User Guide for KITE Availability Tool

National Institute for Health and Welfare

05/2015

Contents

1. Introduction

KITE is an open web-based availability service for researchers to find metadata about existing Finnish sample collections and sample collection variables. KITE is a catalog and a query tool developed by the Institute for Molecular Medicine Finland (FIMM) and the National Institute for Health and Welfare (THL), and funded by the BBMRI.fi.

KITE gives internationally standardized and summarized information about available sample collections, and offers browse and search functionalities for targeted variable search. The user can either browse variables straight from the collection trees or use the search tool with several options to find and select specific variables. The selected variables are collected to a query cart from where they can be easily downloaded.

https://kite.fimm.fi/

2. Front page
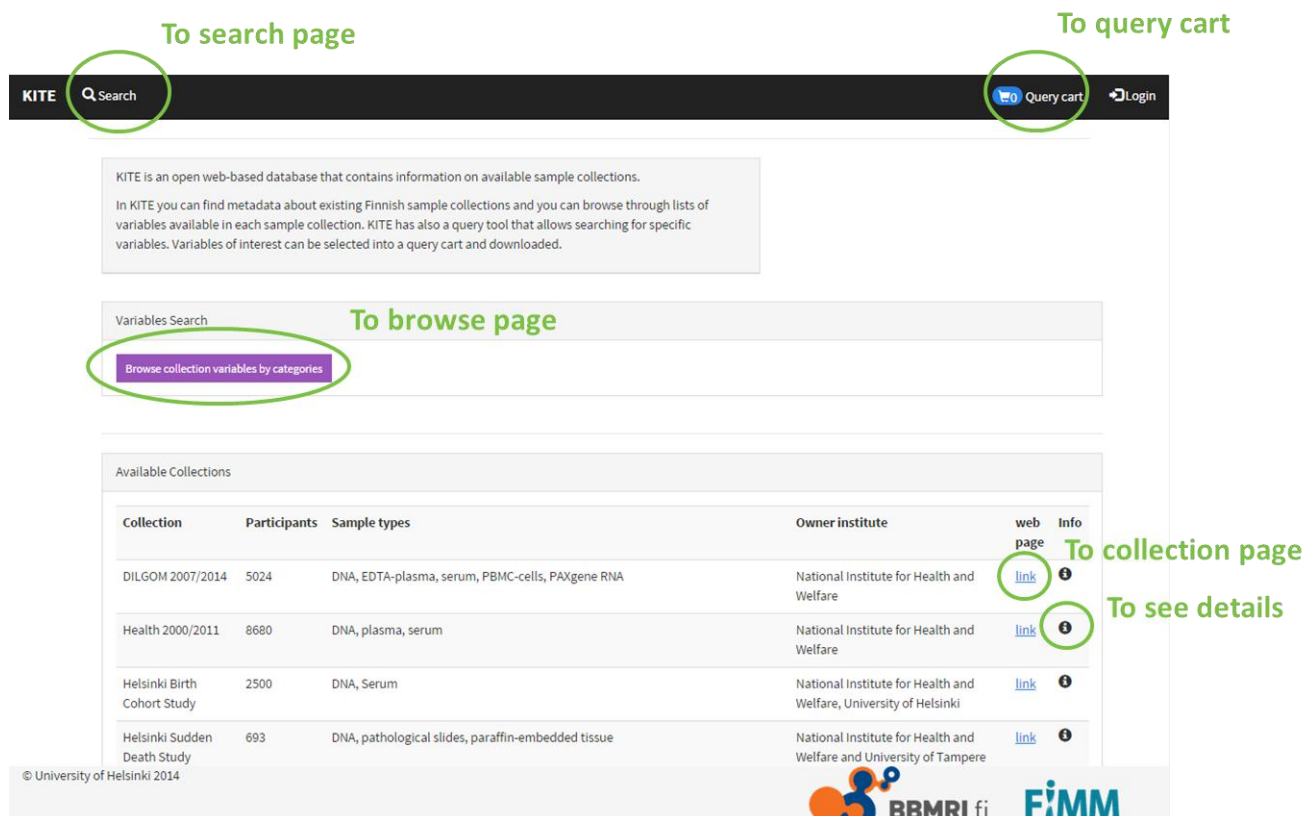
Available Collections

The front page has a table of sample collections showing the collection name, number of participants, sample types, owner institute and link to collection web page.

By clicking link under the 'web page' heading, it will open to a new window to collection's web page. By clicking ❶ you will find summarized details about the collection. The collections are described using MIABIS attributes (http://bbmri-wiki.wikidot.com/en:dataset-collection).

'Browse collection variables by categories' -button will lead to browse page (read more in **page 5**)

'Search' -menu button leads to search page (read more in **page 6**)

'Query cart' -menu button leads to query cart (read more in **page 11**) and also shows the number of variables in the query cart.

Picture 2.1. The KITE front page view and paths to other pages presented.
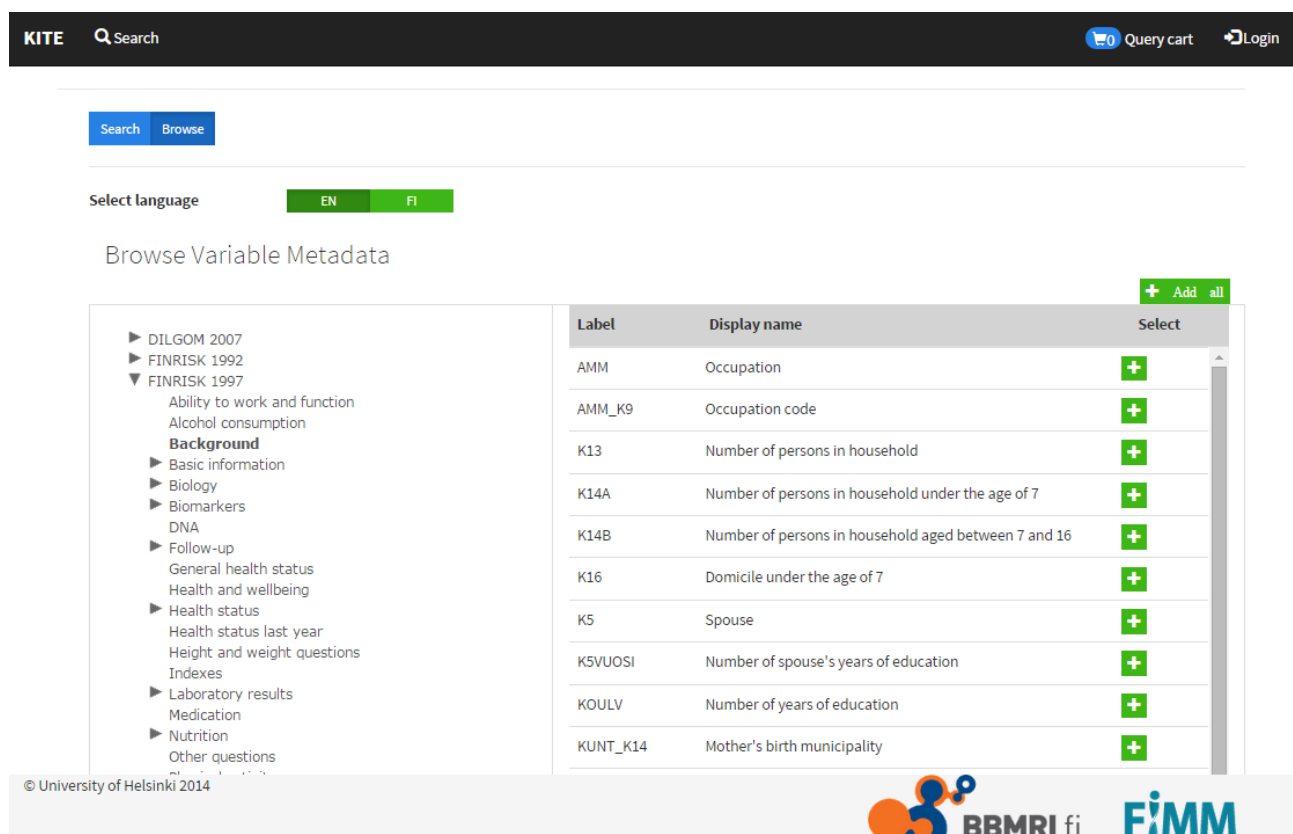
3. Browse page

Browse page is for browsing through lists of variables available in each sample collection. These variables are divided into categories (=topics). The user can see all the projects and the data categories on a list.

Click the arrow to open a collection tree and you will see different categories below. By selecting a category you will see related variables. You can add variables to query cart by clicking the green plus figure ➕ . It is also possible to select all the variables in a category by selecting ➕ Add all .

Language selection button is found above the list (currently variable data available in Finnish and English).

See more information about a variable by placing the mouse cursor on a variable row. An information box will show up.

Picture 3.1. KITE Browse page view. The collection trees on left and the variable table on right.

Definitions for Browse page

(Variable table)

Label: variable-specific original code in a given collection

Display name: a short description of the variable

Select: by clicking from the right side of the table, you can add variables to the query cart. You can also add all.

(Information box)

Description: detailed information about a variable, such as original question in the questionnaire form or detailed information about laboratory measurement.

Question nro: original question number in questionnaire

Protocol: measurement protocol

Enums: enumerations of class variables

Min: minimum possible value for variable

Max: maximum possible value for variable

Datatype: type of data such as integer, date, free text, etc.

UMLS: Unified Medical Language System (UMLS) Metathesaurus reference library code and description for the variable (https://uts.nlm.nih.gov//metathesaurus.html)

Keywords: search words related to the topic of the variables

4. Search page

The Search page offers two options for collection variable search:

- **Basic search** is a quick search with one simple field for search word/s

- **Options search** has more options to limit and control the search results

4.1. Basic search

Basic search is a quick search for finding variables. The user can enter search word/s and find variables matching the specific subject. This search will find variables from all the collections that match the search word/s.

Enter the search word to 'Search for' field.

Search results: See more information about a variable by placing the mouse cursor on a variable row. An information box will show up. Search summary shows terms found per each collection.

Language selection is above the search field (currently variable data available in Finnish and English).

TIP: Use basic search, if you want to search in which collection there is data about specific topic. For example, to find in which sample collection information on 'Smoking' is available.

TIP: If you want to view and download all the variables belonging to one collection (for example, FINRISK 2002) write the whole collection name ("FINRISK 2002") to the search field and add all variables to the car.

Picture 4.1. KITE Basic Search view.

Definitions for Basic search results

(Table of search results)

Collection: name of the collection

Category: Topic of the variable (for example, "Health status" includes variables that are related to health status topic)

Label: variable-specific original code in a given collection

Name: a short description of the variable

Protocol: measurement protocol

Select: by clicking from the right side of the table, you can add variables to the query cart. You can also add all.

(Information box)

Description: detailed information about a variable, such as original question in the questionnaire form or detailed information about laboratory measurement.

Question nro: original question number in questionnaire

Protocol: measurement protocol

Enums: enumerations of class variables

Min: minimum possible value for variable

Max: maximum possible value for variable

Datatype: type of data such as integer, date, free text, etc.

UMLS: Unified Medical Language System (UMLS) Metathesaurus reference library code and description for the variable (https://uts.nlm.nih.gov//metathesaurus.html)

Keywords: search words related to the topic of the variables

4.2. Options search

Options search gives wider options to find variables from specific collections or by using specific variable information fields, such as category, description, variable keywords and reference terms (UMLS) etc. With these options it is possible to limit and control the search results, and to target the search as needed.

Picture 4.2. KITE Options search view.

In 'Select collections for variable search' section you can choose collection/s which you want to include in the search. It is possible to select or deselect all the collections.

In 'Select variable information field for search' section you can choose from different information fields you want to include in the search. It is possible to select or deselect all the information fields.

Enter the search word/s to 'Search for' field.

Search results: See more information about a variable by placing the mouse cursor on a variable row. An information box will show up. Search summary shows variables found per each collection that answer the specific search criteria.

Language selection is above the search field (currently variable data available in Finnish and English).

TIP: If you want to find information about a specific variable for which you have the original variable code, select only 'Label' in the 'Select variable information fields for search' –box.

TIP: If you want to find similar variables from different collection, select the 'Display name' and 'UMLS' in the 'Select variable information fields for search' –box.

Definitions for Options search

(Variable information fields)

Category: search by topic

Description: use detailed variable description for search

Display name: search using short descriptions

Keywords: search includes related keywords and returns more variables

Label: search by variable-code

UMLS: search by reference term

Definitions for Options search results

(Table of search results)

Collection: name of the collection

Category: combining topic for variables (for example, "Health status" includes variables that are related to health status topic)

Label: variable-specific original code in a given collection

Name: a short description of the variable

Protocol: measurement protocol

UMLS: Unified Medical Language System (UMLS) reference library code and description for the variable (https://uts.nlm.nih.gov//metathesaurus.html)

Select: by clicking [+] from the right side of the table, you can add variables to the query cart. You can also add all.

(Information box)

Description: detailed information about a variable, such as original question in the questionnaire form or detailed information about laboratory measurement.

Question nro: original question number in questionnaire

Protocol: measurement protocol

Enums: enumerations of class variables

Min: minimum possible value for variable

Max: maximum possible value for variable

Datatype: type of data such as integer, date, free text, etc.

UMLS: Unified Medical Language System (UMLS) Metathesaurus reference library code and description for the variable (https://uts.nlm.nih.gov//metathesaurus.html)

Keywords: search words related to the topic of the variables

5. Query cart

Query cart lists all the variables selected from the search and browse results. Using the query cart, you can also export variable metadata from KITE as csv-file.

If the query cart does not have any variables selected, the 'Query cart is empty' text is visible.

If the query cart contains selected variables, you can view variables and remove variables that are not needed. The 'Clear' -button clears all the variables from your query cart.

Export the variable list by clicking the 'Export variable list (-csv)' -button.

6. Login

Login is not applicable yet.

Contact information

Email biopankki@thl.fi

National Institute for Health and Welfare, THL

THL Biobank