

UC Davis

UC Davis Previously Published Works

Title

Big Data for Nutrition Research in Pediatric Oncology: Current State and Framework for Advancement

Permalink

<https://escholarship.org/uc/item/2wp8n507>

Journal

JNCI Monographs, 2019(54)

ISSN

1052-6773

Authors

Phillips, Charles A

Pollock, Brad H

Publication Date

2019-09-01

DOI

10.1093/jncimonographs/lgz019

Peer reviewed

ARTICLE

Big Data for Nutrition Research in Pediatric Oncology: Current State and Framework for Advancement

Charles A. Phillips, Brad H. Pollock

See the Notes section for the full list of authors' affiliations.

Correspondence to: Charles Phillips, MD, MSHP, Children's Hospital of Philadelphia, Roberts Center for Pediatric Research, Applied Clinical Research Center, 11th Floor, 11392, 2716 South Street, Philadelphia, PA 19146-2305 (e-mail: phillips2@email.chop.edu).

Abstract

Recognition and treatment of malnutrition in pediatric oncology patients is crucial because it is associated with increased morbidity and mortality. Nutrition-relevant data collected from cancer clinical trials and nutrition-specific studies are insufficient to drive high-impact nutrition research without augmentation from additional data sources. To date, clinical big data resources are underused for nutrition research in pediatric oncology. Health-care big data can be broadly subclassified into three clinical data categories: administrative, electronic health record (including clinical data research networks and learning health systems), and mobile health. Along with -omics data, each has unique applications and limitations. We summarize the potential use of clinical big data to drive pediatric oncology nutrition research and identify key scientific gaps. A framework for advancement of big data utilization for pediatric oncology nutrition research is presented and focuses on transdisciplinary teams, data interoperability, validated cohort curation, data repurposing, and mobile health applications.

Malnutrition, including under- and overnutrition, is common in pediatric cancer and affects up to 78% of patients before or during cancer treatment (1,2). Recognition and treatment of malnutrition is vital because it is associated with poorer outcomes for survival, wound healing, physical and cognitive development, immune function, and quality of life (1-3). Research to prevent or ameliorate malnutrition in children with cancer has been hindered by data limitations including the lack of access to nutrition data and difficulties combining data across studies. Nutrition data are often collected without commonly defined data elements in disparate data structures and storage systems (4). This limits the ability to combine data across studies and establish larger data repositories to address complex nutrition research questions (5). Furthermore, most studies of nutrition support in pediatric oncology predate use of standardized diagnostic indicators for malnutrition defined in 2014 by the Academy of Nutrition and Dietetics and the American Society for Parenteral and Enteral Nutrition (6). These data limitations are at least partially responsible for the failure to close important knowledge gaps for evidence-based nutrition supportive care within pediatric oncology. Moving forward, maintaining the status quo with nutrition and anthropometric data collected

from clinical trials, cancer registries, and nutrition studies will likely be insufficient to drive meaningful nutrition research without augmentation from additional data sources. Incorporating big data into nutrition research can address many of the current data limitations and accelerate future discoveries.

Big data sources exist outside traditional nutrition and supportive care research models for pediatric oncology. *Healthcare big data* is defined as the digital data assets contained as structured or unstructured data generated from experiments, direct patient care by health-care providers, or patients themselves (7). The importance of big data has been recognized through the creation of programs such as the National Institute of Health's Big Data to Knowledge initiative and the Patient-Centered Outcomes Research Institute's (PCORI) Clinical Data Research Network and its PCORnet infrastructure (8). These initiatives were designed to advance the methods and resources available to rapidly improve health-care research and delivery.

As a previously untapped resource, healthcare big data can aid in improving nutrition-related outcomes in pediatric oncology. We discuss advantages, research design considerations, and challenges in using different big data sources to increase

Table 1. Big data sources

Type	Discipline expertise	Methods considerations
Omics	<ul style="list-style-type: none"> • Genetics • Molecular/cellular biology • Bioinformatics • Biostatistics 	Experimental design (translational investigation study design); statistical modeling (time series and nested within-subjects effects); linkage of biology data with clinical data
Administrative	<ul style="list-style-type: none"> • Health services research • Applied epidemiology • Population health science • Health policy research • Health economics • Medical informatics • Health information Technology 	Identification of patient cohorts; reduction of biasing observational data; missing data; secular changes in coding (impacting longitudinal datasets); propensity scores; instrumental variables
Electronic health record*	<ul style="list-style-type: none"> • Medical informatics • Health information technology • Clinical domain experts • Health economics • Finance • Psychometrics 	Definition and validation of patient cohorts and clinical phenotyping (constructing synthetic cohorts with data points over time including entry into a health system); missing and inaccurate data; secular changes in coding; propensity scores; instrumental variables; patient reported outcomes; unstructured data and natural language processing
mHealth (quantified self)	<ul style="list-style-type: none"> • Medical informatics • Health information technology • Communications • Behavioral economics • Electrical engineering 	Validation of patient cohorts; ontologies for mHealth data; linkage between EHRs and research data repositories; missing and miscalibrated data; signal processing and pattern recognition; repeated-measures statistical analyses; precision public health (targeted study populations)

*Electronic health record-derived big data sources include clinical data research networks and learning health systems.

the research pipeline (Table 1). When available, prior nutrition research in oncology is cited. The importance of expanding nutrition research and incorporating big data is partially driven by the dearth of nutrition examples for several of the topics discussed. Finally, we present a framework to incorporate big data into future oncologic nutrition research.

Clinical Big Data Relevant to Pediatric Oncology Nutrition Research

Administrative Data

Administrative data, including population health registries, census, vital statistics, and medical claims data, are often created as a by-product of administering health surveillance programs, reimbursing health-care services, or other governmental and/or regulatory functions. Administrative data are attractive for research because they are often readily available, relatively inexpensive, structured, and continuously curated and cover large segments of the population (9). Administrative data are made available for research by public or private entities (eg, the Research Data Assistance Center from the Centers for Medicare and Medicaid, the Kids Inpatient Database, or the Pediatric Health Information System). The Pediatric Health Information System has been used to conduct pediatric oncology research outside the clinical trials infrastructure to address topics such as racial disparities (10) and off-study immunotherapy use (11). Administrative (group-level) data can be combined with individual data, for example, proximity to grocery stores or food deserts can be integrated with individual data for nutritional studies.

Administrative data pose multiple challenges. When used in isolation, diagnosis codes do not have sufficient specificity for case identification for multiple conditions including pediatric malignancies (12,13). In addition to difficulties creating valid

study cohorts, limitations of administrative data may include more missing and misclassified information (eg, *International Classification of Diseases, Ninth Edition [ICD9]* to ICD10 coding changes) and discrepant follow-up impacting estimates of survival and cancer recurrence. These concerns can be partially mitigated through optimal study design, enhanced technologies to improve data collection, validation of exposure and outcome data elements, and the use of appropriate statistical methods (eg, multivariate modeling, multiple imputation, propensity scores, and instrumental variable estimation).

Electronic Health Record Data

In contrast to protocol-derived research cohort study databases, there is growing use of electronic health records (EHRs) to support clinical and translational research. Meaningful use requirements for EHRs have placed an emphasis on collection of structured data, which should lead to more clinical research opportunities (14). This, coupled with the development of large-scale clinical data research networks (CDRNs) linking EHR data from multiple institutions, represents a massive investment in data infrastructure, which complements administrative and clinical trial data. EHR data integrated into CDRNs can form a critical component for learning health systems (LHS). LHS integrate structured EHR data, research done in routine care settings, and quality improvement processes for the purpose of rapidly advancing new knowledge (15). These research systems are attractive because they can link critical components of the EHR including patient notes, diagnoses, medications, laboratory values, and imaging data.

EHR data, CDRN, and LHS are particularly attractive for nutrition research because they capture anthropometric and nutrition intervention data. Beyond anthropometric data, EHRs often contain other important nutrition indicators such as mid-upper arm circumference, estimated caloric intake and needs, and relevant clinical laboratory biomarkers. Studies of antibiotic exposure and

growth during childhood illustrate how longitudinal EHR-derived data can inform observational nutrition research. First, researchers demonstrated that aggregated anthropometric EHR data mirrored the findings of the reputable National Health and Nutrition Examination Survey (16). Next, single institution EHR data demonstrated a link between antibiotic exposure and obesity (15). Finally, the single institution study results were confirmed using a national sample contained in PCORnet, a PCORI-funded CDRN (17). By exploiting routinely collected data, EHR-based studies can be an economical alternative to large cohort studies. For example, the PCORI-sponsored ADAPTABLE Trial (Aspirin Dosing: A Patient-centric Trial Assessing Benefits and Long-Term Effectiveness) demonstrates how data from EHRs can directly support clinical research by utilizing the EHR data to optimize enrollment and track outcomes for patients participating in this pragmatic clinical trial (18).

Study-specific databases do have some advantages compared to EHR-derived databases, including higher quality data, validation, and strict eligibility requirements, and are assured of collecting data elements vital to addressing the study hypotheses. However, they are typically expensive to construct. In contrast, EHR-derived databases are attractive because they may be more representative of real-world populations (no eligibility requirements), larger (number of individuals represented), and much less expensive.

EHR databases, CDRNs and LHS contain much of the information present in administrative databases as well as a treasure trove of additional data. However, they retain similar limitations, including bias, missing data, and nonuniform diagnostic coding, which render construction of valid historical study cohorts challenging. Accurate cohort construction in CDRN and LHS can be facilitated by utilizing computable phenotypes, which can leverage additional EHR information present in the CDRN and LHS (19). Within this context, a computable phenotype is a machine-evaluable definition for a given condition based on standard terminology (eg, SNOMED diagnosis codes) and clinical features that can be determined from available EHR data. Computable phenotypes should be validated whenever possible.

Previous studies demonstrate the need for rigorous quality assurance practices when utilizing EHR data (20,21). Anthropometric data pertinent to nutrition research are especially prone to erroneous values, and algorithms to clean these data have been developed (22). Perhaps the biggest new challenge of using EHR data for research is wrangling it from various EHR vendors into a harmonized common data model (CDM) that runs across the CDRNs or LHS. Ontologies for clinical research and CDMs have been developed to address these concerns (23), but conforming to these solutions is labor intensive. Other challenges are integrating standards for data that typically live outside the EHR such as patient-reported outcomes. Finally, many EHR-derived networks contain structured and unstructured data. Natural language processing (NLP) can be highly useful to convert unstructured EHR data into structured data. However, NLP approaches may still not yield results sufficiently complete and accurate for clinical and translational research.

Mobile Health and Quantified Self-Data

Mobile health (mHealth), the use of mobile and wireless technologies for health, aims to capitalize on the rapid uptake of information and communication technologies to improve

health-system efficiency and health outcomes (24). This encompasses an incredibly wide range of devices from mobile phones, smart watches, and other wearable devices. mHealth data includes quantified self-data from individuals engaged in the self-tracking of any kind of biological, physical, behavioral, or environmental information (25). These technologies potentially allow researchers to view patients in their “natural environment” and are believed to improve understanding of person-specific disease risk factors and treatment response (26).

mHealth holds particular promise for advancing nutrition research for pediatric oncology. Nutritionally relevant mHealth applications include electronic food diaries, home-health device monitoring such as feeding pump summary information, and continuous glucose monitoring. Furthermore, established dietary assessment methods in epidemiological studies are ripe for adaptation into an mHealth format including duplicate diet approaches, food consumption records, 24-hour dietary recalls, dietary records, dietary histories, and other food frequency questionnaires (27).

Challenges for mHealth data include integration with the EHR (28) and creating mHealth data with an architecture that allows the information to be used appropriately (19,29). Aside from privacy and integrity concerns, validation of mHealth data is required prior to integration; standards for this are not universally defined. These and other challenges posed by mHealth research can be addressed and mitigated by using mHealth reporting guidelines published by the World Health Organization (24).

Omics Data

In addition to clinical big data, -omics data are also important for nutrition research in pediatric oncology. Omics refers to the study of biologic fields ending in *omics* such as genomics including nutrigenomics (genetic variants that influence metabolism of specific nutrients), proteomics, or metabolomics and can be expanded to include informatics or statistical tools germane to biology subfields. For nutrition specifically, the microbiome is likely to have significant relevance to -omics research. Although not detailed here, the incorporation of genetic and other biologic information into patient care as well as clinical and translational research is crucial to advance precision medicine approaches to nutrition care. The framework presented below, with goals such as data interoperability, are also applicable to -omics data.

Framework for Clinical Big Data Utilization for Pediatric Oncology Nutrition Research

The overarching challenge of utilizing big data for nutrition research within pediatric oncology is that data are spread across multiple entities including laboratory data, the EHR, administrative data warehouses, cancer registries, and research protocol data repositories. We propose a framework with the following five key components to address these issues and inform a path forward for nutrition research within pediatric oncology.

Form Transdisciplinary Research Teams

Effective big data research requires transdisciplinary research teams that possess discipline expertise in study design and can also prospectively address challenges such as data linking.

For example, the ADAPTABLE Trial used expertise across the informatics and clinical domains to conduct a prospective-controlled randomized trial that compared the effectiveness of two doses of aspirin for postmyocardial infarction patients. ADAPTABLE used the EHR as both the means of identifying all participants and the principal source of the primary clinical trial data and was supplemented by the collection of patient-reported outcome data (18,30).

Support Big Data Infrastructure and Interoperability

The single most important step toward big data utilization is employing consistent ontologies between studies. A CDM for clinical research has been shown to enable linking of observational data across different studies and clinical research databases (31). However, a universal CDM for nutrition research is likely infeasible because of the large amount of resources this would require. Even if a CDM is not achievable, developing core common data elements that are shared between studies is a realistic long-term goal; for example, data dictionaries, questionnaires, and protocols could be shared with unified formatting. Within the nutrition research and informatics community, work has been done to harmonize terms for nutrition supportive care and diagnoses (32,33). The research community and EHR vendors should work toward utilization of common data elements that could be linked across studies. Important elements on the path toward larger-scale data integration at the patient level include privacy, security, deidentification, anonymization, and informed consent. At the institutional level, these elements include data use agreements, secure data transfers, and collaborative network participation.

In North America, the Children's Oncology Group has the potential to take a lead in data infrastructure as the largest purveyor of pediatric oncology research. Standardizing a common core of anthropometric and nutrition data elements with shared variable names for distinct Children's Oncology Group studies would open the door for powerful nutrition studies in the future. The authors recognize the call for standards is not novel; calls for consistency with study design parameters within pediatric cancer cooperative groups have been made for more than 25 years (34). We are instead advocating for better recognition and integration of digital data standards across oncology studies so that nutrition and supportive care research can be performed more efficiently.

Validate Cohorts

A common challenge for administrative, EHR, and mHealth data is the identification and curation of large, often pooled, validated cohorts. In contrast to single cohorts constructed for research investigations, cohorts created by merging available data are at increased risk for bias. Ideally, data harmonization and the use of common data dictionaries minimize error and reduce heterogeneity around definitions of exposures, covariates, outcomes, and modeling approaches. In the absence of interoperable data, computable phenotypes and probabilistic matching techniques have the potential to link the same, or similar, patient(s) across disparate studies. Providing evidence for nutrition interventions by combining individual-level data across studies (pooled analyses) provides a more powerful approach than combining only study results (eg, meta-analyses). Pooled analyses must use appropriate statistical methods that account for heterogeneity across different data sources such as study

covariate adjustment, modeling within study and between study effects, and splitting methods such as recursive partitioning. When appropriately validated and analyzed, pooled data can provide dramatically increased statistical power and stronger weight of evidence for best nutrition supportive care practices.

Repurpose Data

A significant portion of administrative, EHR, and mHealth data is not collected for nutrition or supportive care research but can often be repurposed. For example, a recent feasibility study demonstrated that anthropometric data in the EHR can be used in real-time as a malnutrition screen for pediatric oncology patients (35). For research database construction, there should be an emphasis on potential reuse and repurposing of big data for pediatric patients; particularly germane for long-term longitudinal tracking of pediatric patients across different stages of development.

Explore mHealth and Quantified-Self Research Opportunities

Nutrition supportive care research is an attractive substrate for quantified self and mHealth research [eg, a mobile phone application helped optimize nutrition behaviors for adult oncology patients (36)]. Similar studies have not been performed for pediatric patients. Pediatric oncology will face additional mHealth challenges in that interventions must be developmentally appropriate (ie, young children may not be able to enter data onto a smart phone or other device).

To date, clinical big data resources are underused for nutrition research in pediatric oncology. Although big data is not a panacea for all that ails nutrition research, addressing the data issues described in this paper are key to promoting data liquidity (37) which will accelerate the conduct of nutrition research in pediatric oncology. Specifically, it would promote more accurate and efficient cohort construction, allow for more accurate estimates of study parameters during planning phases, lead to increased accrual (with increased population representativeness), increase the precision of statistical analyses, and ultimately, more rapidly disseminate study findings to impact outcomes. As a research community, we have an obligation to maximize the use of big data to improve nutrition outcomes for our pediatric oncology patients.

Funding

Brad H. Pollock received support from the Children's Oncology Group NCI Community Oncology Research Program Research Base grant (UM1 CA189955) and the University of California Davis Comprehensive Cancer Center grant (P30 CA093373). Charles A. Phillips received grant support from the Children's Hospital of Philadelphia Foerderer Award Fund.

Notes

Affiliations of authors: Division of Oncology, Children's Hospital of Philadelphia, Philadelphia, PA (CAP); Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA (CAP); Department of Public

Health Sciences, School of Medicine, University of California, Davis, CA (BHP); University of California Davis Comprehensive Cancer Center, Sacramento, CA (BHP).

The authors have no conflicts of interest.

References

- Co-Reyes E, Li R, Huh W, Chandra J. Malnutrition and obesity in pediatric oncology patients: causes, consequences, and interventions. *Pediatr Blood Cancer*. 2012;59(7):1160–1167.
- Iniesta RR, Paciarotti I, Brougham MF, McKenzie JM, Wilson DC. Effects of pediatric cancer and its treatment on nutritional status: a systematic review. *Nutr Rev*. 2015;73(5):276–295.
- Brinksma A, Sanderman R, Roodbol PF, et al. Malnutrition is associated with worse health-related quality of life in children with cancer. *Support Care Cancer*. 2015;23(10):3043–3052.
- Gaynor EP, Sullivan PB. Nutritional status and nutritional management in children with cancer. *Arch Dis Child*. 2015;100(12):1169–1172.
- Zhang FF, Liu S, Chung M, Kelly MJ. Growth patterns during and after treatment in patients with pediatric ALL: a meta-analysis. *Pediatr Blood Cancer*. 2015;62(8):1452–1460.
- Becker P, Carney LN, Corkins MR, et al. Consensus statement of the Academy of Nutrition and Dietetics/American Society for Parenteral and Enteral Nutrition: indicators recommended for the identification and documentation of pediatric malnutrition (undernutrition). *Nutr Clin Pract*. 2015;30(1):147–161.
- Bourne PE. What big data means to me. *J Am Med Inform Assoc*. 2014;21(2):194.
- Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. *J Am Med Inform Assoc*. 2014;21(4):578–582.
- Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Qual Saf Health Care*. 2003;12(suppl 2):ii58–63.
- Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature*. 2009;457(7232):1012–1014.
- DiNofia AM, Salazar E, Seif AE, et al. Bortezomib inpatient prescribing practices in free-standing children's hospitals in the United States. *PLoS One*. 2016;11(3):e0151362.
- Citrin R, Horowitz JP, Reilly AF, et al. Creation of a pediatric mature B-cell non-Hodgkin lymphoma cohort within the Pediatric Health Information System Database. *PLoS One*. 2017;12(10):e0186960.
- Phillips CA, Razzaghi H, Aglio T, et al. Development and evaluation of a computable phenotype to identify pediatric patients with leukemia and lymphoma treated with chemotherapy using electronic health record data [published online ahead of print June 17, 2019]. *Pediatr Blood Cancer*. 2019; e27876.
- Friedman CP, Wong AK, Blumenthal D. Achieving a nationwide learning health system. *Sci Transl Med*. 2010;2(57):57cm29.
- Bailey LC, Forrest CB, Zhang P, Richards TM, Livshits A, DeRusso PA. Association of antibiotics in infancy with early childhood obesity. *JAMA Pediatr*. 2014;168(11):1063–1069.
- Bailey LC, Milov DE, Kelleher K, et al. Multi-institutional sharing of electronic health record data to assess childhood obesity. *PLoS One*. 2013;8(6):e66192.
- Block JP, Bailey LC, Gillman MW, et al. PCORnet antibiotics and childhood growth study: process for cohort creation and cohort description. *Acad Pediatr*. 2018;18(5):569–576.
- ADAPTABLE, the Aspirin Study—A Patient-Centered Trial. <https://theaspirinstudy.org/>. Accessed February 13, 2019.
- Richesson RL, Smerek MM, Blake Cameron C. A framework to support the sharing and reuse of computable phenotype definitions across health care delivery and clinical research applications. *EGEMS (Egms)*. 2016;4(3):1232.
- Khare R, Ruth BJ, Miller M, et al. Predicting causes of data quality issues in a clinical data research network. *AMIA Jt Summits Transl Sci Proc*. 2018;2017:113–121.
- Qualls LG, Phillips TA, Hammill BG, et al. Evaluating foundational data quality in the national patient-centered clinical research network (PCORnet(R)). *EGEMS (Egms)*. 2018;6(1):3.
- Daymont C, Ross ME, Russell Localio A, Fiks AG, Wasserman RC, Grundmeier RW. Automated identification of implausible values in growth data from pediatric electronic health records. *J Am Med Inform Assoc*. 2017;24(6):1080–1087.
- Sim I, Carini S, Tu S, et al. The human studies database project: federating human studies design data using the ontology of clinical research. *AMIA Jt Summits Transl Sci Proc*. 2010;2010:51–55.
- Agarwal S, LeFevre AE, Lee J, et al. Guidelines for reporting of health interventions using mobile phones: Mobile health (mHealth) evidence reporting and assessment (mERA) checklist. *BMJ (Clin Res Ed)*. 2016;352(8049):i1174.
- Swan M. The quantified self: fundamental disruption in big data science and biological discovery. *Big Data*. 2013;1(2):85–99.
- Kumar S, Abowd GD, Abraham WT, et al. Center of excellence for mobile sensor data-to-knowledge (MD2K). *J Am Med Inform Assoc*. 2015;22(6):1137–1142.
- Shim JS, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. *Epidemiol Health*. 2014;36:e2014009.
- Chen C, Haddad D, Selsky J, et al. Making sense of mobile health data: an open architecture to improve individual- and population-level health. *J Med Internet Res*. 2012;14(4):e112.
- Estrin D, Sim I, Health CD. Open mHealth architecture: an engine for health care innovation. *Science (New York, NY)*. 2010;330(6005):759–760.
- Faulkner M, Alikhaani J, Brown L, et al. Exploring meaningful patient engagement in ADAPTABLE (Aspirin Dosing: A Patient-centric Trial Assessing Benefits and Long-term Effectiveness). *Med Care*. 2018;56(suppl 10, suppl 1):S11–S15.
- Voss EA, Makadia R, Matcho A, et al. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. *J Am Med Inform Assoc*. 2015;22(3):553–564.
- Gabler GJ, Coenen M, Bolleers C, et al. Toward harmonization of the nutrition care process terminology and the international classification of functioning, disability and health-dietetics: results of a mapping exercise and implications for nutrition and dietetics practice and research. *J Acad Nutr Diet*. 2018;118(1):13–20.e13.
- Yuill K. *Report on Knowledge and Use of a Nutrition Care Process & Standardised Language by Dietitians in Europe*. Netherlands: European Federation of the Associations of Dietitians (EFAD); 2012.
- Pollock BH. Quality assurance for interventions in clinical trials. Multicenter data monitoring, data management, and analysis. *Cancer*. 1994;74(suppl 9):2647–2652.
- Phillips CA, Bailer J, Foster E, et al. Implementation of an automated pediatric malnutrition screen using anthropometric measurements in the electronic health record [published online ahead of print Oct. 5, 2018]. *J Acad Nutr Diet*. 2018. doi: 10.1016/j.jand.2018.07.014.
- Orlemann T, Reljic D, Zenker B, et al. A novel mobile phone app (oncofood) to record and optimize the dietary behavior of oncologic patients: pilot study. *JMIR Cancer*. 2018;4(2):e10703.
- Kean MA, Abernethy AP, Clark AM, et al. Achieving data liquidity in the cancer community: proposal for a coalition of all stakeholders. *Natl Acad Sci*. 2012.