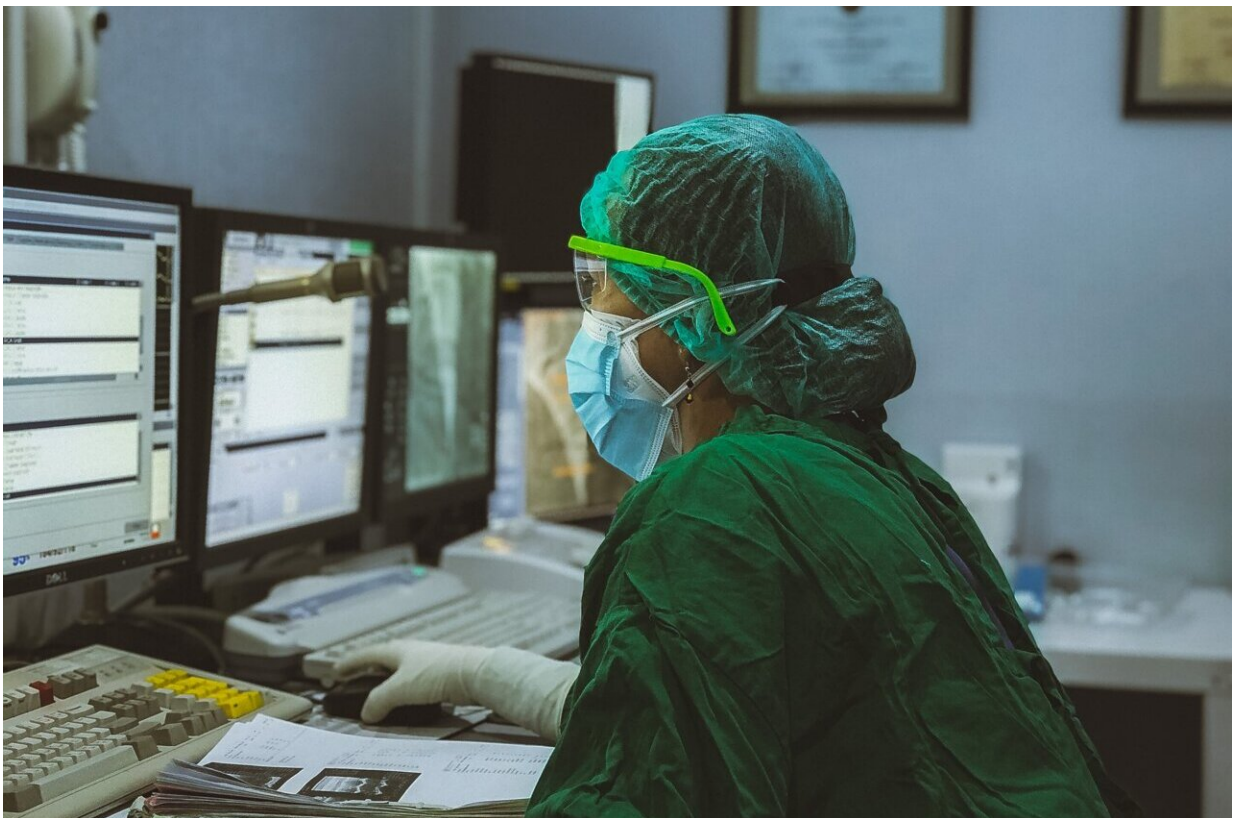


Data scientists develop rare disease phenopacket standard and tools for global use

May 17 2023



Credit: Unsplash/CC0 Public Domain

Researchers in the Department of Biomedical Informatics (DBMI) at the University of Colorado School of Medicine have reached a major

milestone in developing standards and tools for creating phenopackets that may foster more innovation and advancement in the medical field by allowing health professionals to more easily collect and share data.

A newly-released paper in *PLOS ONE* highlights the latest suite of coordinated standards and tools used to collect data related to rare diseases.

The phenopackets, a new Global Alliance for Genomics and Health (GA4GH) schema for sharing disease and phenotype information, characterize an individual patient or biosample. Such phenopackets, which can comprise a variety of data including [genetic information](#), treatments, diagnoses and other information, can be especially useful for researchers because they are designed to be shared between people, clinics, and health systems.

"One of the challenges with clinical data is that it's usually collected for purposes such as billing, quality assurance, insurance or pharmacy orders. It's not designed to describe the patient as a biological subject," explains Melissa Haendel, Ph.D., DBMI professor and chief research informatics officer for the CU Anschutz Medical Campus.

Haendel conceptualized the phenopacket standard as a way to broadly share characteristics of a patient and describe patients in a biologically meaningful way. Haendel is the co-founder of the Monarch Initiative, a GA4GH driver project that is dedicated to improving rare disease diagnostics, which have been hampered by the lack of patient-level data sharing. She also leads the National Human Genome Research Institute (NHGRI) Center of Excellence in Genome Sciences's Phenomics First Resource.

"The phenopacket is an exchange standard to say, 'How do we get patient phenotype information into and out of electronic health records? How do

we send that information to clinical labs? How do we submit it to journals?" Haendel says. "The goal is to take that set of features and be able to move them around in different contexts in a de-identified, non-personal way so that we can use those for understanding disease trajectories without any sort of compromise to patient rights."

A world before phenopacket standards

Monica Munoz-Torres, Ph.D., associate professor in the DBMI and program director for Monarch Initiative and Phenomics First Resource, describes collecting phenotype data up until recently as "the wild West."

"There have been tons of tools that have tried to capture clinical data," says Munoz-Torres, also a co-developer of the standard. "The problem is that none of these previous models have actually been able to represent a general model for representing [clinical data](#) of individual patients with arbitrary diseases or the linkage to genomic and pedigree data for each of these patients."

That realization almost a decade ago is what prompted the work on creating standards.

"We sat down and thought about how we fix that problem, so we partnered with GA4GH because we knew that this was a group that was rethinking how we make all the knowledge that we have about both genomics and health computable, transferable, and understandable."

The starting place, Munoz-Torres says, was reaching out to potential stakeholders that could advise on their own genomics and health data needs and what they'd want to see in a phenopacket. Those conversations eventually led the researchers to creating their own easily implementable schema and the newly released tools.

"We figured the best way to do this in a programmatic manner was to design this schema and put it to the test. We received years of community feedback," Munoz-Torres explains. "Now we've gone the extra mile to create a set of tools that make it easier to create, validate, and work with those phenopackets."

The future for data sharing

An early indicator of how successful the work Haendel, Munoz-Torres, and their colleagues have conducted has been the adoption of the standard.

Almost 1 million phenopackets have been created to date from organizations all over the world, including the European Joint Programme on Rare Diseases and Japan's biobank network. Releasing the phenopackets as a standard of the International Organization for Standards (ISO) ensures that even more governments and organizations around the world will be able to adopt them in the future.

The markers of continued success will be in the research community, which will be able to more easily collect and analyze data about [rare diseases](#).

"We will be able to see people making discoveries that weren't previously possible," Munoz-Torres says. "They will be providing new knowledge."

More information: Daniel Danis et al, Phenopacket-tools: Building and validating GA4GH Phenopackets, *PLOS ONE* (2023). [DOI: 10.1371/journal.pone.0285433](https://doi.org/10.1371/journal.pone.0285433)

Provided by CU Anschutz Medical Campus

Citation: Data scientists develop rare disease phenopacket standard and tools for global use (2023, May 17) retrieved 25 April 2024 from <https://medicalxpress.com/news/2023-05-scientists-rare-disease-phenopacket-standard.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.