

National Cancer Institute (NCI)
Epidemiology and Genomics Research Program (EGRP)



Cohort Metadata Repository (CMR)

CMR User Guide

Version 1



November 2016

Table of Contents

INTRODUCTION	1
USE OF METADATA	1
USE OF HARMONIZATION SCRIPT	1
CMR LANDING OR PUBLIC HOME PAGE	1
NEW ACCOUNT	2
ACCESSING THE SITE	3
REQUEST NEW PASSWORD	3
HELP OR REQUESTING ACCESS	4
USER’S HOME PAGE FEATURES	4
USER MENU	5
Change Password	5
Log Out	6
SEARCH	6
ABOUT CMR	6
COHORTS	6
COHORTS TABLE AND NAVIGATION	7
SELECTED COHORT	8
SEARCH COHORTS	9
COHORT METADATA	9
SELECTED DATASET	10
VARIABLES SEARCH	11
VARIABLES SEARCH PAGE	12
Variable Filter Categories	13
Variable Full Text Search	14
VARIABLES	16
HARMONIZATION DATASETS	17
HARMONIZED DATASET	17
DATASET HARMONIZATION SUMMARY	18
VARIABLE HARMONIZATION DETAIL	20
METADATA/CODE REQUEST	23
DISCUSSION FORUM	23
ADD TOPIC	24
REPLY	24
VIEW/EDIT OWN COMMENTS	25
CONTACT US	26

List of Figures

Figure 1: <i>CMR public Home Page</i>	2
Figure 2: <i>Page-bottom links to About CMR and selected websites</i>	2
Figure 3: <i>Error message examples</i>	3
Figure 4: <i>User’s CMR home page</i>	4

Figure 5: <i>The selected menu option is color-coded in light brown</i>	5
Figure 6: <i>Use these menu items to change the password</i>	5
Figure 7: <i>Results of a search for a key word</i>	6
Figure 8: <i>The Cohorts page – list of cohorts in the CMR database</i>	7
Figure 9: <i>The detailed page of a selected cohort</i>	8
Figure 10: <i>The Cohorts Search page</i>	9
Figure 11: <i>The Cohort Datasets table at the bottom of a detailed page (Figure 9)</i>	10
Figure 12: <i>Selected dataset</i>	10
Figure 13: <i>Top of the database level Variables Search page – no filter selected</i>	11
Figure 14: <i>Variables Search page of a selected cohort (cohort level)</i>	12
Figure 15: <i>Variables Search page of a selected dataset (dataset level)</i>	13
Figure 16: <i>Handling the filter categories</i>	14
Figure 17: <i>Results of a search against the entire database – example</i>	15
Figure 18: <i>Results of a search for variables in a selected cohort – example</i>	16
Figure 19: <i>Detail page of a cohort dataset variable (example: MARSTAT)</i>	16
Figure 20: <i>The Harmonization Datasets page and table</i>	17
Figure 21: <i>Details page of a selected harmonized dataset</i>	18
Figure 22: <i>Harmonization chart</i>	19
Figure 23: <i>Pop-up variable summary</i>	19
Figure 24: <i>Status and script of DMCA_BL_HIPCIRC in CSDLH</i>	20
Figure 25: <i>Detail harmonization data for a selected variable in a cohort</i>	20
Figure 26: <i>Top section of the detail page of a harmonized variable</i>	21
Figure 27: <i>The harmonization detail table of a harmonized variable (truncated)</i>	22
Figure 28: <i>Metadata/Code Request Form</i>	23
Figure 29: <i>Discussion Forum page</i>	23
Figure 30: <i>Create Discussion topic on this form</i>	24
Figure 31: <i>Add, view, or reply to topics from this page</i>	24
Figure 32: <i>History page of a selected topic</i>	25
Figure 33: <i>Top section of the View/Edit form for your own topics</i>	25

INTRODUCTION

This web site is designed for NCI Cohort Consortium investigators participating in harmonization projects utilizing the Cohort Metadata Repository (CMR). The CMR is a tool that documents data harmonization across cohorts. Variables from each cohort can be searched and compared to determine if harmonization is possible. Once harmonization has occurred, the harmonized variables and the specifications used to create them are also documented in the CMR.

The CMR contains only metadata (variable names, formats, codes, descriptions) and no individual level data.

The site displays the following information: (a) descriptions of cohort data collection to provide metadata context and (b) detailed metadata for cohort variables and harmonized variables along with the algorithms used to create the harmonized variables. For more detailed descriptive information about cohorts, go to the [Cancer Epidemiology Descriptive Cohort Database \(CEDCD\)](#) website.

Use of Metadata

The cohort metadata are provided for the purposes of data harmonization for projects in which cohorts have agreed to participate.

Access to the site is read-only to ensure that the information can be searched and viewed. Only the site administrator has access to add or edit cohort or harmonization information. Investigators wishing to edit or add information to the site should contact the site administrator (CMRHelpdesk@westat.com). If an investigator wishes to request any metadata or harmonization code then they can do so with the **Metadata/Code Request** form (see the section on Metadata/Code Request).

Use of Harmonization Script

The harmonization script is provided in “pseudo-code” which states which cohort variables need to be used and how they need to be coded or transformed to arrive at the harmonized variables. The script is not in any particular programming language but is closest to the SAS language. It will need to be formatted slightly to run in SAS (e.g., remove [blank] and replace with the proper syntax for SAS missing codes) and it will need more extensive formatting to run in other programming languages.

The script was derived only on the basis of the metadata provided by the cohorts for the variables requested. It has not been tested on individual level data. As such it does not account for potential logical inconsistencies or other data issues in the individual level data. When individual level data are pooled for harmonization, the individual level data and the harmonization code will undergo thorough QC review and testing. If you identify any issues with the coding of harmonized variables for your cohort and would like to submit corrections, please contact the site administrators (CMRHelpdesk@westat.com).

CMR Landing or Public Home Page

Type, or copy and paste, <http://cmr.nci.nih.gov> in your browser's **Address** field – the CMR landing or public Home Page will appear (Figure 1).

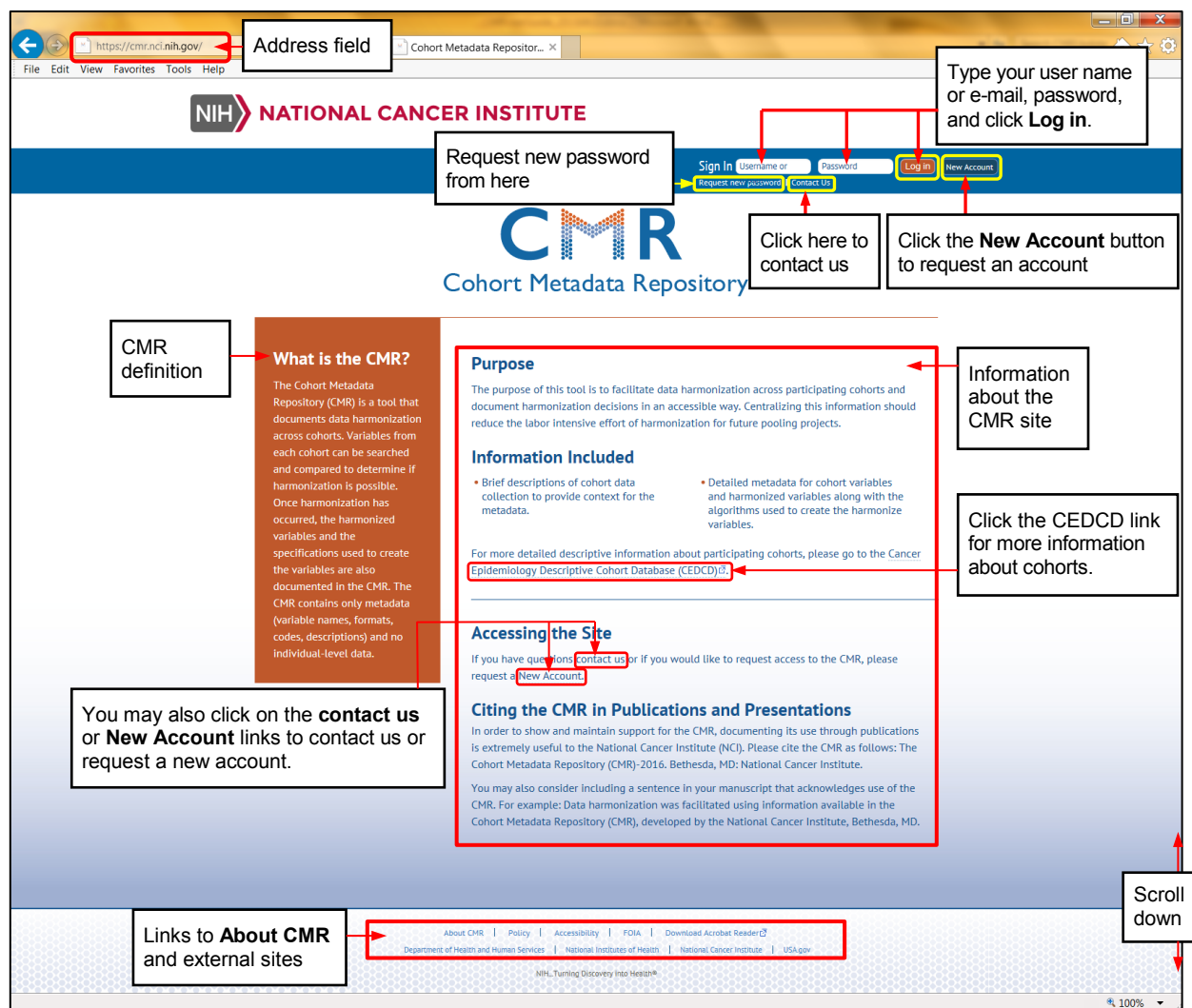


Figure 1: *CMR public Home Page*

This page displays a definition of the CMR; a description of its purpose; of the information included; and of the CMR access policy. At the page bottom there are links to **About CMR** and selected external websites: **Policy** (of CMR), **Accessibility** (website accessibility use and conditions), **FOIA** (Freedom of Information Act home page), **Download Acrobat Reader**, **Department of Health and Human Services**, **National Institutes of Health**, **National Cancer Institute**, and **USA.gov**. These links are visible on all CMR pages.



Figure 2: *Page-bottom links to About CMR and selected websites*

New Account

Access is intended for epidemiologic researchers. To request an account, click on the **New Account** button or on the text link in the **Accessing the Site** section of the home page (see Figure 1). The **CMR Registration Form** will appear.

1. Fill out all fields,
2. Select an **Affiliation** value,
3. Click on the **CMR User Agreement** link to read the agreement and read it,
4. Check the **I have read the CMR User Agreement** box, so you confirm you agree with the CMR conditions,
5. Solve the math question, and
6. Click on the **Register** button.

Your request will be submitted for approval. Many requests are approved. If yours is approved, you will receive email notification by the site administrator. If you do not receive notification within two business days, or have any questions, please contact the site administrator at CMRHelpdesk@westat.com.

Accessing the Site

Type your user name or e-mail, password, and click **Log in**. The home page will appear (Figure 4). If errors are detected, messages will highlight them. Provide correct information and click **Log in** again. On successful login, the error messages disappear. You may also click the gray **X** at the top right corner of the error message to delete an error message before a new login attempt. If you forgot the password and the “**Have you forgotten your password?**” link appears, click on it to have one-time-login instructions sent to your email (the system must have your email address on record).

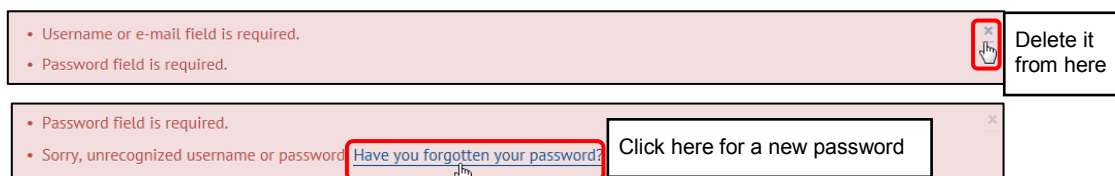


Figure 3: Error message examples

Request New Password

If you forget the password or need a new one, click on the **Request new password** (Figure 1) or the “**Have you forgotten your password?**” link (Figure 3) to have one-time-login instructions sent to your email. The **User account** page will appear. Before changing the password you may still sign in from the **User account** page using the old password if you remember it. Follow these steps to reset the password.

1. Enter your user name or email address and click on the **E-mail new password** button. The CMR will send a message to your email address.
2. Go to your email. The message contains a link that can be used only **once** to log in. It will open a page where you can reset your password. The link expires after one day and nothing happens if it is not used. Click on the link to validate your email and to get the **Reset Password** page.
3. Click on the **Log in** button. This will take you to the user account page on which you may change the password.
4. Type your new password in the **Password** box.
5. Retype the password in the **Confirm password** box.

- Click **Save** – at the bottom of the page. The CMR will confirm that “**The changes have been saved**” and you will be logged in. Use the new password the next time you sign in.

Help or Requesting Access

If your login attempts fail, you forget user name or password, have questions, or do not have an account, follow the available onscreen instructions or contact the **CMR Help Desk** at CMRHelpdesk@westat.com.

USER’S HOME PAGE FEATURES

This section only *enumerates* the user’s home page features. For details see the sections devoted to each feature in turn. This page describes the CMR and directs you to the lists of **Cohorts**, **Variables**, etc. It also provides a menu bar, a **Search** engine, a **User menu**, and other menu options (Figure 4).

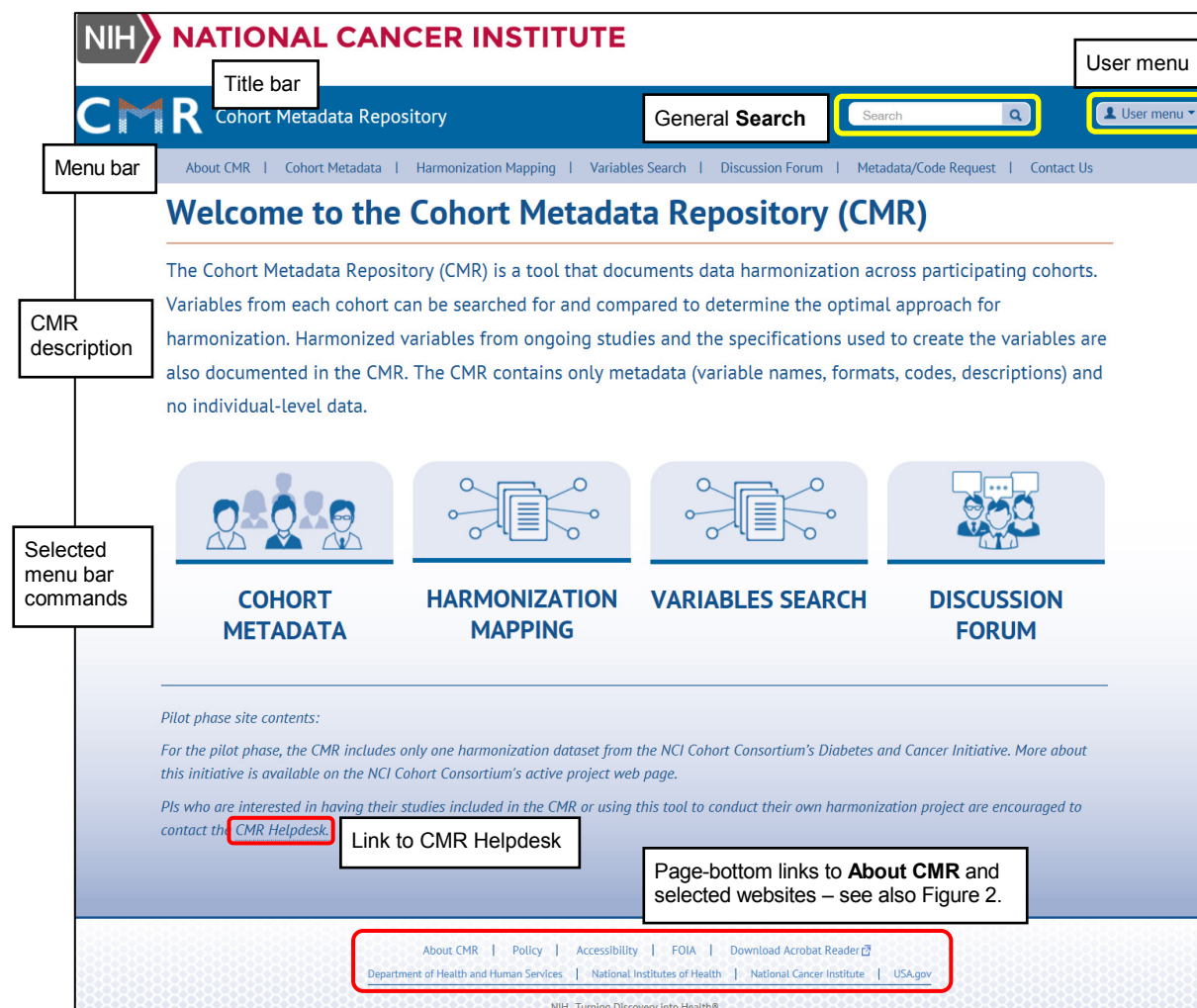


Figure 4: User’s CMR home page

In strict *enumeration* from left to right and top to bottom, the user’s home page displays the following features:

- The **title bar**, **CMR Cohort Metadata Repository**, is visible on all CMR pages. Click on it to return to the home page.

- Use the **Search** engine for general searches that you can later narrow down to a smaller list, or for a specific dataset or variable name that will produce a small list of results. In the latter case the spelling must be **exactly the same** as that used in the database. The more specific **Variables Search** engine provides better context. See and compare the sections on Search and Variables Search.
- The **User menu** button displays two options: **My account** and **Log out**. See the section on User Menu for details.
- The **menu bar** appears on all CMR pages. It provides access to **About CMR**, **Cohort Metadata**, **Harmonization Mapping**, **Variables Search**, **Discussion Forum**, **Metadata/Code Request**, and **Contact Us** pages. **Cohort Metadata**, **Harmonization Mapping**, **Variables Search**, and **Discussion Forum** are also accessible from their middle-page image (Figure 4). Each feature is described later in the guide. Click on any menu option tab or image to display the destination page. The selected menu option tab is color-coded in light brown (Figure 5). The path to the current page is displayed under the menu bar. Click on any path step to display that page.

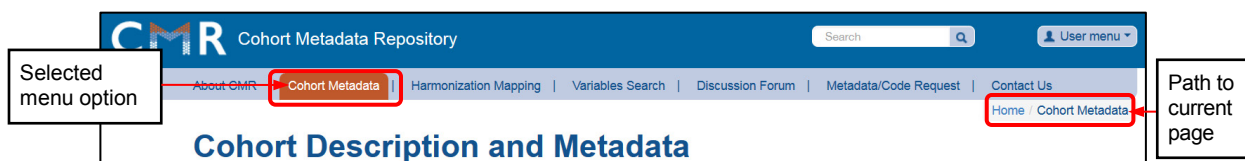


Figure 5: *The selected menu option is color-coded in light brown*

- The **images** in the middle of the page link to areas where you will be doing most of your work. Click on an image or on the corresponding menu option tab to display the selected pages or lists. Each image and menu option is described later in the guide.

User Menu

The **User menu** button displays two options: **My account** and **Log out** (Figure 6). Use **My account** to change your password and **Log out** to properly log out of the CMR.

Change Password

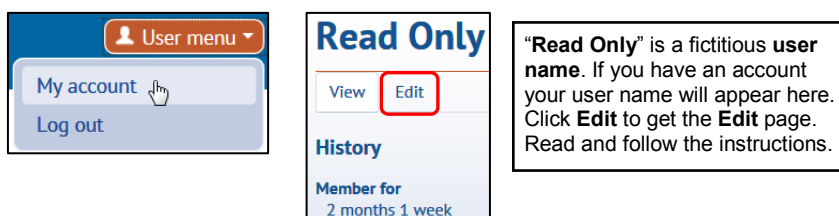


Figure 6: *Use these menu items to change the password*

You may change your password. The steps for changing the password are similar to those for resetting it (see the section on Request New Password):

1. Sign in to the CMR.
2. Select **User menu > My account > Edit** (Figure 6). On the **Edit** page you can change your password or email address. Read and follow the instructions on the **Edit** page.
3. Click **Save** at the bottom of the page to save your changes.

Log Out

Use the **User menu > Log out** command (under **My account** – see Figure 6) to sign out of the CMR and close it properly. The system may keep you logged in if you do not log out properly.

Search

The general **Search** at the top of the home page is visible on all CMR pages. It is a full site search and a useful tool if you know exactly what you are searching for, but the more specific **Variables Search** engines provide better context. See also the section on Variables Search.

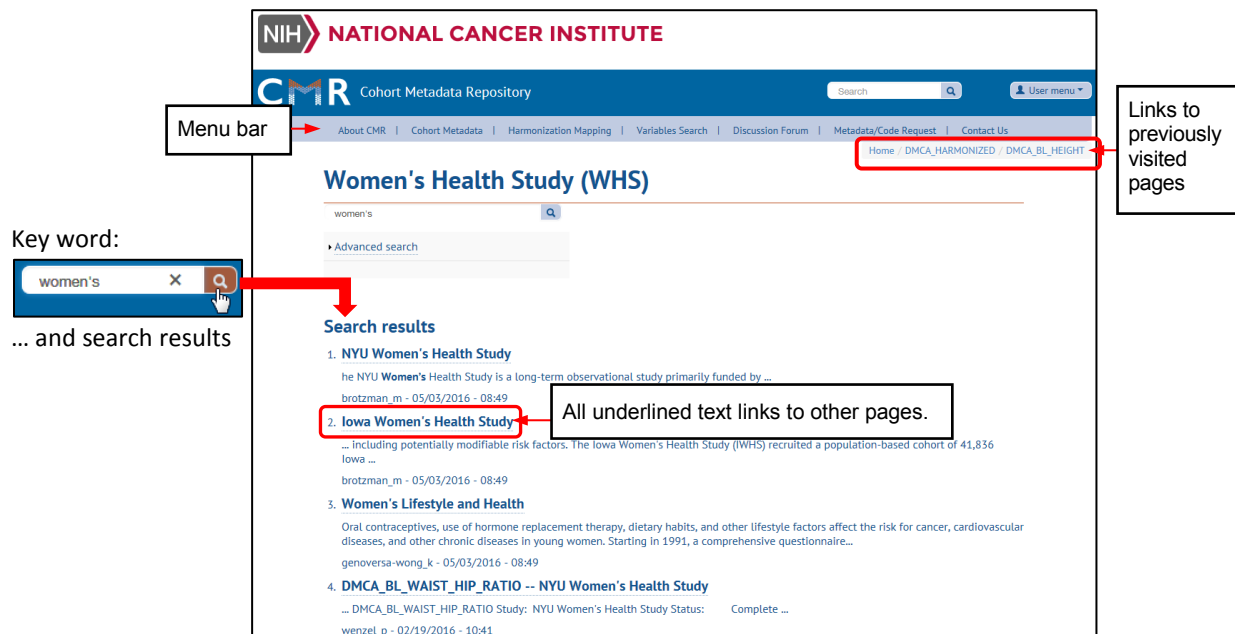


Figure 7: Results of a search for a key word

Use this **Search** to get a list of results and select what you are interested in. You may search for a **key word** (or phrase) or for a specific dataset or variable **name**. The results of key word searches may range from nothing to long lists that you can further filter down to what you need. For example, a search for “women’s” will produce a list of items that contain the key word “women’s” (Figure 7). The numbered items in the list of results link to further pages. If your article is visible on the first page, click on its title to go to its pages. If it is not visible, scroll down or navigate to it and click on its title when you see it.

About CMR

The **About CMR** page displays summaries about the CMR, its purpose, and how the included information is collected. The public can access this page from the **About CMR** link at the bottom of the landing page (Figure 1). CMR users can also access it from the **About CMR** menu option.

COHORTS

This page lists the cohorts that provided metadata to the CMR. It can be accessed from the **Cohort Metadata** image on the home page or the **Cohort Metadata** menu item from any page.

Selected menu option

Links to previously visited pages

Search cohorts from this button. See also Figure 10.

Search Cohorts

Set here the number of records per page. The values are 20, 50, or View all. The default value is 20.

All underlined text links to other pages. Click on a **Name or **Cohort Metadata** entry to view a more detailed description of the selected cohort.**

Navigation buttons

Short Name	Name	Cohort Metadata
AgHealth	Agricultural Health Study	AgHealth Metadata
ATBC	Alpha-Tocopherol Beta-Carotene Cancer Prevention Study	ATBC Metadata
BGS	Breakthrough Generations Study	BGS Metadata
BCGP	British Columbia Generations Project	BCGP Metadata
CSDLH	Canadian Study of Diet, Lifestyle and Health	CSDLH Metadata
COSM	Cohort of Swedish Men	COSM Metadata
EPIC	European Prospective Investigation into Cancer and Nutrition	EPIC Metadata
GCS	Golestan Cohort Study	GCS Metadata
IWHS	Iowa Women's Health Study	IWHS Metadata
JPHC1	Japan Public Health Center-Based Prospective Study Cohort 1	JPHC1 Metadata
JPHC2	Japan Public Health Center-Based Prospective Study Cohort 2	JPHC2 Metadata
NLCS	Netherlands Cohort Study	NLCS Metadata
NSHDS	Northern Sweden Health and Disease Study	NSHDS Metadata
NYUWHS	NYU Women's Health Study	NYUWHS Metadata
PHSI	Physicians' Health Study Cohort I	PHSI Metadata
PHSII	Physicians' Health Study Cohort II	PHSII Metadata
PLCO	Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial	PLCO Metadata
AHS-2	Seventh-day Adventist Cohort Study-2	AHS-2 Metadata
SIS	Sister Study	SIS Metadata
SMC	Swedish Mammography Cohort	SMC Metadata

1 2 next >

Figure 8: *The Cohorts page – list of cohorts in the CMR database*

Cohorts Table and Navigation

The **Cohorts** table (Figure 8) displays as many records as you select in the **Items per page** field. The default is **twenty** (20) records per page. **Fifty** (50) or **View all** are the other options. Navigation buttons are displayed at the bottom of the table. The number of the current page is grayed out. Click on any blue-text button to go to that page (**2**, **next**, **previous**, etc.). The table displays the following columns: **Short Name**, **Name**, and **Cohort Metadata** descriptions. The **Name** and **Cohort Metadata** entries link to the more detailed cohort description page. In the CMR all underlined text links to other pages. Click on a **Name** or **Cohort Metadata** entry to view the detailed page of the selected cohort.

Click on the **Search Cohorts** button to display a more detailed cohort search engine that allows filtering of the results from a left-side panel. See the section on Search Cohorts for details.

Selected Cohort

On the **Cohorts** page click on a **Name** or **Cohort Metadata** entry to view the detailed page of that cohort.

The screenshot shows the detailed page for the Iowa Women's Health Study (IWHS). The page is organized into several sections, each with a corresponding label on the left side of the image:

- Cohort summary:** This section contains a paragraph about cancer incidence and a list of research goals. A callout box points to the text "Click here to view or search for all variables in the datasets of the selected cohort." which links to a "Variables" button.
- General information:** This section includes fields for Name, Acronym, Website, and Investigators.
- Cohort design:** This section includes a field for Cohort design.
- Populations:** This section includes fields for Populations, Selection criteria, and Number of participants.
- Data collection events:** This section includes a table of data collection events. A callout box points to the "IWHS Vital Status Followup" row, stating: "Click on an event name to call up the description page of the selected event."
- Datasets:** This section includes a table of datasets. A callout box points to the "IWHS_BI_1986" row, stating: "Click on a dataset name to call up the description page of the selected dataset. Click then on the Variables button of the selected dataset to view or search for the variables in the selected dataset. See also Figure 12."

Additional annotations include a box at the top right labeled "Links to previously visited pages" pointing to the breadcrumb trail, and a box at the bottom right labeled "Scroll down" pointing to the bottom of the page.

Figure 9: The detailed page of a selected cohort

This page describes the cohort, also including abbreviated general information, cohort design, populations, data collection events, and datasets. All underlined text links to other pages.

Search Cohorts

On the **Cohorts** page (Figure 8) click on the **Search Cohorts** button to display the **Cohorts Search** page (Figure 10). This page lists the CMR cohorts. It allows filtering of the results by the categories in the left side panel and full text search of the cohort descriptions. Use the **filters** on the left panel, **full text search**, or a combination of the two to narrow the list of cohorts to those of interest.

Current search
Search found 24 studies

COHORT DESIGN

☐ Cohort study (23)

☐ Clinical trial / Intervention trial (4)

☐ Other (1)

COUNTRY OF RESIDENCE

☐ United States (11)

☐ Sweden (5)

☐ Canada (3)

☐ Japan (2)

☐ Netherlands (2)

Show more

MINIMUM AGE

From To

MAXIMUM AGE

From To

GENDER

☐ Women only (7)

☐ Men only (4)

DATA SOURCES

☐ Others (24)

☐ Questionnaires (24)

☐ Physical measures (4)

Cohorts Search

Use the left side to filter or the search box to search for cohorts included in the CMR. Click on a cohort name to see limited descriptive information about the cohort and to navigate to that cohort's metadata datasets or variables.

Items per page

20

Set here the number of records per page. The values are 20, 50, or View all. The default is 20.

Short Name	Name	Cohort Metadata
AgHealth	Agricultural Health Study	AgHealth Metadata
	Cancer Prevention Study	ATBC Metadata
		BGS Metadata
BCGP	British Columbia Generations Project	BCGP Metadata
CSDLH	Canadian Study of Diet, Lifestyle and Health	CSDLH Metadata
COSM	Cohort of Swedish Men	COSM Metadata
EPIC	European Prospective Investigation into Cancer and Nutrition	EPIC Metadata
GCS	Golestan Cohort Study	GCS Metadata
IWHS	Iowa Women's Health Study	IWHS Metadata
JPHC1	Japan Public Health Center-Based Prospective Study Cohort 1	JPHC1 Metadata
JPHC2	Japan Public Health Center-Based Prospective Study Cohort 2	JPHC2 Metadata
NLCS	Netherlands Cohort Study	NLCS Metadata
NSHDS	Northern Sweden Health and Disease Study	NSHDS Metadata
NYUWHS	NYU Women's Health Study	NYUWHS Metadata
PHSI	Physicians' Health Study Cohort I	PHSI Metadata
PHSII	Physicians' Health Study Cohort II	PHSII Metadata
PLCO	Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial	PLCO Metadata
AHS-2	Seventh-day Adventist Cohort Study-2	AHS-2 Metadata
SIS	Sister Study	SIS Metadata
SMC	Swedish Mammography Cohort	SMC Metadata

1 2 next >

Navigation buttons

Use these filters to search for CMR cohorts.

Full text search – type text and click on the search icon or press **Enter** on the keyboard.

Click on a **Name** or **Cohort Metadata** entry to view the detailed page of that cohort (Figure 9) and to navigate from there to that cohort's datasets or variables.

Figure 10: The Cohorts Search page

COHORT METADATA

This table (see Figure 9 and Figure 11) lists the cohort datasets in the CMR. They contain metadata of variables provided by the cohort, but do not contain individual level data. They are named and organized for the purposes of the CMR. They are not the actual dataset names used by the cohort.

CMR cohort datasets are named according to a standardized convention to allow for ease of locating and searching datasets. The naming convention includes the cohort acronym, the data collection event (e.g., **BL** for baseline) and the years of data collection.

To view the **Datasets** table select a cohort **Name** or **Cohort Metadata** entry on the **Cohorts** (Figure 8) or **Cohorts Search** (Figure 10) page and then navigate to the bottom of the detailed page of the selected cohort (Figure 9). The default table order is by **Name**, alphabetical and ascending. Each cohort also has a dataset for cancer incidence and another for mortality outcomes.

Click on a name to view a more detailed description of the selected dataset (Figure 12).

Datasets	
Name	Description
DMCA_HARMONIZED	Diabetes and Cancer Initiative harmonized dataset. This dataset contains the meta-data of a...
IWHS_BL_1986	IWHS metadata for subset of baseline variables requested for the Diabetes and Cancer...
IWHS_CancerIncidence	
IWHS_VitalStatus	IWHS mortality outcome metadata for subset of variables requested for Diabetes and Cancer...

Figure 11: The Cohort Datasets table at the bottom of a detailed page (Figure 9)

Selected Dataset

This page describes the selected dataset and allows viewing and searching its variables. The dataset names are created for the CMR using a defined naming convention of cohort acronym, data collection event and time generated at data collection.

About CMR | Cohort Metadata | **Harmonization Mapping** | Variables Search | Discussion Forum | Metadata/Code Request | Contact Us
Home / Cohort Datasets / AgHealth_BL_1993-1997

AgHealth_BL_1993-1997

AgHealth baseline meta-data for limited subset of variables abstracted for the Diabetes and Cancer Harmonization project. Note that the cohort has different baseline questionnaires for participants (private and commercial applicators) and their spouses so the meta-data are slightly different. Variables beginning with 'A' are for participants and variables beginning with 'S' are for spouses.

Dataset Type:
Cohort dataset

Cohorts
Items per page: 20

Name
Agricultural Health Study

[Variables](#)

Click on the name of the parent cohort to go to the cohort page (Figure 9).

Figure 12: Selected dataset

On the **Cohort Datasets** page (Figure 11) click on a dataset name to select it. Its page will appear (Figure 12). It displays the name and a description of the dataset, its type, and the cohort to which it belongs. The variables can be viewed or searched from the **Variables** button. Click on the cohort name to go to the detailed page of the parent cohort (Figure 9) where information is provided about the cohort and **all** its datasets. The **Variables** button on the cohort page displays all the variables in the datasets of the selected **cohort**. The **Variables** button on a dataset page displays the variables in the selected **dataset**.

VARIABLES SEARCH

Variables Search can be accessed from the **Variables** button – if displayed on any page –, from the **Variables Search** tab on the menu bar (from any page), or from the **Variables Search** icon on the user's home page. The **Variable Search** page (Figure 13) lists the variables that match the results of the search filters or path used to display it.

The screenshot shows the 'Variables Search' page with a navigation bar at the top containing links: About CMR, Cohort Metadata, Harmonization Mapping, Variables Search (highlighted), Discussion Forum, Metadata/Code Request, and Contact Us. Below the navigation bar is a breadcrumb trail: Home / Variables Search.

Current search
Search found 2350 variables

Variables Search

Use the left side to filter or the search box to search the variables in the CMR. Click on a variable name to see the metadata for that variable.

Using the filter will more comprehensively locate all variables in the content domain you select. Using the search locates variables by text search of the variable names and labels and may not be as comprehensive.

- **Name** is the variable name.
- **Label** is a short description of the variable. This may be the actual variable label provided by the cohort. If a label was not provided by the cohort, a brief description was assigned by the site administrator. These can be updated upon request of the Cohort PIs.
- **Dataset** is the CMR-defined Cohort Dataset or Harmonized Dataset that includes the variable.

Questionnaires and codebooks provided by the cohorts were abstracted to compile the metadata into this format for CMR display and search capabilities.

CMR PILOT PHASE INCLUSION:
Metadata are currently limited to a subset of content domains and variables selected for a pilot project NCI Cohort Consortium's Diabetes and Cancer Initiative. More about this initiative is available on the NCI Cohort Consortium's active projects webpage.

The dietary data harmonization project, led by Dr. Stephanie Smith Warner of the Harvard T.H. Chan School of Public Health, is currently ongoing. There will be harmonized definitions for 11 food groups (red meat, processed meat, poultry, fish, fruits, vegetables, dairy products, milk, coffee, tea, and sugar-sweetened beverages). Please contact the [CMR Helpdesk](#) for more information.

PIs who are interested in having their studies included in the CMR or using this tool to conduct their own harmonization project are encouraged to contact the CMR Helpdesk.

Full text search

Type text here and click on the search icon or press **Enter** on the keyboard.

Name	Label	Dataset
ALC_GRAMS	Total grams of ethanol per day at baseline	CSDLH_FFQ_BL_1992-1998
Alive	Vital status	COSM_VitalStatus
Alive	Vital status	SMC_VitalStatus
alive	Vital Status	PHS I_VitalStatus
alive	Vital Status	PHS II_VitalStatus
alive	Alive=1, dead=0	WHS_VitalStatus

Annotations:

- Left sidebar:** Use the filters to search for variables. Check boxes to (un)select filtering criteria. Click on a variable name to see information about it.
- Search box:** Type text here and click on the search icon or press **Enter** on the keyboard.
- Right side:** Read **Variables Search** explanations and instructions here.
- Bottom right:** Scroll down to view the entire page and the navigation buttons.

Figure 13: Top of the *database level* Variables Search page – no filter selected

If your search path has been the menu bar tab or home page icon, then the **Variables Search** page shows *all the variables and variable filters in the entire CMR database* – and no filter is preselected. Figure 13 shows the top section of the CMR database level **Variables Search** page.

If you click on the **Variables** button of a selected **cohort** (see Figure 9), then the **Variables Search** page shows *all the variables in the datasets of the selected cohort* – and the corresponding filters are automatically preselected.

If you click on the **Variables** button of a selected **dataset**, then the **Variables Search** page shows *all the variables in the selected dataset* – and the corresponding filters are automatically preselected.

Once you have landed on the **Variables Search** page, the page gives you access to all the variables in the CMR database, depending only on the search text and selected filters. The layout and functionality of the **Variables Search** page is similar to that of the **Cohorts Search** page (see also Figure 10).

The **Variables Search** page has the same layout and functionality regardless of the search level.

Variables Search Page

Search summaries show the level of the result lists.

Current search
Search found 187 variables - [Clear all criteria](#) [X]
Study: Iowa Women's Health Cohort [X]

Use the filters to search for variables. Check boxes to (un)select filtering criteria. Click on a variable name to see information about it.

Dataset Type
☐ Harmonization dataset (98)
☐ Cohort dataset (89)

Selected cohort
Iowa Women's Health Study

Type text and click on the search icon or press Enter on the keyboard.

Full text search

Datasets in the selected cohort

Scroll down to view the entire page and the navigation buttons.

Variables Search

Use the left side to filter or the search box to search the variables in the CMR. Click on a variable name to see the metadata for that variable.

Use all variables in the content domain you select. Using the variable names and labels and may not be as

- Name** is the variable name.
- Label** is a short description of the variable. This may be the actual variable label provided by the cohort. If a label was not provided by the cohort, a brief description was assigned by the site administrator. These can be updated upon request of the Cohort PIs.
- Dataset** is the CMR-defined Cohort Dataset or Harmonized Dataset that includes the variable.

Questionnaires and codebooks provided by the cohorts were abstracted to compile the metadata into this format for CMR display and search capabilities.

CMR PILOT PHASE INCLUSION:
Metadata are currently limited to a subset of content domains and variables selected for a pilot project NCI Cohort Consortium's Diabetes and Cancer Initiative. More about this initiative is available on the NCI Cohort Consortium's active projects webpage.

The dietary data harmonization project, led by Dr. Stephanie Smith Warner of the Harvard T.H. Chan School of Public Health, is currently ongoing. There will be harmonized definitions for 11 food groups (red meat, processed meat, poultry, fish, fruits, vegetables, dairy products, milk, coffee, tea, and sugar-sweetened beverages). Please contact the [CMR Helpdesk](#) for more information.

PIs who are interested in having their studies included in the CMR or using this tool to conduct their own harmonization project are encouraged to contact the [CMR Helpdesk](#).

Name	Label	Dataset
allcanc	Any cancer (earliest Prmsite)	IWHS_CancerIncidence
DMCA_BL_HEIGHT	Height at baseline	DMCA_HARMONIZED
ID	ID number	IWHS_BL_1986
tot_d	Death	IWHS_VitalStatus
DATEOD	Date of death	IWHS_VitalStatus
diag_age	Age at diagnosis (earliest)	IWHS_CancerIncidence
DMCA_BL_HEIGHT_SOURCE	Source of height at baseline	DMCA_HARMONIZED

Figure 14: Variables Search page of a selected cohort (cohort level)

Depending on the path taken to reach the **Variables Search** page, the resulting list will display all the variables in the CMR database, selected cohort, or selected dataset. This page allows full text search on variable name and label or filtering of the results by the categories in the left-side panel. You may combine full text searches and filtering.

The **Variables Search** table has the following columns: **Name** – the name of the variable, **Label** – a short description of the variable, and **Dataset** – the datasets in which the variable is included. At **database** level (Figure 13), the page shows *all the variables in the CMR database* – and no filter is preselected. The **Dataset** column lists all the datasets in the database.

At **cohort** level (Figure 14), the page shows *all the variables in the datasets of the selected cohort* and the variables from any harmonized dataset in which that cohort participated. The corresponding filters are automatically preselected. You can limit the results to just cohort variables and exclude the harmonized by selecting the corresponding cohort subset **Dataset Type**.

At **dataset** level, the page shows *all the variables in the selected dataset* – and the corresponding filters are automatically preselected (Figure 15).

Search summaries show the *level* of the result lists.

Use the filters to search for variables. Check boxes to (un)select filtering criteria. Click on a variable name to see information about it.

Type text and click on the search icon or press **Enter on the keyboard.**

Current search
Search found 53 variables - [Clear all criteria](#) [X]
Study: [European Prospective Investigation into Cancer and Nutrition](#) [X]
Dataset: [EPIC_BL_1991_1999](#) [X]

Variables Search
Use the left side to filter or the search box to search the variables in the CMR. Click on a variable name to see the metadata for that variable.
Using the filter will more comprehensively locate all variables in the content domain you select. Using the search locates variables by text search of the variable names and labels and may not be as comprehensive.

Selected dataset
EPIC_BL_1991_1999

Selected cohort
European Prospective Investigation into Cancer and Nutrition

Full text search

The selected dataset

Name	Label	Dataset
Idepic	Unique EPIC subject ID	EPIC_BL_1991_1999
D_Lsq	Date of lifestyle questionnaire	EPIC_BL_1991_1999
Sex	Sex	EPIC_BL_1991_1999
Age_Recr	Age at recruitment	EPIC_BL_1991_1999
D_Recrui	Date of recruitment	EPIC_BL_1991_1999
Mar_Stat	Marital status	EPIC_BL_1991_1999

Variables in the selected dataset

Scroll down to view the entire page and the navigation buttons.

Figure 15: Variables Search page of a selected **dataset** (dataset level)

Variable Filter Categories

Note: All CMR variables have been included into categories by content domain. This allows filtering on like variables across cohorts to explore feasibility of harmonization. The categories are hierarchical, so you can broaden or narrow your criteria with your selection of filters. For finding variables by content area **the use of the filters is recommended** rather than full text search.

Current search
Search found 10 variables - [Clear all criteria \[X\]](#)
Anthropometric Structures: [Self-reported waist circumference \[X\]](#)

Collapsed filters

Current search
Search found 10 variables - [Clear all criteria \[X\]](#)
Anthropometric Structures: [Self-reported waist circumference \[X\]](#)

Expanded filter category.
Click a down arrow (▼) to collapse a filter category. Click a right arrow (▶) to expand a filter category. Click **Show more** to expand filters in a category.

Select this one to select all above it. Uncheck it to deselect it and all above it. Uncheck **Body circumferences** to deselect all under it.

Search summary – shows which filters have been selected. The selected filters can also be deselected from summary – click on any underlined text or on an [X] to deselect a filter or all filters.

Using the filter will more comprehensively locate all variables in the content domain you select. Using the search locates variables by text search of the variable names and labels and may not be as

This may be the actual variable label provided by the report, a brief description was assigned by the site request of the Cohort PIs.

Harmonized Dataset that includes the variable.

Shorts were abstracted to compile the metadata into

ent domains and variables selected for a pilot project

NCI Cohort Consortium's Diabetes and Cancer Initiative. More about this initiative is available on the NCI Cohort Consortium's active projects webpage.

The dietary data harmonization project, led by Dr. Stephanie Smith Warner of the Harvard T.H. Chan School of Public Health, is currently ongoing. There will be harmonized definitions for 11 food groups (red meat, processed meat, poultry, fish, fruits, vegetables, dairy products, milk, coffee, tea, and sugar-sweetened beverages). Please contact the CMR Helpdesk for more information.

Pls who are interested in having their studies included in the CMR or using this tool to conduct their own harmonization project are encouraged to contact the CMR Helpdesk.

Full text search

Name	Label	Dataset
bodysize ^{waist}	waist circumference (cm), at recruitment	BGS_BL_2003-2011
Waist_C	Waist circumference (computed)	EPIC_BL_1991-1999
WAIST	Waist circumference	WLH_BL_1991-1992
	umference	COSM_BL_1997
	t measurement (inches)	IWHS_BL_1986
	d measurement (inches)	IWHS_BL_1986
F14	Waist circumference	SMC_BL_1997
WAIST_CIR	Waist circumference at or nearest to baseline	CSDLH_BL_1992-1998
WAIST_CIR_YR	Year waist circumference ascertained	CSDLH_BL_1992-1998
S_PM_WAIST_SR_AVG	Average waist circumference - self-reported	BCGP_BL_2008-Ongoing

Click in check boxes to (de)select filtering categories.

Figure 16: Handling the filter categories

Each filter category (Figure 16) can be collapsed or expanded from the preceding arrow (▼, ▶). When the arrow points down the category is expanded; click on it to collapse the category. Click in a filter check box to select it. The results will be filtered by that category and the number of filters will be adjusted accordingly. Within one and the same higher filter category all higher categories are automatically selected by the selection of the lowest categories. Uncheck the lowest or the highest to deselect all those categories. Click **Show more** to expand the list of filters in a category. The **Show more** link becomes **Show fewer**. Click **Show fewer** to return to the original list.

Variable Full Text Search

Note: The full text search currently has some limitations: it searches only the variable name and label and not the more detailed variable description.

To find a variable, if you know its name, type it (full text) in the search box and click on the search icon or press **Enter** on the keyboard. The table will display the records that contain the search text. You may search on name or label entries. Full text searches and filtering may be combined. Full text search will

search within the *level* established by the selected filter categories – that is, if no filter is selected, full text search will look in the entire CMR database, if a filter category is selected, full text search will look in the corresponding category. Compare Figure 17 and Figure 18 below.

The screenshot shows the 'Variables Search' page. The left sidebar contains filter categories: DATASET TYPE (Cohort dataset (18), Harmonization dataset (1)), DATASET (Complete (1), Impossible (1)), DATA COLLECTION EVENT, POPULATION, DATA SOURCES, COHORT (Agricultural Health Cohort (3), Alpha-Tocopherol Beta-Carotene Cancer Prevention Cohort (2), Breakthrough Generations Cohort (2), British Columbia Generations Project (2), Canadian Cohort of Diet, Lifestyle and Health (2)), DATA SOURCE, and SOCIODEMOGRAPHIC/SOCIOECONOMIC CHARACTERISTICS (Household (19), Marital/Partner status (19)). The main content area has a search bar with 'marital' entered. Below the search bar is a table of results.

Current search
Search found 19 variables

Variables Search
Use the left side to filter or the search box to search the variables in the CMR. Click on a variable name to see the metadata for that variable.
Using the filter will more comprehensively locate all variables in the content domain you select. Using the search locates variables by text search of the variable names and labels and may not be as comprehensive.

- Name** is the variable name.
- Label** is a short description of the variable. This may be the actual variable label provided by the cohort. If a label was not provided by the cohort, a brief description was assigned by the site administrator. These can be updated upon request of the Cohort PIs.
- Dataset** is the CMR-defined Cohort Dataset or Harmonized Dataset that includes the variable.

Questionnaires and codebooks provided by the cohorts were abstracted to compile the metadata into this format for CMR display and search capabilities.

CMR PILOT PHASE INCLUSION:
Metadata are currently limited to a subset of content domains and variables selected for a pilot project NCI Cohort Consortium's Diabetes and Cancer Initiative. More about this initiative is available on the NCI Cohort Consortium's active projects webpage.

The dietary data harmonization project, led by Dr. Stephanie Smith Warner of the Harvard T.H. Chan School of Public Health, is currently ongoing. There will be harmonized definitions for 11 food groups (red meat, processed meat, poultry, fish, fruits, vegetables, dairy products, milk, coffee, tea, and sugar-sweetened beverages). Please contact the [CMR Helpdesk](#) for more information.

Pls who are interested in having their studies included in the CMR or using this tool to conduct their own harmonization project are encouraged to contact the [CMR Helpdesk](#).

Click on a name to view the detail page of the selected variable. See example in Figure 19.

Name	Label	Dataset
MARSTAT	Marital Status	IWHS_BL_1986
MARITAL	Marital status	ATBC_BL_1985-1988
BACMARIZ	Current marital status	VITAL_BL_2000-2002
civil	Marital status	NSHDS_BL_1992
F1327	Marital Status	COSM_BL_1997
MARITAL	Marital status	WHS_BL_1993-1996
MARRIAGE	Marital status	NIH-AARP_BL_1995-1996
SE_Marital	CALC: Current marital status [SE12,SE13,SE14]	SIS_BL_2003-2009
Mar_Stat	Marital status	EPIC_BL_1991-1999
maritalstatus	marital status at recruitment	BGS_BL_2003-2011
marital	Marital Status	PLCO_BL_1993-2001
MaritalStatus	specifies the patient marital status	GCS_BL_2004-2008
A_SDC_MARITAL_STATUS	Marital status	BCGP_BL_2008-Ongoing
MARITAL	MARITAL	AHS2_BL_2002-2007
MARSTAT	MARITAL STATUS [BURGST]	NLCS_BL_1986
amarital	E38: Marital status	AgHealth_BL_1993-1997
SMARITAL	What is your current marital status?	AgHealth_BL_1993-1997
DMCA_BL_MARITAL	Marital status	DMCA_HARMONIZED
Q9	Marital Status	CSDLH_BL_1992-1998

Figure 17: Results of a search against the entire database – example

In Figure 17 the CMR found variables for the search string “marital” in the *entire database*. The left-pane filter categories have been reduced as it is logical. Full text search is available at all levels. It may be combined with left-pane filters. If filters have been selected, leave them selected to stay within the same *cohort* or *dataset*; unselect them to do a search against the entire dataset.

Current search
Search found 2 variables - Clear all criteria [X]
Study: Iowa Women's Health Cohort [X]

Variables Search
Use the left side to filter or the search box to search the variables in the CMR. Click on a variable name to see the metadata for that variable.

Using the filter will more comprehensively locate all variables in the content domain you select. Using the search locates variables by text search of the variable names and labels and may not be as comprehensive.

- Name** is the variable name.
- Label** is a short description of the variable. This may be the actual variable label provided by the cohort. If a label was not provided by the cohort, a brief description was assigned by the site administrator. These can be updated upon request of the Cohort PIs.
- Dataset** is the CMR-defined Cohort Dataset or Harmonized Dataset that includes the variable.

Questionnaires and codebooks provided by the cohorts were abstracted to compile the metadata into this format for CMR display and search capabilities.

CMR PILOT PHASE INCLUSION:
Metadata are currently limited to a subset of content domains and variables selected for a pilot project NCI Cohort Consortium's Diabetes and Cancer Initiative. More about this initiative is available on the NCI Cohort Consortium's active projects webpage.

The dietary data harmonization project, led by Dr. Stephanie Smith Warner of the Harvard T.H. Chan School of Public Health, is currently ongoing. There will be harmonized definitions for 11 food groups (red meat, processed meat, poultry, fish, fruits, vegetables, dairy products, milk, coffee, tea, and sugar-sweetened beverages). Please contact the CMR Helpdesk for more information.

PIs who are interested in having their studies included in the CMR or using this tool to conduct their own harmonization project are encouraged to contact the CMR Helpdesk.

Search box:

Name	Label	Dataset
MARSTAT	Marital Status	IWHS_BL_1986
DMCA_BL_MARITAL	Marital status	DMCA_HARMONIZED

Click on a name to view the detail page of the selected variable. See example in Figure 19.

Figure 18: Results of a search for variables in a selected cohort – example

VARIABLES

MARSTAT

Description
Label: Marital Status
Dataset: IWHS_BL_1986
Cohorts: Iowa Women's Health Study
Value Type: Integer
Repeatable: No
Description: Q122. What is your current marital status?

Domains
Data Source: Baseline Questionnaire
Sociodemographic/Socioeconomic Characteristics: Marital/Partner status

Categories

Name	Label	Missing
1	Never Married	
2	Currently Married	
3	Separated or Divorced	
4	Widowed	
.	Missing	✓

Variable Label is provided by the cohort.

Value Type can be integer, decimal, text, Boolean, etc.

Description is the wording of questions in the questionnaire or the cohort's derivation if the variable was derived by the cohort. The question number is shown if available (example: Q122).

Domains are categories to which variables are assigned for filtering purposes.

Categories are coded for categorical variables (such as 1 = Never Married, etc.). For continuous variables, the table shows Valued or Missing (see example in Figure 26).

Figure 19: Detail page of a cohort dataset variable (example: MARSTAT)

Variables are included in **Cohort** and **Harmonized Datasets**. The detail page of a variable (Figure 19) describes the variable including its known description, cohorts, value types, categories (i.e., code values and labels), domains (categories that have been assigned to the variable), etc. Cohort and harmonized dataset variable pages are slightly different. If you are viewing harmonized dataset variables, harmonization data is also included at the bottom of the variable page. When the searched-for variable has been found, click on its name to view its detail page.

HARMONIZATION DATASETS

This table lists the harmonized datasets that have been created and stored in the CMR (Figure 20). The page features are the same as those of the **Cohorts Datasets** page (Figure 11). The page can be accessed from the **Harmonization Mapping** menu option or home page icon.

When there are more than one harmonized datasets, the table will display twenty records per page, sorted by **Name**. For the time being there is one harmonized dataset, the **DMCA_HARMONIZED** – the Diabetes and Cancer Initiative (DMCA) dataset. Select a harmonized dataset to view the list of cohorts included in it, to search for the included variables, to view the metadata for the harmonized variables, to display a chart showing which cohorts were successfully mapped to the harmonized variables, and to see the script that maps the cohort variables to the harmonization variable.

About CMR | Cohort Metadata | **Harmonization Mapping** | Variables Search | Discussion Forum | Metadata/Code Request | Contact Us
Home / Harmonization Mapping

Harmonization Datasets

Each of the harmonized dataset consists of metadata for a set of variables resulting from data harmonization efforts for an individual project. The processing algorithms used to map cohort-specific metadata to the variable as defined in the Harmonization Dataset can be viewed at the variable level.

Select a Harmonized Dataset to:

- View the list of cohorts included
- Query to find variables
- View the metadata for the harmonized variables
- View a chart displaying which cohorts were successfully mapped to the harmonized variables
- View the algorithms used to harmonize the variables from each cohort

CMR PILOT PHASE INCLUSION:
 For the pilot phase, the CMR includes only one harmonization dataset from the NCI Cohort Consortium's Diabetes and Cancer Initiative. More about this initiative is available on the NCI Cohort Consortium's active projects webpage.

PIs who are interested in having their studies included in the CMR or using this tool to conduct their own harmonization project are encouraged to contact the CMR Helpdesk.

Name	Description
DMCA_HARMONIZED	Diabetes and Cancer Initiative harmonized dataset. This dataset contains the meta-data of a subset of target variables harmonized for the Diabetes and Cancer Initiative. Meta-data includes the...

Click here to view the details of the selected harmonized dataset.

Figure 20: *The Harmonization Datasets page and table*

Harmonized Dataset

On the **Harmonization Datasets** page (Figure 20) click on the name of the desired harmonized dataset. The details page of the selected harmonized dataset will appear (Figure 21). This page describes the harmonized dataset, also showing the cohorts that are included in the project. Click on the **Harmonization** or **Variables** button to view the harmonization chart (summary) or, respectively, the **Variables** page of the harmonized dataset.

The screenshot shows the 'DMCA_HARMONIZED' dataset details page. The page has a navigation bar at the top with links: About CMR, Cohort Metadata, Harmonization Mapping (highlighted), Variables Search, Discussion Forum, Metadata/Code Request, and Contact Us. Below the navigation bar is the breadcrumb 'Home / DMCA_HARMONIZED'. The main heading is 'DMCA_HARMONIZED'. Below this is a paragraph describing the dataset: 'Diabetes and Cancer Initiative harmonized dataset. This dataset contains the meta-data of a subset of target variables harmonized for the Diabetes and Cancer Initiative. Meta-data includes the variable names, descriptions, and formats. No individual level data are included on this site. 20 cohorts have provided meta-data for harmonization mapping. The project lead is Marc Gunter of the Imperial College of London.' Below this is a section 'Content domains harmonized on this site include:' followed by a bulleted list of domains: demographics, prevalent cancer, baseline anthropometry, baseline smoking, baseline alcohol, baseline selected reproductive factors, baseline family history of cancer in first degree relatives (overall and breast, colorectal, prostate), prevalent diabetes, cancer outcomes, and mortality. Below this is a section 'Dataset Type:' with the value 'Harmonization dataset'. Below this is a section 'Cohorts' with a dropdown menu 'Items per page' set to '20'. Below this is a table of cohorts with the following entries: Agricultural Health Study, Alpha-Tocopherol Beta-Carotene Cancer Prevention Study, Breakthrough Generations Study, British Columbia Generations Project, Canadian Study of Diet, Lifestyle and Health, Cohort of Swedish Men, European Prospective Investigation into Cancer and Nutrition, Golestan Cohort Study, Iowa Women's Health Study, Japan Public Health Center-Based Prospective Study Cohort 1, Japan Public Health Center-Based Prospective Study Cohort 2, Netherlands Cohort Study, Northern Sweden Health and Disease Study, NYU Women's Health Study, Physicians' Health Study Cohort I, Physicians' Health Study Cohort II, Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, Seventh-day Adventist Cohort Study-2, Sister Study, and Swedish Mammography Cohort. At the bottom of the table are pagination links: 1, 2, next. There are four callout boxes with arrows pointing to specific elements: 1. 'Click here to see a harmonization chart of variables by cohort (Figure 22).', 2. 'Click here to see the Variables page of the harmonized dataset.', 3. 'Set here the number of records per page. The values are 20, 50, or View all. The default value is 20.', 4. 'Click on a Name entry to view a more detailed description of the selected cohort.'

Click here to see a harmonization chart of variables by cohort (Figure 22).

Click here to see the Variables page of the harmonized dataset.

Set here the number of records per page. The values are 20, 50, or View all. The default value is 20.

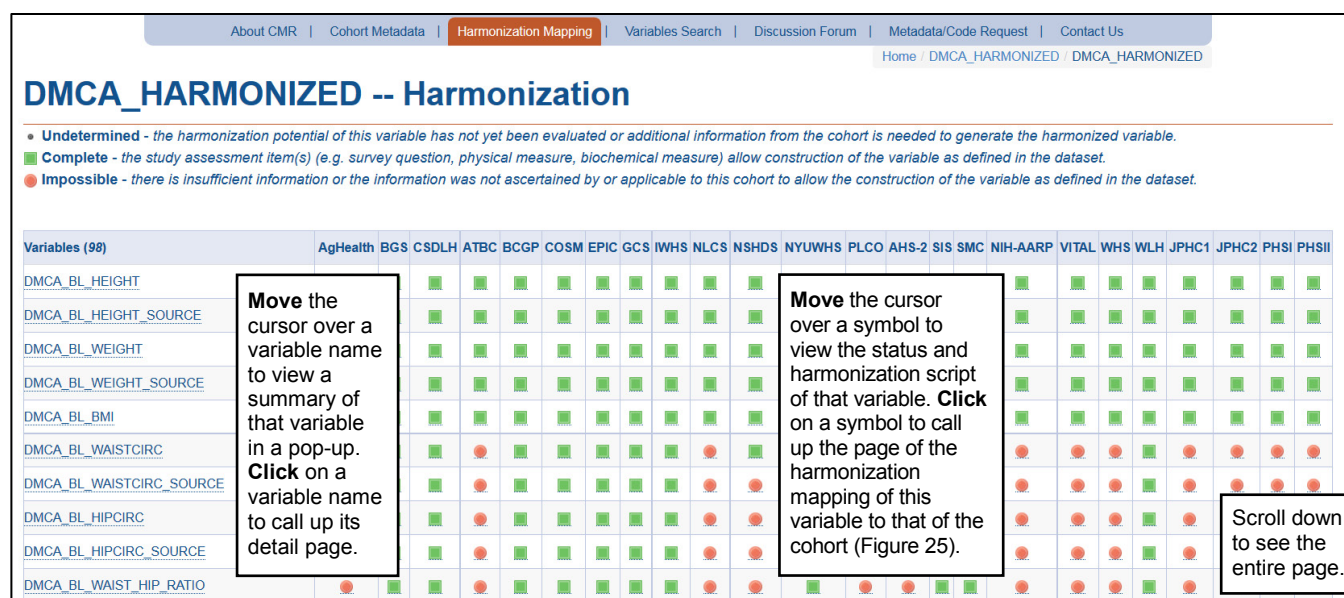
Click on a Name entry to view a more detailed description of the selected cohort.

Scroll down to see the navigation buttons.

Figure 21: Details page of a selected harmonized dataset

Dataset Harmonization Summary

The harmonization chart is a summary of the harmonization of the variables by cohort. Click on the **Harmonization** button (Figure 21) to view the chart (Figure 22).



The legend is the following:

- **Undetermined** – the harmonization potential of this variable has not yet been evaluated.
- **Complete** – the cohort assessment item(s) (e.g. survey question, physical measure, biochemical measure) allow construction of the variable as defined in the dataset.
- **Impossible** – there is no information or insufficient information or the information was not ascertained or applicable for this cohort to allow the construction of the variable as defined in the dataset.

Undetermined states may also be because the metadata have not yet been received, or because the cohort needs to be contacted directly to derive the variables.

In the chart, move the cursor over a variable name (the first column in Figure 22) to view a brief summary of that variable in a pop-up (Figure 23). Click on that variable name to call up its detail page (Figure 19). These variables have been harmonized as indicated in the chart summary (Figure 22), so that their detail page (Figure 19) will also include harmonization metadata and mapping. See also the section on Variable Harmonization Detail.

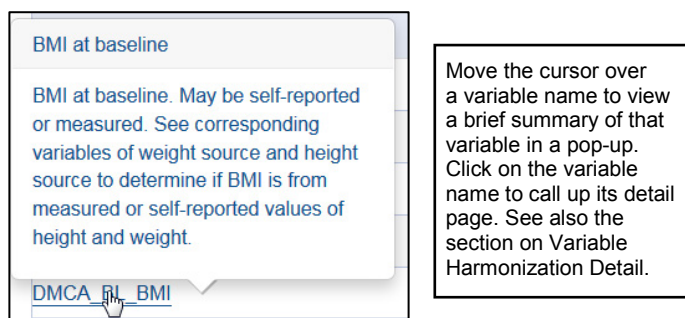


Figure 23: Pop-up variable summary

Move the cursor over a legend symbol to view a brief summary of the status and harmonization script of that variable in a selected cohort (see example in Figure 24). Click on that symbol to call up the detail page of that variable in the selected cohort (Figure 25).

Variables (98)	AgHealth	BGS	CSDLH	ATBC	BCGP	COSM	EPIC
DMCA_BL_HEIGHT							
DMCA_BL_HEIGHT_SOURCE							
DMCA_BL_WEIGHT							
DMCA_BL_WEIGHT_SOURCE							
DMCA_BL_BMI							
DMCA_BL_WAISTCIRC							
DMCA_BL_WAISTCIRC_SOURCE							
DMCA_BL_HIPCIRC							
DMCA_BL_HIPCIRC_SOURCE							
DMCA_BL_WAIST_HIP_RATIO							

Status: Complete

Script:

```
IF (HIP_CIR = . OR HIP_CIR = U)
  THEN DMCA_BL_HIPCIRC = 999;
ELSE DMCA_BL_HIPCIRC = HIP_CIR
;
```

Move the cursor over a symbol to view the status and harmonization script of that variable in a selected cohort. Click on that symbol to call up the detail page of that variable in the selected cohort (Figure 25).

Figure 24: Status and script of DMCA_BL_HIPCIRC in CSDLH

Home / DMCA_HARMONIZED / DMCA_BL_HIPCIRC

DMCA_BL_HIPCIRC -- Canadian Study of Diet, Lifestyle and Health

Variable: [DMCA_BL_HIPCIRC](#)

Cohort: Canadian Study of Diet, Lifestyle and Health

Status: Complete

Script:

```
IF (HIP_CIR = . OR HIP_CIR = U) THEN DMCA_BL_HIPCIRC = 999;
ELSE DMCA_BL_HIPCIRC = HIP_CIR;
```

Click on the variable name to call up its detail page.

Figure 25: Detail harmonization data for a selected variable in a cohort

Variable Harmonization Detail

This page is similar to the detail page of any other variable (as shown in Figure 19), but, as these variables have been harmonized, the page also includes the list of cohorts for which the variable has been harmonized as well as a table that shows harmonization details for the selected variable across cohorts (see Figure 26 and Figure 27).

About CMR | Cohort Metadata | Harmonization Mapping | Variables Search | Discussion Forum | Metadata/Code Request | Contact Us

Home / DMCA_HARMONIZED

DMCA_BL_HIPCIRC

Description

Label:
Hip circumference at baseline

Cohorts:
[Agricultural Health Study](#)
[Breakthrough Generations Study](#)
[Canadian Study of Diet, Lifestyle and Health](#)
[Alpha-Tocopherol Beta-Carotene Cancer Prevention Study](#)
[British Columbia Generations Project](#)
[Cohort of Swedish Men](#)
[European Prospective Investigation into Cancer and Nutrition](#)
[Golestan Cohort Study](#)
[Iowa Women's Health Study](#)
[Netherlands Cohort Study](#)
[Northern Sweden Health and Disease Study](#)
[NYU Women's Health Study](#)
[Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial](#)
[Seventh-day Adventist Cohort Study-2](#)
[Sister Study](#)
[Swedish Mammography Cohort](#)
[The National Institutes of Health AARP Diet and Health Study](#)
[VITamins And Lifestyle](#)
[Women's Health Study](#)
[Women's Lifestyle and Health](#)
[Japan Public Health Center-Based Prospective Study Cohort 1](#)
[Japan Public Health Center-Based Prospective Study Cohort 2](#)
[Physicians' Health Study Cohort I](#)
[Physicians' Health Study Cohort II](#)

Dataset:
DMCA_HARMONIZED

Value Type:
Decimal

Repeatable:
No

Unit:
cm

Description:
Hip circumference at baseline. May be self-reported or measured. See variable DMCA_BL_HIPCIRC_SOURCE to distinguish self-report vs. measured.

Categories

Name	Label	Missing
[Value]	Valued	
999	Unknown	✓

Domains

Anthropometric Structures:
Hip circumference

List of cohorts participating in this harmonization project. Click on any to view its detailed description page (Figure 9).

Value type of the variable

Unit of the variable

Description of the harmonized variable may contain important notes about definition and derivation.

Categories are coded for *categorical* variables. For *continuous* variables, the table shows **Valued** or **Missing** (see also Figure 19).

The harmonization data table follows under the **Categories** table. Scroll down to view it and see also Figure 27.

Figure 26: Top section of the detail page of a harmonized variable

The harmonization data table follows under the **Categories** table (Figure 27). Scroll down to view it.

Harmonization

- **Undetermined** - the harmonization potential of this variable has not yet been evaluated.
- **Complete** - the study assessment item(s) (e.g. survey question, physical measure, biochemical measure) allow construction of the variable as defined in the dataset.
- **Impossible** - there is no information or insufficient information or the information was not ascertain or applicable by this cohort to allow the construction of the variable as defined in the dataset.

Cohort	Status	Comment	Script	
AgHealth	●	Not ascertained by cohort at baseline.		View
BGS	■		IF (bodysizehip = 888 OR bodysizehip = 999) THEN DMCA_BL_HIPCIRCUMFERENCE = 999; ELSE DMCA_BL_HIPCIRCUMFERENCE = bodysizehip;	View
CSDLH	■		IF (HIP_CIR = . OR HIP_CIR = U) THEN DMCA_BL_HIPCIRCUMFERENCE = 999; ELSE DMCA_BL_HIPCIRCUMFERENCE = HIP_CIR;	View
ATBC	●	Not ascertained by cohort at baseline.		View
BCGP	■		DMCA_BL_HIPCIRCUMFERENCE = S_PM_HIP_SR_AVG;	View
COSM	■		IF (F15 = .) THEN DMCA_BL_HIPCIRCUMFERENCE = 999; ELSE DMCA_BL_HIPCIRCUMFERENCE = F15;	View
EPIC	■		IF (DMCA_BL_WAIST_HIP_RATIO-whr_adj) THEN DMCA_BL_HIPCIRCUMFERENCE=(waist_adj/whr_adj); ELSE IF (DMCA_BL_WAIST_HIP_RATIO-whr_C) THEN DMCA_BL_HIPCIRCUMFERENCE=(Waist_C/Whr_C);	View
GCS	■		IF (BodyIndicesHips = 999) THEN DMCA_BL_HIPCIRCUMFERENCE = 999; ELSE DMCA_BL_HIPCIRCUMFERENCE = BodyIndicesHips;	View
HPFS	●			View
IWHS	■		IF BUTT1 = 99.99 OR BUTT2 = 99.99 THEN DMCA_BL_HIPCIRCUMFERENCE = 999; ELSE DMCA_BL_HIPCIRCUMFERENCE = (((BUTT1 + BUTT2)/2) * 2.54);	View
NLCS	●	Not ascertained by cohort at baseline.		View
NSHDS	●	Not ascertained by cohort at baseline.		View
NHS I	●			View
NYUWHS	■		IF (HIP = [blank]) THEN DMCA_BL_HIPCIRCUMFERENCE = 999; ELSE DMCA_BL_HIPCIRCUMFERENCE = HIP;	View
PLCO	●	Not ascertained by cohort at baseline.		View
AHS-2	●	Not ascertained by cohort at baseline.		View
SCS	●			View
SCHS	●			View
SIS	■		IF (Ex_Hip = .) THEN DMCA_BL_HIPCIRCUMFERENCE = 999; ELSE DMCA_BL_HIPCIRCUMFERENCE = (Ex_Hip * 2.54);	View
SMC	■		IF (F15 = .) THEN DMCA_BL_HIPCIRCUMFERENCE = 999; ELSE DMCA_BL_HIPCIRCUMFERENCE = F15;	View
NIH-AARP	●	Not ascertained by cohort at baseline.		View
VITAL	●	Not ascertained by cohort at baseline.		View
WHS	●	Not ascertained by cohort at baseline.		View

Click on a **View** link to call up the detail harmonization data page for the selected variable in the specified cohort (Figure 25).

Comment provides additional explanations to supplement the script for variables that were harmonized and also reasons for the **Impossible** or **Undetermined** status.

Scroll all the way down to see the **View DMCA_HARMONIZED Harmonization** link. Click on it to view the entire harmonization chart for the selected harmonized dataset (Figure 22).

Figure 27: The harmonization detail table of a harmonized variable (truncated)

The variable harmonization detail table lists first the cohort acronyms in the **Cohort** column. **Status** is the second column and it uses the same symbols as the dataset harmonization summary chart (Figure 22). Then there are **Comment**, **Script**, and **View** columns. **View** links to the detail harmonization data page for a selected variable in the specified cohort (Figure 25). Click on the **View DMCA_HARMONIZED Harmonization** link at the bottom of the page to view the entire harmonization chart for the selected harmonized dataset (Figure 22). **Comment** provides additional explanations to supplement the script for variables that were successfully harmonized and also reasons for the **Impossible** or **Undetermined** status.

METADATA/CODE REQUEST

Use this menu option to request metadata or harmonization code. The command can be accessed from the **Metadata/Code Request** tab on the menu bar. Click on it to call up the **CMR Metadata and Harmonization Code Request Form**. Complete all fields on the form, attach a completed **CMR Request Checklist** with the specified cohorts and/or harmonization code you are requesting, specify the purpose of the request, and click **Submit**. You can download the **CMR Checklist** from the **CMR Request Checklist** link. If you have questions about this request, click on the **Contact Us form** link to submit your questions.

The screenshot shows the 'CMR Metadata and Harmonization Code Request Form'. Callouts include:

- Your name and email are prefilled in these fields.** (points to the top input fields)
- Attach CMR Request Checklist** (points to the 'Browse...' button)
- Attach and upload files in this area.** (points to the 'Upload' button)
- Specify purpose of request here.** (points to the 'Purpose of the Request' text area)
- Submit** (points to the 'Submit' button)
- Complete all fields on this form to request metadata or harmonization code from the CMR.** (points to the form instructions)
- Attach a completed CMR Request Checklist with the specified cohorts and/or harmonization code you are requesting.** (points to the 'Attach CMR Request Checklist' section)
- Download the CMR Checklist here: CMR Request Checklist** (points to the 'CMR Request Checklist' link)
- If you have questions about this request, submit a question through the Contact Us form.** (points to the 'Contact Us form' link)
- Click here to contact us.** (points to the 'Contact Us' link)
- Click here to download the CMR Request Checklist.** (points to the 'CMR Request Checklist' link)

Figure 28: Metadata/Code Request Form

DISCUSSION FORUM

The screenshot shows the 'Discussion Forum' page. Callouts include:

- Add topic from here** (points to the 'Add new Discussion topic' button)
- Links to other pages** (points to the 'Welcome to the CMR Discussion Forum' link)
- Diabetes and Cancer Pooling Project** (points to the topic title in the table)

Discussion(s)	Topics	Posts	Last post
Welcome to the CMR Discussion Forum All comments appearing in the CMR Discussion Forum are strictly the views of the individual posters. Neither NCI nor site administrators have verified any data descriptions, or code corrections posted, nor do they endorse statements posted here. The purpose of this forum is to provide a space for CMR users to interact and share tips for successful data harmonization. Site administrators will monitor these boards periodically. If you have questions, please visit contact us to communicate with site administrators. Welcome to the CMR Discussion Forum!	0	0	n/a
Diabetes and Cancer Pooling Project	2	2	By neagu_g 1 min 24 sec ago

Figure 29: Discussion Forum page

The **Discussion Forum** page is self-explanatory. The table has the following columns: **Discussion(s)**, number of **Topics**, number of **Posts**, and **Last Post**.

Add Topic

Click on the **Add new Discussion** topic link to add a topic.

Figure 30: Create Discussion topic on this form

Type in a **Subject**, select a **Discussion** value, enter comments in the **Body** section, and click **Save**. The new topic will increase the number of **Topics** and **Posts** (Posts include replies) on the **Discussion Forum** page (Figure 29). Your user name and time elapsed since the new post was made will be entered in the **Last post** column.

If you are the author of a post, you may view or edit your comments. Click on a **Discussion(s)** link (see Figure 29) to open that page. On the new page (Figure 31) you may add a topic, view/edit your comments, or reply to others. This table can be sorted by clicking on any column header: **Topic**, **Replies**, or **Last reply**.

Add topic from here

Discussion Forum

[Add new Discussion topic](#)

Topic	Replies	Last reply
Welcome to the CMR Discussion Forum By miller_a 1 month 1 week ago	1	By neagu_g 22 min 52 sec ago
New variables By neagu_g 1 hour 35 min ago	0	n/a

Click on a link to view/edit your comments or to reply to others.

(Re)sort the table by clicking on a column header: **Topic**, **Replies**, or **Last reply**.

Figure 31: Add, view, or reply to topics from this page

Reply

Click on a **Topic** entry (Figure 31). The history page of the selection will appear (Figure 32).

Home / Discussions / Diabetes and Cancer Pooling Project

Welcome to the CMR Discussion Forum

Submitted by miller_a on Tue, 09/20/2016 - 12:04

Discussions:

- Diabetes and Cancer Pooling Project

All comments appearing in the CMR Discussion Forum are strictly the views of the individual posters. Neither NCI nor site administrators have verified any data descriptions, or code corrections posted, nor do they endorse statements posted here. The purpose of this forum is to provide a space for CMR users to interact and share tips for successful data harmonization. Site administrators will monitor these boards periodically. If you have questions, please visit [contact us](#) to communicate with site administrators.

Welcome to the CMR Discussion Forum!

[Add new comment](#)

Submitted by neagu_g on Wed, 11/02/2016 - 15:49 [Permalink](#)

Test Reply

Test reply.

[reply](#)

Add new comment

Your name
neagu_g

Subject

Comment *

Format [Source](#) **B** **I** **U** **L** **Code** **Image** **Table** **Link**

[body p](#)

[Save](#)

Click here to return to the table in Figure 31.

Reply from here

View and Edit tabs appear under the topic title of your own entries (Figure 33). View is the default. Click Edit to edit your comments.

Add new comment to the selected topic

Figure 32: History page of a selected topic

View/Edit Own Comments

Your own comments, to which no reply is attached, can be viewed or edited from a screen similar to Figure 32 with the difference that **View** and **Edit** tabs appear under the topic title of your own entries. **View** is the default. Click **Edit** to edit your comments.

About CMR | Cohort Metadata | Harmonization Mapping | Variables Search | Discussion Forum | Metadata/Code Request | Contact Us

Home / Discussions / Diabetes and Cancer Pooling Project

New variables

[View](#) [Edit](#)

Submitted by neagu_g on Wed, 11/02/2016 - 14:36

Discussions:

- Diabetes and Cancer Pooling Project

New variables have been added.

To edit your comments, click **Edit** and follow the onscreen instructions.

Figure 33: Top section of the View/Edit form for your own topics

Contact Us

Click on the **Contact Us** menu option or link (a link to the **Contact Us** form is included on many pages) to display the **Contact Us** form. Fill out the form (follow the onscreen instructions to do so) and click **Submit** to send your questions to **CMR Help Desk**.