



Client Phenotypes

Client phenotyping is a machine learning-enhanced approach that identifies best-fit differentiated client care models to improve testing yields and client retention.

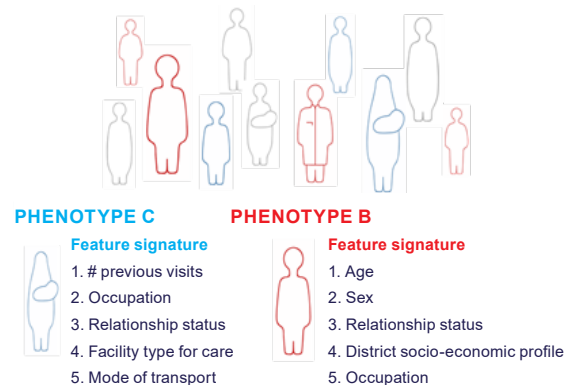
THE OPPORTUNITY

There is no average client; there is no average care experience. Each client has different constraints and assets that drive their treatment adherence; as such, clients require different care models. Clinical programs use data to drive testing and adherence strategies. However, the data available often describe clients and circumstances in the aggregate, masking individual-level characteristics needed to more precisely target services. To provide client-centric care and achieve epidemic control, it is crucial to leverage individual-level data on client attributes and experiences, and design and test care strategies that meet clients where they are, and with what they need.

THE DATA.FI SOLUTION

Data.FI works with U.S. Agency for International Development (USAID) Missions, host country governments, and implementing partners to create **Client Phenotypes** that leverage machine learning techniques and a variety of data sources to create highly informative profiles reflecting a person's characteristics and localized geographic environment. Client phenotyping goes beyond traditional market segmentation approaches that use simple heuristics, such as age and gender. Phenotypes are tailored to the use case, whether identifying patients for testing or customizing interventions to promote adherence. Models are then integrated with local systems to provide decision support for health workers in making client-by-client decisions on administering tests and selecting client support interventions. At the national level, these models can define optimal testing strategies to meet yield targets and design and direct precision intervention strategies to improve patient retention.

Creating Phenotypes: Phenotypes are groups of clients with multiple dimensions of common characteristics that are machine learned directly from data on clients' medical histories, health facilities, and locational attributes derived from unstructured satellite imagery using deep learning, among other sources. The solution provides each client with a unique set of ranked data features predictive of their behavior—a "feature signature"—or a set of features that emerges across datasets and geographies (see figure). Each client phenotype (group) experiences a unique set of constraints and assets. Understanding each signature enables precision deployment of interventions and resources to maximize impact.



Targeting Interventions to Improve Cascade Outcomes: Phenotypes are tailored to predict individual undiagnosed PLHIV, or those who will become lost to follow-up (LTFU) or virally unsuppressed without enhanced, customized care, enabling providers to improve clients' clinical outcomes and allows decision-makers to make the best use of resources. Precision targeting using risk assessment tools integrated with existing systems and informed by client phenotypes helps frontline health workers decide on a client-by-client basis whether they should administer a test or intervention. Moreover, phenotypes are not static. Rather, they learn over time from testing and intervention outcomes to become more accurate in their predictions and differentiated care models.

WHAT IS THE IMPACT?

Although Client Phenotyping is a new approach offered by Data.FI that is not yet widely applied within the global HIV epidemic, it has been deployed in the U.S. with promising results. In building this capability, Data.FI partner macro-eyes analyzed more than 6,000,000 patient records, along with satellite imagery and other data, to create phenotypes that describe clients in the U.S. The insights generated from phenotyping, such as quantifying the risk of LTFU, were then integrated into local health systems to optimize client scheduling, reduce wasted health worker time, and maximize the number of clients seen. In the U.S., phenotyping successfully predicted LTFU with 85% accuracy and informed interventions to optimize appointment scheduling, reducing client no-shows by 20%. The insights generated from phenotyping can have transformative effects on achieving the U.S. President's Emergency Plan for AIDS Relief (PEPFAR) testing and treatment targets.



PUTTING THE SOLUTION INTO ACTION

In each country, the approach to client phenotyping relies on the same model; however, each country application integrates country-specific client, facility, and geographic data into a machine learning pipeline. Although access to country data is required, the solution can be built on whatever data are available and does not rely on any single data source or variable. The approach also does not require cleaned or deduplicated data, as data cleaning and client-level deduplication are inherent in the phenotyping process. This flexibility means the approach can be applied in any context, regardless of what data are available.

Data.FI can support implementation of the solution through the following:

- Provision of cleaned datasets with associated data dictionaries, suitable for analysis
- Deduplication of individual-level client records
- Development of client phenotypes that drive differentiated care models to, for example:
 - Inform high-yield testing strategies
 - Predict LTFU and customize LTFU interventions
- Integration of predictive models within local infrastructure for ongoing and automated client-centric approaches at the point of care

Data for Implementation (Data.FI) is a five-year cooperative agreement funded by the U.S. President's Emergency Plan for AIDS Relief through the U.S. Agency for International Development under Agreement No. 7200AA19CA0004, beginning April 15, 2019. It is implemented by Palladium, in partnership with JSI Research & Training Institute (JSI), Johns Hopkins University (JHU) Department of Epidemiology, Right to Care (RTC), Cooper/Smith, IMC Worldwide, Jembi Health Systems, and Macro-Eyes, and supported by expert local resource partners.

This publication was produced for review by the U.S. President's Emergency Plan for AIDS Relief through the United States Agency for International Development. It was prepared by Data for Implementation. The information provided is not official U.S. Government information and does not necessarily reflect the views or positions of the U.S. President's Emergency Plan for AIDS Relief, U.S. Agency for International Development, or the United States Government.

February 2020

FOR MORE INFORMATION

Emily Harris, Data.FI AOR
emharris@usaid.gov
+1-571-309-1266

Jenifer Chapman, Project Director
jenifer.chapman@thepalladiumgroup.com
+1-202-775-9680