Making Public Health Data Linkable

John J. Grefenstette PhD, Willem G. van Panhuis MD PhD, Donald S. Burke MD

Public Health Dynamics Laboratory, University of Pittsburgh

Email: gref@pitt.edu

Abstract

Project Tycho was established in 2009 at the University of Pittsburgh Graduate School of Public Health to improve access to detailed public health data for analysis and policy making. The Tycho database currently contains the digital version of the entire US weekly nationally notifiable disease surveillance system, including data from 1888 to the present. The Tycho project hopes that by providing this example of US surveillance data, health officials around the world will be convinced of the value of this resource and will contribute additional datasets or make their own data available in a linkable form.

Many analyses of public health data require other data types, including demographic and climate data. While additional datasets can be incorporated into the Tycho database itself, remote datasets can also be made accessible through the use of database linkages. This presentation will describe the Tycho database infrastructure and public interfaces, including web-based search features and a public automatic program interface (API). Opportunities and challenges in linking the database to other large databases will be discussed.