

Glossary of Common Terms

- [Abbreviations](#)
- [Terms](#)

Abbreviations

| Abbreviation | Full Text | Description |
|--------------|---|---|
| ACCE Project | Analytic validity, Clinical validity, Clinical utility, and Ethical, legal, and social implications | ACCE project (est. in 2000 by the CDC) created a model for evaluating the use of genetic tests, applicable across multiple tests and testing scenarios. |
| ACH | All Children's Hospital | A pediatric hospital located in St. Petersburg, Florida that is an academic health system providing treatment, education, research and advocacy in child health and care in more than 50 specialties. Part of the JHHS. |
| ACT | | |
| ADF | Azure Data Factory | |
| ADT | Admit, Discharge, Transfer | Extremely common in HL7 processing and among the most widely used of all message types. They communicate patient demographic and visit information, as well as the reason why the message is being sent. ADT messages are typically initiated by the EMR or a registration application and are used to keep ancillary systems in sync regarding the state of a patient. When a patient's record is updated, an ADT message is sent. This way, all systems can maintain the patient's current contact information, insurance, and next of kin, as well as their current location and attending doctor. |
| ADLS | Azure Data Lake Service | A highly scalable public cloud service that allows developers, scientists, business professionals, and other Microsoft customers to gain insight from large, complex data sets. Comprised of a data storage and data analytics component. |
| AKS | | cluster |
| API | Application Programming Interface | An interface or communication protocol between different parts of a computer program intended to simplify the implementation and maintenance of software. |
| APL | Applied Physics Lab | The nation's largest university-affiliated research center. |
| CAMP | Center of Excellence Analytics Medicine Program | A 12-week program that trains participants in the Precision Medicine Analytics Platform (PMAP). The program is a John Hopkins initiative to accelerate biomedical research by combining EMR, medical imaging, physiological monitoring, and genomics into a cloud-based big data platform. |
| CCDA | Core for Clinical Research Data Acquisition | |
| CDM | | |
| CER | Comparative Effectiveness Research | CER is the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers, and policy makers to make informed decisions that will improve health care at both the individual and population levels. (IOM) |
| CDSS | Clinical Decision Support System | |

| | | |
|----------------|---|---|
| CoE (or PMCoE) | Center of Excellence | Formally designated units within JHM, generally engaged in clinical treatment and research of a specific disease. |
| COPD | | |
| CRISP | Chesapeake Regional Information System for our Patients | A regional health information exchange (HIE) serving Maryland and the District of Columbia. Formally designated as Maryland's statewide health information exchange by the Maryland Health Care Commission. |
| CRBO | Clinical Research Billing Orientation | |
| CRMS | Clinical Research Management System | |
| CSV | Comma Separated Values | A delimited text file that uses a comma to separate values. Each line of the file is a data record. Each record consists of one or more fields, separated by commas. |
| DUA | Data Use Agreement | Legally binding contracts between a party and another executing party that provide the terms for transferring data from the provider organization to the recipient organization. |
| EDW | Epic Data Warehouse | |
| EHR | Electronic Health Record | EHRs are designed to reach out beyond the health organization that originally collects and compiles the information. They are built to share information with other health care providers, such as laboratories and specialists, so they contain information from all the clinicians involved in the patient's care. |
| EMR | Electronic Medical Record | The digital version of the paper charts in the clinician's office. An EMR contains the medical and treatment history of the patients in one practice. |
| EVGAPP | Evaluation of Genomic Applications in Practice and Prevention | Working Group of CDC |
| ETL | Extract, Transform, Load | A process that extracts the data from different RDBMS source systems then transforms the data (like applying calculations, concatenations, etc.) and finally loads the data into the PMAP system. |
| FHIR | Fast Healthcare Interoperability Resource | A standard describing data formats and elements (known as "resources") and an application programming interface (API) for exchanging electronic health records. |
| HDFS | Hadoop Distributed File System | A distributed file system designed to run on commodity hardware. It is highly fault-tolerant and is designed to be deployed on low-cost hardware. HDFS provides high throughput access to application data and is suitable for applications that have large data sets. |
| HIE | Health Information Exchange | Allows clinical information to move electronically among disparate health information systems. The goal of an HIE is to deliver the right health information to the right place at the right time - providing safer, timelier, efficient, effective, equitable, patient-centered care. |
| HIPAA | Health Insurance Portability and Accountability Act | <ul style="list-style-type: none"> • Provides the ability to transfer and continue health insurance coverage for millions of American workers and their families when they change or lose their jobs; • Reduces health care fraud and abuse; • Mandates industry-wide standards for health care information on electronic billing and other processes; and • Requires the protection and confidential handling of protected health information |
| HL7 | Health Level Seven | A set of international standards for the transfer of clinical and administrative data between software applications used by various healthcare providers |
| HTE | Heterogeneity of Treatment Effect | A key definition in precision medicine. Def: The nonrandom, explainable variability in the direction and magnitude of individual treatment effects, including both beneficial and adverse effects. |
| ICF | Informed Consent Form | A paper that explains the course, risks, benefits, rights and other considerations of participating in a clinical study. |

| | | |
|-------|--|---|
| IOMT | Internet of Medical Things | |
| IRB | Institutional Review Board | Responsible for protecting the rights and welfare of the human subjects of research conducted by faculty and staff at the Institutions. |
| JHHS | Johns Hopkins Health System | A not-for-profit organization dedicated to providing the highest quality patient health care in the treatment and prevention of human illness. JHHS is an academically based health system consisting of Johns Hopkins Hospital, All Children's Hospital, Bayview Medical Center, Howard County General, Johns Hopkins Community Physicians, Johns Hopkins Medical Services Corp., and Johns Hopkins Medical Management Corp. |
| JHMI | Johns Hopkins Medical Institute | <i>see JHM</i> |
| JHM | Johns Hopkins Medicine | A collaboration between JHHS and JHU is a vehicle for internal operational coordination and a united voice for external initiatives. The main operational entities of JHM are Johns Hopkins HealthCare LLC and The Johns Hopkins Home Care Group, Inc. |
| JHU | Johns Hopkins University | A private research university located in Baltimore, MD. |
| JSON | JavaScript Object Notation | A lightweight data-interchange format. It is easy for humans to read and write. It is easy for machines to parse and generate. A common output of many RESTful APIs. |
| M2M | Minute-to-Minute | |
| MRN | Medical Record Number | Johns Hopkins Patient Identifier. Similar to an SSN. |
| MYO | <i>Myositis?</i> | |
| NLP | | |
| OCR | Optical Character Recognition | The process of transforming an image of text into an editable format. |
| PACE | Program to Accelerate Clinical Research | PACE at JH was established to help clinical researchers take full advantage of tools within the EPIC Electronic Health Record system for their research projects. |
| PACS | Picture Archiving & Communication System | The source for all medical imaging (eg. MRIs, X-Rays, etc...) |
| PCORI | Patient-Centered Outcomes Research Institute | PCORI helps people make informed healthcare decisions, and improves healthcare delivery and outcomes, by producing and promoting high-integrity, evidence-based information that comes from research guided by patients, caregivers, and the broader healthcare community. |
| PI | Program Increment | Pertains to product development cycle and tracking incremental progress |
| PINE | | |
| PMAP | Precision Medicine Analytics Platform | Brings together revolutions in measurement, connectivity, and data science to enable the discovery of clinically-relevant and biologically-anchored subgroups at scale. |
| RC | Redcap | <i>See Terms below</i> |
| RDBMS | Relational Database Management System | A database management system designed specifically for relational databases. |
| REST | Representation al State Transfer | A RESTful web application exposes information about itself in the form of information about its resources. It also enables the client to take action on those resources, such as create new resources (i.e. create a new user) or change existing resources (i.e. edit a post). |
| RTF | | |

| | | |
|---------|---------------------------------|--|
| SaaS | Software as a Service | A software distribution model in which a third-party provider hosts applications and makes them available to customers over the Internet. |
| SABER | | |
| SNP | Single Nucleotide Polymorphisms | The most common type of genetic variation among people. Researchers have found SNPs that may help predict an individual's response to certain drugs, susceptibility to environmental factors such as toxins, and risk of developing particular diseases. SNPs can also be used to track the inheritance of disease genes within families. |
| SQL | Structured Query Language | A domain-specific language used in programming and designed for managing data held in a relational database management system (RDBMS), or for stream processing in a relational data stream management system (RDSMS). It is particularly useful in handling structured data, i.e. data incorporating relations among entities and variables. |
| TriNetX | Alternative to Slicer Dicer | https://ctr.johnshopkins.edu/programs_resources/programs-resources/i2c/trinetx/ - TriNetX is an alternative to Slicer Dicer, a more accurate tool to estimate patient population sizes. Can help PMCoEs answer the 'estimated size of the repository' question in the IRB. |
| VCF | Variant Call Format | A text file used in bioinformatics for storing gene sequence variations. |
| vCOE | Virtual Center of Excellence | Responsible for the general interests of the broader PMAP program instead of the more focused disease-specific COEs. |
| XML | Extensible Markup Language | A markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable. |

Terms

| Term | Definition |
|----------------|---|
| Clarity | EPIC's database of EMR data. It is updated every 24 hours. |
| Cohort | A collection of patients representing a particular area of research or treatment. |
| Crunchr | An online solution for workforce reporting and people analytics. |
| Data Analyst | Expertise in building dashboards and turning data into actionable insights. |
| Data Commons | Co-locates data, storage, computing infrastructure, and commonly used tools for analyzing and sharing data to create a resource for the research community. |
| Data Engineer | Builds infrastructure to enable backup, indexing, capacity, search, analysis, computation, and scale. |
| Data Lake | A storage repository that holds a vast amount of raw data in its native format until it is needed. |
| Data Scientist | Expertise in making predictions with data. |
| Data Trust | |
| Data Warehouse | A system used for reporting and data analysis. Considered a core component of business intelligence. DWs are central repositories of integrated data from one or more disparate sources. They store current and historical data in one single place that is used for creating analytical reports for workers throughout the enterprise. |
| Delivery | |
| Discovery | |
| Docker | A tool designed to make it easier to create, deploy, and run applications by using containers. Containers allow a developer to package up an application with all of the parts it needs, such as libraries and other dependencies, and ship it all out as one package. |
| Drager | A supplier to ventilator machines that are able to transmit data into a database. |
| eForm-A | For studies involving research-specific interventions or interactions. |

| | |
|------------------------|---|
| eForm-R | The main IRB form all CoE's are required to submit unless they are utilizing research data already covered under an existing eForm-R. eForm-Rs (used for research resource protocols) are required for all protocols designed to cover the creation or set-up of research resources. Common examples of research resources are a data or biospecimen repository and a recruitment database. Generally the purpose of a research resource is to help support current and future research rather than to answer any discrete research questions. If the intent is to create a research resource, a separate IRB application is required for creation of the resource itself. This form is used to describe the resource a researcher is creating, how it will be managed and how it will be accessed for future research use. |
| eForm-E | A form used for projects thought to qualify for an exemption from obtaining patient consent. This may include, interviews, surveys, public observations, educational tests, research conducted in normal educational settings or some behavioral interventions, if the intervention is considered to be benign. The only exception is secondary research activities. While some secondary research activities may qualify for an exemption, for these projects, the eForm-S should be used. |
| eForm-S | This form is used for secondary research using identifiable private information and/or identifiable biospecimens. Data or samples are not required to be in existence at the time of the submission to the IRB in order to use this form. Projects involving secondary research of identifiable private information and/or identifiable biospecimens may qualify for an exemption or may require expedited review. |
| Ensemble | Intersystems |
| EPIC | The preferred electronic medical record system used by more than 250 health care organizations nationwide. To date, 45 percent of the US population have their medical records in an Epic system. |
| Galaxy | An open-source workflow engine enabling rapid analysis of computing infrastructure. Galaxy itself enables researchers to do reproducible analyses, an absolute necessity in the realm of science. Additionally, Galaxy allows the design and execution of complex MapReduce-type workflows, giving power users command of complex pipelines, and the ability to remix analyses. |
| Hadoop | A collection of open-source software utilities that facilitate using a network of many computers to solve problems involving massive amounts of data and computation. |
| HBase | The Hadoop database, a distributed, scalable, big data store. Its goal is the hosting of very large tables -- billions of rows X millions of columns -- atop clusters of commodity hardware. |
| Healthis | An end-to-end genetic information management system built to integrate directly with EHRs. |
| Hive | A data warehouse software project built on top of Apache Hadoop for providing data query and analysis. |
| Iguana | An HL7 Data aggregator |
| Interconnect | |
| Jupyter | Open Data Science at Scale. Web-based computational scientific notebooks. |
| Kafka | Used for building real-time data pipelines and streaming apps. |
| Leaf | An open-source, model-agnostic, data-driven web application for cohort discovery and translational biomedical research. |
| Learning Health System | A learning health system is a health system designed to generate and apply the best evidence for the collaborative healthcare choices of each patient and provider; to drive the process of discovery as a natural outgrowth of patient care; and to ensure innovation, quality, safety, and value in health care. (Institute of Medicine) |
| Longitudinal Viewer | A clinical application that provides an integrated, read-only display of health data in a common data viewer. |
| MariaDB | A popular open-source relational database. Stores ADT data for Johns Hopkins. |
| Mulesoft | |
| Pentaho | A data integration and analytics platform which enables organizations to access, prepare, and analyze all data from any source, in any environment. |
| Precision Medicine | Precision medicine is "an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person." It is in contrast to a one-size-fits-all approach, in which disease treatment and prevention strategies are developed for the "average" person, with less consideration for the differences between individuals. |
| Python | An interpreted, high-level, general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant whitespace. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects. |
| Redcap | A secure web application for building and managing online surveys and databases. A major source of JH research data. |
| Rstudio | |
| Smart Forms | An EMR-based, clinical workflow tool designed for organized data review for specific conditions, effective and efficient facilitated data capture, documentation of a clinical visit, and integrated, dynamic, actionable decision support in a single environment. |

| | |
|--------------|--|
| Tableau | A powerful data visualization tool used in the Business Intelligence Industry that simplifies raw data into a very easily understandable format. |
| Vent Data | |