Family History Data from Clarity and Other Sources

Authors: Sharon Fuller, Deborah J Seger; Group Health Research Institute

Acknowledgements: NCI Grant U54CA163261, NCI Grant P01CA154292, and Group Health Radiology

GroupHealth RESEARCH INSTITUTE

Introduction

Information about a patient's family history is an important tool to identify potential risks for many diseases like cancer or diabetes. Group Health Cooperative (GHC) collects family history using several applications, including visits and surveys, to manage effective screening programs. In 2004, GHC adopted Epic's electronic medical record system. In this poster, we investigate Clarity, Epic's reporting database, to determine its usefulness as a new source of structured family history data.

Methods

We compared several structured data sources for family history of four conditions: colon cancer, colon polyps, breast cancer, and ovarian cancer.

We examined all patients with a known birth date and sex who had any patient encounter recorded in Clarity between Jan 1, 2004 and Jun 30, 2012. This denominator is restricted to females only for breast and ovarian cancer.

Results

Among patients with any evidence of family history of a condition, Clarity's FAMILY_HX table independently contributes as many as 75% of patients with available family history, GHRI's VDW DX dataset adds an additional 2-20%, and Clarity's PROBLEM_LIST table adds no more than 2-3%.

For female patients at GHC, the Breast Screening Recruitment and Reminders survey (BSRR) remains a key source, being the sole contributor of 25% of patients with a family history of breast cancer and 77% for ovarian cancer. The BSRR is completed by women age 40 and over at every screening mammogram.

When comparing information available from different data sources, it is important to be aware of varying definitions. For example, ICD-9 family history diagnosis codes are general, specifying only the primary site with none of the restrictions on relationship degree or age that may be intended in Epic.

A key advantage of Clarity family history data and the BSRR survey over ICD-9 codes is that they provide detailed information about exactly which relatives are affected. On the other hand, VDW data are available from both internal and external providers – the remaining sources are available only for GHC group practice patients.

Conclusion

Examining records from Clarity's FAMILY_HX table may significantly augment, in both numbers and detail, conventional clinical and survey sources of family history data.

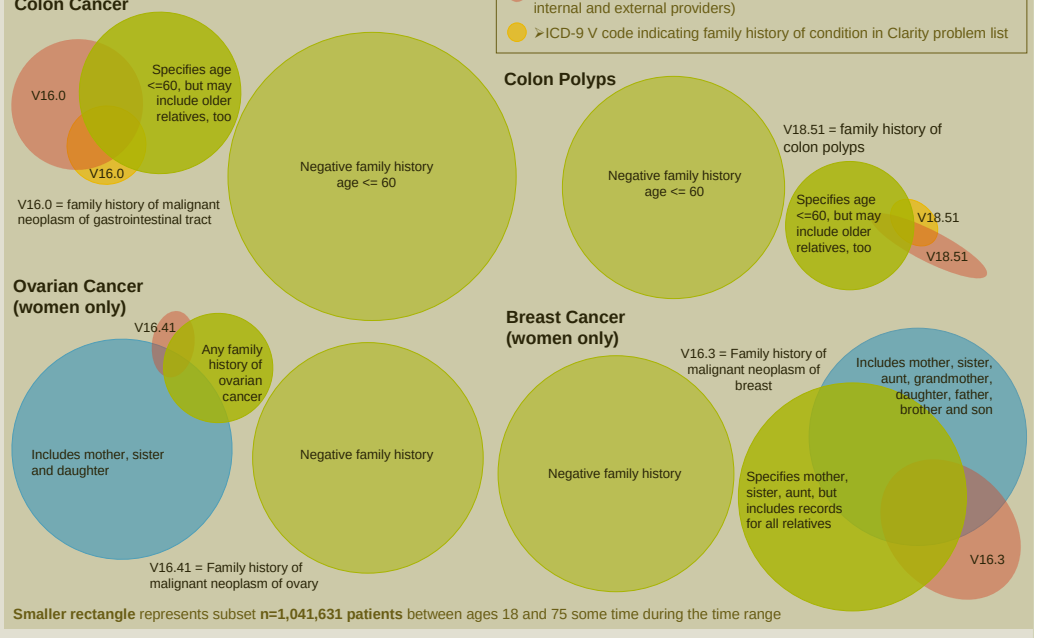
Conditions in Clarity Family History

First Use	Condition
2003	
	Alcohol/Drug
	Autoimmune disease
	Breast CA (mother, sister, aunt)
	CHD (male<55,female<60)
	Colon Cancer <= 60
	Colon Polyps <= 60
	Depression
	Diabetes
	Hypertension
	Inheritable anemias
	Other
2004	
	Asthma/Atopy
	Cancer, Other
	Osteoporosis
	Ovarian Cancer
2006	
	Prostate Cancer
	Stroke
2007	
	Graves disease
	Hashimoto's disease
	High triglycerides
	Hypercholesterolemia
	Hyperthyroid
	Hypothyroid
	Lipidemia
	Low HDL
2011	
	Anesthesia problems
	Arthritis
	Autism
	Bleeding problems
	Cancer (note type)
	Child with birth defects
	Congenital heart defect
	Cystic Fibrosis
	Down syndrome
	DVT
	Fragile X
	Hemophilia
	High blood pressure
	Huntington's chorea
	Mental retardation
	Metabolic disorder
	Miscarriages (many)
	MRSA
	Muscular dystrophy
	Neural tube defect
	Other genetic risks
	Other inherited disorder
	Sickle cell disease or trait
	Stillbirth
	Tay-Sachs
	Thalassemia
2012	
	Glaucoma

In reality, there is **overlap** between people with family histories of these four conditions – the conditions are displayed separately for ease of visualization.

Diagrams for each condition are **proportional** in size to the denominator. Each area represents counts of distinct patients.

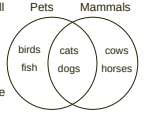
- ➢Condition recorded for at least one family member in Clarity family history
- ➢Condition confirmed not present in Clarity family history
- ➢Patient responses on the GHC Breast Screening Recruitment and Reminders survey (BSRR) – completed at mammogram appointment
- ➢ICD-9 V code indicating family history of condition in VDW dx file (both internal and external providers)
- ➢ICD-9 V code indicating family history of condition in Clarity problem list



Smaller rectangle represents subset **n=1,041,631 patients** between ages 18 and 75 some time during the time range

Larger rectangle represents **n=1,292,799 patients** with a known birth date and sex who had any 'encounter' in Clarity between 1/1/2004 and 6/30/2012

Venn Diagram How-To



Venn diagrams are commonly used to show all possible **relationships** between sets, but the classic 2 or 3-circle diagrams often do not speak to the **size** of each set.

Creating area diagrams that accurately reflect the size of each set and overlap turns out to be an interesting mathematical problem when dealing with more than two sets (two sets can always be represented proportionately using circles).

For a start, you can almost never achieve proportional accuracy with three **circles**. However, if you do not require absolute accuracy, you may find the Google Charts API[1] useful. It is not difficult to write a SAS macro[2] to call the API and choose your own colors and labels. For more customization options, the R program, VennDiagram is quite flexible.[3]

Convex **polygons** (frequently, simple rectangles) can be used for proportional depictions of most 3-way comparisons of clinical interest. These diagrams can be rather handsome, but lack the instant recognizability of the traditional Venn circles. The Java program DrawVenn[4] makes it easy to create 3-way rectilinear diagrams, while its companion, DrawEuler[5], or an online applet[6] can be used for more complex polygons.

To create an area diagram that is easily recognizable but still reasonably accurate in its proportions, you can turn one or more of the circles into an **ellipse**. Use the drawing tools in Microsoft PowerPoint or Excel to create circles with the correct radii. Start with a layout that mimics that produced in Google Charts, and change one or more of your circles to non-circular ellipses to create a diagram that more closely matches your data. For greater accuracy, a drafting program such as AutoCAD (used for the diagrams in this poster) allows you to precisely measure and adjust each overlap area.

[1] http://imagecharteditor.appspot.com/ (deprecated, perhaps because of its inherent inaccuracy)
[2] http://support.sas.com/resources/papers/proceedings11/109-2011.pdf
[3] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3041657/
[4] http://theory.cs.uvic.ca/euler/DrawVenn/index.html
[5] http://theory.cs.uvic.ca/euler/DrawEuler/
[6] http://www.cs.kent.ac.uk/people/staff/pjr/ConvexVenn3/diagrams2010.html