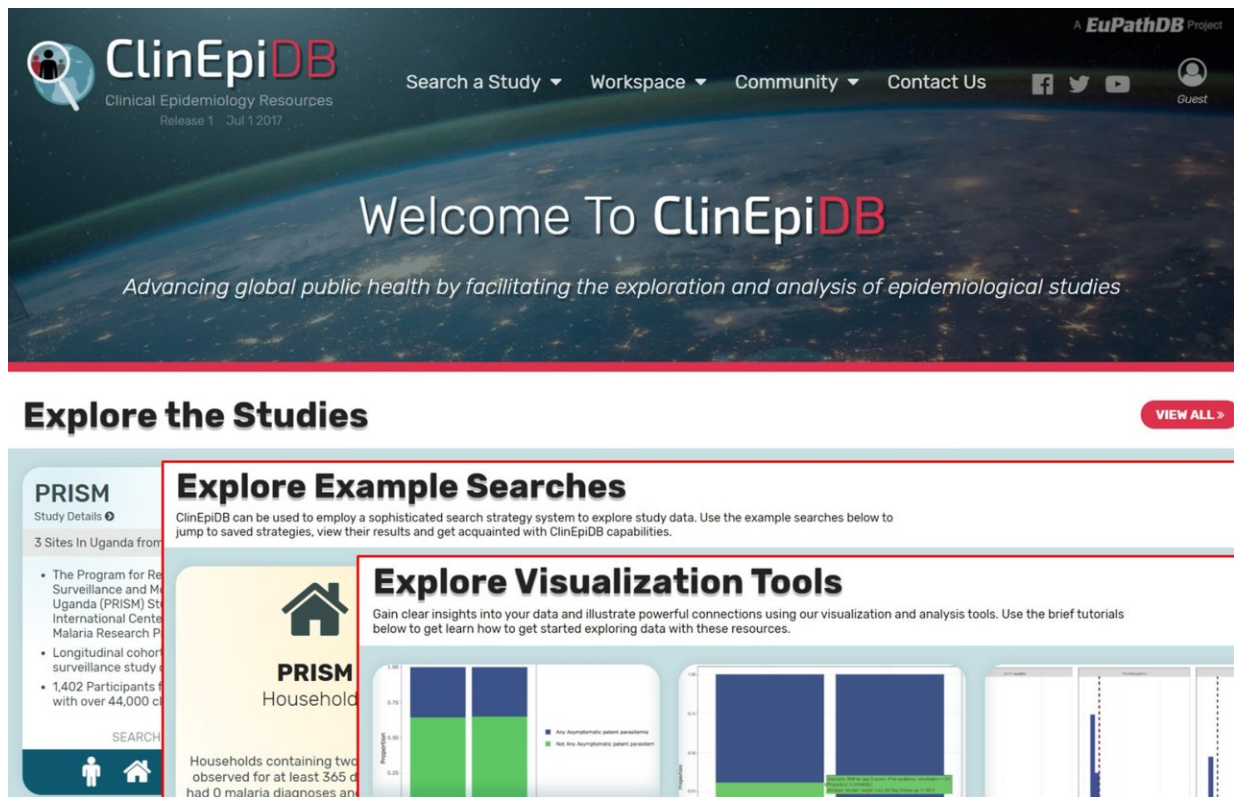


# Open-access data resource aims to bolster collaboration in infectious disease research

February 21 2018, by Katherine Unger Baillie



An international team of researchers has launched the Clinical Epidemiology Database, an open-access online resource enabling investigators to maximize the utility and reach of their data and to make optimal use of information released by others. Credit: University of Pennsylvania

Population-based epidemiological studies provide new opportunities for



innovation and collaboration among researchers addressing pressing global-health concerns. As with the vast quantities of information emerging in other fields, from economic modeling to weather surveillance to genomic medicine, the technical challenges of sharing and mining gigantic datasets can hamper such efforts. A single epidemiological study—tracking the acquisition of functional resistance to malaria, or the relationship of diarrheal disease to developmental outcomes—may involve tens of thousands of clinical observations on thousands of participants from multiple countries.

To overcome these hurdles, an international team of researchers has launched the [Clinical Epidemiology Database](#), an open-access online resource enabling investigators to maximize the utility and reach of their data and to make optimal use of information released by others.

The development of ClinEpiDB has been led by the University of Pennsylvania's David Roos, the E. Otis Kendall Professor of Biology in the School of Arts and Sciences, and Christian Stoeckert, research professor of genetics in Penn's Perelman School of Medicine, along with Jessica Kissinger, distinguished research professor of genetics at the University of Georgia's Institute of Bioinformatics, and Christiane Hertz-Fowler, professor at the University of Liverpool's Institute of Integrative Biology.

ClinEpiDB uses computational infrastructure established during the past 20 years for the [Eukaryotic Pathogen Database](#), one of four national Bioinformatics Resource Centers for Infectious Disease supported by the U.S. National Institute of Allergy and Infectious Diseases, part of the National Institutes of Health, with additional support from The Wellcome Trust (UK) and others. EuPathDB is a thriving genomics resource for integrative analysis of microbial eukaryotes, such as the parasites that cause malaria, sleeping sickness, and other diseases. EuPathDB is currently accessed by more than 70,000 unique visitors



monthly, from 100-plus countries around the world, and has been cited more than 13,000 times in the scientific literature to date.

"It is increasingly possible to generate spectacularly valuable, large-scale datasets, but how to store and manage this information so that people can make sensible use of it is arguably the overriding challenge of our day," says Roos. "The EuPathDB project has demonstrably helped translate the promise of infectious-disease genomics into practice, and with ClinEpiDB we are providing a resource to help get the information from large patient studies into the hands of those who can do the most good with it, while also protecting the confidentiality of [study participants](#)."

Bioinformaticist Brian Brunk oversees the EuPathDB as senior project manager, and molecular epidemiologist Brianna Lindsay is responsible for coordinating the ClinEpiDB initiative.

Many journals and funders encourage, and often require, scientists to make their study data available, but doing so in a useful way can be difficult for data-providers and users alike. ClinEpiDB aims to mitigate these issues by creating standardized processes for accessing and exploring complex clinical data. This new web resource introduces an intuitive interface, enabling users to explore data using point-and-click filtering, simple queries and more complex "search strategies" and a suite of exploratory statistical-analysis tools. The site also provides documentation of study design and background, contact information for data contributors, and links to study-related publications and resources.

According to Stoeckert, "establishing formal definitions of and relationships between data variables is one key to the success of this initiative. EuPathDB uses an OBO Foundry based ontology, aiding integration across datasets and establishing common, user-friendly terms for study details."



The ClinEpiDB launch presents as its inaugural study data from the Program for Resistance, Immunology, Surveillance and Modeling of Malaria project, or PRISM, led by Grant Dorsey, professor of medicine at the University of California, San Francisco, and Moses Kamya, professor and dean, School of Medicine, Makerere University College of Health Sciences, Kampala, Uganda. PRISM includes data from more than 40,000 clinical observations of 1,400 study participants, as one of several NIAID-funded International Centers of Excellence for Malaria Research

"The goal of PRISM project is to improve our understanding of malaria, and measure the impact of population-level control interventions," Dorsey notes. "This study represents seven years of work to date, from scores of researchers, with contributions from many hundreds of Ugandan kids at risk for malaria, as well as their families. It is exciting that ClinEpiDB makes it easy for anyone to browse and analyze the data and to quickly test parameters that may be associated with increased or decreased risk of serious malaria."

Further studies in the pipeline for release on the ClinEpiDB platform include additional ICEMR projects, and two large global enteric disease datasets funded by the Bill & Melinda Gates Foundation: the Global Enteric Multicenter Study, or GEMS, and the MAL-ED study on etiology, risk factors and interactions of enteric infections and malnutrition, and the consequences for child health and development.

Steve Kern, deputy director for quantitative sciences at the Gates Foundation says: "Our mission is to improve global health and reduce inequality, and achieving these goals depends on accessing and interrogating the wealth of available information produced by the global scientific community. We are optimistic that resources like ClinEpiDB will help make information produced by the foundation and its global partners available to all and enable us to take advantage of information



from others, expediting scientific discovery and evidence-driven translation to improve human health worldwide."

Provided by University of Pennsylvania

Citation: Open-access data resource aims to bolster collaboration in infectious disease research (2018, February 21) retrieved 26 April 2024 from <https://medicalxpress.com/news/2018-02-open-access-resource-aims-bolster-collaboration.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.