Tycho Brahe got his nose lopped off over an argument about a math problem. He once refused to get up from a dinner party to relieve himself because he thought it rude, resulting in, probably, a burst bladder that killed him. Yet, his peculiar brand of determination helped give rise to the scientific revolution. The Danish nobleman was among the last great “naked-eye” observers of the cosmos. Before his death in 1601, Brahe passed along his life’s work—30 years of detailed observations of the night sky—to his assistant, Johannes Kepler, urging him not to let the fruit of his labors languish.

They did not. Brahe’s careful observations became the basis for Kepler’s laws of planetary motion, which would, in turn, contribute to Isaac Newton’s law of universal gravitation.
Four centuries later, the Pitt researchers who created Project Tycho, a digital database that provides open access to U.S. disease surveillance data, hope they have created a similar foundation for discovery. The newly built epidemiological archive chronicles reports of 56 infectious diseases in every state before, during, and after vaccination licensure from 1888 to recent times.

It took almost three years and more than 200 million keystrokes to create the Project Tycho archive. Many of those workers were University of Pittsburgh undergrads as well as students from Digital Divide Data, a social enterprise that provides jobs and education to young people in Cambodia, Laos, and Kenya. These clerks standardized and organized almost 90 million cases from weekly public health records (paper and PDFs) from all U.S. states and territories, including more than 3,000 American cities. What they wrought: the largest centralized bank of digitized disease surveillance data ever assembled.

And access to it is free, says Wilbert van Panhuis, an MD/PhD professor of epidemiology at Pitt’s Graduate School of Public Health and lead investigator for the project. “Our vision was that not only us, but everybody should be able to use this public data for analysis and models.” For instance, anybody with enough interest and access to the Internet—a scientist at a university or pharmaceutical company, a journalist, an undergrad—can easily track where and when the polio vaccine was implemented and its efficacy in those cities.

“We hope there are epidemiological, disease-curing Keplers today who will be able to use these data to derive important laws and insights on how epidemics arrive, leave, and interact,” says coinvestigator Donald Burke in the project’s promotional video. Burke is an MD professor of medicine and of infectious diseases (among other appointments) and dean of Pitt Public Health.

The field of public health data compilation has been fraught with redundancies. Most projects are focused on specific questions; a researcher might toil for years answering a question like, What effects do condom distribution programs have on the rate of HIV infection in the rural United States? In search of answers, investigators painstakingly build data sets that often are not shared. And it can be difficult to get funding to create archives with no specific research questions in mind.

Happily, both the National Institutes of Health and the Bill and Melinda Gates Foundation saw value in creating a massive digital archive and funded Project Tycho.

The Project Tycho team has also been inventing new methods to process and analyze public health data. In a November 2013 New England Journal of Medicine paper, Project Tycho researchers (from Pitt’s public health, medicine, and information sciences schools with collaborators from Johns Hopkins) revealed that vaccination programs
for polio, measles, mumps, rubella, hepatitis A, diphtheria, and pertussis (whooping cough) have prevented more than 100 million cases of serious childhood infectious diseases since 1924. Still, some of these pathogens are reemerging. Pertussis vaccines, for example, have been available since the 1920s, but the worst whooping cough epidemic since 1959 occurred in 2012, with more than 48,000 cases nationwide reported by December of that year.

"Parents who question the risk-benefit balance of vaccination may refuse or delay immunization of their children," the Project Tycho team reports, "which leads to local variations in vaccine coverage and increased risk of disease outbreaks." Van Panhuis admits he hopes the project "will introduce new evidence into the debate about vaccination." The next big step for Project Tycho is to go global. But, Van Panhuis says, technological, economic, and political barriers can hinder cooperation. For instance, developing countries that rely on tourism might be wary of releasing information about epidemics. And they may not even have the means to collect data, let alone analyze them. What's in it for us?, the gatekeepers may wonder.

Well, perhaps the lives of millions.

Van Panhuis remains optimistic. He says understanding a disease’s narrative, locally and globally, can help move the scientific field forward in developing theories about causation—and then, ways to control or prevent disease.

Elaine Vitone contributed to this report.

To take a peek at Project Tycho, visit: www.tycho.pitt.edu