The New ExportXLS plugin for the i2b2 (MiCARD) webclient

**Abstract:** This plugin tabulates unidentified patient data, and applicable diagnoses from specified concepts, of a Patient Set; as well as provides convenient exportation of these data to an Excel spreadsheet. When used appropriately, it should help reduce the latency a researcher, using MiCARD, must endure between identifying the proper cohorts and the actual start of a study.

**Background**

The following is excerpted from the MICARD (Massachusetts Integrated Clinical Academic Research Database) website (http://micard.umassmed.edu/):

> MICARD is the UMass implementation of the i2b2 (http://www.i2b2.org) informatics platform for clinical research. MICARD helps researchers overcome one of the greatest problems in population-based research: Rapidly compile large groups of well-characterized patients. Qualified investigators may use the MICARD web-based query tool to determine the aggregate total number of patients at participating UMMHC hospitals who meet a given set of inclusion and exclusion criteria (currently demographics, diagnoses, medications, and laboratory values). Because counts are aggregate, patient privacy is protected.

These data will be most useful for investigators interested in:

- Generating new research hypotheses
- Preparing grant applications that would benefit from pre-identification and/or characterization of a potential research cohort
- Identifying potential cohorts for clinical trials
- Identifying consented biosamples from the UMMS Conquering Diseases Biorepository

Through the years of its existence, MICARD has been continuously proving its importance to the researchers as its user base grows.

The current MICARD / i2b2 (Informatics for Integrating Biology & the Bedside) framework, as well as the IRB requirements, however, limits the information readily available to a researcher. In essence, the following are the steps a researcher must follow:

1) Run a relevant query on MICARD to obtain a list of potential cohorts.
2) Submit an IRB request to obtain permission for further data (both non-identifying and identifying) concerning those cohorts.
3) Wait for IRB permission.
4) Submit request (with IRB permission) to MICARD administration for the data (both non-identifying and identifying).
5) Wait for those data.
6) Obtain the additional (both non-identifying and identifying) data.
7) Sieve through those data to
   a. further narrow down potential cohorts, and
   b. **repeat all these steps** if the resulting cohort population is insufficient for further study.
8) Commence study with the final cohort set.

It has become increasingly obvious that some improvement needs to be introduced to minimize the likelihood of step 7b above, such that researchers may conduct their study with minimum delays.

A way to speed things up for the researchers is to provide them with as much access to the non-identifying data of the potential cohorts as technically feasible, within the constraints of the i2b2 framework that MICARD is based. This has been an on-going effort under the NIH CTSA (Clinical Translational Sciences Award) Grant.

Recently, a new i2b2 webclient plugin, developed in Universita’ di Pavia, Italy, and adapted here in University of Massachusetts Medical School, may be just one of expediting tools for the researchers here.

### Introduction

At a high level, the i2b2 framework (that MICARD is based on) can be said as consisting of a hive of functional cells that manage, manipulate, and convey the payload biomedical data. As a user, a researcher relies on the i2b2 webclient to interface and specify queries to access the relevant data available.

As its name infers, the i2b2 webclient can be accessed using a web browser like Firefox.

The i2b2 webclient contains panels for its user to issue queries through its handy drag-and-drop graphic user-interface (GUI). It also comes with a few “plugins” that are accessible under “Analysis Tools”.

### The ExportXLS Plugin

This plugin provides the following functions:

1. Tabulates any selected Concepts (Ontology terms) relevant (observed on) to a set of selected Patients.
2. Optionally displays relevant unidentified Patient data.
3. Exports aforementioned tabulated data into an Excel spreadsheet file conveniently.

The first function allows researchers to compare the occurrence of concepts in their hypotheses in a set of patients, therefore narrowing down the subset to further focus on.

The second function provides researchers with common, readily available, non-identifying data, such as vital status, birth date, gender, age, language, race, religion, marital status, zip code, state & city of residence, etc. for the patient set of interest to help further narrow down the potential cohorts.

The third function provides a convenient way for a researcher to save these data.

**Usage Instructions**

The high level navigational steps for using the ExportXLS plugin are as following:

1. Click the “Analysis Tools” menu (near the top right corner) to display the set of plugins available (lower right panel).
2. Click the “ExportXLS” plugin from the plugin list to display its tabs.
3. Navigate to the "Specify Data" tab. Then, drag and drop a Patient Set and one or more Concepts (Ontology Term) onto the input boxes.
4. Check or uncheck the **Include Patient Data** box if you want to include the relevant non-identifying information for each patient.
5. Finally, click the "View Results" tab to view the table of the observations.

The following figure shows in better details the key steps involved:
In the example above, the fictitious researcher has a hypothesis that there may be some sort of causal-effect relationship among “hypertensive disease”, “Nutritional deficiencies”, and “Stomach and duodenum diseases” in patients. So, he generated a “Patient Set” based on a query on “Stomach and duodenum diseases”. He then uses that Patient Set so generated in the ExportXLS plugin, while specifying all three of the aforementioned concepts in his analysis run.

A sample resulting table is as shown in the following figure:

Figure 1: Key steps involved in starting & navigating the ExportXLS plugin in the webclient.
The New ExportXLS plugin for the i2b2 (MICARD) webclient

Figure 2: Sample Result Table.

This sample resulting table indicates that there is no occurrence of “Nutritional deficiencies” among the fictitious researcher’s “Stomach and duodenum diseases” cohorts. Furthermore, he may not even have a large enough subset of patients to pursue his study of causal-effect relationship between “hypertensive disease” and “Nutritional deficiencies” in sufferers.

From the above fictitious example, with the ExportXLS webclient plugin, the fictitious researcher can modify his hypothesis immediately without having to endure the long delays outlined before (in the Background section).

Now, suppose our fictitious researcher wants to keep a record of this invaluable quick analysis and move on to try something else, he can then export / download the table to an Excel spreadsheet.

After clicking the [Export Excel] button, the following dialog is displayed:
The New ExportXLS plugin for the i2b2 (MiCARD) webclient

Figure 3: Dialog for opening / saving the resulting Excel spreadsheet.

After clicking the [OK] button, the following dialog simply shows that the status of the download:

Figure 4: Download status dialog from Browser.

Click the [Clear List] button, and then close this dialog after opening / saving the resulting Excel spreadsheet.

An Excel warning dialog like the following may be display:
The New ExportXLS plugin for the i2b2 (MiCARD) webclient

Figure 5: Potential Excel warning dialog to be ignored.

Just click the [Yes] button to proceed displaying the resulting spreadsheet content:

Figure 6: Sample resulting spreadsheet generated.

Browser Compatibilities

- The v2.0 of this plugin only works on the following browsers:
  - Firefox (v.4.0.1),
  - Chrome (v.12.0.742.122), and
  - Safari (v.5.0.5),
  - Internet Explorer (v.8.x or v.9.0.1),
- The i2b2 Web client itself does not work on the following browser:
  - Opera (v.11.50).
Discussions

The labeling of so-called non-identifying data may be subjective. That is, to some, information such as birth date may be too sensitive. In addition, zip code may become identifying data if a particular area happens to be very sparsely populated (such that only small handful of patients share the same zip code) and the other non-identifying data like gender, age, marital status, etc. do not overlap.

These are, of course, valid concerns. The following alterations have been implemented in the deployed version of this plugin locally, for MICARD:

- Dropped zip code.
- Replaced Birth Dates with Birth Years only.

The aforementioned alterations have been evaluated and approved by the UMass Medical School IRB. Indeed, the IRB signoffed on the deployment of the ExportXLS (with the said alterations) on MICARD to be used by researchers to readily access those non-identifying patient data.

Conclusions

MICARD has been and continue to be invaluable in expediting researches. Additional ongoing efforts to further assisting researchers to more efficiently identify their study cohorts are being pursued under the NIH CTSA Grant.

As a result, ExportXLS, a webclient plugin suitable for enhancing the MICARD user experience and reducing the cohort-gathering cycle has been identified and adapted for deployment. Each of the three functions ExportXLS offers should be invaluable additions to a MICARD user’s toolkit. It should also contribute to further expanding the MICARD user base.

In addition, we at UMass Medical School have collaborated with the Universita' di Pavia, Italy, to provide this version (v2.0) of the ExportXLS webclient plugin to the rest of i2b2 community, so that they may benefit from the enhancements and flexibilities introduced to this latest release.

One of those enhancements is the ability for each deploying institution to specify which, if any, column(s) of data (e.g. “Zip code” & “Birth Dates” for UMMS here) they want to exclude from the output table.