

BCCDC Informatics Activities



BC Centre for Disease Control

An agency of the Provincial Health Services Authority



Environmental Health Surveillance Workshop

February 26, 2013



Public Health Informatics

- Application of key disciplines to Public Health
 - information science
 - computer science
 - cognitive or social sciences
 - supplemented as needed by other domains including mathematics, engineering, biostatistics, and GIS
- A multidisciplinary approach to solving complex informational, analytical, and technical problems, guided by scientific principles and applied systematically.



Data Warehouse



BCCDC Context

- 2007 MOU identifies data linkage as key BCCDC function
- 2008 BCCDC Surveillance Workshop identifies centralization of data as priority theme
 - Common data standards
 - Centralized data cleaning
 - Linkage of all data sets within BCCDC



Public Health Reporting Data Warehouse

- Automated and standardized:
 - Data integration
 - Patient-matching; case-matching
 - Geo-coding (health boundaries and derivation of address at time of event)
 - Episode date derivation
- Patient- and specimen-centric data views
- Data provided to end users anonymized or audited



Principles

- Analyses run on real-time data to support surveillance
- Data source cooperation
 - Joint access to marts
- Improved Privacy
 - De-identified data in marts; line lists audited
 - Copies of data are minimized
 - Analyses exported, but data is not

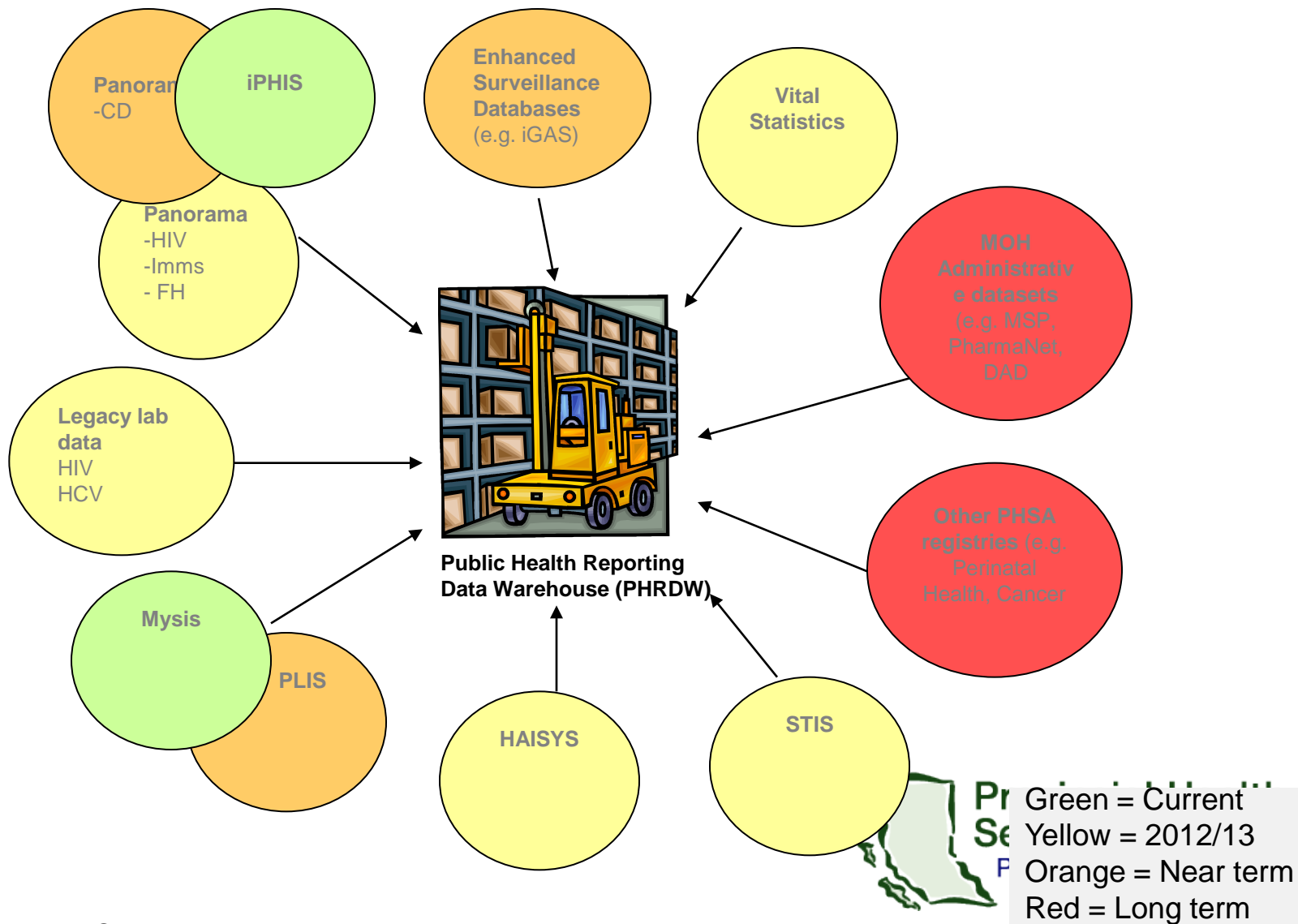


Advantages

- Time savings from automated data reconciliation
- Standardization and data quality
- On-going access to health outcomes data to evaluate PH programs



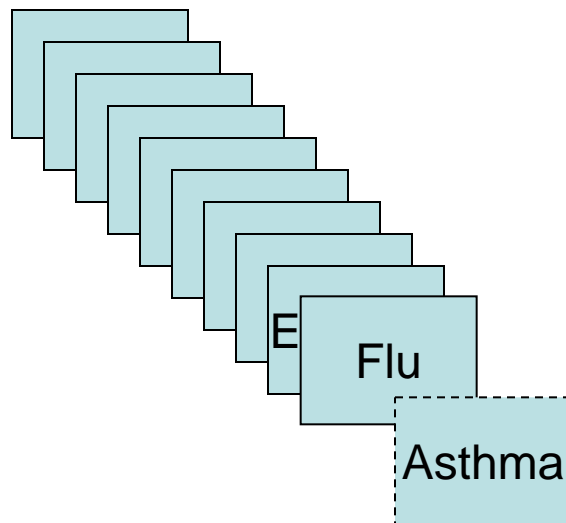
Data Sources



Data Marts



Public Health Reporting
Data Warehouse (PHRDW)

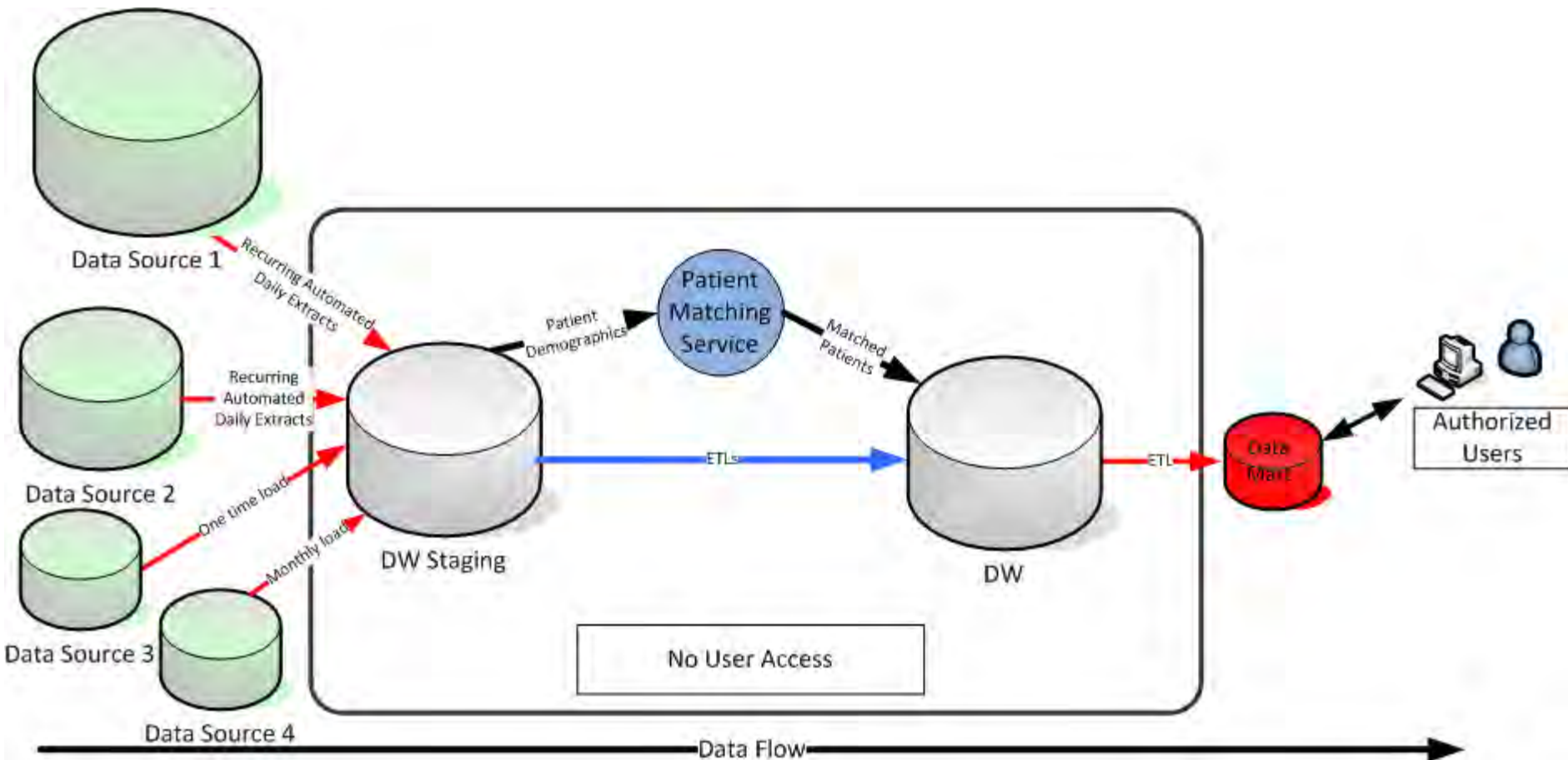


Subset of data for a specific surveillance purpose

- Automated algorithms applied (e.g. lab interpretation, case definitions)
- Mart-specific dimensions created (e.g., age groups)



Architecture



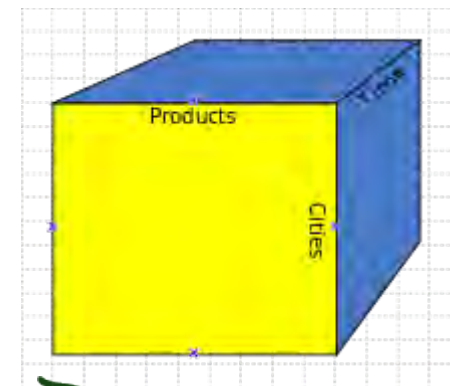
What is a Cube?

Data cube

- a three- (or higher) dimensional array of values
<http://en.wikipedia.org/wiki/Data_cube>

OLAP cube

- a set of data organized in a way that facilitates non-predetermined queries for aggregated information
- a tool for Online Analytical Processing
<http://en.wikipedia.org/wiki/OLAP_cube>



Transformation of source data for cube creation

Variables are grouped into dimensions

Gender	Age	Health Authority	HSDA
F	15	Vancouver Coastal	Vancouver
M	25	Fraser	Fraser North
M	19	Vancouver Coastal	Vancouver
F	22	Fraser	Fraser South

Values become their members (categories)

Data presented in a Cube format – Example 1

Gender	Age	Health Authority	HSDA
F	15	Vancouver Coastal	Vancouver
M	25	Fraser	Fraser North
M	19	Vancouver Coastal	Vancouver
F	22	Fraser	Fraser South



	15-19	20-24	25-29	Age Group
F	1	1		2
M			1	2
Gender	2	1	1	4

Data presented in a Cube format – Example 2

Gender	Age	Health Authority	HSDA
F	15	Vancouver Coastal	Vancouver
M	25	Fraser	Fraser North
M	19	Vancouver Coastal	Vancouver
F	22	Fraser	Fraser South



	Vancouver	Fraser South	Fraser North	HSDA
F	1	1		2
M	1		1	2
Gender	2	1	1	4

The Structure of a Cube

- Measures
e.g. # of patients, # of cases, # of tests, Rate of cases,
Average age, Median turn-around-time
- Dimensions
e.g. Patient, Reported Date, Diagnosis, Health Authority
 - Attributes
e.g. Gender, Ethnicity
 - Hierarchies
e.g. Year-Month-Day, HA-HSDA-LHA

Access Methods for Data Mart Cubes

1. Pivot Tables
 - Client identifiers removed
2. Line List Reports
 - Identifiers viewable; users audited
3. Direct SAS script
 - Analytic code



Pivot Table

Patients included in the Flu Concise Line List report

Result Date.CDC Flu Season: 2011
 Result Date.Year: (Multiple Items)
 Proficiency Test.Yes No: No
 Flu Result: Yes
 Result As of Date: Current
 Patient Health Authority.Health Se: (Multiple Items)

Tests

CDC Flu Season: 2011

Virus Type: Flu A

Virus Subtype: pH1N1, sH3N2

	2011:40	2011:48	2011:49	2011:50	2011:51	2011:52	2012:01	2012:02	2012:03	2012:04	2012:05	2011 Total	Grand Total
*Not In Hierarchy									1			1	1
Coronavirus					1		1		1	2	1	6	6
Flu A									2		5	7	7
Flu A Total	3	5	12	32	26	28	33	38	21	21	6	249	249
Flu B		1			1	2	1	4		6	1	17	17
Human MetaPneumoVirus					1							1	1
Respiratory Syncytial virus							1					1	1
Rhinovirus or Enterovirus	1			1			2	1			1	6	6
Grand Total	3	6	12	32	27	30	34	43	23	27	12	274	274

Online Data Visualization



Data Visualization

- The main goal is to communicate quantitative information clearly and effectively through graphical means.
- As data becomes more widely available and online tools evolve to manipulate, compare, and overlay data, visualizations offer opportunities for:
 - stakeholder engagement
 - situational awareness
 - rapid decision-making
 - advocacy
 - policy support
 - open data



Impact

Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813. Dressée par M. Mimarzd, Inspecteur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en traits des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Léguir, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre. L'on me vint faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davoust qui avaient été détachés sur Minsk et Mabilow et ont rejoint vers Orscha et Witebsk, avaient toujours marché avec l'armée.

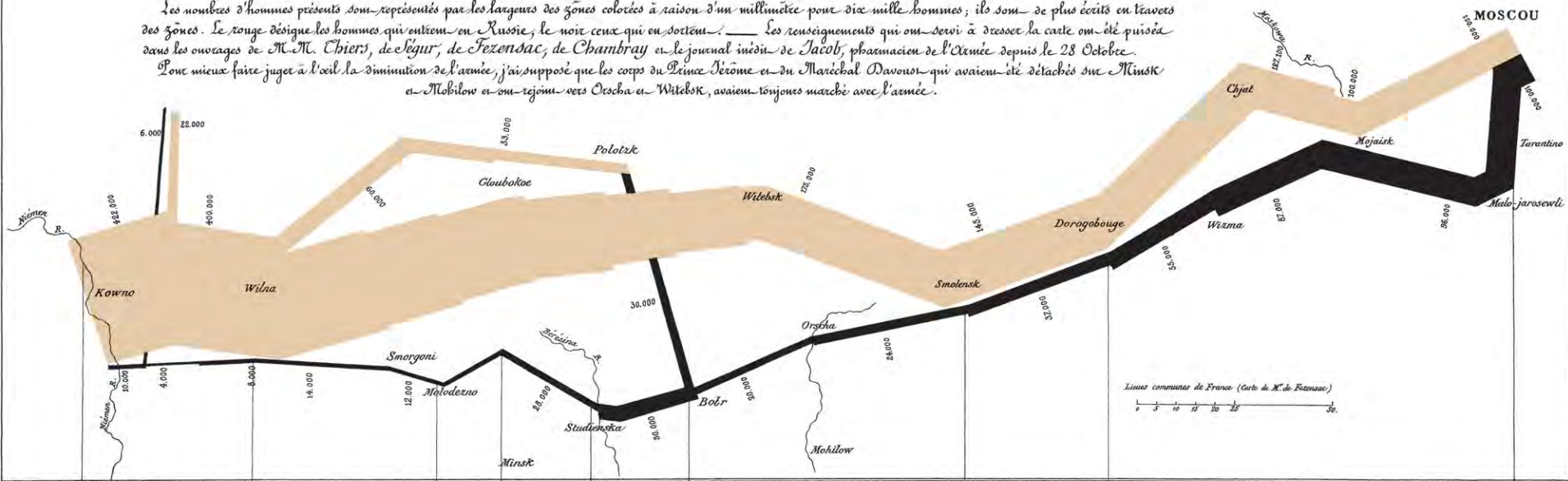
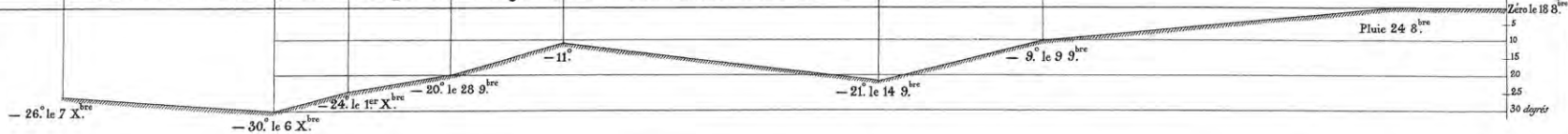


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.



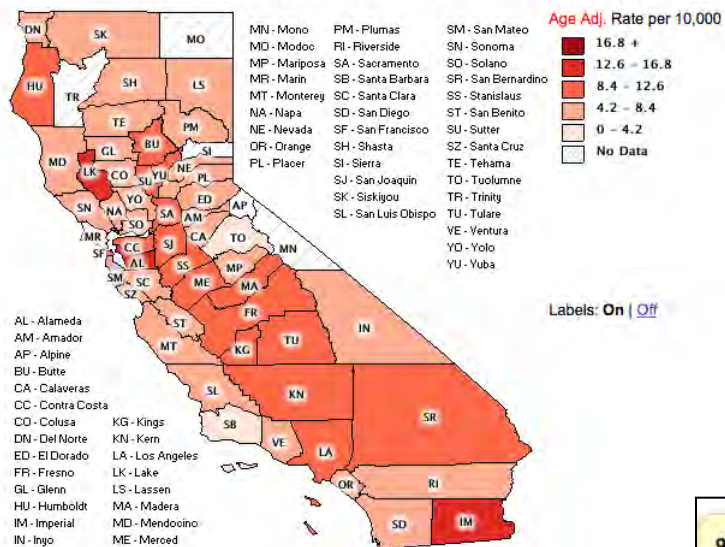
Les Cosaques passent au galop le Niémen gelé.

Imp. Lith. Regnier et Douv. det.



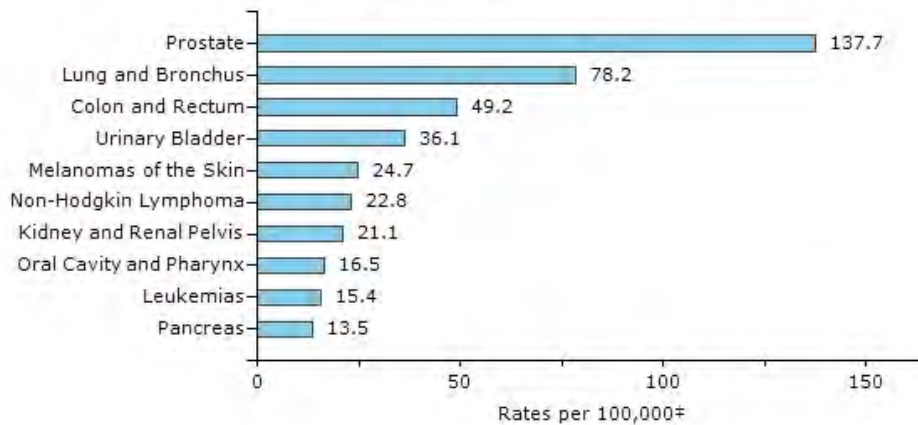
Exploratory vs. Explanatory

Hospitalizations due to asthma in California by County, All Races/Ethnicities, All Ages, Both Sexes, 2009



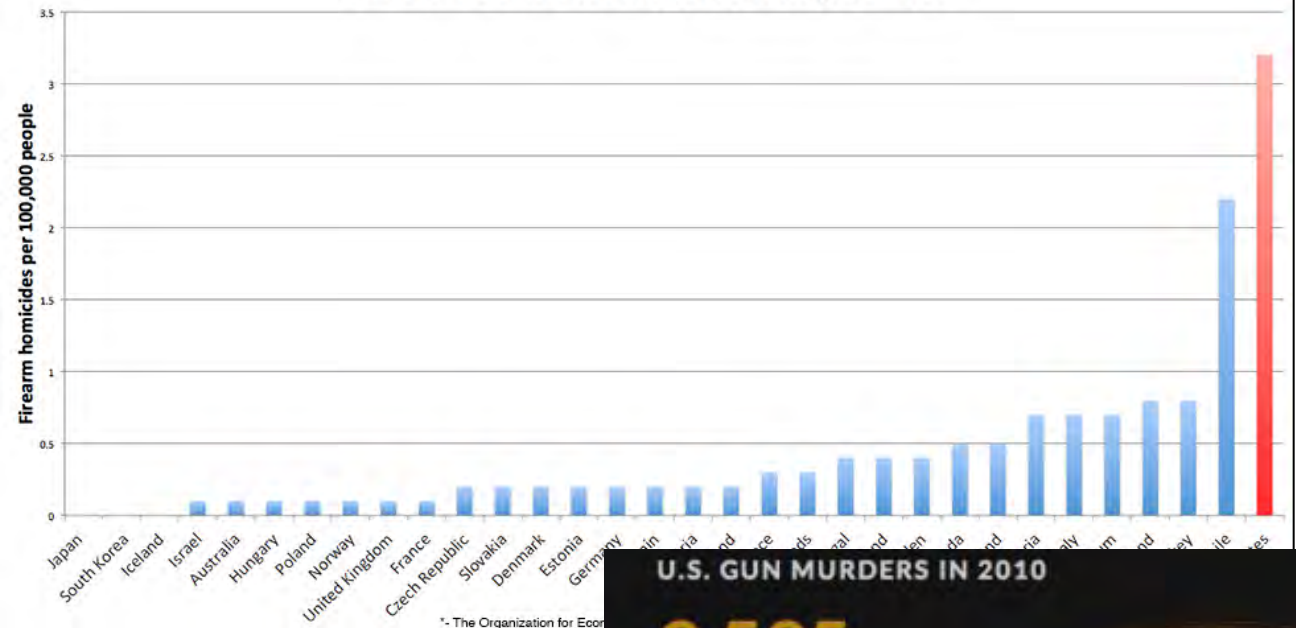
95% Confidence Interval : Hide Show

Top 10 Cancer Sites: 2009, Male, United States—All Races



Informative vs. Persuasive

Gun-related murder rates in the developed world*



*- The Organization for Economic

U.S. GUN MURDERS IN 2010

9,595
PEOPLE KILLED

409,280
STOLEN YEARS

This black man was shot in December in Mississippi at the age of 36 by his son during an argument.

Had he not been killed with a handgun, he might have lived to be 84 and died of respiratory disease.

Dashboards



SCHOOL OF PUBLIC HEALTH
LI KA SHING FACULTY OF MEDICINE
THE UNIVERSITY OF HONG KONG

Home About us Highlights Recommendations Contact us Links 中文

Influenza Surveillance Dashboard

Influenza activity is **moderate** and **stable** this week.

Overall index: 0.49



* click figure for details

Export Data

Select Category
all

next >



HKU Influenza surveillance dashboard is a specific illustrative application to prospective influenza surveillance in Hong Kong

Last Updated: 4 May, 2011

Data Visualization Concepts

- Data dimensions, for example
 - disease
 - time
 - age
 - sex
 - geography
- Visual encodings, for example
 - position
 - colour
 - length
 - size

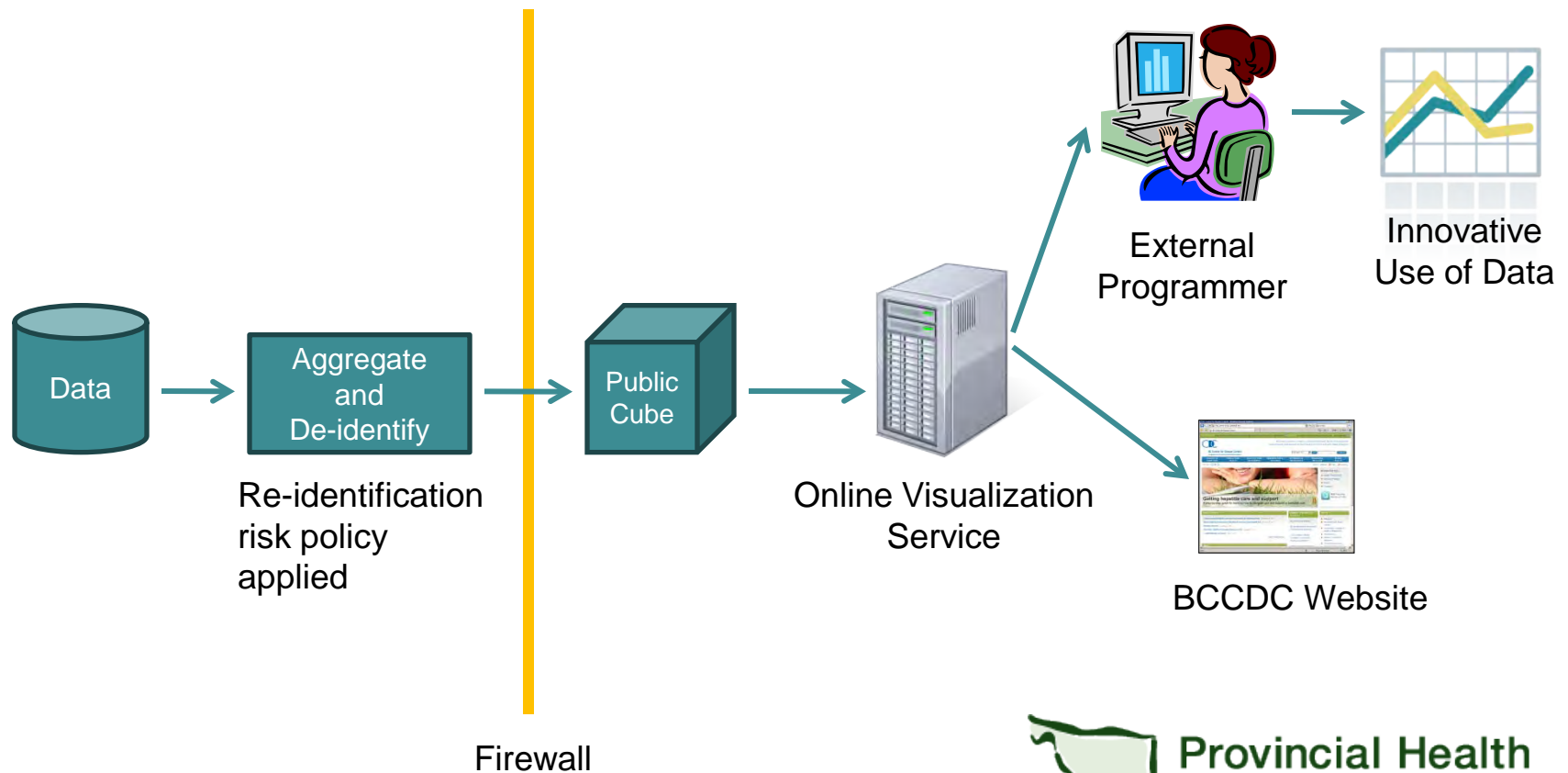


Project Overview

- Goal: To facilitate members of the public, including health care professionals based outside of the BCCDC, to prepare tables and figures that communicate information regarding diseases and public health interventions in British Columbia in a manner that is clear and meaningful.
- Audiences: public, media, researchers, students
- Phase 1 scope: CD Annual Report



Online Data Architecture



Privacy

- A public-use data file of aggregate information will be created and will adhere to the re-identification risk policy (comparable to data available in annual reports).
- No personal information will be stored on the site. Only the public-use file will be uploaded.
- The site has no access to internal resources or line-level data.
- A process will be developed to ensure all data is de-identified and that no identifying or quasi-identifying information is ever posted.



Proof of Concept

