnant?

Follow the pink boxes to see follow a sample cohort query.

Once you find the concept in the **Navigate Terms** or **Find Terms** tabbed pane,

1. Click, hold and drag the term from the **Navigate Terms** pane.
2. Drop it into the next available **Group panel** (e.g. Group 1 below) This will add it to the active query criteria.



- The search will turn up cases that meet any one of the criteria dropped into a single Group. In logical terms, these items are grouped by “OR”. A green informational message will pop up to remind you of this relationship; it will say “one or more of these”. In the example below, a patient will be included if she has any of the migraine diagnoses listed in Group 2.
- The criteria in one Group will be treated as an “AND” with items in another Group. In other words, both Groups will have to be true. In the example, the patients will have to be female and have one of the diagnoses in Group 2.



Remember: Criteria placed in the same Group are evaluated to see if any one of them is true. In most cases, each demographic category should get its own group. So, for example, you would not place Gender criteria and Language criteria in the same group, or else you will get –say – patients who are either Female or speak Dutch.

Query Tool

Query Name: _____

Temporal Constraint: Treat all groups independently

Group 1	Group 2	Group 3
Occurs > 0x	Occurs > 0x	Occurs > 0x
Age >= 18 Years	Female	346.00 Migraine with aura, without mention of intractable n 346.01 Migraine with aura, with intractable migraine, so st 346.02 Migraine with aura, without mention of intractable n 346.03 Migraine with aura, with intractable migraine, so st

one or more of these AND one or more of these AND one or more of these

Run Query Clear Print Query 3 Groups New Group

Here we have added a second demographic criterion. Note that it must be put in a new group in order to identify cases for which “age>=18 years” is true AND “Female” is true.

These are the diagnostic criteria for patients with migraine headaches with auras. Only one condition in this group has to be true for a patient to be included..

Modifiers: Principal vs. Secondary diagnosis

For each diagnosis you choose, you can select Principal Diagnosis, Secondary Diagnosis, or neither (the selection of Principal or Secondary is considered a “modifier”). If you select Principal Diagnosis, only patient encounters for which the selected diagnosis was the principal diagnosis will be included. Likewise, if you select Secondary Diagnosis, only patient encounters for which the selected diagnosis was a secondary diagnosis will be included. If you don’t select a modifier, the search will default to include both principal and secondary diagnoses.

Navigate Terms Find Terms

- Secondary Diagnosis
 - 339.00 Cluster headache syndrome, unspecified
 - 339.01 Episodic cluster headache
 - 339.02 Chronic cluster headache
 - 339.03 Episodic paroxysmal hemicrania
 - 339.04 Chronic paroxysmal hemicrania
 - 346.00 Migraine with aura, without mention of intractable migraine
- Principal Diagnosis
- Secondary Diagnosis
 - 346.01 Migraine with aura, with intractable migraine, so stated, with
 - 346.02 Migraine with aura, without mention of intractable migraine
 - 346.03 Migraine with aura, with intractable migraine, so stated, with

Query Tool

Query Name: _____

Temporal Constraint: _____

Group 1

Occurs > 0x

346.00 Migraine with aura, without mention of intractable migraine

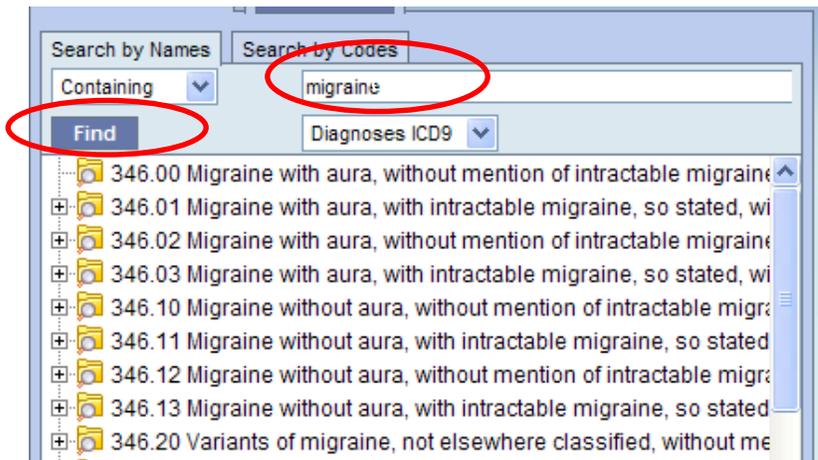
In the example above, the Principal Diagnosis option was chosen for “346.00 Migraine with aura, without mention of intractable migraine ...”. Though the whole item description doesn’t show in the Query Tool on the right side of the image, if you were to scroll to the right in the i2b2 interface you would see “[Principal Diagnosis]” appended to the end of the diagnosis description.

For folders that includes many diagnoses in sub-folders,

- if you choose the folder itself and drag to your query space, you will be automatically choosing both principal and secondary diagnoses.
- if you expand the folder and select a modifier, you will be choosing that modifier for all diagnostic codes contained with that folder.

Using Find Terms to Help Develop Your Query

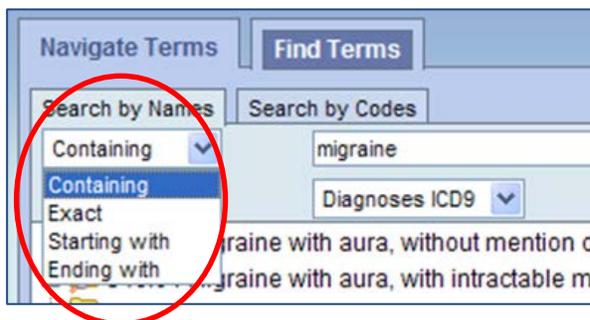
Although **Navigate Terms** can often be used as the primary tool to build a query, **Find Terms** offers additional options for finding the criteria you need. These options include **Search by Names** and **Search by Codes**. Using **Search by Names**, you can find your query terms by searching for full words, partial words or phrases.



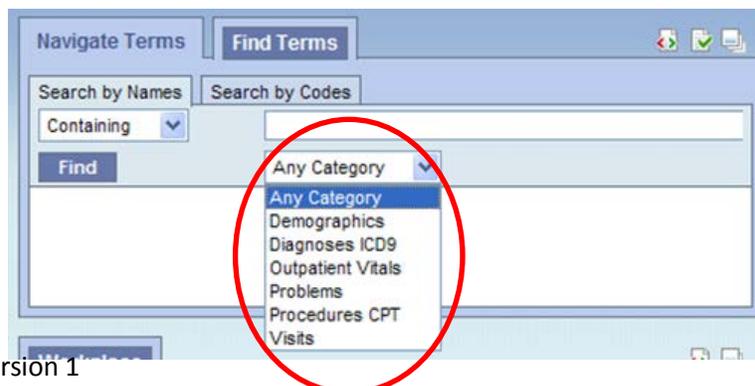
After selecting Find Terms, make sure you select **Search by Names** (it will be shown in lighter-greener shade of blue). Once you enter the search terms, click the **Find** button.

Here we have found the ICD9 codes for migraine headaches by using: **Find Terms > Search by Name**.

A pull-down menu allows you to choose what sort of match you want between the terms that are found and the words/phrase you have entered. You can choose: Containing, Exact, Starting with, or Ending with.

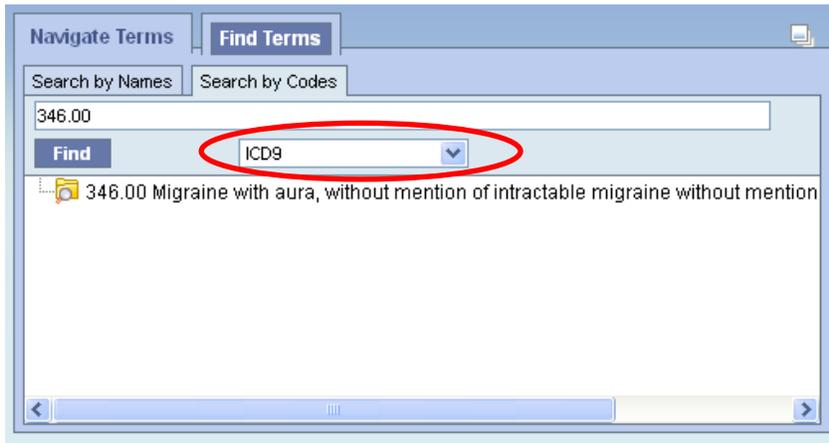


A second pull-down menu allows you to indicate the category of the term you are seeking. For example, in the search above for terms related to "migraine", the option "Diagnoses ICD9" was chosen.



If your initial search turns up no matches, try using different names for the condition or therapy you are seeking. For example, many cancers are referred to “malignant neoplasms” in the Diagnoses ICD9 vocabulary.

If you know the code you are looking for and don’t want to have to navigate the hierarchy in the **Navigate Terms** tab, you can use the **Search by Codes** option under **Find Terms**. Make sure you click on **Search by Codes** after selecting **Find Terms**.



As with searching by names, when searching by codes, you have the option of limiting the search to certain categories.

Find Terms > Search by Codes only works if you provide the exact code. If you would like to search for terms using a code, but want to find all ICD9 or CPT codes that include a string (for example, you want all codes that have “346.” as part of the code including 346.00, 346.01, 346.02, etc.), you can use the **Search by Names** option and enter the code string. This works because all the terms that are searched by name include the code as part of the name. For example, “346.03 Migraine with aura, with intractable migraine, so stated, with status migrainosus”.



Note that the **Find Terms** mode of finding criteria for your search does not allow you to specify modifiers (i.e. principal vs. secondary diagnosis) for your diagnostic criteria. However, if you mouse-over the folder icon to the left of a term once you have found it, you’ll see the full pathway for the term within the code hierarchy. You can use that information to go back to the **Navigate Terms** mode, navigate to the terms you are interested in, and select the modifiers you want.

Demographics

The age field searches for the patient’s age today (at the time of the search) as opposed to the age at the time of the clinical encounter.

Diagnoses vs. Problems

Diagnoses are associated with a particular encounter, whereas Problems, which come from the problem list in the patient’s medical record, are meant to provide a succinct overview of the patient’s most important health information. The Problems will include chronic medical conditions, family history, past tests, and past procedures.

Editing Your Query



1. To delete a criterion you have already dropped into a group, right-click on the icon to the left of the term and select **Delete**. (See number 1 in red circle below.)
2. To delete a whole group, click on the **X** on the upper right-hand corner of the Group box.
3. To create a new Group, which will appear to the right of the existing Groups, click **New Group**.
4. To clear all search information that is currently in the Groups of the **Query Tool** pane, click **Clear**.

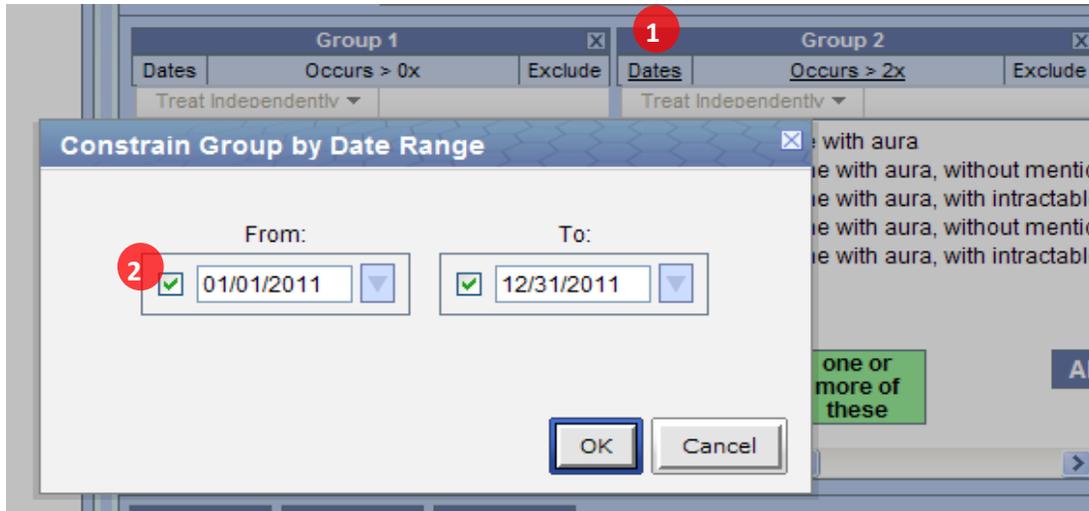
The screenshot shows the Query Tool interface with three groups of criteria. The interface includes a 'Query Name' field, a 'Temporal Constraint' dropdown set to 'Treat all groups independently', and three group boxes labeled 'Group 1', 'Group 2', and 'Group 3'. Each group has a header with 'Dates', 'Occurs > 0x', and 'Exclude' options. Group 1 contains the term 'Female'. Group 2 contains several terms related to 'Migraine with aura'. Group 3 is empty. A 'Delete' context menu is open over a criterion in Group 2, with a red circle '1' highlighting it. A red circle '2' highlights the close 'X' button on Group 1. A red circle '3' highlights the 'New Group' button at the bottom right. A red circle '4' highlights the 'Clear' button at the bottom left. Green boxes with the text 'one or more of these' are placed below Group 1 and Group 2. A yellow box with the text 'drop a term on here' is placed below Group 3. The bottom status bar shows '2 Groups' and buttons for 'Run Query', 'Clear', 'Print Query', and 'New Group'.

Refining Your Query

Constraining Dates of Encounter

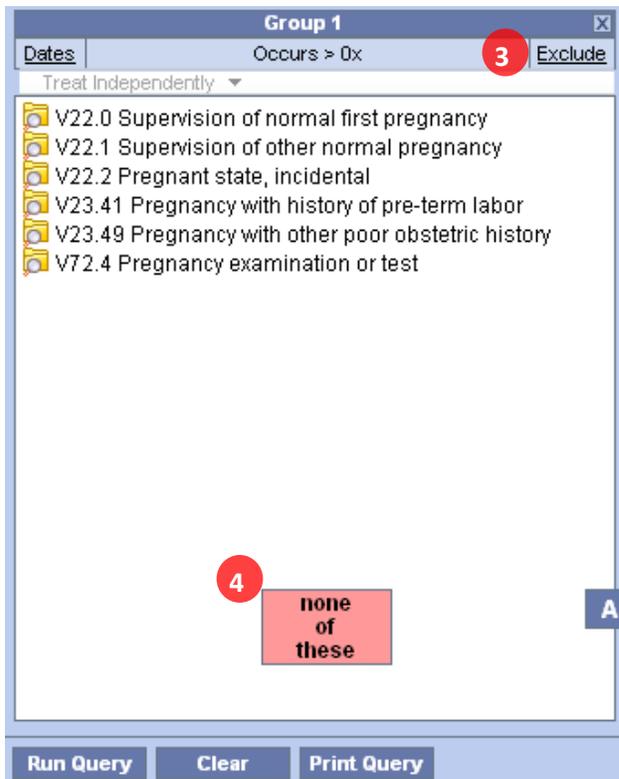


In some cases you will choose to restrict your search to encounters that took place within a certain time period. To include only certain dates, click on **Dates** at the top left of the given Group panel (see 1 in figure). A pop-up window will allow you to specify the date range (2) and once the dates are selected the word Dates will be underlined in the panel (1). **Note that the date constraint function works best when applied to the diagnoses or procedures.**



Adding Exclusion Criteria

You can create a group of criteria and then click **Exclude** (3) to exclude all patients that meet any of the criteria in the list. A pink informational message (4) reminds you that “none of these” will be included.



Here we exclude patients with pregnancy in their medical record.

If we exclude all patients with pregnancy-associated diagnoses, we would actually be excluding all patients who ever had a pregnancy in their medical record. For our example, we're interested in patients who are currently pregnant. To find that out, let's restrict the search for pregnancy codes to the last eight months (here, November 1, 2012 until July 1, 2013). To do this, we make use of both the **Dates** option and the **Exclude** option.

The screenshot shows two windows, 'Group 1' and 'Group 2', with a 'Constrain Group by Date Range' dialog box overlaid. Group 1 contains a list of pregnancy-related ICD-9 codes (V22.0, V22.1, V22.2, V23.41, V23.49, V72.4). Group 2 contains 'Female' and 'Age 2'. The dialog box has 'From:' and 'To:' labels, each followed by a date picker set to 11/01/2012 and 07/01/2013. There are 'OK' and 'Cancel' buttons at the bottom right of the dialog.

Multiple Occurance



In some cases you may choose to make your selection criteria more conservative (more specific) by choosing only to include patients who have a particular diagnosis (or treatment, etc.) in their record more than once. To do so, click on **Occurs > 0x** (see 1 in figure). The constraint window will pop up where you can specify the number of occurrences (2). The label on the group will change to reflect this additional constraint (1).

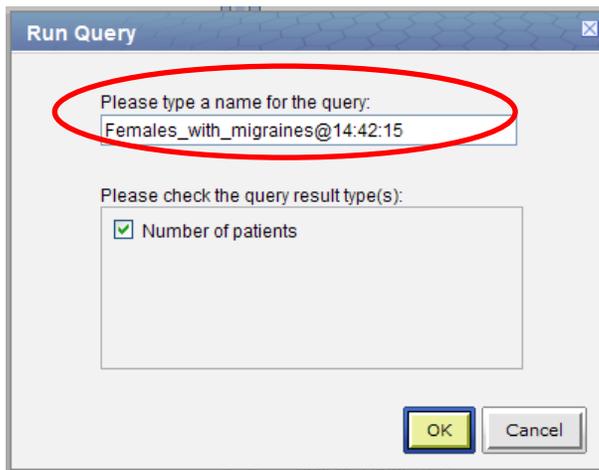
The screenshot shows two windows, 'Group 1' and 'Group 2', with a 'Constrain Group by Number of Occurances' dialog box overlaid. Group 1 contains 'Female'. Group 2 contains 'Occurs > 2x' and a list of migraine-related ICD-9 codes (346.0, 346.00). The dialog box has a text field 'Event(s) within the group occur more than' followed by a dropdown menu set to '2'. There are 'OK' and 'Cancel' buttons at the bottom. A green box with the text 'one or more of these' is visible in the background.

Running Your Query

Click on **Run Query** at the bottom of the Query Tool pane.

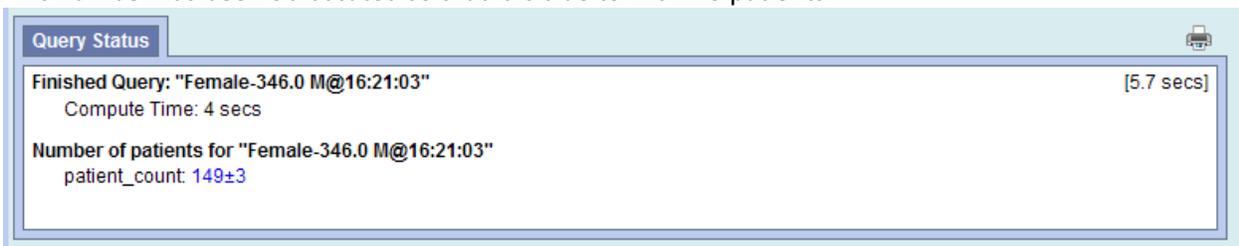


A **Run Query** window will pop-up. Before clicking **OK** you can rename the query. You may want to do this for easier reference when you retrieve the query at a later time. We recommend that you leave the time of the query (to the right of the “@”).

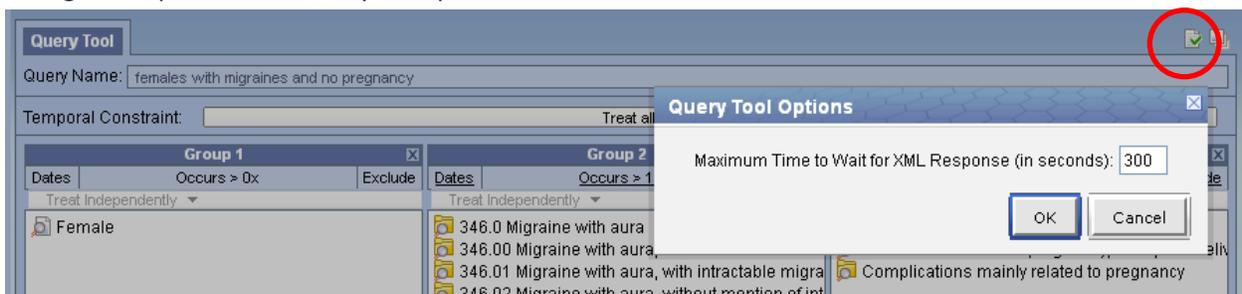


Query Status Tab

The **Query Status** panel shows the result of the query for the aggregate number of potential patients. This number has been obfuscated so that it is true to within 3 patients.



If your query is taking a long time to run, it may time out. The default time allocated to a query is 180 seconds, or 3 minutes. To increase the amount of time that i2b2 will allow your query to run, you can change the option in the Query Tool panel.



IV. Advanced Query Tools

Previous Queries

Every time you run a query it is listed in the **Previous Queries** tab of your workbench screen. It will be assigned a default name that includes the time it was run, unless you specified a new name before running the query. You can rename a query once it is logged in the list of **Previous Queries**.



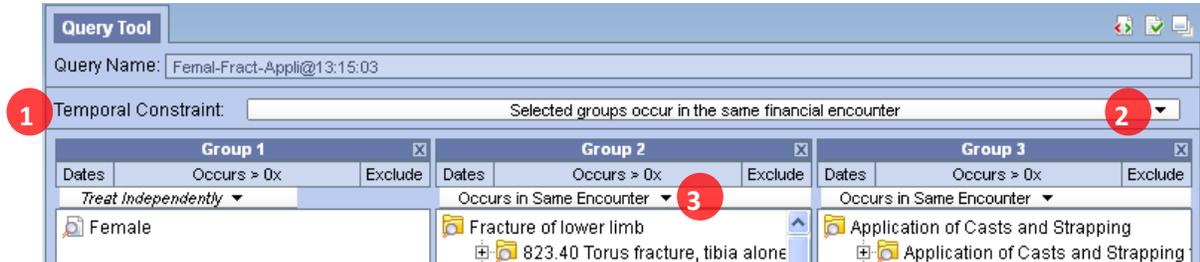
You may want to retrieve a previous query to see the results again or to tweak some parameters and run the modified query. To do either of those, select the query in the **Previous Queries** tab, drag it to the text box labeled **Query Name**, and click the **Run Query** button. See screen shot below.

Temporal Constraint

In the default situation, all of the groups have to be true for a patient in order for the patient to be counted towards your cohort (this is the “AND” between the group panels) and the truth value of each group is found independent of the other groups. However, you have the option to constrain the query so that the groups will only be true if their events coincide with the timing of events in another group. So, for example, if you have entered a list of diagnosis codes in Group 1 and a list of procedure codes in Group 2, you can constrain the search for patients to those who have documented procedures in Group 2 during the same encounters when the diagnoses in Group 1 were documented in their record.

The option to constrain the criteria to common encounters, can be found in the pull-down menu for the **Temporal Constraint** (see 1 in the figure below) field in the Query Tool pane. The default is “Treat all groups independently.” You can choose “Selected groups occur in the same financial encounter” (2). Once you select the “same financial encounter” option, the headings of all the Group boxes will change to, “Occurs in Same Encounter” (3). Note that there is a bug in the software that sometimes causes problems when “Occurs in Same Encounter” is applied to a group that contains Demographics criteria.

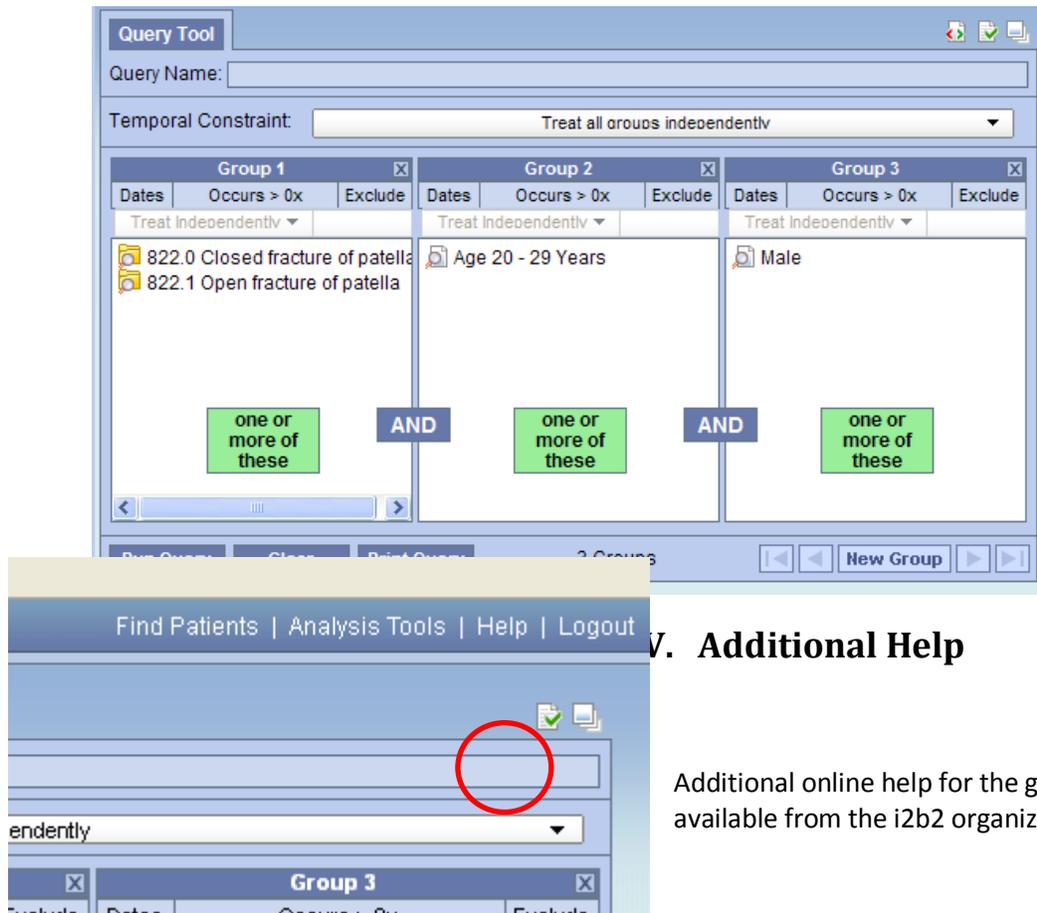
Therefore, we advise that you go to the heading of groups that include Demographics criteria and turn the heading (in the white bar) back to "Treat independently". See, for example, Group 1 in the figure below. We are working on having this problem fixed.



Tips for Optimizing your Query

To optimize the search time for your query, choose the most restrictive criteria as your first Group and move to increasingly broad categories as you move to the right.

For example, if you are searching for males in their 20's who have broken their knee caps, the best way to construct the query is shown in the figure below. Group 1 narrows the cohort to only patients who have broken their patellas (knee caps), Group 2 selects those aged 20-29 within that group, and Group 3 selects only the males from within that cohort.



V. Additional Help

Additional online help for the general i2b2 tool is available from the i2b2 organization and can be

accessed by clicking Help in the upper right-hand side of the workbench.

If you would like an in-person demonstration of i2b2 or have further questions, please contact the Informatics Consulting Service (a service line of the CTSI) at ics@umn.edu. We will do our best to respond to your inquiry within 2-3 business days.