Data Discovery and Data Use Toolkit

This toolkit provides users an opportunity to explore the Iowa Public Health Tracking (IPHT) Portal and learn about the data resources included. These resources will assist in accurate interpretation and appropriate use of the data and measures available.

The resources focused on include the Content Landing Pages, Indicator Profiles, and Metadata. A FAQ for each of these resources is included in the packet. The exercises also make use of the IPHT User Manual as an important reference document for users.

The included exercises are intended to guide new users in finding and using the data and information available on the IPHT Portal. The exercises are structured to walk the user through the content and data included on the portal to better acclimate users.

The IPHT Portal is online at https://pht.idph.state.ia.us

Questions about this toolkit or about the IPHT Portal should be sent to EPHT@idph.iowa.gov.

This toolkit was originally developed for use at the 2013 University of Iowa, College of Public Health Colloquium.

Content Landing Pages – Frequently Asked Questions

1. **What are the Content Landing Pages?**
   The Iowa Public Health Tracking (IPHT) Portal includes data from a range of environmental and health topic areas. Each topic area on the portal includes a Content Landing Page that provides background information. The information includes what measures are available on the portal, the relationship between health and environment, exposure, risk, and prevention.

2. **Why are the Content Landing Pages important to IPHT?**
   The Content Landing Pages are important because it provides an overview of what data is available on the portal. It helps users understand why each content area has been included, and provides a platform for actionable messaging to prevent or reduce risk. The Content Landing Pages also include links to the related Indicator Profiles, Dashboards, and external resources for additional information.

3. **How do I access Content Landing Pages on the IPHT Portals?**
   The landing pages can be accessed through the 'Environment' and 'Health Effects' tabs on the Navigation Bar on the IPHT Portal.

Indicator Profiles – Frequently Asked Questions

1. **What are NCDM Indicators?**
The Iowa Public Health Tracking (IPHT) Portal includes a core set of nationally consistent data and measures (NCDM) concerning health, exposures, and environmental hazards. The purpose of these NCDMs was to ensure compatibility and comparability of data and measures useful for understanding the impact of our environment on our health.

2. **How were the NCDMs developed?**
NCDMs are the result of collaboration with partners and data stewards at the national, state, and local levels as part of the National Environmental Public Health Tracking Network (EPHTN). To determine data needs for the EPHTN, a group of experts, including data stewards, evaluated data to determine whether they were suitable for environmental public health tracking. Critical elements from these data contributed to the development of recommendations for nationally consistent EPHTN data. The groups evaluated the quality and completeness of existing data resources, research and evidence on linkages between environmental hazards and health outcomes, and the availability of standardized measures used to:
   - Quantify the magnitude of a public health problem
   - Detect unusual trends in health, exposures, and hazards
   - Identify populations at risk of environmentally related diseases or of exposure to hazards
   - Generate hypotheses about the relationship between health and the environment
   - Direct and evaluate control and prevention measures and individual actions
   - Facilitate policy development

3. **What are Indicator Profiles?**
An indicator profile is the result of the work to develop NCDMs. The indicator profile provides basic information about the indicator including:
   - Definitions of the measures
   - Derivations of the measures
   - Geographic scope and scale
   - Time period and scale
   - Rationale for use of the measure
   - Use of the measure
   - Limitations of the measure
   - Data sources
   - Limitations of data sources
   - References

4. **Why is the Indicator Profile important to IPHT?**
   Indicator Profiles are important because it provides descriptive information about measures available on the Network. Documentation of the measures allows users to understand and interpret the data presented on the IPHT Portals. It helps a user to decide if a measure is appropriate for the intended use.

5. **How do I access Indicator Profiles on the IPHT Portals?**
   Indicator Profiles are available as linked PDF documents on the landing page for each content area. The landing pages can be accessed through the 'Environment' and 'Health Effects' tabs on the Navigation Bar on the IPHT Portal.
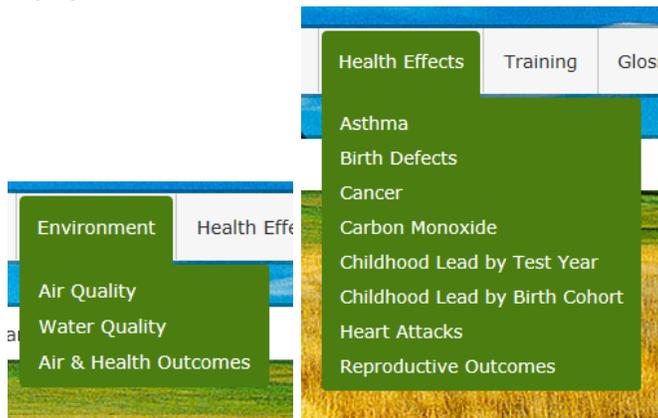
Asthma Measures on the Tracking Network
Hospitalizations for Asthma

This indicator uses data collected by hospitals. It can be used to identify trends and patterns in the occurrence of asthma hospitalizations across time and space. The data are organized by different variables to help estimate the number of asthma hospital admissions in different time periods, age groups, and geographic areas, such as states and counties. Asthma hospital admissions tend to be for more severe asthma attacks and do not include asthma among individuals who do not receive medical care, who are not hospitalized, or who are treated in outpatient settings. Differences between geographic areas may be the result of differences in the underlying population or in the diagnostic or coding techniques used by the reporting hospital.

📄 Hospitalizations for Asthma.pdf

Asthma Data on the Tracking Network

Metadata – Frequently Asked Questions

1. **What is Metadata?**
   Metadata is commonly referred to as "data about data". There are several different types of metadata. These can be broadly defined under the categories of Descriptive, Structural, and Administrative.

   - **Descriptive Metadata**: Information that describes the content, quality, and context of a data resource for the purpose of facilitating identification and discovery. It may reference additional information like quality assurance documents and data dictionaries. Through descriptive metadata a user can learn the what, why, when, who, where, and how for a data resource.

   - **Structural Metadata**: Information about how the item is put together or arranged such as the table of contents page, individual page numbers, or illustration. It basically describes the structure of an item, such as a book, so that all of the pages of that item can be displayed in the correct order. In the electronic world it facilitates navigation and presentation of electronic resources.

   - **Administrative Metadata**: Includes information about resolution, bit depth, type of equipment used to produce the file, storage format, and file name and location. It can also include basic facts on ownership, rights, and reproduction information.

   The Iowa Public Health Tracking Network (IPHT) makes use of Descriptive Metadata.

2. **What are the benefits for creating Descriptive Metadata?**
   As more data are being created and stored, there is a need to document data resources for future use and improve accessibility. The benefits of Descriptive Metadata for users include:

   - Accessibility - Provides the availability of data resources via online services.
   - Data Discovery - Provides information about an organization's data holdings so users can locate available resources relevant to an area of interest or study.
   - Appropriate Use - Supplies the means to document limitations about the data resource or disclaimers that are important for potential users to be aware of.

3. **Why is Descriptive Metadata important to the IPHT?**
   Descriptive Metadata is important for two key aspects of network functionality.

- It allows for the discovery of data resources on the Network. Network users can locate resources through a variety of means including keywords, geographic boundaries, and date and time. All of these elements are part of the Descriptive Metadata entry.

- It is also important because it provides descriptive information about data resources available on the Network. Through Descriptive Metadata, a Network user can determine the content of the resource, why it was created, how it was created, any limitations, access and use restrictions, data quality, and contact information. It helps a user to decide if a resource is appropriate for the intended use.

4. **Are there standards for the creation and maintenance of Descriptive Metadata?**
Yes, there are several standards available for the creation and maintenance of Descriptive Metadata. The National Environmental Public Health Tracking Network (EPHTN) reviewed and selected the Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata was the most appropriate and adaptable for use within the EPHTN. Developed in 1994 (revised in 1998) to describe geospatial data, the standard is unique in that it contains elements from all three metadata types (Descriptive, Structural, and Administrative). As an EPHTN participant the Iowa Department of Public Health (IDPH) has adopted these standards for use throughout IPHT.

5. **What is the EPHTN Metadata Profile?**
The EPHTN Metadata Profile outlines the format and content for describing data resources on the Network. The profile complies with the FGDC Content Standard for Digital Geospatial Metadata. It contains 52 of the 195 total data elements available for completion in the FGDC Standard. This represents the minimum description necessary for providing data to the Network. The profile consists of a template and guidance to assist in the completion of each element. It was developed to make the creation of Descriptive Metadata easier for data stewards. IPHT used the EPHTN profile in developing tools to manage and display metadata on the IPHT Portals.

6. **How do I access Descriptive Metadata on the IPHT Portals?**
Descriptive Metadata for those data resources available on the IPHT Portals can be accessed through the 'Metadata' tab on the Navigation Bar on the IPHT Portal.
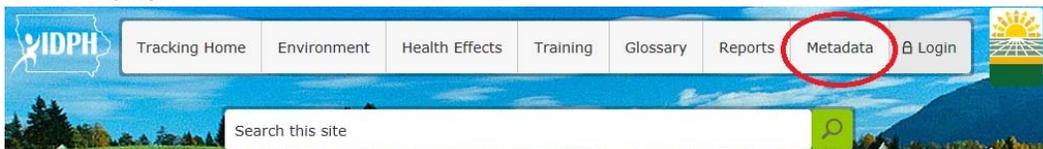
Section 8 of the Tracking Portal User Manual has additional details on how and where to access Metadata on the IPHT Portals.

The purpose of this series of exercises is to guide new users in finding and using the data and information available on the IPHT Portal. The exercises are structured to walk the user through the content and data included on the portal to better acclimate users.

Data Exercise 1: Air Quality Data Discovery

Scenario: You are interested in the potential exposures to high levels of Ozone in 2005.

Goals:

- Learn about the Ozone measures available on the IPHT portal
- Use the Metadata and Indicator Profiles to better understand the Air Quality data
- Modify a Dashboard Scorecard to facilitate data use

1. What are two ways to find Ozone air quality information on the portal from the IPHT home page?

2. What are the Air Quality measures related to Ozone available on the portal?

The IPHT portal includes both Monitored and Modeled data. Use the Indicator Profiles for both data sets to answer the following questions.

3. What are the years of Data available for the Modeled Air Quality Data?

4. Would it be appropriate to use the Monitored Air Quality data to summarize the level of exposure to Ozone at the county level throughout Iowa?   Why or Why not?

The next set of questions can be answered using the scorecard on Air Quality Monitored Data dashboard.

5. How many counties had days in exceedance of the regulatory standard for Ozone in the year 2005? (Use the filters on the monitored data dashboard to show data for 2005 in all monitored counties)

6. Which County or Counties had the most days in exceedance of the regulatory standard for Ozone? (Right click on the "Ozone # of days" column and sort it largest to smallest)

7. Which 2 counties had larger populations at risk of exposure to Ozone levels in exceedance of the regulatory standard than Clinton county in 2005?   Why?
(Right click on the "Ozone Person days" column and sort it largest to smallest)

The purpose of this series of exercises is to guide new users in finding and using the data and information available on the IPHT Portal. The exercises are structured to walk the user through the content and data included on the portal to better acclimate users.

Data Exercise 2: Asthma Data Use

Scenario: You are interested in the burden of Asthma in Iowa in 2010.

Goals:

- Learn about the Asthma measures available on the IPHT portal
- Use the Metadata and Indicator Profiles to better understand the Asthma Hospitalization data
- Modify a Dashboard Chart to drill down into data

For this exercise we will begin on the Asthma State Data dashboard, which can be found through the Top Navigation bar under Health Effects > Asthma > State Data

1. How many Hospitalizations for Asthma occurred in Iowa in 2010? (Use the filters on the Asthma State data dashboard to show data for 2010)

2. Was this an increase or decrease from 2009? What is the percent change? (Each cell of a scorecard includes three values: the measure value, an arrow indicating change in the measure from previous year, and the percent change in the value from the previous year)

The next set of questions can be answered using the Crude Rate by Age Group and Admission Count by Age Group bar graphs on the Asthma State Data dashboard.

3. Which age group had the highest number of hospitalizations for asthma?

4. Which age group had the highest crude rate of hospitalizations for asthma?

Many charts on the IPHT portal can be modified by the user to dig deeper into the data. The next question can be answered by modifying the Admission Count by Age Group bar graph on the Asthma State Data dashboard.

5. Among 15 to 34 year olds; did males or females have a higher number of hospitalizations for asthma? (Right click on the bar representing the 15 -34 age group, hover over the "Drill Down to" item to see the pop out menu and select gender to modify the bar chart)

6. How many Males ages 0-4 and How many Females ages 0-4 were hospitalized for Asthma? (Right click on the background of the bar chart, hover over the "Report Type" item to see the pop out menu and select "Grid" to change the bar chart into a data table)

The next set of questions can be answered using the Iowa Asthma Hospitalizations metadata. A link to this metadata can be found in the Metadata Library under the item Metadata on the main navigation menu.

7. According to the metadata, this dataset does not include data from federal hospitals, non-acute care hospitals, and patients treated outside of their state of residence. What impact does this have on measures reported on IPHT?

8. If were you were interested in obtaining a custom dataset for further research on Asthma Hospitalizations in Iowa, how would you make a request?

The purpose of this series of exercises is to guide new users in finding and using the data and information available on the IPHT Portal. The exercises are structured to walk the user through the content and data included on the portal to better acclimate users.

Data Exercise 3: Cancer Data Use

Scenario: You are interested in the cancer rates in Iowa.

Goals:

- Learn about the different ways Cancer measures are presented on the IPHT portal
- Drive Dashboard Charts by select a row in the scorecard
- Use the Metadata and Indicator Profiles to better understand the Cancer data

For this exercise we will begin by using the User Guide linked from the How Do I page under Training in the Main Navigation bar. Section 3.6 is the section related to Cancer, but make sure to read section 2.12 as well.

1. County level Cancer data is presented as 5 year aggregate measures, but the dashboard filters only display single year values. For five year aggregate measures what does the year selected in the dashboard filter represent? (Section 2.12 of the user manual talks about the year filter)

2. Cancer site data is organized in a hierarchical structure that groups individual cancer sites by anatomic or functional organ systems. If we wanted to view Gallbladder Cancer we would have to expand which top level group in the Cancer Site Filter? (Section 3.6 of the User Manual includes a Note about the Cancer Site Filter)

The next question can be answered using the Counties by Cancer Site dashboard, which can be found through the Top Navigation bar under Health Effects > Cancer > Counties by Cancer Site

3. What county has the highest 5 Year average Age Adjusted Rate of Lung and Bronchus Cancer for the year 2000 -2004? (Use the filters on the Counties by Cancer Site dashboard to select Year = 2004, Counties = Select All, Cancer Site = Lung and Bronchus. Once applied a scorecard column can be sorted by right clicking on a column header.)

On dashboards where the scorecard can have multiple rows, the charts on that dashboard are filtered by selecting a row on the scorecard. Try clicking on a couple different rows in the scorecard and see how the Age Adjusted Rate over Time chart updates to display the trend for the county selected.

The next question can be answered using the Cancer Sites by County dashboard, which can be found in the sibling dashboard links above the filters on any Cancer Dashboard.

4. What is the 5 Year Annual Average number of cases of Leukemia in Polk County from 2006 – 2010? What is the 5 Year Annual Average number of cases of Other Lymphocytic Leukemia? (Use the filters on the Counties by Cancer Site dashboard to select Year = 2010, County = Polk. Notice how on this dashboard the rows of the scorecard show the different cancer sites. Use the arrows to expand and collapse cancer site groups.)

5. What is the 5 Year Annual Average number of cases of Other Lymphocytic Leukemia in Polk County for 2006-2010?

6. How many actual cases of Other Lymphocytic Leukemia occurred in Polk County from 2006-2010?

The next question can be answered using the Iowa Cancer metadata. A link to this metadata can be found in the Metadata Library under the item Metadata on the main navigation menu.

7. Where did IDPH obtain the Cancer data used on the IPHT Portal?

The purpose of this series of exercises is to guide new users in finding and using the data and information available on the IPHT Portal. The exercises are structured to walk the user through the content and data included on the portal to better acclimate users.

Data Exercise 4: Childhood Lead Poisoning Data Discovery

Scenario: You are interested in the risk of a child under 3 years of age being Lead Poisoned in Iowa.

Goals:

- Learn about the Childhood Lead Poisoning measures available on the IPHT portal
- Use the Metadata and Indicator Profiles to understand the differences between Cohort and Annual data
- Read charts that overlay health outcome and socioeconomic datasets

These first two questions can be answered using the information published on the content landing pages of the Childhood Lead Poisoning sections of the IPHT portal.

1. Identify and briefly describe the two ways Childhood Lead Poisoning data are available on the portal?

2. What are two risk factors for exposure to Lead in the environment?

The next questions can be answered using the indicator profiles for either Blood Lead Level by Test Year or Blood Lead Level by Birth Cohort.

3. When is a Blood Lead Level (BLL) considered to be a confirmed elevated result?

4. Iowa has a universal testing requirement that all children must be tested for elevated BLL before age 6; how could this requirement affect comparisons to childhood lead poisoning data from other states?

The next set of questions can be answered using the Childhood Lead Poisoning by Birth Cohort County data dashboard.

5. How many children were born in 2005 and tested before age 3 in Polk County? What percent had a confirmed elevated BLL? (Use the filters to select Year = 2005, Counties = Polk, Age Group = Under 3)

6. Scroll down the page to the Children Tested by Year chart. What was the Cohort Tested Percent in 2005 for children under age 3 in Polk County (Don't forget to select Polk county row in the scorecard to drive the charts. Values will appear on the chart when hovering over the point, or Right click on the background of the bar chart, hover over the "Report Type" item to see the pop out menu and select "Grid" to change the bar chart into a data table.)

The next set of questions can be answered using the Childhood Lead Poisoning by Test Year County data dashboard.

7. How many children under age 3 were tested in 2005 in Polk County? What percent had a confirmed elevated BLL? (Use the filters to select Year = 2005, Counties = Polk, Age Group = Under 3)

8. Using the Children Tested by Year chart, what was the Percent of children under 5 living in poverty in Polk County? (Don't forget to select Polk county row in the scorecard to drive the charts. Values will appear on the chart when hovering over the point, or Right click on the background of the bar chart, hover over the "Report Type" item to see the pop out menu and select "Grid" to change the bar chart into a data table.)

9. Why is this Indicator value the same for all years? (This indicator is fully described in the Age of Housing and Child Poverty Indicator profile, linked on the Content landing page for Childhood Lead Poisoning by Test Year.)

The purpose of this series of exercises is to guide new users in finding and using the data and information available on the IPHT Portal. The exercises are structured to walk the user through the content and data included on the portal to better acclimate users.

Data Exercise 5: Reproductive Health Outcomes Data Use

Scenario: You need data for a report comparing Infant Mortality rates in Linn and Johnson County.

Goals:

- Learn about the Infant Mortality measures available on the IPHT portal
- Use the Metadata and Indicator Profiles to understand how Infant Mortality is measured
- Export tables and charts from a dashboard to use in a report

These first questions can be answered using the Infant Mortality Indicator Profile available on the Reproductive Outcomes section of the IPHT portal.

1. Identify the four measures of Infant Mortality available on the portal, and the age at death range each uses?

2. How can these measures be used?

The next questions can be answered using the Reproductive Outcomes County 5 Year Data dashboard.

Use the Filters on the Dashboard to have the scorecard show the measures for Linn and Johnson Counties in 2011.

3. What is the year range of data represented in the values on the Scorecard? (Section 2.12 of the user manual includes a note about the Year Filter.)

4. Export the Scorecard to both Excel and Powerpoint by clicking on the Triangle in the upper right of the scorecard to access the Web Part Menu.

Scorecards with more than 15 counties included do not export well to Powerpoint because of the slide size limitations. It is typically best to export scorecards to Excel which allows you to modify the table or copy and paste values into other table formats for use.

5. Use the Filters on the Dashboard to change the year displayed to 2010. Export this scorecard to Excel.

Exporting separate scorecards covering multiple years of data to Excel allows users to extract the measure values and create a new table with multiple years of data for use in a report.

6. Back on the dashboard, scroll down to the 5 Year Average Mortality Rate chart, is the Infant Mortality rate more stable over time in Linn County or Johnson County? Briefly explain the reason you chose your answer? (Remember to change which county the chart shows data for by selecting a county row in the scorecard.)

7. Export the chart for Linn County into Excel and the chart for Johnson County into Powerpoint by clicking on the Triangle in the upper right of the Chart to access the Web Part Menu.

Choosing which format to export a chart can depend on how you intend to use the chart. Powerpoint is good if you want to insert the chart directly into a presentation, while using Excel gives more flexibility for later use of the chart in other documents or reports.