Brigham and Women's Hospital

Harvard T.H. Chan School of Public Health

MEDIANS Workbench

User Guide (V 1.2)

Tianrun Cai, M.D.

Document Version: V 1.2

Software Version: 1.2.2

Contents at a Glance

Cover 1
Welcome 6
Credits 7
About the Author 8
Section I. Set Environment Variables 8
I-1. Open Environment Variable Editor 11
I-1-1. From “Start Menu” 11
I-1-2. From “This PC” 11
I-2. Create a New Environment Variable 12
I-3. Launch CHANL 12
Section II. Load Source Data 13
II-1. Load Data from Local Files 13
II-2. Load Data from Database 13
II-2-1. Database Connection 13
II-2-1-1. MSSQL 13
II-2-1-2. MySQL 15
II-2-1-3. XML Database 16
II-2-1-3-1. Generating XML Database File 16
II-2-1-3-2. Load the Folder of XML 17
II-2-1-3-3. Load XML Files 17
II-2-1-4. Access Database 18
II-2-2. Data Selection 19
II-2-2-1. Query Result Display 19
II-2-2-2. Sort Data 19
II-2-2-3. Data Grouping 19
II-2-2-4. Select Data to Load 19
II-2-2-5. Data Information Display 20
II-2-2-6. Load Data to Main Panel 20
II-2-2-7. Automatic Hiding 20
II-2-2-8. Recall the Data Frame 21
II-2-2-9. Index Data on the Main Panel 21
Section III. Data Search Options 22
III-1. Project 22
III-1-1. Create a Project 22
III-1-2. Load a Project 22
III-1-3. Save a Project 22

III-1-4. Delete a Project 22
III-2. CUI Search 23
III-2-1. Load a Dictionary 23
III-2-2. Enter CUIs 23
III-2-3. Select Color 23
III-2-4. Run Search 23
III-3. Term Search 24
III-3-1. Enter Terms 24
III-3-2. Delete Saved Terms 24
III-3-3. Choose Term Search Type 24
III-3-3-1. Full Word or Phrase Search 24
III-3-3-2. String or Stem Search 24
III-3-3-3. Concept Search 25
III-3-3-3-1. Child Concept 25
III-3-3-3-2. Sibling Concept 25
III-3-3-3-3. Negation Information 25
III-3-3-3-4. Certainty Information 25
III-3-3-3-5. FamilyHX 25
III-3-3-3-6. FromCHANLdb (On testing) 25
III-3-3-4. Regular Expression Search 25
III-4. Semantic Type Search (In development) 25
III-5. Similarity Search (In development) 25
III-6. Color Customization 25
III-7. Term and Concept Exclusion 26
III-7-1. Enter Exclusion CUIs or Terms 26
III-7-2. Choose CUIs or Terms 26
III-7-3. Delete CUIs or Terms 26
Section IV. Searching Result Representation 27
IV-1. Main Panel 27
IV-1-1. Run a Search 27
IV-1-2. Reset Highlighted Terms 27
IV-1-3. Column Name Highlighting 27
IV-1-4. Zoom In and Zoom Out 27
IV-1-5. Matched Location 27
IV-1-6. Format Report 28
IV-1-7. Manually Highlighting 28
IV-1-8. Recall Query Result Frame 28
IV-1-9. About CHANL 28

IV-1-10. Switch Report 28
IV-1-10-1. Locate Report Switching Bar 28
IV-1-10-2. Display Report Header 29
IV-1-10-3. Switch Report 29
IV-2. Detailed Matching Information 29
IV-2-1. Information Listed 29
IV-2-2. Collapse and Expand Matching Information 30
IV-2-3. Display Matching Sentence 30
IV-3. Search Result Summary 30
IV-3-1. Summary Data Format 30
IV-3-2. Color 30
IV-3-3. Reflect to the Main Panel 30
IV-4. Concept Dictionary 30
IV-4-1. Data Format 30
IV-4-2. Sort data 31
Section V. Perform Review 32
V-1. Review Model 32
V-1-1. Create a New Review Model 32
V-1-1-1. Variable Name 32
V-1-1-2. Description 32
V-1-1-3. Choices 33
V-1-1-4. Variable Type 33
V-1-1-5. Comment 33
V-1-1-6. Default Value 34
V-1-1-7. Save a Model 34
V-1-2. Edit a Review Model 34
V-1-3. Import a Review Model 35
V-1-4. Delete a Review Model 35
V-2. Record Review Results 35
V-2-1. Record Results 35
V-2-2. Enter or Select Comment 35
V-2-3. Review Time 35
V-2-3-1. CHANL Timer and Clock 35
V-2-3-2. Time Records 35
V-2-4. Append Results 35
V-3. Review Result Files 36
V-3-1. Create a Review Result File 36
V-3-2. Save Change to a Review Result File 36

V-3-3. Export a Review Result File 37
V-3-4. Delete a Review Result File 37
Section VI. Appendix 38
VI-1. Create XML database for CHANL (Python) 38
VI-2. SQL query example 40
VI-2-1. Create a View in MSSQL 40
VI-2-2. Create a View in MySQL 40
VI-2-3. Query Data from a Table 40
VI-3. Example Review Model 41
VI-3-1. Model Creating 41
VI-3-2. Review Model Generated 41
VI-4. Brief Introduction of Regular Expression 41

Welcome

CHANL, Chart Review Powered by Natural Language Processing, is designed to improve the efficiency and accuracy of chart review of narrative notes from electronic medical records. Using natural language processing technology, CHANL can perform semantic analysis for narrative notes and intelligent search for thousands of concepts simultaneously and can automatically identify and highlight other terms and concepts related to keywords in notes for helping chart review. This tool has been used by multiple research groups from different institutions including Partner's Healthcare System, Harvard School of Public Health, Harvard Medical School, Boston Children's Hospital, and VA Healthcare System.

Credits

Software Designer

Tianrun Cai, M.D.

Software Developer

Tianrun Cai, M.D.

Software Graphics

Tianrun Cai, M.D.

Software Contributor

Tianxi Cai, ScD.

Bryan Cai, PhD (Stanford University**)**

User Guide Author

Tianrun Cai, M.D.

User Guide Editor

Amanda King, MS (T.H. Chan Harvard School of Public Health**)**

Jacqueline Honerlaw, MPH (VA Boston Healthcare System**)**

Splash Graphics

Amanda King, MS

Tianrun Cai, M.D.

About the Author

Tianrun Cai, M.D., is a faculty member serving as a health-informatician in the Division of Rheumatology, Immunology and Allergy at Brigham and Women's Hospital. After a career as a general surgeon, he transitioned to full-time research in the field of health informatics. He has been designing and writing tools for extracting clinical data from narrative reports using Python and Java for many years. His projects have focused on developing approaches using natural language processing (NLP) across different disciplines including cardiology, radiology and rheumatology to improve the efficiency of medical research.

Section 0. Environment Requirement

Windows 7,8,10 64bit OS

MicroSoft Visual C++ Redistributable 2015-2019

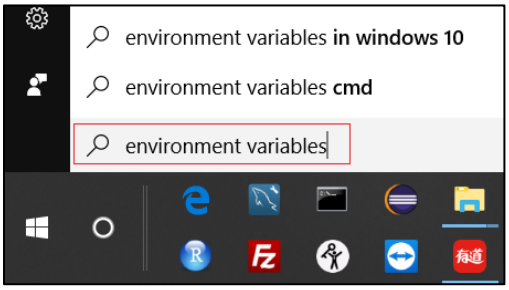
Section I. Set Environment Variables

I-1. Open Environment Variable Editor

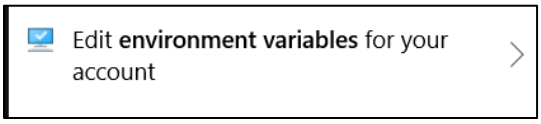
Note: The icons and menus are from Windows 10. In different versions of Windows such as Windows server or Windows 7, 8 some icons, menus or positions might be different from those displayed below.

I-1-1. From "Start Menu"

1. Click the search button on the "Start" menu or taskbar.
2. Search for "environment variables"

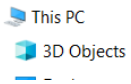


3. Click on the option "edit environment variables for your account" to open the environment variable editor.

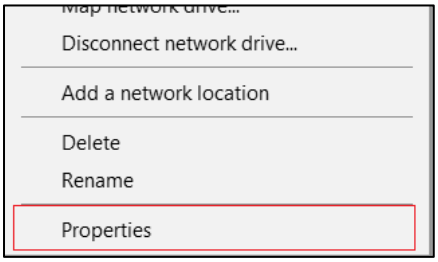


I-1-2. From "This PC"

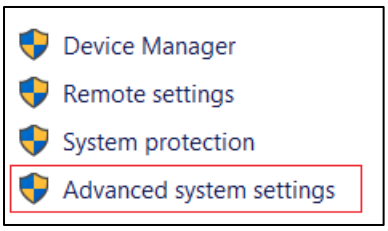
1. Right click "This PC",



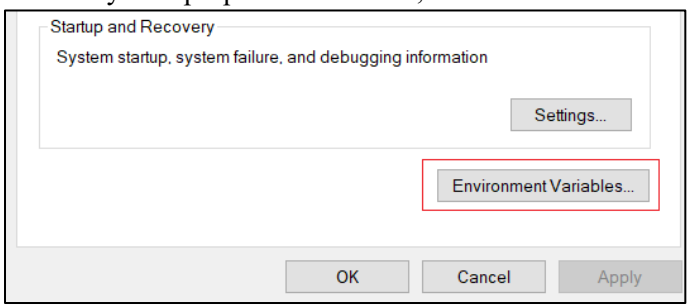
2. choose "Properties" on the pop-up menu



3. then click "Advanced system settings" in the "System" window.

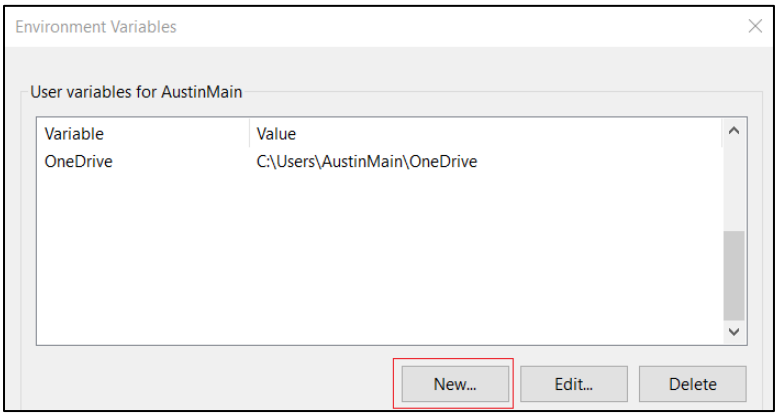


4. In the “System properties” window, on the “Advanced” tab, click “Environment Variables”.

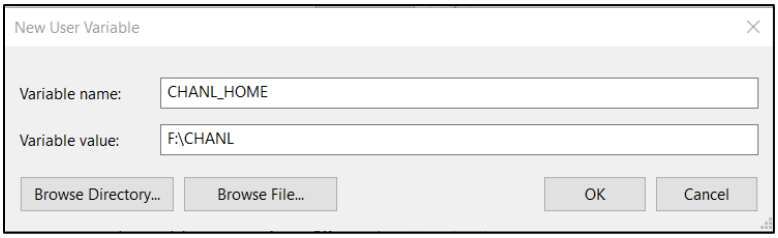


I-2. Create a New Environment Variable

1. Click the button “New”



2. Then enter “CHANL_HOME” for the variable name and enter in the file path of the CHANL folder in the “Variable value” field. E.g. “F: \CHANL”



3. Log out or sign out of the current Windows user (**instead of “disconnect”**), then log-in again.

I-3. Launch CHANL

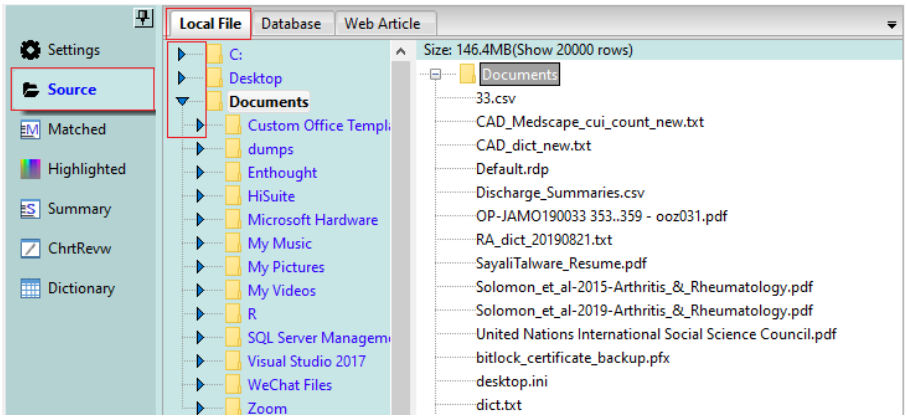
Navigate to the CHANL folder and double click the icon CHANL.exe, or click on the shortcut on the desktop to launch CHANL.

Section II. Load Source Data

Note: Currently, there are two options available: Local files and Databases.

II-1. Load Data from Local Files

1. Select "Source" in the middle of the CHANL interface, click the "Source" tab, and click the "Local file" tab, browse folders by clicking or double click folder names.



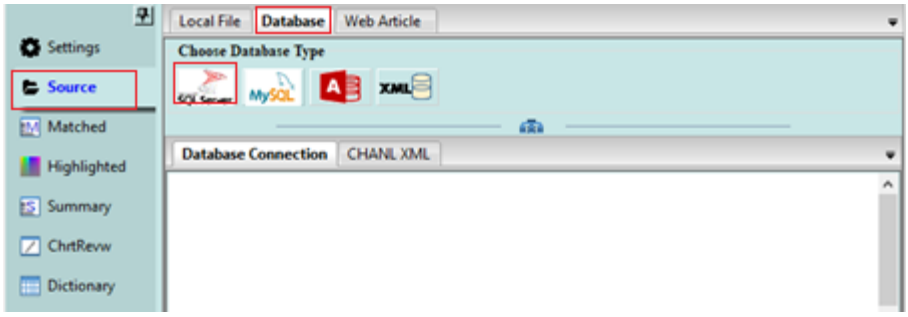
2. In a target folder choose a file to load to CHANL.
   Currently, CHANL supports plain text file such as '.txt', '.csv', '.py' etc. and PDF files. CHANL does not support images, '.doc' files and zipped files.
   The file size allowed is 40Mb. If the size of a file is greater than 40Mb, just the top 20,000 rows will be displayed. This feature could be used to check a sample of huge txt files.
   For PDF files, CHANL will automatically ignore non-text content and non-ascii characters.

II-2. Load Data from Database

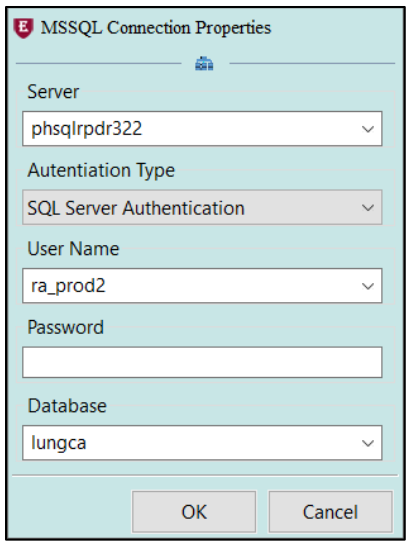
II-2-1. Database Connection

II-2-1-1. MSSQL

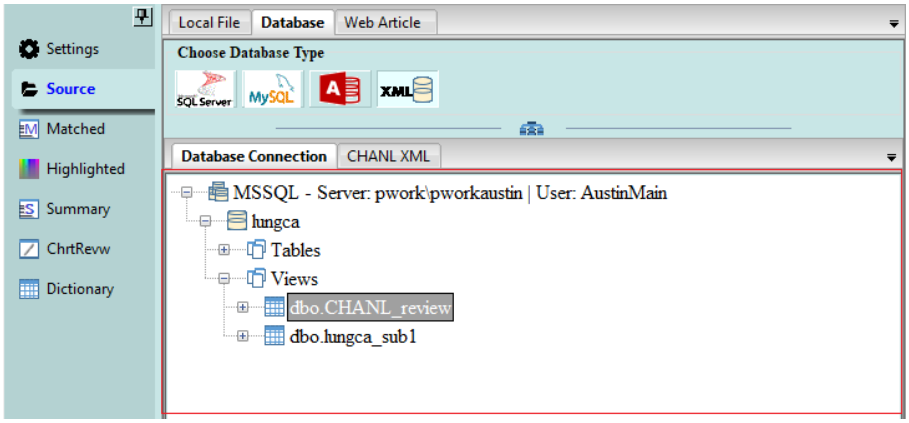
1. Select the "Source" tab in the middle of the CHANL interface, click the "Database" tab, then choose a database type.



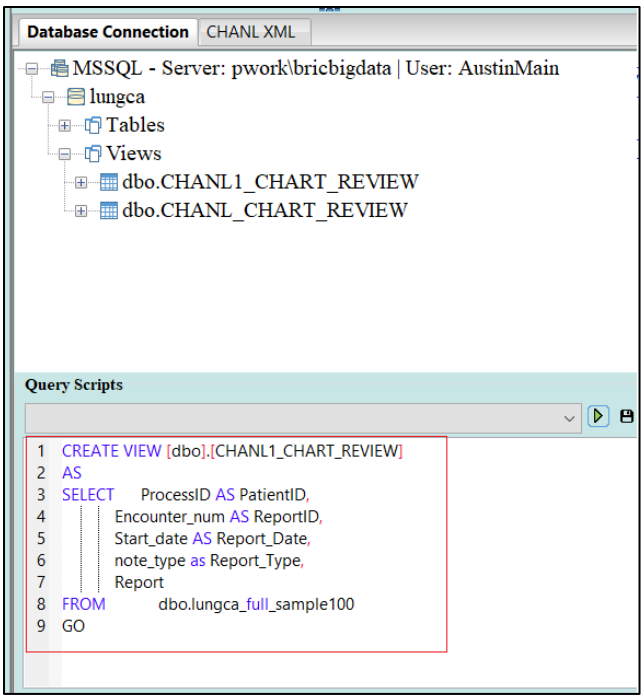
2. Click , and the MSSQL database connection window will pop up.



3. Provide "Server" name, then choose "Authentication Type". If "SQL Server Authentication" is chosen, please enter "User Name", "Password" and "Database" name. If "Windows Authentication" is chosen, "User name" and "Password" are not needed.
4. Click "OK" to launch the connection to the target database. Once the connection is built, see the example below:



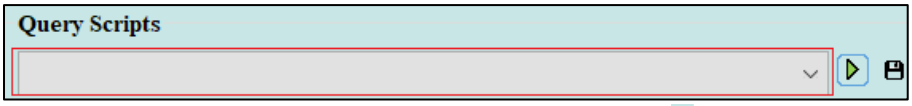
5. In order to have better performance, we recommend building a view to customize the order of column names by "PatientID", "ReportID", "Report_Date", "Report_Type", "Report". We can create a view in CHANL directly using the example SQL script below (also see example SQL scripts in Appendix 2.), when a connection to a specific database is built. Please re-connect the database after creating a new view. (**Please put the date of notes on the 3rd column because CHANL need to recognize the date information when do summary for matched information**)



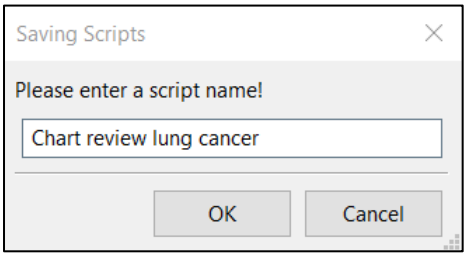
6. Enter or choose SQL scripts in the "Query Scripts" area then click to query patient data to load to CHANL. See an example below:

   *select * from dbo.CHANL_CHART_REVIEW where processID in (1,3,100)*

   A few query examples are listed in the dropdown menu below.

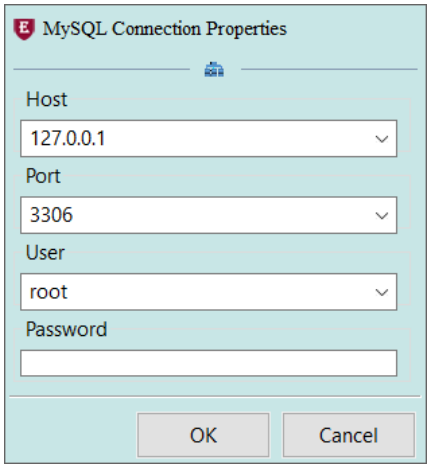


7. Scripts can be saved to CHANL for later use by clicking , then enter a name for scripts on the popup window and click "OK":

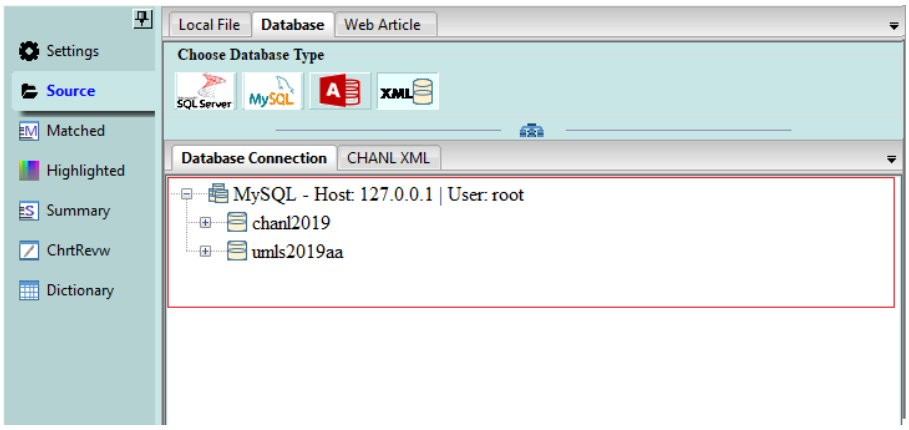


II-2-1-2. MySQL

1. The MySQL database connection window will pop up if we click the button .



2. Enter information for “Host”, “Port”, “User” and “Password”.
3. Click “OK” to launch the connection to MySQL database. For MySQL connection, we don’t need to specify the database name.
4. After the connection is built, unlike the MSSQL database connection in CHANL, the MySQL connection will show different databases. See below:

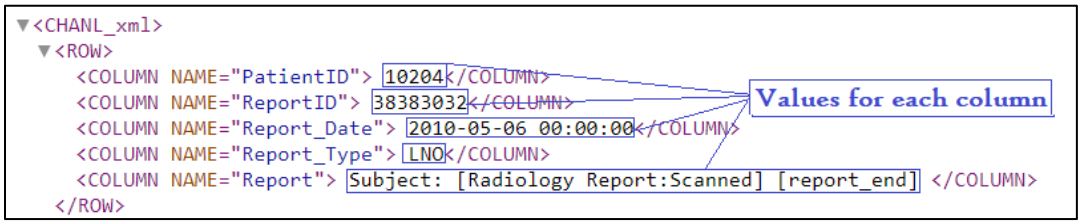


5. Browse a database and choose a target table to query data and load to CHANL. Note: the gramma of query scripts in MySQL is different from MSSQL. Please change the scripts to fit the relevant database.

II-2-1-3. XML Database

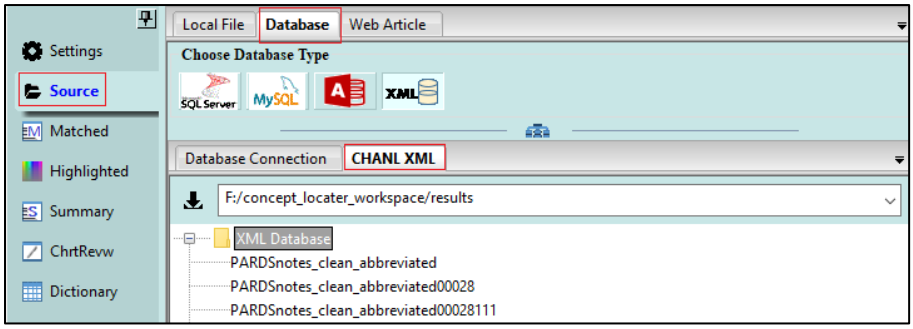
II-2-1-3-1. Generating XML Database File

Here XML databases are XML files saved in a local or shared drive. In order to import XML files to CHANL interface as data sources, we need to convert plain text files to a specific format of XML files shown below:

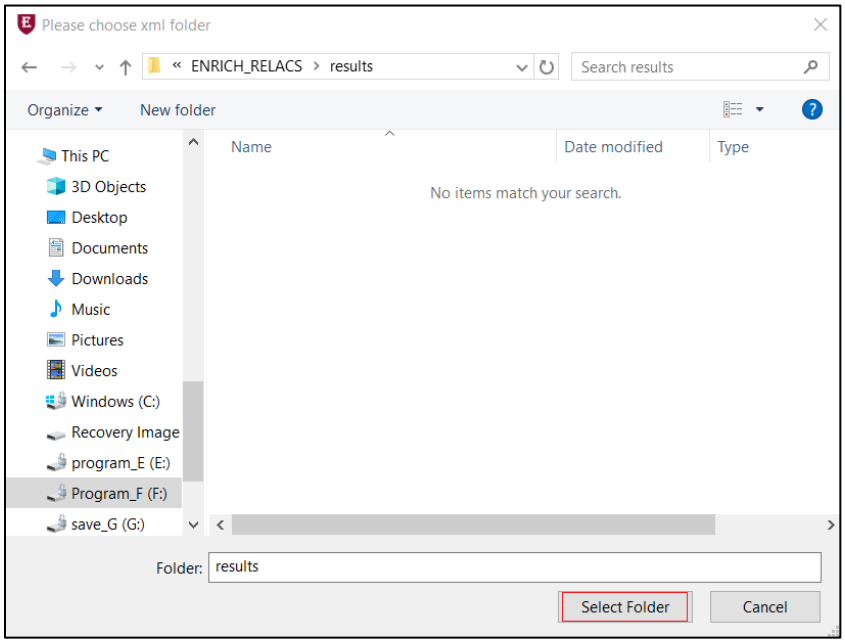


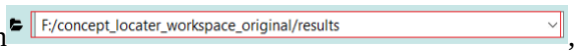
The format of xml files should include the root "CHANL_XML", the first level of node "ROW" and the second level of node "COLUMN NAME". Each "ROW" contains information for each note. Five "COLUMN NAME" values are needed for CHANL including "PatientID" (patient identifier), "ReportID" (note identifier), "Report_Date" (date for the note), "Report_Type", and "Report" (the content of a note). Please see the script in Python in **Appendix 1**. to create XML files using database or local txt files.
When generating XML database, we recommend not exceeding 100 patients for each file.

II-2-1-3-2. Load the Folder of XML

1. Select "Source" in the middle of the CHANL interface, choose the tab "Database" and click the "CHANL XML" tab.
2. 

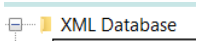
3. Click "Select Folder" on the popup window to browse the computer to select a folder which contains xml files.



4. The path of the folder will be shown 

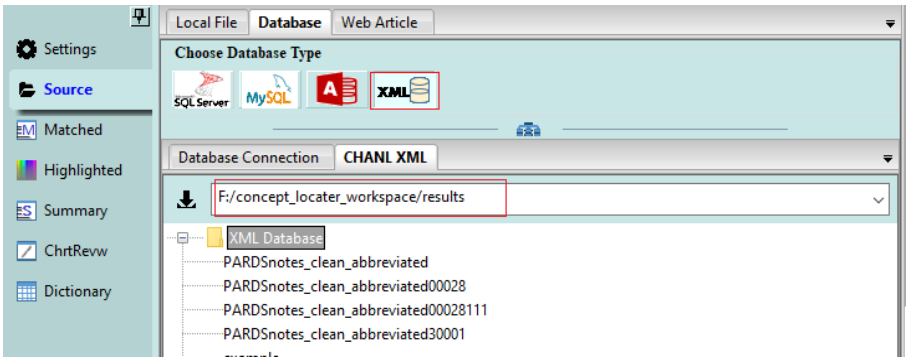
,

II-2-1-3-3. Load XML Files

Choose a file under 

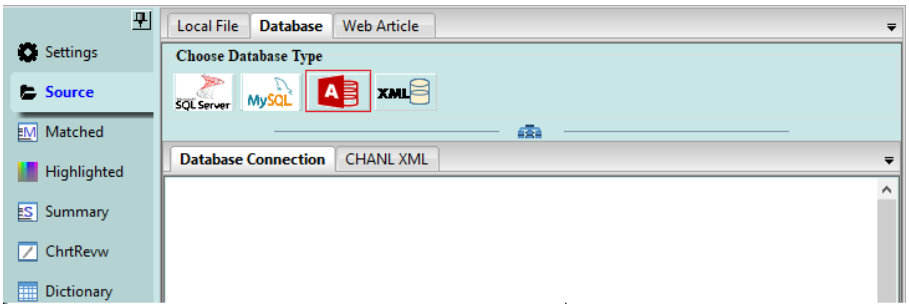
by clicking the file name and clicking

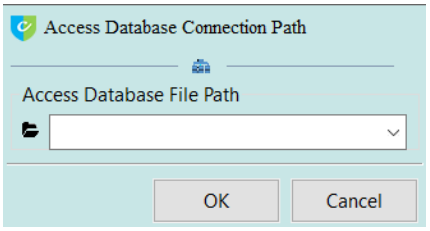
to build the connection to the XML database.



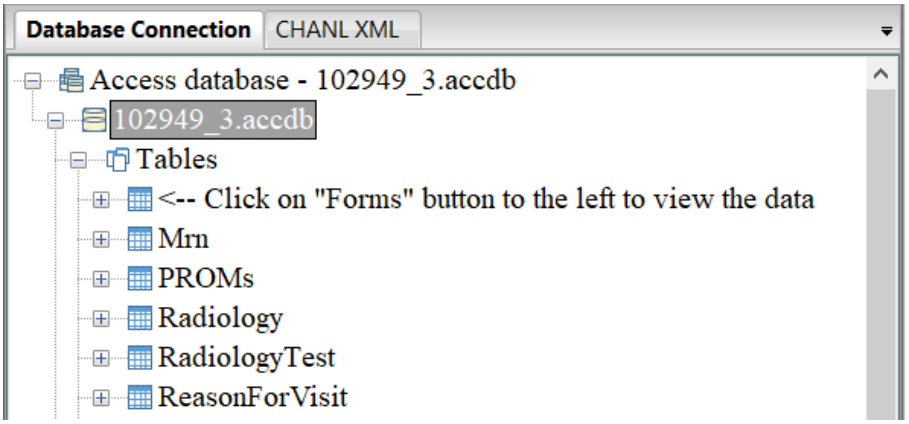
II-2-1-4. Access Database

1. Click the button to open a window for choosing an Access database file path.





2. Click the button on the popup window to choose Microsoft Access database file with extension name ".mdb" or ".accdb".
3. Click "OK" to connect the database. The tables and columns for each table will be listed in the tab "Database Connection"

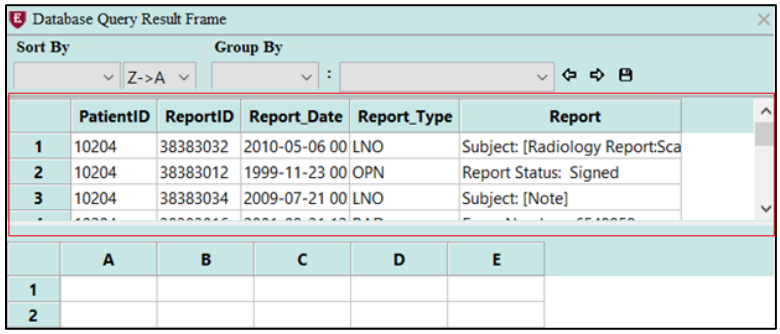


4. Data query gramma is similar as MSSQL. Please read the relevant part for MSSQL above.

II-2-2. Data Selection

II-2-2-1. Query Result Display

The data will be displayed in the upper grid in a popup window when it's loaded from a database. Note: this popup window will always stay on top of opened windows.

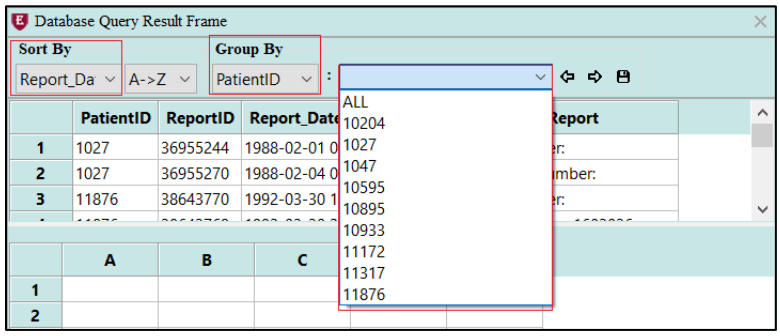


II-2-2-2. Sort Data

Data can be sorted by one of the columns loaded by ascending (A→Z) or descending(Z→A) order.

II-2-2-3. Data Grouping

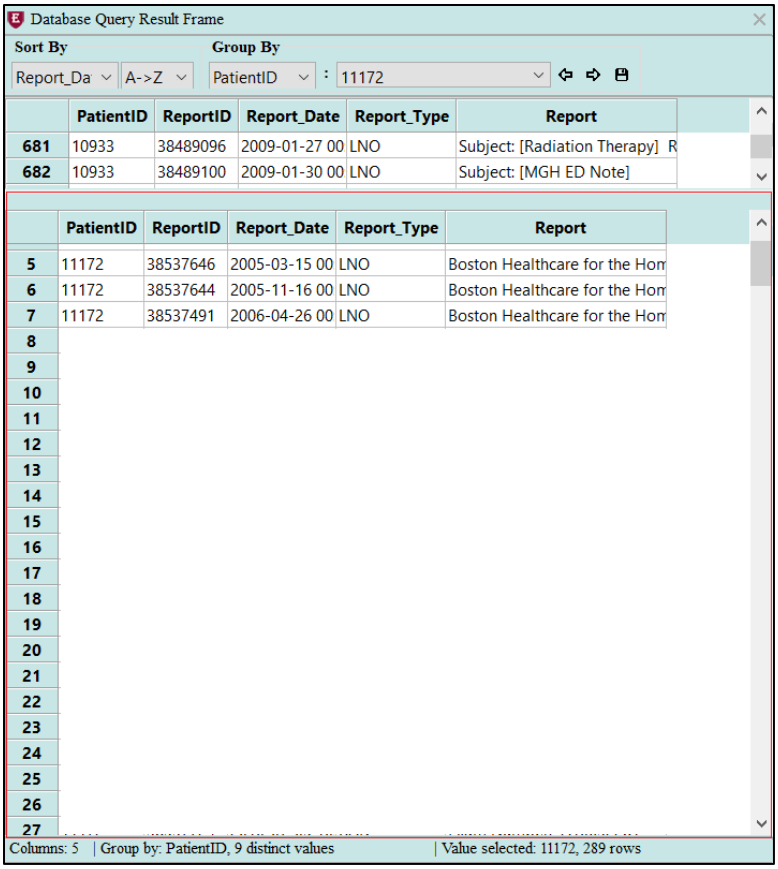
Before loading data to the display panel, we need to group data by a distinct value of a certain column. For example, group data by distinct "PatientID" values.



II-2-2-4. Select Data to Load

Choose a value of "PatientID" from a dropdown menu beside "PatientID"

. Then all data for the specific value chosen above will be loaded to the lower grid in the popup window.

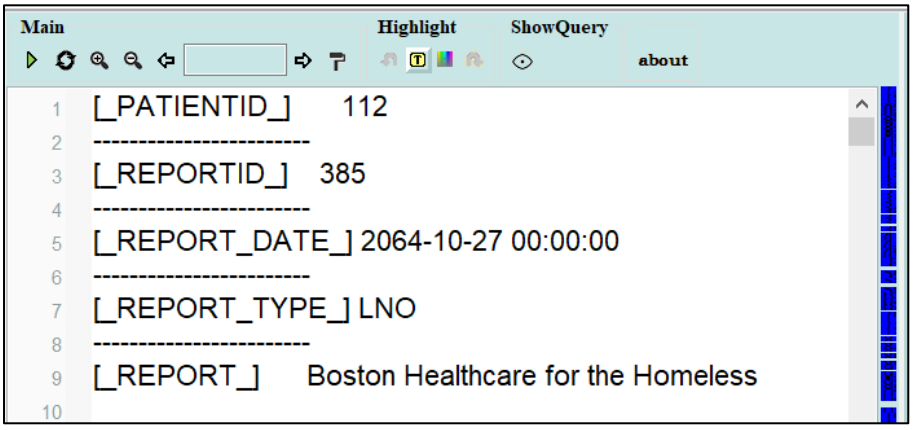


II-2-2-5. Data Information Display

With column number in the database table chosen, “Group by” column name. The count of the distinct value of the “Group by” column name, the selected value and the count of the rows will be displayed on the bottom of the window.

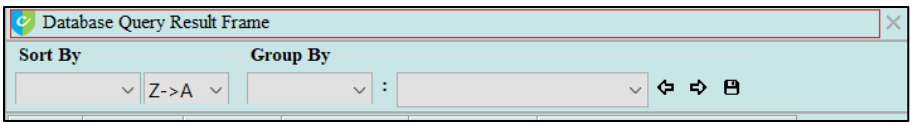
II-2-2-6. Load Data to Main Panel

When the data is selected in step II-2-2-4, the data will be reformatted and loaded to CHANL.



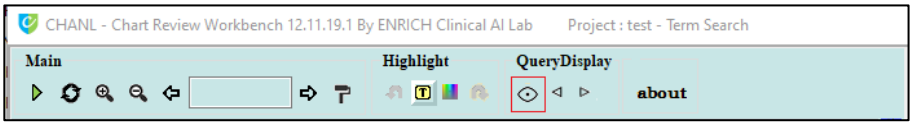
II-2-2-7. Automatic Hiding

Drag the title bar of the “Database Query Result Frame” to the edge of the screen (left, right and top), then move the mouse off. The frame will be automatically hidden.



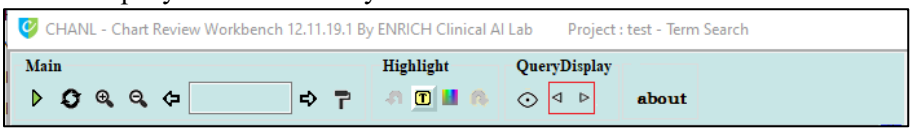
II-2-2-8. Recall the Data Frame

Put the mouse on the edge of the screen when the frame is hidden or click the button on the main panel, and the frame will be displayed again.



II-2-2-9. Index Data on the Main Panel

Click the arrows displayed below to index data if data is already showed on the query result frame and 'Group By' value is already chosen.

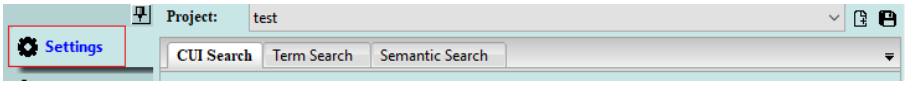


Section III. Data Search Options

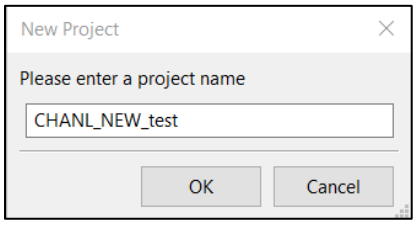
III-1. Project

III-1-1. Create a Project

1. Select the "Option" tab in the middle of the CHANL interface, then click on the top right.

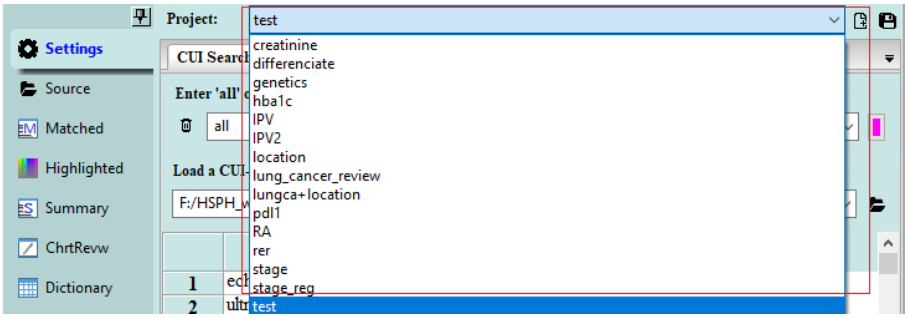


2. Enter a project name in the popup window to create a new project.



III-1-2. Load a Project

1. From the project dropdown menu select a project to load.

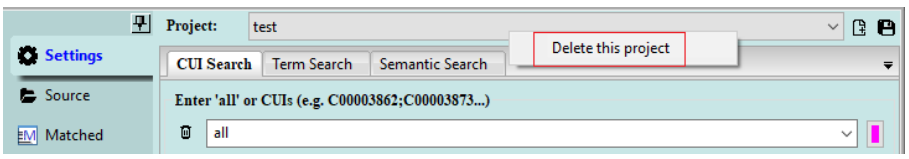


III-1-3. Save a Project

1. A project will be saved automatically after running any search, but you can save a project with another name by clicking on the top right.

III-1-4. Delete a Project

1. Select a project, right click the project name, then click "Delete this project" on the popup menu to delete an existing project.

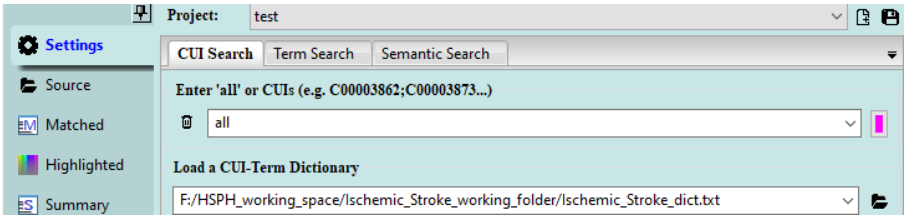


Note: Before doing any type of search, notes need to be loaded and a project needs to be selected or created.

III-2. CUI Search

III-2-1. Load a Dictionary

1. On the "CUI search" tab, click to browse the computer to load a pre-built CUI-term dictionary.

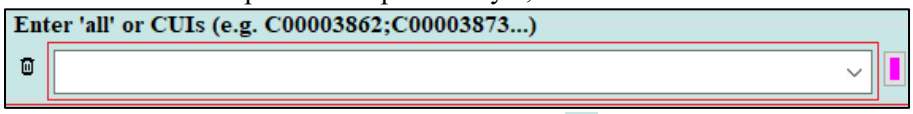


2. The format of the dictionary should be a term and a CUI in each line separated by " | ". See the example below:

```
malignant hilar lung neoplasm | c2607931
malignant hilar lung tumor | c2607931
malignant histiocytosis of lung | c2200137
malignant lung hilum neoplasm | c2607931
malignant lung hilum tumor | c2607931
malignant lung neoplasm | c0242379
malignant lung tumor | c0242379
malignant lymphoma of lung | c2205908
malignant mastocytosis of lung | c2205913
malignant mesenchymoma of lung | c2205920
```

III-2-2. Enter CUIs

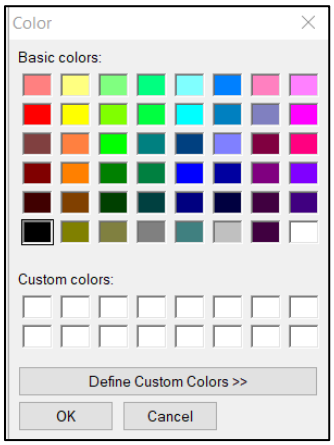
1. Enter a CUI or multiple CUIs separated by ";" or enter "all" to search all concepts in the dictionary.



2. Select a CUI from the dropdown list, then click on the left to delete the CUI from the list.

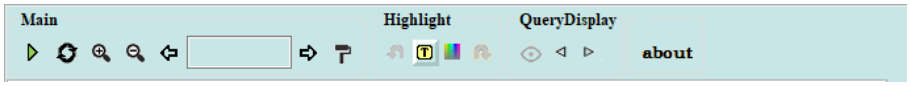
III-2-3. Select Color

1. Click on the "CUI Search" tab to select a color from a popup window below for CUI Search.



III-2-4. Run Search

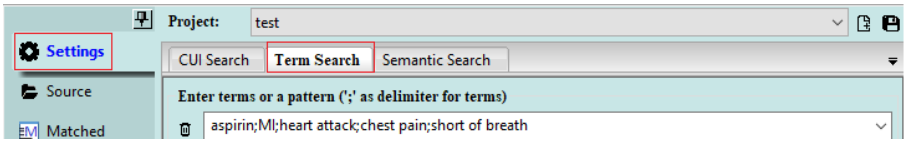
Click on the "Main" panel to run search.



III-3. Term Search

III-3-1. Enter Terms

1. Enter a term or pattern or multiple terms separated by ";" in the space below.

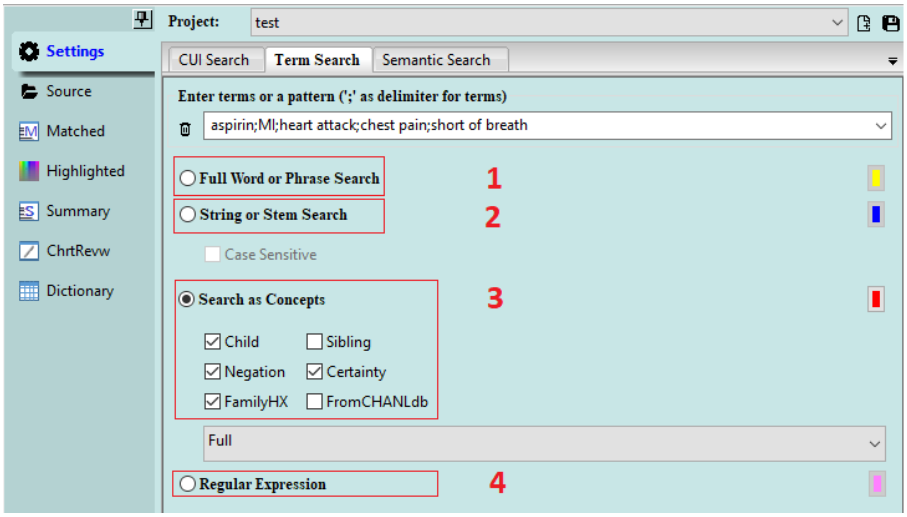


III-3-2. Delete Saved Terms

1. Choose a record from the dropdown list then click to delete the record from the list.

III-3-3. Choose Term Search Type

There are four options for performing term searches: Full Word or Phrase Search, String or Stem Search, Concept Search and Regular Expression.

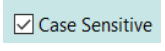


III-3-3-1. Full Word or Phrase Search

This setting will search a term entered as a word with a boundary and ignore case.
E.g. If we search a word "**arthralgia**" in two sentences:

a. *he had arthralgias yesterday.* (No match)
b. *he had **arthralgia** yesterday.* (1 match)

III-3-3-2. String or Stem Search

This setting will not consider boundaries. Case will be considered if  is checked.
E.g. If we search a word "**arthralgia**" in two sentences below:

c. *he had **arthralgia**s yesterday.* (1 match)
d. *he had **arthralgia** yesterday.* (1 match)

III-3-3-3. Concept Search

With this setting, CHANL will search a term entered as well as all the synonyms of the term. Concept Search is not case sensitive. E.g. If we search a word "**arthralgia**" as a concept, CHANL will include all the synonyms and lexical variants such as "joint pain", "painful joint", "painful joints", "arthralgias", etc. in the search list. There are more options when performing Concept Search.

III-3-3-3-1. Child Concept

If ☐ Child is checked, all the sub-concepts of the concept entered will be included in the search list. E.g. If we search a word "**arthralgia**" as a concept including sub-concepts, in addition to all synonyms of arthralgia, all synonyms of sub-concepts such as "shoulder pain", "hip pain", "joints stiff" will also be included in the search list and highlighted.

III-3-3-3-2. Sibling Concept

If ☐ Sibling is checked, all the direct sibling concepts of a concept entered will be included in the search list and highlighted.

III-3-3-3-3. Negation Information

If ☐ Negation is checked, all the terms or phrases with negative meanings such as "no", "hasn't", and "deny" will be included in the search list and highlighted.

III-3-3-3-4. Certainty Information

If ☐ Certainty is checked, all terms or phrases about certainty such as "likely" and "possible" will be included in the search list and highlighted.

III-3-3-3-5. FamilyHX

If ☐ FamilyHX is checked, all terms or phrases about family members will be included and highlighted.

III-3-3-3-6. FromCHANLdb (On testing)

We can check ☐ FromCHANLdb to query data with high speed if a database optimized for CHANL is installed.

III-3-3-4. Regular Expression Search

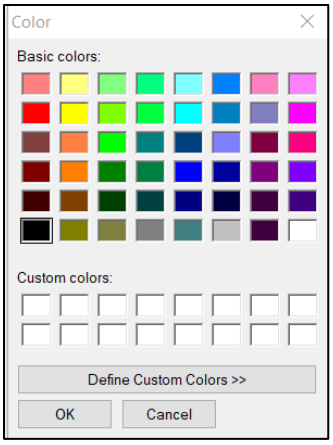
CHANL can also search terms using Regular Expression. Please check the menu here for generating a regular expression.

III-4. Semantic Type Search (In development)

III-5. Similarity Search (In development)

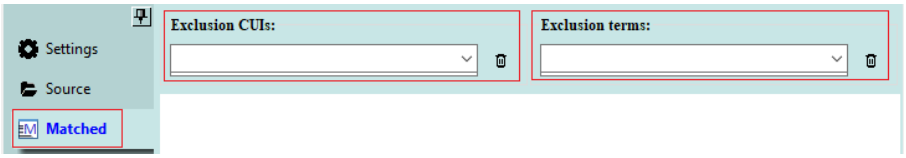
III-6. Color Customization

For each search option, there is a button ▮ for choosing a color for highlighting.



III-7. Term and Concept Exclusion

When performing a concept search or CUI search, we can exclude certain terms or concepts.



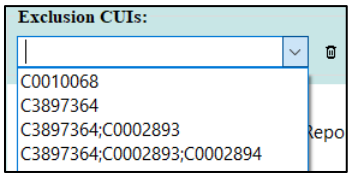
III-7-1. Enter Exclusion CUIs or Terms

Before running a search, please choose the tab “Matched”

in the middle of the CHANL interface. Enter CUIs separated by “;” or terms separated by “;”. All the terms of the CUIs will be excluded from the search list.

III-7-2. Choose CUIs or Terms

Click the relevant dropdown list to choose a record.



III-7-3. Delete CUIs or Terms

Choose a record from the dropdown list, then click follow to delete a record from the list.

Section IV. Searching Result Representation

IV-1. Main Panel

IV-1-1. Run a Search

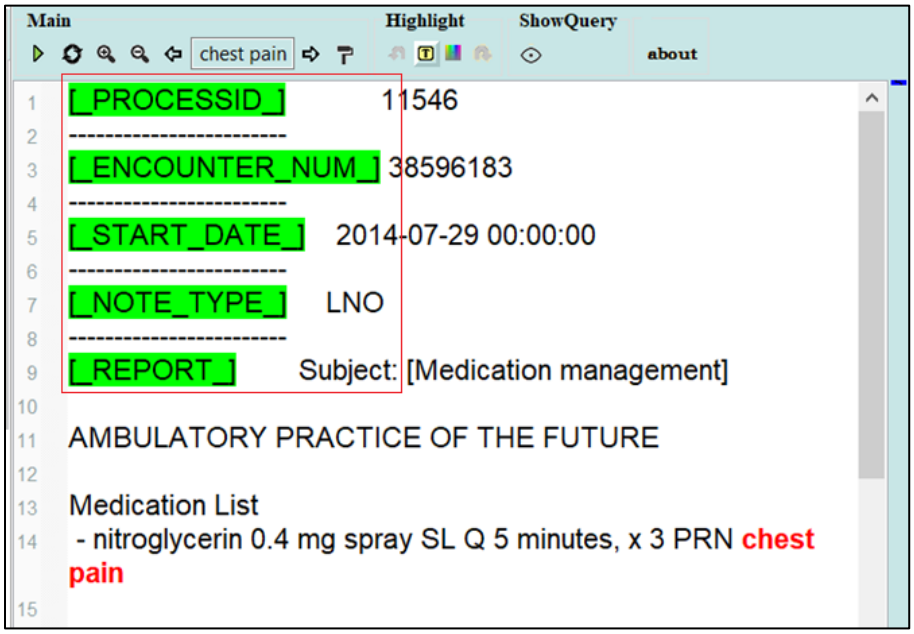
After choosing or creating a project and loading notes to the main panel, click on the main panel to run a search to highlight target terms.

IV-1-2. Reset Highlighted Terms

Click to remove the highlighting for terms.

IV-1-3. Column Name Highlighting

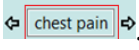
When loading data from a database query or correctly formatted XML file, the column names will be automatically highlighted with green.



IV-1-4. Zoom In and Zoom Out

Click to increase the font size of the loaded notes. Click to decrease the font size of the loaded notes.

IV-1-5. Matched Location

Click to iterate the matched terms in the notes in a backward direction and click to iterate the matched terms in the notes in a forward direction. The notes will scroll to where the current term is located. The current term will be shown like this: 

.

IV-1-6. Format Report

Click the button to format the report by replacing multiple spaces with a single space, multiple returns by a single return. (Note: reformatting is not applied to the leading spaces of each line).

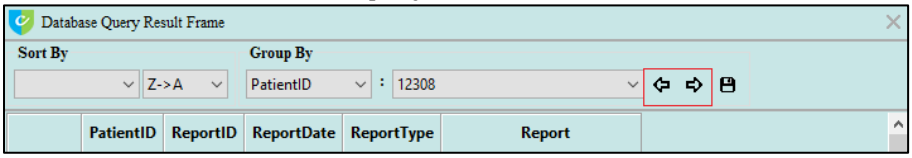
IV-1-7. Manually Highlighting

1. Click the toggle button to enable “Manual Highlighting”. Then click the button to choose a color for manual highlighting.

2. On the note, hold the left button on the mouse and select text to highlight.
3. Click the button to undo a highlight and click to redo a highlight when they are not greyed out.

IV-1-8. Recall Query Result Frame

1. Click the button to recall the data query frame if the data is loaded from a database query or the CHANL format of XML files.
2. Click the button or to index data from the query result frame in the main panel. This will be the same as the buttons on the query result frame:



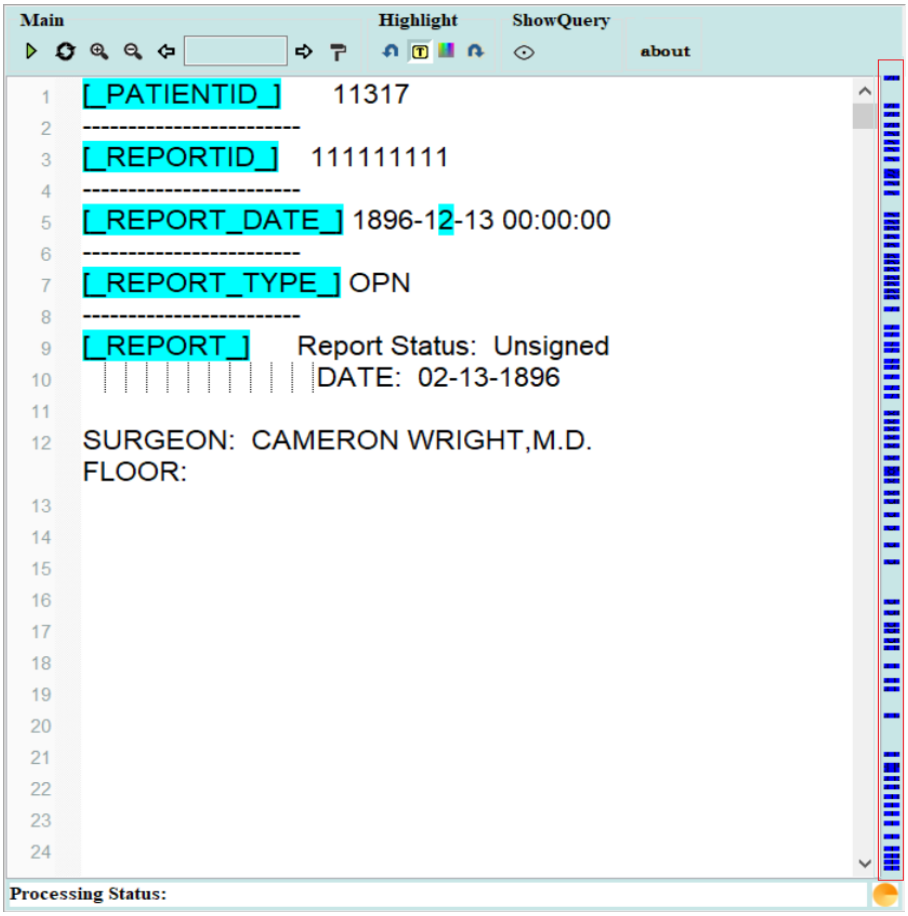
IV-1-9. About CHANL

Click the button to view the version and Author information of the software.

IV-1-10. Switch Report

IV-1-10-1. Locate Report Switching Bar

When data is loaded from a database query or XML file, there will be an extra information bar beside the scroll bar on the main panel. The information bar consists of many blue squares . Each blue square stands for a report.



IV-1-10-2. Display Report Header

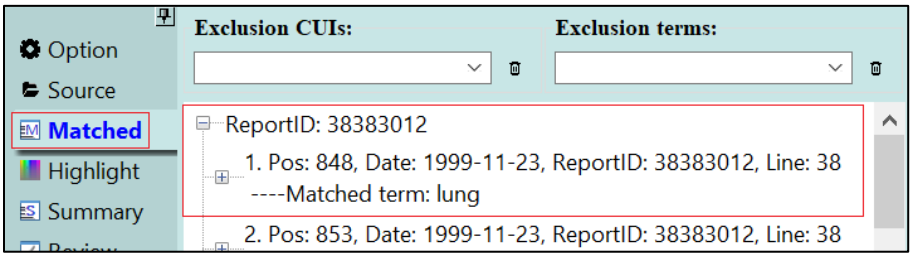
Hover the mouse on a blue square , and there will be a tooltip window displaying the header information of the report.

IV-1-10-3. Switch Report

When clicking any of the blue squares, the displayed content will switch to the report clicked.

IV-2. Detailed Matching Information

The detailed information for matching terms in notes will be displayed in the tab Matched. The order of the matched information listed in this tab will be the same as the information highlighted in the main panel. When clicking on a certain matched item, the content in the main panel will switch to the matched position.



IV-2-1. Information Listed

The information in each item listed in the tab Matched includes index, the matched sequence position, report date, reportID, line number and matched term. Matching information for different reports will be grouped for each report and listed under different nodes if the initial data is from a query or XML file.

IV-2-2. Collapse and Expand Matching Information

Click ⊞ to expand or ⊟ to collapse the report node.

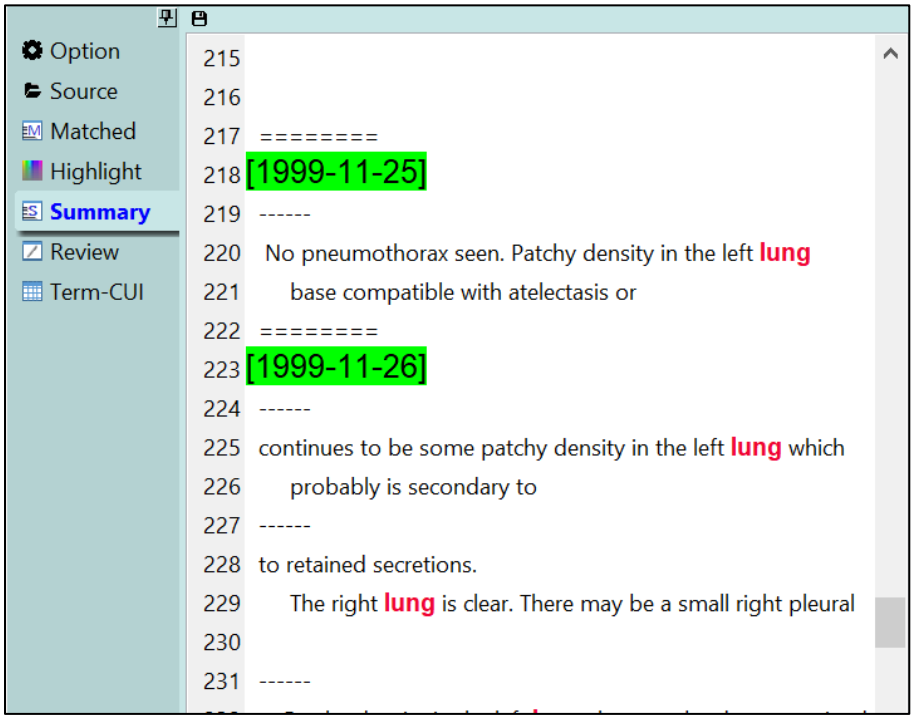
IV-2-3. Display Matching Sentence

Click ⊞ to expand an item to display the matched sentence. Click ⊟ to collapse the item.

IV-3. Search Result Summary

IV-3-1. Summary Data Format

In the tab Summary, the matched sentences will be listed as a group for a distinct date. The dates will be ascending. Different dates are separated by `========` and different sentences are separated by `------`.



IV-3-2. Color

Dates are highlighted with green. Matched terms will be highlighted with red. Negation terms will be highlighted with blue and family members will be highlighted with brown.

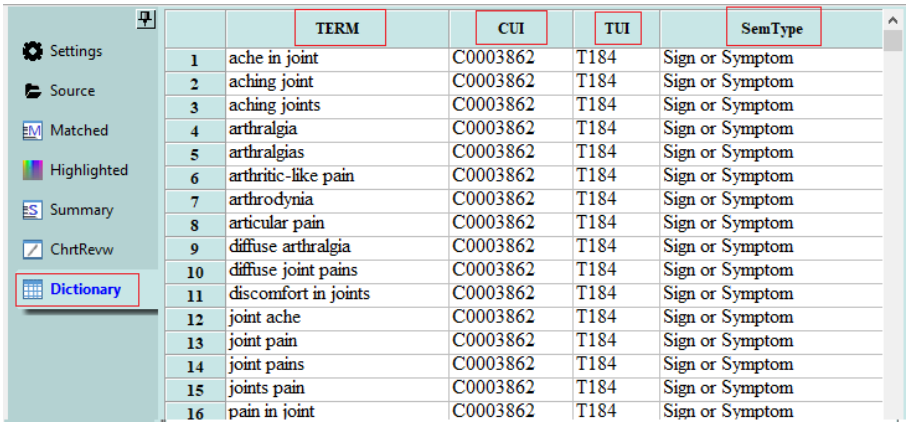
IV-3-3. Reflect to the Main Panel

Double click each sentence in the summary tab, the main panel will switch to the relevant position where the sentence is located.

IV-4. Concept Dictionary

IV-4-1. Data Format

When doing a concept search, all the potential terms as well as relevant concept unique identifier (CUI), semantic type unique identifier (TUI) and semantic type will be listed in the tab Dictionary.



	TERM	CUI	TUI	SemType
1	ache in joint	C0003862	T184	Sign or Symptom
2	aching joint	C0003862	T184	Sign or Symptom
3	aching joints	C0003862	T184	Sign or Symptom
4	arthralgia	C0003862	T184	Sign or Symptom
5	arthralgias	C0003862	T184	Sign or Symptom
6	arthritic-like pain	C0003862	T184	Sign or Symptom
7	arthrodynia	C0003862	T184	Sign or Symptom
8	articular pain	C0003862	T184	Sign or Symptom
9	diffuse arthralgia	C0003862	T184	Sign or Symptom
10	diffuse joint pains	C0003862	T184	Sign or Symptom
11	discomfort in joints	C0003862	T184	Sign or Symptom
12	joint ache	C0003862	T184	Sign or Symptom
13	joint pain	C0003862	T184	Sign or Symptom
14	joint pains	C0003862	T184	Sign or Symptom
15	joints pain	C0003862	T184	Sign or Symptom
16	pain in joint	C0003862	T184	Sign or Symptom

IV-4-2. Sort data

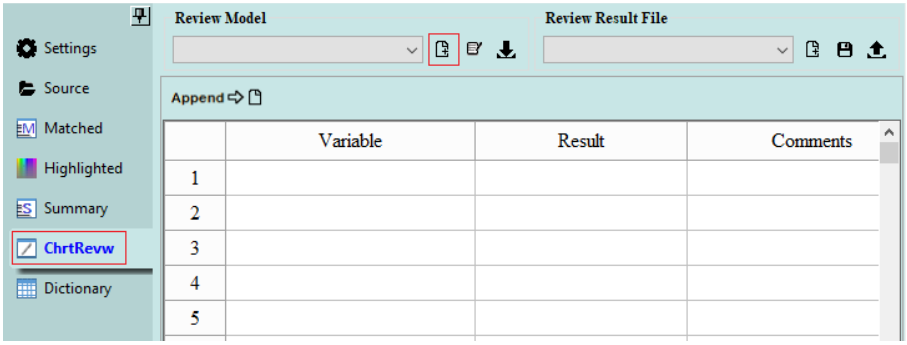
Click on each column name, and data will be re-sorted by relevant column ascendingly.

Section V. Perform Review

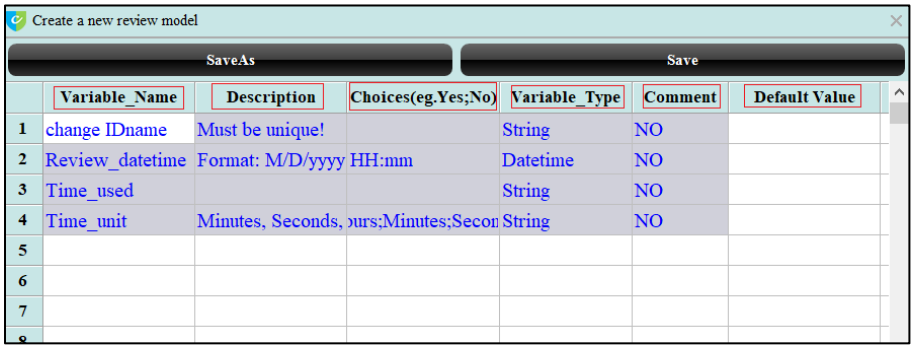
V-1. Review Model

V-1-1. Create a New Review Model

Choose the tab ChrtRevw in the middle of the CHANL interface, click to start creating a new review model.



A popup window will appear:



V-1-1-1. Variable Name

In the review model popup window, please enter variable names (e.g. “PatientID”, “Diagnosis”) for chart review.

	Variable_Name
1	PatientID
2	Review_datetime
3	Time_used
4	Time_unit
5	Diagnosis
6	

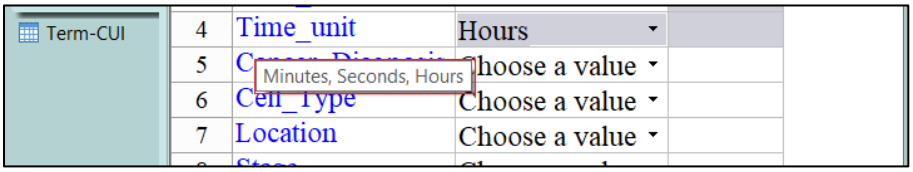
There are 3 default variable names in each model: Review_datetime, Time_used, Time_unit. The values for these variables will be filled by CHANL automatically if users don’t actively enter values.

V-1-1-2. Description

A value in the column “Description” will be displayed when the mouse is hovered over a variable to remind reviewers about the detail of a variable. E.g. Minutes, Seconds, Hours in the “Description”

	Variable_Name	Description	Choices(eg.Yes;No)	Variable_Type	Comment	Default
1	PatientID	Must be unique!		String	NO	
2	Review_datetime	Format: M/D/yyyy HH:mm		Datetime	NO	
3	Time_used			String	NO	
4	Time_unit	Minutes, Seconds, Hours	ours;Minutes;Secon	String	NO	

will be shown in a review model when the mouse is hovered over the variable “Time_unit”.

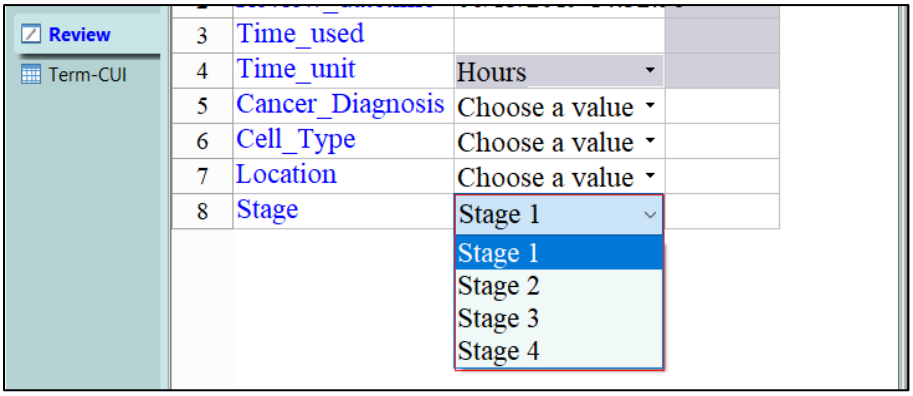


V-1-1-3. Choices

We can enter choices in the column “Choices” if there is a list of predefined values for a variable. For example, choices for cancer stage could be “Stage 1;Stage 2;Stage 3;Stage 4”. The delimiter must be “;”.

8	Stage	Could be TNM stage	Stage 1;Stage 2;Stage 3;Stage 4	String	YES

In the review model loaded, we can choose one of the stages for the variable “Stage”.



V-1-1-4. Variable Type

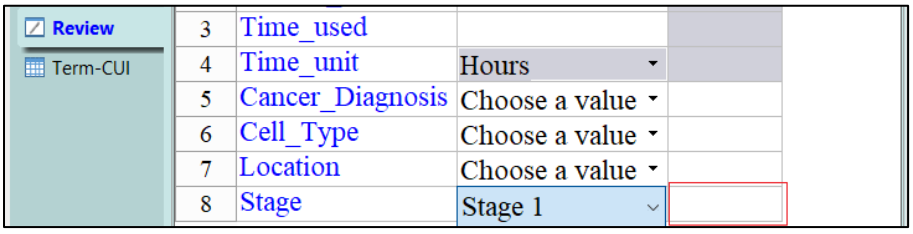
Please define a variable type for each variable. When a variable type is defined as “Integer”, a string such as “stage 1” will be invalid. Please define the type as “String” if a string value is a variable. There are 8 types on the list including 'Integer', 'Float', 'String', 'Date', 'Datetime', 'Boolean', 'currency' and 'String(unit)'. Please choose one from the dropdown list.

V-1-1-5. Comment

“Comment” is enabled by choosing “YES” in the dropdown list, a variable of “*_Cmt” will be automatically generated in the result file for the relevant variable and the “Comment” area will be editable. E.g. In the model creation window, “YES” is chosen for the “Comment” for the variable “Stage”.

	Variable_Name	Description	Choices(eg.Yes;No)	Variable_Type	Comment	Default
1	PatientNum	Must be unique!		String	NO	
2	Review_datetime	Format: M/D/yyyy	HH:mm	Datetime	NO	
3	Time_used			String	NO	
4	Time_unit	Minutes, Seconds,	urs;Minutes;Secon	String	NO	Hours
5	Cancer_Diagnosis	0 - Negtive;1 - De	0;1;9	Integer	YES	
6	Cell_Type		l;Adenocarcinoma;	String	YES	
7	Location		ht upper;Right mid	String	YES	
8	Stage	Could be TNM sta	1;Stage 2;Stage 3;S	String	YES	

The cell of the "Comment" for the variable "Stage" will be editable on the review model.



A variable "Stage_cmt" will be generated in the review result file automatically.

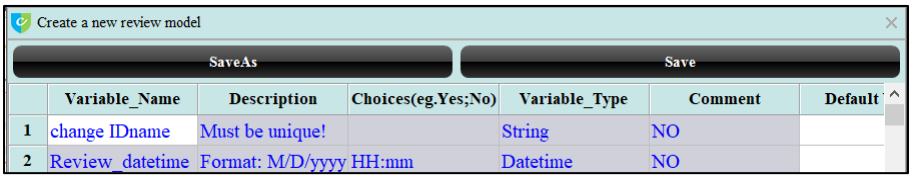
A	B	C	D	E	F	G	H	I
PatientNun	Review_da	Time_used	Time_unit	Cancer_Di	Cell_Type	Location	Stage	Stage_cmt
999	06/13/201	27.0	Seconds	null	null	null	null	null

V-1-1-6. Default Value

If a value appears often for a variable, we can set a default value for this variable. Eg. If "Stage 1" is very common for the variable "Stage", we can set "Stage 1" as the default value for the variable "Stage". However, we can still select another value from the dropdown list.

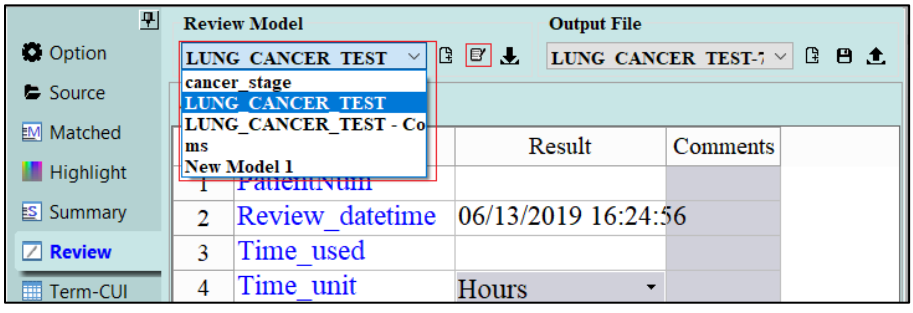
V-1-1-7. Save a Model

Click the button "Save" to save a model.



V-1-2. Edit a Review Model

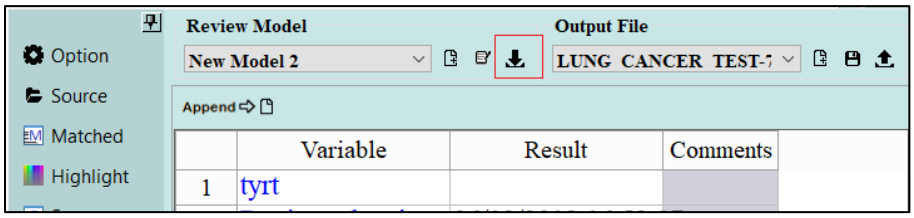
Choose a model from the dropdown list and click to enter the review model editor window.



After modification, click the button "Save" to save the change or click "Save As" to save as another model name. The review model needs to be re-loaded to make it effective by selecting the model from the dropdown list.

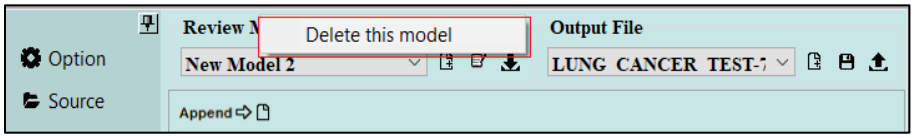
V-1-3. Import a Review Model

Click the button to browse the computer to import a CHANL review model.



V-1-4. Delete a Review Model

Choose a review model from the dropdown list and right click on the model name. Click the popup menu "Delete this model" to delete an existing model from the list.



V-2. Record Review Results

V-2-1. Record Results

The first row is for PatientID which must be filled. All other fields will be filled as 'null' automatically if there are no values entered. Please enter values with the correct value type.

V-2-2. Enter or Select Comment

If the field of comment is not disabled or greyed out, we can enter or select comments using the CHANL manual highlighting tool.

Before using the CHANL highlighting tool, we need to select the field of comment, then click the toggle button on the main panel to enable "Manual Highlighting". The text we select in the notes will be copied to the comment field.

V-2-3. Review Time

V-2-3-1. CHANL Timer and Clock

Every time we append results to a file, there will be a CHANL clock generated

. It displays the real computer time and the time when the result appending occurred. Double clicking the CHANL icon can close the clock. It will appear again when clicking the button "Append". The clock can be moved by dragging.

V-2-3-2. Time Records

Review_timeadate will always be filled automatically. The field for "Time_used" will be calculated and filled automatically along with the time unit if no values are entered. The automatic time will start from the time a new review model was loaded or last time the result was successfully appended.

V-2-4. Append Results

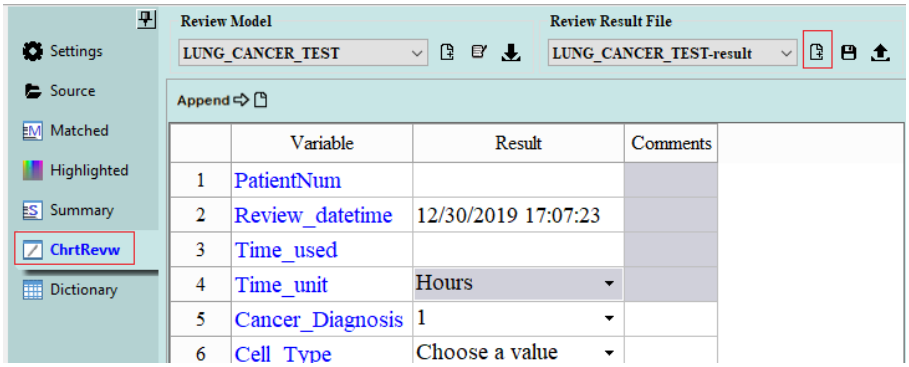
Click the button

to append results to a review result file. The result will be displayed in the "Review Result Display" sheet which will be described later.

V-3. Review Result Files

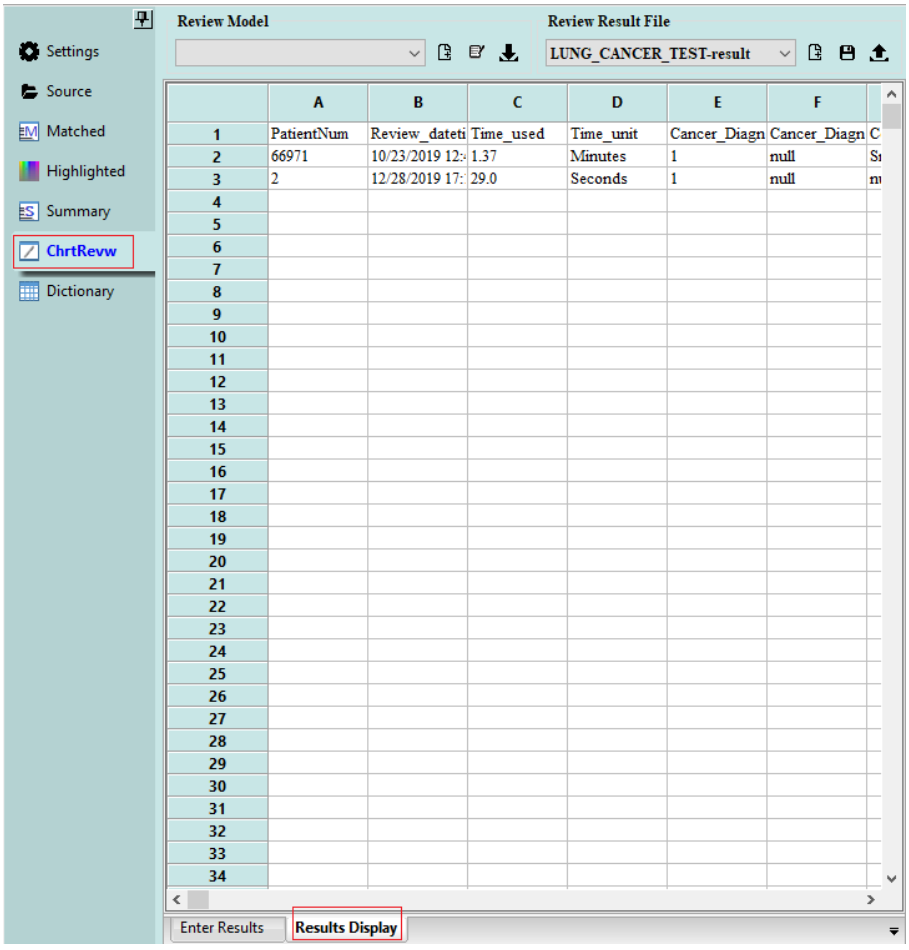
V-3-1. Create a Review Result File

Choose a Review model first. Click the tab “Review” then click the button in the area “Review Result File” to create a new review result file for the model chosen. If the review model is “Lung_Cancer_test” and a name entered for result file is “11”, the final result file name is “lung_cancer_test_11”.



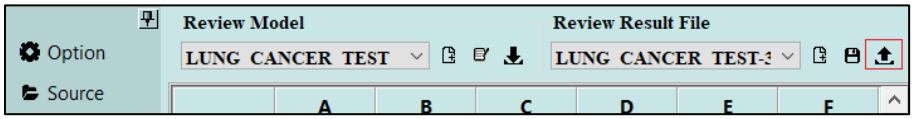
V-3-2. Save Change to a Review Result File

When successfully appending data to a result file, data will be automatically saved. But we can also go to the “Review Result Display” tab to manually change the result and click the button to save the modification.



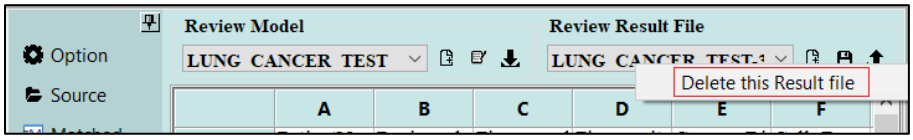
V-3-3. Export a Review Result File

We can export CHANL review result to a CSV file with delimiter "|" by clicking the button .



V-3-4. Delete a Review Result File

Load a result file, then right click the name of the file loaded to delete it.



Section VI. Appendix

VI-1. Create XML database for CHANL (Python)

```
# -*- coding: utf-8 -*-

import xml.etree.ElementTree as ET
import os,string,pyodbc,sys,traceback
from dateutil import parser
import re

def filter_content(data):
    printable = set(string.printable)
    return filter(lambda x: x in printable, data)
def create_new_xml(rootname,path):
    root=ET.Element(rootname)
    tree = ET.ElementTree(root)
    tree.write(path)
def formalizenumber(n):
    return '0'*(8-len(str(n)))+str(n)
def standardizecolumnnames(line):
    colnames = line.strip().split('|')
    colnames[-1]='Note'
    colnames.insert(1,'Index')
    for i in xrange(len(colnames)):
        if 'date' in colnames[i].lower():
            datepos=i
            colnames[i]='Note_date'
            tmp=colnames[2]
            colnames[2]=colnames[i]
            colnames[i]=tmp
            break
    return datepos,colnames
def connect_SQL(server,database,Autivationtype='SQL Server
Authentication',user='',pwd=''):
        """
        connect to MSSQL database, return a connection
        """

        if Autivationtype=='sql':
            try:
                cnxn = pyodbc.connect('DRIVER={SQL Server Native
Client 11.0};SERVER=%s ;DATABASE=%s;UID=%s;PWD=%s'
%(server,database,user,pwd))
                print 'database connected'
                return cnxn
            except:
                print 'database not connected'
                return None
        elif Autivationtype=='windows':
            try:
```

```
                cnxn = pyodbc.connect('DRIVER={SQL Server Native
Client 11.0};SERVER=%s ;DATABASE=%s;Trusted_Connection=yes;'
%(server,database))
                print 'database connected'
                return cnxn
            except:
                print 'database not connected'
                return None

def formalizedate(dt):
    d = parser.parse(dt)
    return (d.strftime("%Y-%m-%d"))
def CDATA(text=None):

    element = ET.Element('![CDATA[')
    element.text = text
    return element
ET._original_serialize_xml = ET._serialize_xml
def _serialize_xml(write, elem, encoding, qnames, namespaces):

    if elem.tag == '![CDATA[':

        write("\n<%s%s]]>\n" % (elem.tag, elem.text))

        return

    return ET._original_serialize_xml(

        write, elem, encoding, qnames, namespaces)

ET._serialize_xml = ET._serialize['xml'] = _serialize_xml

def getleadingspacecount(line):
    res = re.match(r'^ +',line)
    if res:
        return len(res.group())
    else:return 0
def writexml_from_databbase(columnnames,script,conn,xmlpath):
    cursor = conn.cursor()
    cursor = cursor.execute(script)
    records = cursor.fetchall()
    conn.close()
    if not os.path.exists(xmlpath):
        create_new_xml('CHANL_XML',xmlpath)

    tree = ET.parse(xmlpath)
    root = tree.getroot()
    try:
        for line in records:
            row = ET.SubElement(root,'ROW')
            for i in xrange(len(columnnames)):
                column = ET.SubElement(row,'COLUMN',NAME =
columnnames[i])
                content = '\n'.join([getleadingspacecount(ele)*'
'+ele.strip() for ele in str(line[i]).split('\n')])
```

```
                line[i]= filter_content(content).strip()
                cdata = CDATA(line[i])
                column.append(cdata)

    except:
        type1, value, tb = sys.exc_info()
        error_information=''.join(traceback.format_exception(type1,
value, tb))
        print error_information
        return
    tree.write(xmlpath, "utf-8")
```

VI-2. SQL query example

VI-2-1. Create a View in MSSQL

```
CREATE VIEW [dbo].[CHANL_CHART_REVIEW]
AS
SELECT        ProcessID AS PatientID,
              Encounter_num AS ReportID,
              Start_date AS Report_Date,
              note_type as Report_Type,
              Report
FROM            dbo.lungca_full_sample100
GO

## Note: Please follow the order of the column names
```

VI-2-2. Create a View in MySQL

```
CREATE VIEW CHANL_CHART_REVIEW
AS
SELECT
    ProcessID AS PatientID,
    Encounter_num AS ReportID,
    Start_date AS Report_Date,
    note_type as Report_Type,
    Report
FROM    Database_Name.Table_Name
GO

## Note: Please follow the order of the column names
```

VI-2-3. Query Data from a Table

```
##Please select 5 columns with order below to query data for CHANL.

SELECT
    ProcessID, ## PatientID
    Encounter_num, ## ReportID,
    Start_date, ## Report_Date,
    note_type, ## Report_Type,
    Report
FROM    Database_Name.Table_Name
```

VI-3. Example Review Model

VI-3-1. Model Creating

Variable_Name	Description	Choices	Variable Type	Comment	Default
PatientNum	Must be unique!		String	NO	
Review_datetime	Format: M/D/yyyy HH:mm		Datetime	NO	
Time_used			String	NO	
Time_unit	Minutes, Seconds, Hours	Hours;Minutes;Seconds	String	NO	Hours
Cancer_Diagnosis	0 - Negative; 1 - Definite; 9 - Unclear	0;1;9	Integer	YES	1
Cell_Type		NA;Small Cell; Non-small cell;Adenocarcinoma;squamous cell;large cell;Other	String	YES	
Location		Left upper;Left lower;Right upper;Right middle;Right lower;NA;other	String	YES	
Stage		1;2;3;4	String	NO	2

VI-3-2. Review Model Generated

	Variable	Result	Comments
1	PatientNum		
2	Review_datetime	06/13/2019 23:53:47	
3	Time_used		
4	Time_unit	Hours	
5	Cancer_Diagnosis	Choose a value	
6	Cell_Type	Choose a value	
7	Location	Choose a value	
8	Stage	Choose a value	

VI-4. Brief Introduction of Regular Expression

From (https://www.oreilly.com/openbook/cgi/appb_01.html)

/abc/

Matches *abc* anywhere within the string

/^abc/

Matches *abc* at the beginning of the string

/abc$/

Matches *abc* at the end of the string

/a|b/

Matches either *a* or *b* Can also be used with words (i.e., /*perl*|*tcl*/)

/ab{2,4}c/

Matches an *a* followed by *2-4 b*'s, followed by *c*. If the second number is omitted, such as /*ab* {2,}*c*/, the expression will match two or more *b*'s.

/ab*c/

Matches an *a* followed by zero or more *b*'s, followed by *c*. Expressions are greedy--it will match as many as possible. Same as /*ab*{0,}*c*/.

/ab+c/

Matches an *a* followed by one or more *b*'s followed by *c*. Same as /*ab*{1,}*c*/.

/ab?c/

Matches an *a* followed by an optional *b* followed by c Same as /*ab*{0,1}*c*/. This has a different meaning in Perl 5. In Perl 5, the expression: /ab*?c/matches an *a* followed by as few *b*'s as possible (non-greedy).

/./

Matches any single character except a newline (\n) /*p..l* / matches a p followed by any two characters, followed by l, so it will match such strings as perl, pall, pdgl, p3gl, etc.

/[abc]/

A character class--matches any one of the three characters listed. A pattern of /[abc]+/ matches strings such as abcab, acbc, abbac, aaa, abcacbac, ccc, etc.

/\d/

Matches a digit. Same as /[0-9]/Multipliers can be used (/\d+/ matches one or more digits)

/\w/

Matches a character classified as a word. Same as /[a-zA-Z0-9_]/

/\s/

Matches a character classified as whitespace. Same as /[\r\t\n\f]/

/\b/

Matches a word boundary or a backspace/test\b/ matches test, but not testing. However, \b matches a backspace character inside a class (i.e., [\b])

/[^abc]/

Matches a character that is not in the class/[^abc]+/ will match such strings as hello, test, perl, etc.

/\D/

Matches a character that is not a digit. Same as /[^0-9]/

/\W/

Matches a character that is not a word. Same as /[^a-zA-Z0-9_]/

/\S/

Matches a character that is not whitespace. Same as /[^ \r\t\n\f]/

/\B/

Requires that there is no word boundary/hello\B/ matches hello, but not hello there

/*/

Matches the * character. Use the \ character to escape characters that have significance in a regular expression.

/(abc)/

Matches *abc* anywhere within the string, but the parentheses act as memory, storing *abc* in the variable $1.

Example 1:

/name=(.*)/ will store zero or more characters after name= in variable $1.

Example 2:

/name=(.*)&user=\1/ will store zero or more characters after name= in $1. Then, Perl will replace \1 with the value in $1, and check to see if the pattern matches.

Example 3:

/name=([^&]*)/ will store zero or more characters after name= but before the & character in variable $1.

Example 4:

/name=([^&]+)&age=(.*)$/ will store one or more characters after name= but before & in $1. It then matches the & character. All characters after age= but before the end of the line are stored in $2.

/abc/i

Ignores case. Matches either abc, Abc, ABC, aBc, aBC, etc.

CHANL

help you to save time for the art of life