Data in public health

I n 1854, physician John Snow helped curtail a cholera outbreak in a London neighborhood by mapping cases and identifying a central public water pump as the potential source. This event is considered by many to represent the founding of modern epidemiology. Data and analysis play an increasingly important role in public health today. This can be illustrated by examining the rise in the prevalence of autism spectrum disorders (ASDs), where data from varied sources highlight potential factors while ruling out others, such as childhood vaccines, facilitating wise policy choices.

The global prevalence of ASDs has been estimated to be 62 cases per 10,000 people, with higher estimates of 147 cases per 10,000 in the United States. These rates have increased substantially since the 1990s, although it is not clear how much of this reflects a rise in diagnosis as opposed to true increases in the frequencies of these disorders.

Autism has a substantial heritable component based on evidence from studies of families, including those with monozygotic twins. A series of large studies using whole-genome sequencing methods have identified over 100 genes that appear to contribute to ASD susceptibility. A collaboration between the research community, a patient advocacy group, and a technology company (www.mss. ng) seeks to sequence the genomes of 10,000 well-phenotyped individuals from families affected by ASD, making the data freely available to researchers. Studies to date have confirmed that the genetics of autism are extremely complicated—a small number of genomic variations are closely associated with ASD, but many other variations have much lower predictive power. More than half of siblings, each of whom has ASD, have different ASD-associated variations. Future studies, facilitated by an open data approach, will no doubt help advance our understanding of this complex disorder.

Sociological data have also been used to gain insight into autism prevalence. By using data from California for individuals with ASD diagnoses in combination with

"...data sharing...can drive research progress on major public health challenges."

birth records, information (including likely residential locations) was inferred for more than 300,000 children covering 6 years. These data enabled studies of spatial clustering of ASD diagnoses and allowed estimation of the contributions of variations in diagnostic criteria and other social factors to the increase in ASD diagnoses. The estimates indicate that unknown environmental triggers play a substantial contributing role.

One "environmental" factor that has been suggested as contributing to the increase in autism prevalence is vaccination, particularly vaccines for measles, mumps, and rubella (MMR). This connection has been studied extensively in epidemiological studies around the world. A meta-analysis that integrated the results of five cohort studies and five case-control studies found no evidence for an association between ASDs and vaccination or any of the components of vaccines suggested to have a role in ASD development.

A new data collection strategy was reported in 2013 to examine contagious diseases across the United States, including the impact of vaccines. Researchers digitized all available city and state notifiable disease data from 1888 to 2011, mostly from hard-copy sources. Information corresponding to nearly 88 million cases has been stored in a database that is open to interested parties without restriction (www.tycho.pitt.edu). Analyses of these data revealed that vaccine development and systematic vaccination programs have led to dramatic reductions in the number of cases. Overall, it is estimated that ~100 million cases of serious childhood diseases have been prevented through these vaccination programs.

These examples illustrate how data collection and sharing through publication and other innovative means can drive research progress on major public health challenges. Such evidence, particularly on large populations, can help researchers and policy-makers move beyond anecdotes—which can be personally compelling, but often misleading—for the good of individuals and society.

–**Jeremy Berg**

Editor-in-Chief, Science *Journals.* jberg@aaas.org

Science

Data in public health

Jeremy Berg

ARTICLE TOOLS	http://science.sciencemag.org/content/355/6326/669
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service