CITIZEN SCIENTIST PROGRAM

APPENDIX B: GLOSSARY OF COMMONLY USED TERMS

Algorithm: A step-by-step set of operations (mathematical equations) used to perform reasoning, or a set of rules that defines a sequence of operations. An example of an algorithm applied to health care research is the computable phenotype.

Bioinformatics: biomedical informatics; information and data related to patients and their medical records. This is similar to "IT" or information technology, but the ideas, techniques, hardware and software are used with health information and applied to asking and answering questions about health care delivery and patient outcomes.

Common Data Element: Standardized pieces of data that are expected to be collected and are then used across networks. Common data elements are collected and formatted in the same way in different setting so the data that are collected may be pooled together "aggregated". For example, without creating a "common data element" with rules for how to format, sex might be called gender, M and F might be used instead of male and female. When someone puts together data elements that are not standardized, it becomes much more difficult to understand the information. When you have thousands of data elements to aggregate for millions or billions of people, the information may be impossible to understand.

Common Data Model: Standardizes the definition, content, and format of data across networks to enable a single standardized view that can be used for querying. A common data model builds from the idea of a common data element and applies to the entire dataset, providing rules for what information should be included in the dataset, along with how each common data model should be collected and formatted. Common data models are used to combine information from a variety of settings into one cohesive, understandable dataset. These datasets can be very large and powerful, allowing researchers to use information from millions of people to ask and answer questions that might be relevant to populations. Additionally, these datasets may be able to find enough people in sub-groups or with very rare conditions to be able to carry out research that is normally impossible.

Comparative Effectiveness Research: The direct comparison of existing health care interventions to determine which work best for which patients and which pose the greatest benefit or harm. A common example might be determining whether chemotherapy or radiation is the more effective treatment for a specific kind of cancer.

Computable Phenotype: An electronic algorithm used to identify certain conditions or characteristics. This process can be used to identify cohorts of individuals for studies. The algorithm uses a series of logical steps ("computable") to decide whether someone is in the group you are searching for (the "phenotype"). Asthma is a good example because it can be difficult to identify a patient with an asthma diagnosis from a medical record or insurance claims and encounter data. A patient with only one diagnosis code of asthma may not actually have asthma when you ask the patient's treating provider. An algorithm would use claims and encounter data to (1) find all patients with at least one diagnosis code of asthma, THEN (2) exclude patients that have asthma diagnosis codes from longer than 5 years ago, THEN (3) exclude patients that have less than two diagnosis codes of asthma, THEN (4) include only patients that have diagnosis codes of asthma spanning at least two years. The patients who are left in the group after running the algorithm are much more likely to have asthma when you ask their doctor than the patients in the group before you ran the algorithm. This is a computable phenotype, meaning it is possible to create a series of mathematical and logical steps to run on health record data ("computable") that will end up with a result similar to what a doctor would report in the exam room (the "phenotype"). [This is just a simple example that is not exactly the same as the algorithm that would actually be used with real patient data].

Consent2Share: A database of people who are willing to have their medical records flagged as someone who is interested in hearing about research opportunities, and then allow UF Health to share their name and contact information with UF researchers when their medical records show they might qualify for a future research study.

Continuity of Care Document: An electronic document exported to summarize patient care (Similar to a CCR or Continuity of Care Record)

Claims Data: Payment information provided by your insurance carrier for payable medical, dental, prescription, (and other) benefits.

False-Positives: Data showing up as true/positive when test and experiences prove it as false. Example is that there are more kids with asthma diagnoses than kids who actually have asthma. This is due to one episode of some sort of respiratory distress, which gets diagnosed as asthma when there is no established pattern of respiratory distress. Please see the example for "computable phenotype" for more contextual information about the implication of this when looking at large datasets of health information (health information used could be claims data from an insurance company or health record data from a health care provider).

Genotype: The inherited genetic instructions of an organism. Your actual DNA code is your genotype. The way that is "expressed" is your phenotype, meaning how the genes translate into how your body actually looks or works.

Grant Proposals: A response to a call for proposals issued by a funding agency in which the researcher explains the process of how they will gather the information the funding agency is seeking

Grant Reviewers: Individuals who are qualified, based on knowledge, education or experience, to evaluate submitted proposals for scientific and technical merit in relation to the call for proposals, which outline the information the funding agency seeks.

Heterogeneity: Refers to genetic and other differences between people.

Homogeneous: The opposite of heterogeneous. Composed of parts that are all the same.

Honest Broker: An appointed entity that keeps private information, but distributes parts of those sets of data to other entities (researchers) who do not need all of the information. The purpose of this is to protect identity, and this person (or persons) is also a decision maker in whether to allow use of the existing data or materials.

Hypertension: High Blood Pressure

i2b2: NIH funded National Center for Biomedical Computing. A scalable informatics framework that enables researchers to use existing clinical data for discovery research and facilitate designs of targeted therapies. Universities use i2b2 to get "aggregate" information about the patients in their integrated data repositories for "preparatory to research queries". In other words, this is a software tool that helps researchers understand whether there are enough patients within the health system to make their proposed research feasible. For example, a research team wants to conduct a study that needs 500 patients who are at least 18 years old, have Type 1 diabetes, and have consented to be re-contacted for research. The team can use i2b2 to ask how many patients have these characteristics (i.e. meet the "eligibility criteria") to find out whether it is possible to carry out the research successfully in the proposed clinical setting. In this case, the research team gets only "aggregate" data. They will get approximately (but not exactly to protect patient privacy) the total number of patients that meet the criteria, but the team does not receive any "patient-level" or "person-level" data. That is, the research team will not get information on any specific person. In cases where there are fewer than 10 people who meet the criteria, the research team will be told only that the number is too small to report. These protections are in place to protect patient privacy and confidentiality and are built into the way the i2b2 software works automatically.

medical decision making, practices, and/or products being tailored to the individual patient are the basis for personalized medicine. This concept is also sometimes referred to as “precision medicine”.

Pharmacogenetics: Pharmaceutical (“drug”) treatment choices are tailored or customized to the individual patient using genetic information.

Phenotype: An organism’s observable characteristics or traits resulting from the expression of genetic material coupled with the environmental factors the organism is exposed to. Blue eyes and blond hair are observable phenotypical characteristics, for example.

PopMedNet: Software application that enables creation, operation and governance of distributed health data networks. This system allows other user networks to collaborate on public health research projects while allowing the information owners to protect their confidential information.

Pragmatic Clinical Trials: Pragmatic trials are designed to evaluate the effectiveness of interventions in real-world settings and routine practice conditions, whereas explanatory trials aim to test whether an intervention works under optimal situations.

Query: The mechanism for getting information from a database. They are questions that are asked of the database in a predefined format, written in a language that the computer understands in order to produce Information that is valuable to the questioner.

Review Panels: Groups of individuals who actively engage in discussing various portions of a grant proposal. They identify strengths and weaknesses of an application in relation to the call for proposals.

RxNorm: Normalizes names for clinical drugs; links names to commonly used drug vocabulary. This is another example of an “ontology” like LOINC. LOINC is used to name and define laboratory data. RxNorm is used to name and define prescription drug information.

Scalable: This term refers to ability of the concept to be repeated or adopted in settings other than those in which it was first tested. For example, implementation science is interested in conducting “pragmatic” research, or research carried out in real world settings that is also “scalable”, meaning that if the research study is found to be successful, it can be translated into similar clinical settings across the region, state or country. Thinking about how scalable the research study is from the design phase will maximize its impact on clinical care and patient outcomes.

State Agency Directors: The Directors of the Agencies of the Department of Health, Centers for Medicare and Medicaid Services, the Centers for Disease Control and Prevention, etc.

Tobacco Cessation: The act of quitting the use of tobacco, whether through the use of cigarettes, snuff, or other tobacco products.