

1-1-2012

Statistical Insight: A Review

Emily Vardell

University of Miami Miller School of Medicine, evardell@med.miami.edu

Yanira Garcia-Barcena

University of Miami Miller School of Medicine, YGarcia6@med.miami.edu

Recommended Citation

Vardell, E., Garcia-Barcena, Y. (2012). Statistical Insight: A Review. *Medical Reference Services Quarterly*, Vol. 31(3), in press.

This Article is brought to you for free and open access by the Department of Health Informatics at Scholarly Repository. It has been accepted for inclusion in Faculty Research, Publications, and Presentations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.

ONLINE UPDATES: A COLUMN FOR SEARCH ANALYSTS

Emily Vardell, Column Editor

Statistical Insight: A Review

Emily Vardell

Yanira Garcia-Barcena

ABSTRACT. Statistical Insight is a database that offers the ability to search across multiple sources of data, including the federal government, private organizations, research centers, and international intergovernmental organizations in one search. Two sample searches on the same topic, a basic and an advanced, were conducted to evaluate the database.

KEYWORDS. Database evaluation, online databases, Statistical Insight, statistics

AUTHORS.

Emily Vardell, MLS (evardell@med.miami.edu) is Director for Reference and Education, Department of Health Informatics, Calder Memorial Library, Leonard M. Miller School of Medicine, University of Miami, P.O. Box 016950 (R-950), Miami, FL 33101.

Yanira (Jenny) Garcia-Barcena, MSLS (jennygarcia@miami.edu) is Senior Reference Librarian, Department of Health Informatics, Calder Memorial Library, Leonard M. Miller School of Medicine, University of Miami, P.O. Box 016950 (R-950), Miami, FL 33101.

Comments and suggestions should be sent to the Column Editor, Emily Vardell
(evardell@med.miami.edu).

BACKGROUND

Statistical Insight searches across statistical information from more than 200 federal agencies, all 50 state governments, 100 international intergovernmental organizations, and 500 private and academic sources. Statistical Insight digitally archives its content, ensuring access to publications after they are no longer available on agency websites. The documents include GIF images of all tables, PDF versions of the full text of statistical sources, as well as spreadsheets for approximately 20% of the tables, allowing for user manipulation of the data.

Statistical Insight is fee-based and comprised of different modules that must be subscribed to, some in addition to the database subscription. The Tables Collection™, which is required for Statistical Insight to function,¹ enables access to tables via table-level indexing. The optional Abstract Modules include the American Statistics Index, which indexes and abstracts federal government statistical data; the Statistical Reference Index, which includes statistical data from magazines, private organizations, state government agencies, businesses, independent research organizations, and other research centers; and the Index to International Statistics, covering statistical publications of international, intergovernmental organizations.

ProQuest Statistical Insight was formerly named ProQuest Statistical Universe. The official ProQuest/LexisNexis guidebook lists improvements to the new Insight database including: one-search solution, faceted search results (i.e., the data is filtered by topic), improved relevance ranking, unrestricted search size, user accounts, easy access to prior editions, batch export options, and more frequent updates (i.e., new data is loaded weekly rather than monthly). Statistical Universe users will note that the search interface itself has not changed much in the new Statistical Insight and, therefore, will not require additional training in order to use

Statistical Insight. Users who are interested in additional help can select “Wiki” in the upper right-hand corner of the database to access knowledge base articles, a glossary of statistical terms, and technical support.

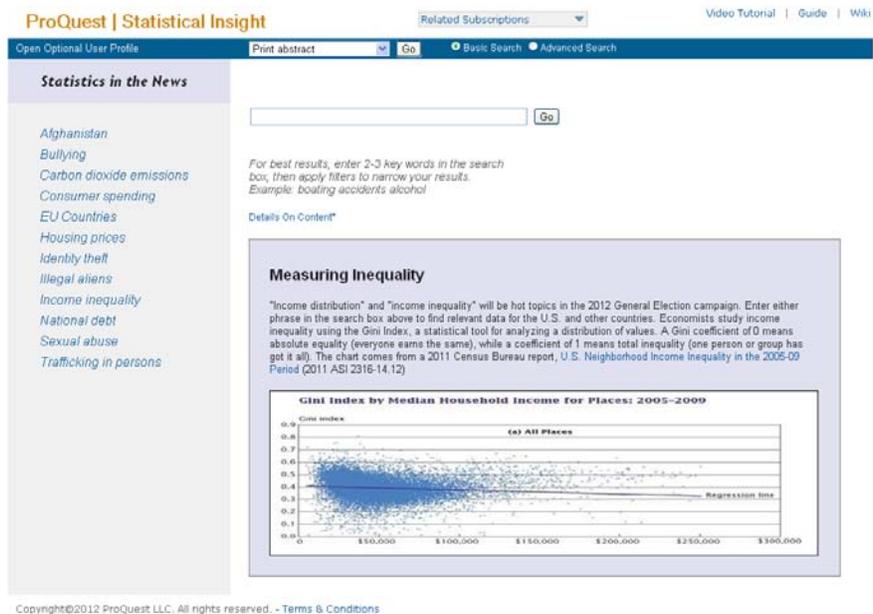
HOME PAGE

Statistical Insight has a simple interface with a single search box in the center of the screen and filtered elements in the left-hand column (see Figure 1). The major sections on the home page include:

- **(Open Optional) User Profile:** offering the ability to save searches and return to previous searches
- **Output (labeled “Print abstract”):** drop-down box in the top center of the screen allows users to print, e-mail, or export records
- **Related subscriptions:** drop-down box above the output option provides access to other LexisNexis/ProQuest subscriptions
- **Guide and Wiki:** links in the upper right-hand corner with additional information on the database
- **Basic Search/Advanced Search:** users can toggle between the two search interfaces

To help users find relevant data more quickly, Statistical Insight recently added the “Statistics in the News” section on the left hand side of the home page. This section highlights statistics that may be of highest interest to users. At the time of publication, the topics included statistics on Afghanistan, bullying, carbon dioxide emissions, housing prices, identity theft, sexual abuse, and trafficking in persons. These highlighted topics also demonstrate the wide variety of statistics

available in Statistical Insight. In addition, the home page features a selected graph on a popular topic; at the time of publication the featured graph was on measuring inequality in income distribution (see Figure 1 for details).



Legend: FIGURE 1. Statistical Insight home page

BASIC SEARCH

A sample search was conducted to obtain worldwide statistics on multidrug-resistant tuberculosis, with incidence and prevalence rates divided by country, sex, race, age, and occupations. To start with a basic search, the default “Basic Search” radio button in the top blue bar was left selected and *tuberculosis multidrug resistan** was entered in the search box. Statistical Insight allows for truncation, or adding a star at the end of a term, to retrieve multiple endings (e.g., *resistant* as well as *resistance*).

Users will note that when typing search terms, auto-complete or subject/index terms appear, allowing them the ability to re-define query terms. In the case of multidrug-resistant tuberculosis, neither feature was utilized because there are no phrases or subject terms that could be applied specifically to multidrug-resistant tuberculosis.

Additionally, as with a Google search, there is no need to type AND between search terms, unless a Boolean statement is used. In this case, *tuberculosis multidrug resistan** retrieves the same results as *tuberculosis AND multidrug AND resistan**. Also similar to a Google search, users can enclose exact phrases in quotation marks to get the most focused results. In this sample search, the three key terms were not enclosed as an exact phrase to keep a broader search (see Figure 2).

The screenshot displays the ProQuest Statistical Insight search interface. At the top, the search bar contains the query "tuberculosis multidrug resistan*" and a "Go" button. Below the search bar, it indicates "13 Results" and offers sorting options: "Relevance" (selected) and "Date Published".

The left sidebar shows "Applied Filters: None" and a list of filter categories including Date Published, Date Covered, Full Text Type, Source, Countries & Regions, Local Area, Data Breakdowns By..., and Subject. The Subject filter is expanded to show "Tuberculosis (13)", "Drugs (9)", and "Acquired immune deficiency syndrome (7)".

The main content area lists four search results:

- 1. Surveillance of Anti-Tuberculosis Drug Resistance in the World: An Updated Analysis, 2007-10** (p. 111-119D) Published: 2012, Source: World Health Organization (United Nations), Record Number: 2012 IIS 4640-P1.5.2
Tuberculosis (multidrug - resistant) cases and trends, by drug and selected country or WHO world region, 2000s-10, WHO jmi article
- 2. Antituberculosis Drug Resistance Worldwide** Published: 1998, Source: World Health Organization (United Nations), Record Number: 1998 IIS 4640-P3.27.1
Drug (multidrug) resistance prevalence, for tuberculosis, by selected country, 1990s-97, WHO ipt article
- 3. Directly Observed Treatment, Short-course Strategy and Multidrug-Resistant Tuberculosis: Are any Modifications Required?** Published: 2000, Source: World Health Organization (United Nations), Record Number: 2000 IIS 4640-P1.5.5
Multidrug resistant tuberculosis prevalence, by selected E. Euro country/former Soviet oblast, 1990s-1997/98; and cost of selected treatment regimens, by drug; 2000 WHO jmi article
- 4. Barriers to Successful Tuberculosis Treatment in Tomsk, Russian Federation: Non-Adherence, Default and the Acquisition of Multidrug Resistance** (p. 703-711) Published: 2008, Source: World Health Organization (United Nations), Record Number: 2008 IIS 4640-P1.5.4
Russia tuberculosis treatment barriers in Tomsk, determinants, regression analysis, 2001, WHO jmi article

Each result includes a small icon for "ABSTRACT" and "PDF DOCUMENT".

Copyright©2012 ProQuest LLC. All rights reserved. - Terms & Conditions

Legend: FIGURE 2. Statistical Insight basic search

The results documents can be sorted by relevance or date of publication by selecting the appropriate radio button beneath the search box. The basic *tuberculosis multidrug resistan** search returned 13 results, displayed as a document list. Clicking on a particular title displays the ProQuest Summary with information about the report, including title, date, source, ISBN, key terms included in tables (e.g., TB incidence, prevalence, multi-drug resistant cases, detection rates, etc.), index terms from the thesaurus, listing of countries covered in report, and more (see Figure 3).

The screenshot shows the ProQuest Statistical Insight interface. At the top, there is a search bar with the query "(tuberculosis multidrug resistan*) AND (incidence OR preva)" and a "Go" button. Below the search bar, it indicates "1 Results" and offers sorting options: "Relevance" (selected) and "Date Published".

The search results list includes one entry:

- 1. **Global Tuberculosis Control: Surveillance, Planning, Financing. WHO Report 2008** (304 p.) Published: 2008, Source: World Health Organization (United Nations), Record Number: 2008 IIS 4640-S17
Tuberculosis situation and control strategies, with incidence and prevalence, by age and sex; sputum smear detection; treatment outcomes; HIV concurrence; and costs, by country/region, 1980s-2008, WHO annual rpt

Below the search results, there is a "PROQUEST SUMMARY" window. The summary provides the following details:

- Report Title: Global Tuberculosis Control: Surveillance, Planning, Financing. WHO Report 2008
- Date: 2008
- Source: World Health Organization (United Nations)
- Length: 304 p.
- Periodicity: Annual
- Record Number: 2008 IIS 4640-S17
- ISBN: 978 92 4 156354 3
- Available on IIS microfiche

 The summary also includes sections for "Data Sources" (National government responses to WHO questionnaire) and "Format and Data Presentation" (Contents listing, acknowledgments, and abbreviations. (p. iii-viii); Report, with maps, charts, and 25 tables. (p. 1-78); Annex 1, with text statistics, 1 map, charts, and 1 table repeated for 22 high-burden countries. (p. 79-169)).

At the bottom of the page, there is a copyright notice: "Copyright©2012 ProQuest LLC. All rights reserved. - Terms & Conditions".

Legend: FIGURE 3. Statistical Insight filtered basic search

From the results obtained, users can choose to view them in various formats, including PDF documents, abstracts, tables, or earlier editions of documents, depending on what formats are available for the particular report. Users have one-click access to these formats by clicking on the icon of the desired format. If a user selects the “tables” format, the option exists to view only those specific “tables that match search,” or to view “all tables” available for the particular document. The full table can be viewed by clicking XLS or GIF, depending on the desired table format (XLS downloads the chart, whereas a GIF of the chart will open in a new tab).

To add multiple dimensions to a broad search, results can be limited by factors including country, sex, race, age, and occupations. Once a search is conducted, the filters column automatically appears on the left-hand side of the database (in place of the “Statistics in the News” section). In the Applied Filters column, users can filter by “Countries & Regions.” For this search, the plus sign next to “Worldwide” was selected, which displayed the option to filter “By Country,” filtering down to four results which present the statistics by country.

To limit the initial 13 results by sex, race, age, and occupations, the plus sign next to “Data Breakdowns By...” was selected. Then each section was individually selected, “By Age,” “By occupation,” “By race and ethnic group,” and “By sex,” retrieving the same three documents each. Users may also filter by publication date by selecting “Date Published” at the top of the Applied Filters column, and then sliding the “Date Published Range Selector” to include the desired publication years. The same can be done for coverage years by selecting the “Date Covered” option next to “Date Published.”

The initial search was additionally refined by adding Boolean statements in the strategy: *(tuberculosis multidrug resistan*) AND (incidence OR prevalence OR rate)*, returning seven results. Statistical Insight uses “stemming,” which automatically retrieves plurals (i.e., searching

“rate” will include “rate” and “rates”). The search was narrowed further by viewing the “Subject” section under Applied Filters. This subsection displays how many documents contain specific key terms that could be relevant to a given topic and can be viewed alphabetically or by frequency. Two filters were applied to the search; first “Age-sex distribution” was selected, retrieving two documents. The Subject section displayed a new list of topics specific to this filtered subject. Finally, “Geographic rankings” was selected, resulting in one document covering both age-sex distribution and geographic rankings of multidrug-resistant tuberculosis incidence (see Figure 3). The selected filters can be removed by selecting the red X icon in the upper left-hand corner of the Applied Filters column.

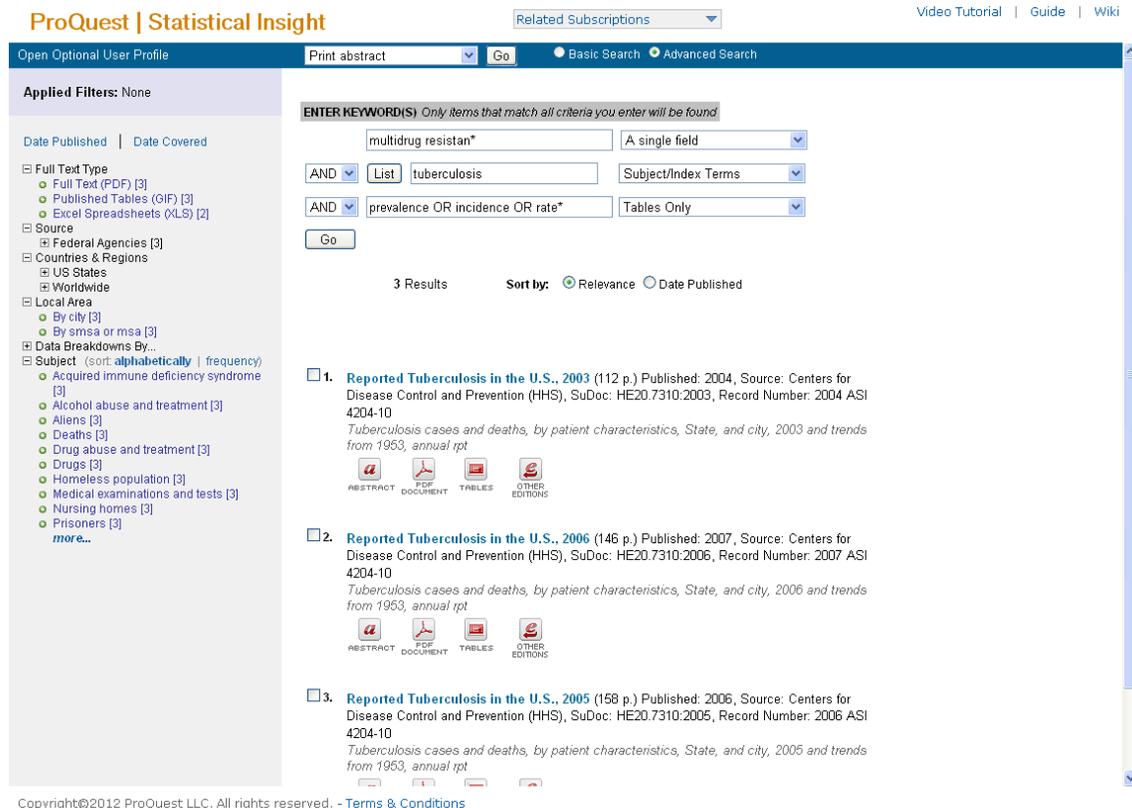
Documents of interest can be printed, e-mailed, and exported. Users must simply select the check box next to the desired record(s) and then select the preferred option from the drop-down menu at the top center of the screen. Options include printing the abstract (which prints the ProQuest Summary) or citations, e-mailing search results (which e-mails a link to a list of the citations), and/or exporting to RefWorks or EndNote. These options are available with or without an individual user account. Printing is accessible from within an abstract, PDF document, or table.

ADVANCED SEARCH

To toggle from the basic to the advanced search, a user can select the Advanced Search radio button at the top of the screen. The Advanced Search options can be used to craft a more sophisticated search. Several approaches with specific strengths are available through the drop-down menus next to the three search boxes on the Advanced Search page. The options include

searching by a single field, all fields except Full Text, all fields including Full Text, document title, table title, table only, source/publisher, document number, and subject/index terms.

For the first sample advanced search, *multidrug resistan** was entered in the first search box and “A single field” was selected from the drop-down field box. In the second search box, *tuberculosis* was entered and “Subject/Index Terms” was selected from the drop-down box. Finally, *prevalence OR incidence OR rate** was entered in the third search box, and “Tables Only” was selected. If “Subject/Index Terms” is selected from the drop-down box, Statistical Insight offers the option to select “List,” which will display a list of the databases’ subject terms organized alphabetically. Users will also note that “AND” is the default Boolean operator for combining search terms in the Advanced Search screen (see Figure 4).



Legend: FIGURE 4. Statistical Insight advanced search

This search retrieved three results, displayed as described in the Basic search. Searching “A Single field,” as was done with *multidrug resistan**, searches for the phrase in the title of documents, the paragraphs of abstracts, notation of contents, and in the description or record of tables. This approach will typically retrieve the phrases/word combinations (e.g., multidrug resistance, prevalence rates, tuberculosis infection) in the same data field. The “Subject/Index Terms” search allows for a very focused search with targeted results, because the database identifies the main statistical subject in every document, retrieving only those documents indexed to one of the subject terms in ProQuest’s index thesaurus.

WHY STATISTICAL INSIGHT?

Statistical Insight offers a one-stop solution for those interested in searching for statistics. Because it combines statistics from the United States federal government, as well as private organizations, research centers, and international intergovernmental organizations, it can be argued that there is no other database that offers the same ability. Users interested in federal government health statistics may refer to FastStats <<http://www.cdc.gov/nchs/fastats/>>, and those interested in international intergovernmental organizations may rely on the statistics available in databases from the World Health Organization, including the Global Health Atlas <<http://apps.who.int/globalatlas/>> or the Global Health Observatory Data Repository <<http://apps.who.int/ghodata/>>. However, Statistical Insight appears to be the only resource that combines multiple types of data sources into a single search engine. In the time of reliance on easy, one-stop searching, Statistical Insight offers a service that researchers and students may find quite appealing.

FOR MORE INFORMATION

For more information on Statistical Insight, please direct queries to:

ProQuest Sales Representatives

Tel : +1 800 521 0600

E-mail: academicinfo@proquest.com

REFERENCES

1. LexisNexis. "LexisNexis Statistical Insight Guidebook." Available:

<http://www.lnstatistical.com/help/statistical/LN_guidebook.pdf>. Accessed: March 27, 2012.