ScienceDirect

Contents lists available at **sciencedirect.com**
Journal homepage: **www.elsevier.com/locate/jval**

Themed Section: Data, Privacy, and Health

A Rapid Review on the Value of Biobanks Containing Genetic Information

Elisabet Rodriguez Llorian, PhD, Nicola Kopac, MPH, Louloua Ashikhusein Waliji, BA, Kennedy Borle, MSc, Nick Dragojlovic, PhD, GenCOUNSEL, Alison M. Elliott, PhD, Larry D. Lynd, PhD

ABSTRACT

Objectives: Increasing access to health data through biobanks containing genetic information has the potential to expand the knowledge base and thereby improve screening, diagnosis, and treatment options for many diseases. Nevertheless, although privacy concerns and risks surrounding genetic data sharing are well documented, direct evidence in favor of the hypothesized benefits of data integration is scarce, which complicates decision making in this area. Therefore, the objective of this study is to summarize the available evidence on the research and clinical impacts of biobanks containing genetic information, so as to better understand how to quantify the value of expanding genomic data access.

Methods: Using a rapid review methodology, we performed a search of MEDLINE/PubMed and Embase databases; and websites of biobanks and genomic initiatives published from 2010 to 2022. We classified findings into 11 indicators including outputs (a direct product of the biobank activities) and outcomes (changes in scientific and clinical capacity).

Results: Of 8479 abstracts and 101 gray literature sources were reviewed, 96 records were included. Although most records did not report key indicators systematically, the available evidence concentrated on research indicators such as publications and gene-disorder association discoveries (63% of studies), followed by research infrastructure (26%), and clinical indicators (11%) such as supporting the diagnosis of individual patients.

Conclusions: Existing evidence on the benefits of biobanks is skewed toward easily quantifiable research outputs. Measuring a comprehensive set of outputs and outcomes inspired by value frameworks is necessary to generate better evidence on the benefits of genomic data sharing.

Keywords: biobanks, genetic disease, value framework.

VALUE HEALTH. 2023; 26(9):1286–1295

Introduction

Health-related data repositories containing human genetic information (and more broadly big data) have experienced substantial growth in the last decade.[1,2] The global healthcare data storage market had grown to > $3 billion USD in 2020,[3] and it is estimated that approximately 25 petabytes of genomic data will be produced worldwide annually by 2023.[1] Biobanks, which are defined as a collection of biological and linked health information with the goal of supporting research in medicine and health,[4,5] are a common form of data repository that varies in size, research topic, specimens collected, users, procedures for sample collection and storage, and other attributes.[4,6] There is an implicit assumption that there is substantial value in integrating this type of data and making it broadly accessible to researchers and clinicians. Nevertheless, a comprehensive description of the expected benefits from advances in scientific knowledge derived from accessing and integrating data through biobanks had not been systematically addressed until recently.[7] In contrast, the data privacy concerns and risks associated with greater genomic data access and sharing through the establishment and maintenance of biobanks are well documented in the literature.[8–10] Therefore, the paucity of empirical evidence for the hypothesized direct and indirect benefits of biobanks[7] makes demonstrating the value of biobanks (in broad terms, their efficiency to deliver benefits with respect to financial and nonfinancial costs[11]) challenging, even though balancing the benefits with privacy risks and other costs is crucial to fostering their sustainability and societal acceptance.[12–14]

Although the value framework proposed by Rush et al[7] provides some initial guidance on characterizing potential benefits by outlining a set of indicators primarily focused on outputs such as publications, patents, grants, collaborations, conference presentations, and projects supported by the biobank, privacy concerns and risks should ultimately be balanced against concrete outcomes (eg, scientific discoveries, changes in clinical management). From a program evaluation perspective, *outputs*, defined as the direct products or services derived from the activities of the biobank, differ from *outcomes*, which generally describe measurable changes in capacity (scientific, clinical, or other) that are derived from outputs or other outcomes and that represent

concrete results[15] such as a measurable increase in the scientific knowledge base, or changes in clinical processes such as diagnosis and treatment.

Therefore, the objective of this study was to summarize the available evidence on the outputs and outcomes resulting from biobanks containing genetic information, in an effort to complement and build upon the conceptual literature on value frameworks for biobanks[7,16-18] and the emerging evidence on biobank outputs and outcomes for specific conditions (eg, rare diseases[19] and cancer care[20]). Note that, throughout this study, "outcome" is used interchangeably with "impact" (which in program evaluation often refers to "ultimate outcomes" that reflect changes in the state, conditions, or wellbeing of the ultimate beneficiaries of an initiative). Instead, we emphasize the differences between outputs and outcomes (whether intermediate or ultimate) and use the term "indicator" to refer to a unit of measurement for both outcomes and outputs. Given the high degree of heterogeneity among relevant studies (which precluded the use of critical appraisal checklists or quantitative evidence synthesis), as well as the existing paucity of collected evidence in this area, we identified a rapid review methodology as the most appropriate and resource efficient approach. Results from this review will provide an overarching perspective on the current evidence landscape on the benefits of biobanks and highlight those areas with the most need for further evidence generation.

Methods

This study followed a rapid review methodology using recommendations provided by the Cochrane Rapid Review Methods Group,[21] with the aim of summarizing the literature in a systematic and reproducible manner while being resource efficient and streamlining the review process.[22,23] Before initiating the study, a scan was conducted on PROSPERO and Cochrane on July 19, 2021 and no similar protocols were found.

Biobank Definition

We used a broad definition of a biobank that included all forms of data repositories in which patient-level human genetic information is housed. We defined genetic information as both deoxyribonucleic acid and ribonucleic acid data from any human tissue sample (including tumor biopsies) obtained using any sequencing or other genetic testing technology. This definition also includes genetic data that has been obtained from a biological sample, even if the biological sample is no longer stored within the biobank.

Search Strategy

The search strategy was jointly developed by a subgroup of coauthors (E.R.L., N.K., K.B., N.D., L.D.L.) with experience conducting systematic and scoping literature reviews. We conducted our search on July 23, 2021 (updated February 8, 2022), using the PubMed/MEDLINE and Embase databases and limited the search to articles published between January 1, 2010 and February 8, 2022. There have been dramatic changes in the technology used for genetic sequencing in the last decade, and limiting the search to 2010 onward captures the more recent and applicable genetic testing technology and associated stored data. Database searches were performed using Medical Subject Headings terms and keywords related to biobanks and biorepositories as well as terms related to outputs and outcomes. Full search terms are listed in the Supplemental Methods found at https://doi.org/10.1016/j.jval.2023.02.017. Furthermore, we used a list of biobanks and genomic data initiatives from the Global Alliance for Genomics and Health (accessed July 19, 2021 available from https://www.ga4gh.org) and directly checked their websites to identify relevant gray literature sources. Given the a priori expectation of limited evidence on outputs and outcomes based on the existing literature7 and aiming to assess all available evidence, we allowed for the inclusion of abstracts (ie, even when no full-text was available) if information of interest was provided within the abstract. We also allowed for the inclusion of team suggestions, based on the same inclusion criteria. The multidisciplinary background of the team includes medical genetics, genetic counseling, epidemiology, pharmacy, and health economics. Given the heterogeneity of our search strategy, we will refer to *records*, which is defined as any unique source, whether peer-reviewed publication, biobank website or resource, or conference abstract.

Study Selection

Eligibility criteria included records published in English, French, or Spanish that contained information on outputs and outcomes that could result from the use of a biobank data repository. Information on outputs/outcomes needed to be evidence based rather than potential/hypothetical indicators, or isolated examples that were not demonstrative of a tangible, measurable output/outcome. To capture the impacts on scientific knowledge, we included reviews or compilations that summarized the biobank contribution to gene and gene-related associations discoveries. Nevertheless, due to the volume of records and the inability to quantify the value of a single discovery, we chose not to include records reporting on an individual gene- or variant-related (specific gene change/mutation) finding. All patient types and all therapeutic areas were considered as long as their data were collected with intended clinical application related to human disease. Additional details on the inclusion and exclusion criteria are described in the Supplemental Methods found at https://doi.org/10.1016/j.jval.2023.02.017.

Article Screening

Four reviewers (E.R.L., N.K., K.B., N.D.) conducted title and abstract screening, such that 2 reviewers screened each article. We performed 2 pilot exercises of 100 articles each to calibrate and adjust our screening criteria. Any disagreements were resolved through consensus between the 2 reviewers. Four reviewers (E.R.L., N.K., K.B., L.W.) conducted full-text review using standardized criteria, such that 2 reviewers assessed each article. Disagreements were resolved through consensus between the 2 reviewers and, if needed, team discussion.

Data Extraction and Synthesis

Data extraction was performed on all of the included records by one of 3 reviewers (E.R.L., L.W., N.K.). We piloted the data extraction form on 4 records before implementing full data extraction. A second reviewer verified the data extraction sheets to check that they were complete. Discrepancies were resolved by consulting the original data source and through consensus between the 2 reviewers. The fields for data extraction included 3 main areas: (1) record information, (2) descriptive information about the biobank, and (3) reported output and outcome indicators. An initial set of individual indicators was developed following items provided by the value framework proposed by Rush et al[7] and the limited systematized evidence existing on biobank impacts.[19] Further indicators were created when outputs/outcomes found could not be fit into the existing indicators.

Evidence on outputs and outcomes was synthesized using the Synthesis Without Meta-Analysis reporting guidelines.[24] Given that extracted indicators refer to a specific moment in time and are not always comparable across records, we sought to describe

the variety, frequency, and direction of change of indicators reported. Consequently, following Synthesis Without Meta-Analysis guidelines, rather than providing summary statistics for quantitative indicators, we used direction of effect as a standardized metric and vote counting (ie, providing the number of records with positive and negative effects) as our synthesis method.[24] Finally, we did not conduct a critical appraisal of the included records given their heterogeneity and because the majority were descriptive (ie, reporting general results rather than testing a hypothesis).

Results

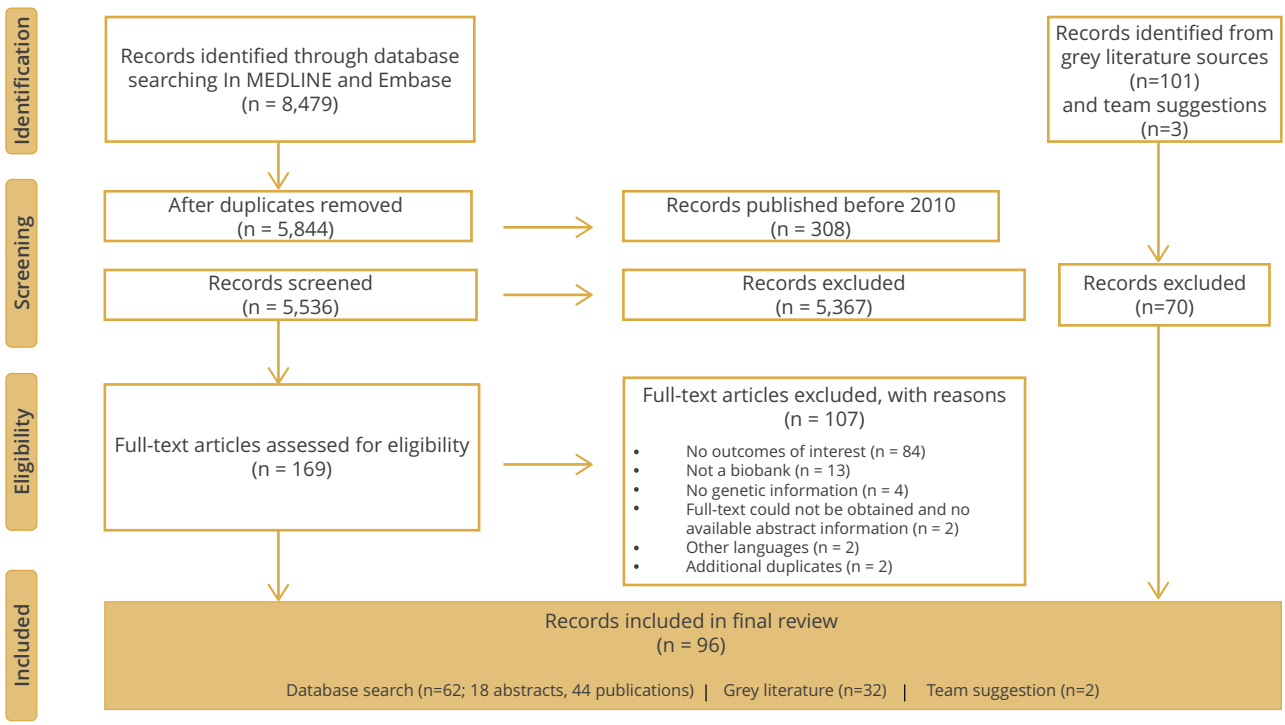
From 8479 records initially identified from databases and 101 from gray literature sources, 96 unique records were included in the final review (46 articles, 32 gray literature resources, and 18 abstracts). The selection process is outlined in Figure 1. We framed our results as the number of records rather than the number of biobanks as the unit of analysis with the aim to characterize the variability and size of the body of evidence around outputs and outcomes. Consequently, several records can be associated with 1 biobank, and 1 record can include and aggregate outputs/outcomes across >1 biobank.

The characteristics of the 96 unique records are detailed in Table 1. Most records were published after 2019 and the most frequent region/country of publication was the United States (41%), followed by Europe (31%). We found 5 literature reviews, 3 meta-analyses, and 2 other records that reported the analysis of several biobanks grouped in a network, with the remaining records referring to a single biobank, which also includes research projects or centers associated with a biobank. Although 84 records addressed the complete biobank data, 12 records included only a subset of the biobank data. Disease categories were mostly nonspecific (49%) or centered on cancer (23%).

As a way of organizing and presenting the results of the review, we developed a taxonomy of outputs and outcomes (described in Fig. 2), which groups 11 types of indicators into 3 categories: (1) research infrastructure, (2) research studies, and (3) clinical outputs and outcomes. The "research infrastructure" category includes outputs such as grant funding and collaborations as well as outcomes such as increased racial, ethnical, or ancestral diversity in genomic databases. These can provide the foundation for specific "research studies," whose outputs and outcomes are measured with indicators such as the number of gene and variant-related discoveries, publications, projects, presentations, and patents. Finally, the "clinical outputs and outcomes" category contains indicators related to patient care. Additionally, the classification structure incorporates the translational research process for population biobanks (converting data into knowledge, followed by turning knowledge into clinical practice).[25]

Overall, we found 208 indicator data points across all 96 records. Most indicators identified were in the research study category, followed by research infrastructure and then clinical (Fig. 3A). Additionally, most records (52 of 96, 54%) reported indicators in only 1 category, and 7 records reported at least 1 indicator in all 3 categories (Fig. 3B). Below we expand on records found for each category and indicator (Table 2[19,20,26-119]). Further details on included records and indicators are presented in Appendix Table 1 found at https://doi.org/10.1016/j.jval.2023.02.017, and external resources can be consulted for additional background on genetic terminology.[120]

Figure 1. PRISMA diagram.



PRISMA indicates Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Table 1. Record characteristics (96 unique records).

Characteristics	Number of records (%)
Year of publication*	
2010-2014	7 (10.9)
2015-2016	9 (14.1)
2017-2018	8 (12.5)
2019-2020	26 (40.6)
2021-2022	14 (21.9)
Region/country of publication†	
North America	31 (48.4)
United States	28 (41)
Canada	3 (4.5)
Europe	20 (31.3)
Australia	5 (7.8)
Asia	7 (10.9)
Africa	1 (1.6)
Type of study	
Literature reviews	7 (7.3)
Meta-analyses‡	3 (3.1)
Description of biobank networks	2 (2.1)
Description of a single biobank	84 (87.5)
Disease categories	
Not specific	47 (48.9)
Cancer	22 (22.9)
Rare diseases (excluding cancer)	7 (7.3)
Cardiovascular disease	2 (2.08)
Others	15 (15.63)
Scope of data used in analysis	
Complete biobank	84 (87.5)
Subcohort§	12 (12.5)

*Excludes gray literature sources, which were all accessed in 2022.
†Refers to first author and excludes gray literature sources.
‡Meta-analyses refers to the combination of data from a group of biobanks.
§When the analysis provided includes a subcohort (ie, not the entirety of information available at the moment from the biobank).

Research Infrastructure

Grants and collaborations

Grants and collaborations can facilitate research, improve biobank information, and, overall, enhance research and clinical indicators. This review found 18 records that reported grants or any type of funding obtained to support the work of the biobank. Sources of funding, uses (eg, initial funding vs funds obtained by leveraging biobank data), or specific amounts were not always specified. In addition, collaborations were reported in 29 records, and they were established in the form of specific studies, research groups, or companies using the biobank information or expanding it through additional data linkages.

Increased representation

One key outcome that facilitates genomics research is the increased representation enabled by accessing large racially, ethnically, and/or ancestrally diverse data repositories with the potential for continuous addition of information and reanalysis. Four records reported contributing to reference genetic data for individuals not of European background. Specific examples include making accessible genetic information for the Taiwanese population not previously used in genetic research,[44] the creation of a customized single nucleotide polymorphism array optimized for the Han Chinese population (Taiwan Biobank),[61] a Japanese reference genome (Tohoku University Tohoku Medical Megabank),[62] and a Qatari gene chip containing gene variants specific to the Qatari population (Qatar Genome Project Biobank).[27] Additionally, the benefits of larger samples of individuals were highlighted in a meta-analysis uncovering the genetics of idiopathic pulmonary fibrosis, which allowed the researchers to study population specific, rare, and sex-dependent variant effects.[60]

Research Studies

Publications, projects, presentations, and patents

Notably, the most frequently reported indicator (included in 54 records) among all 3 categories was a list or the number of scientific publications that used samples or data from the biobank. A subset of these 54 records also included impact factors[28,64,66] and citations of the publications that used biobank information.[28-31,45,46] The number of publications using biobank information varied greatly across records and biobanks, and reporting was not always consistent in terms of including timelines or summary statistics. Among those 28 records that reported publications with clear timelines, the average number of publications was 123 per year (SD 323). The UK Biobank[29] and the database of Genotypes and Phenotypes[67] reported the most with an average of 1224 and 1275 annual publications using their data, respectively, whereas some disease specific biobanks reported only one publication per year, on average (the Genetics of Type 2 Diabetes Consortium[47] and the Cancer Cell Line Encyclopedia project[48]). In addition, 38 records reported research projects but only 1 disaggregated them by progress (ie, approved vs completed).[27] As with publications, the number of projects varied across records, and time spans for conducting the projects were not always reported, making comparisons between records difficult. A fewer number of records included conference presentations/abstracts or patents (5 and 4 in each category, respectively).

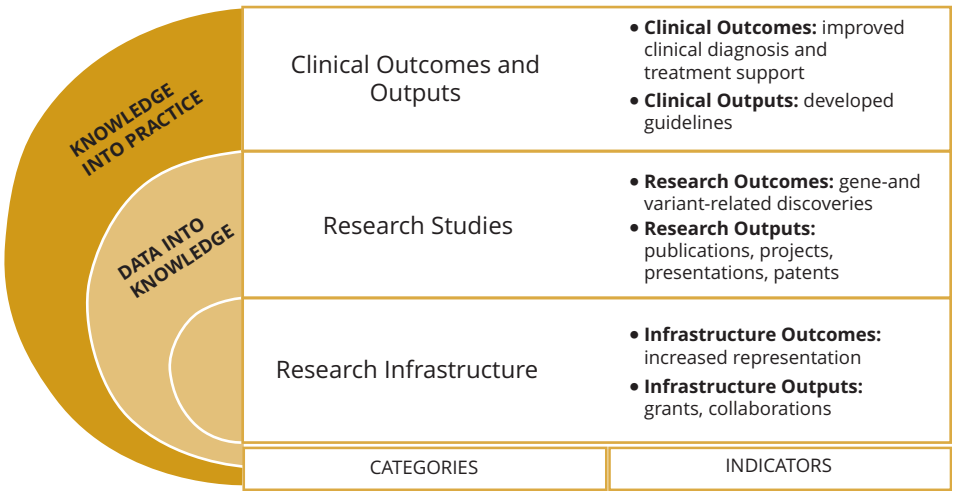
Gene- and variant-related discoveries

We included 32 records that reported multiple gene- and variant-related discoveries. Some common areas included the development of risk scores, pharmacogenomics, newly discovered genes and/or variants, genome-wide significant loci, and gene-disorder, and gene-trait associations.

In particular, 3 records[96,99,100] reported on the development of risk scores—a predictive measure of disease susceptibility—from genetic data in biobanks. One meta-analysis used several biobanks to show how polygenic risk scores can play a role in prioritizing risk factors that are modifiable through medical treatment.[96] Four additional records focused on pharmacogenomics, studying how genetic variants affect patients' responses to drugs.[61,101-103] Notably, a study using the UK Biobank[103] found that when 14 genes with known pharmacogenetic variants were examined, biobank participants were predicted to have an atypical response to an average of 10 drugs, and 24% of individuals had been prescribed 1 of these drugs.

Other frequently reported examples of this type of outcome were newly described genes, variants, or genome-wide significant loci (ie, new regions within the genome that have been found to be associated with a condition). For example, a review of findings discovered using the 100 000 Genomes Project data

Figure 2. Outputs and outcomes synthesized by indicators and categories.



reported 47% of variants not previously observed in other large scale publicly available data sets and > 30 novel genes and disease-causing variants in the noncoding space of the genome.[104] Notably, all 3 meta-analyses and records on biobank networks reported results related to newly described genes, variants, or genome-wide significant loci. The largest one, a meta-analysis of 19 biobanks,[97] detected 188 novel genome-wide significant loci across 14 conditions and uncovered 30 novel loci specifically for gout. Two meta-analyses[60,97] also reported on the advantages of incorporating non-European samples (a research infrastructure indicator) as a strategy to increase the identification of novel genome-wide significant loci. Although most records reported new discoveries, others used the biobank to replicate known associations and indirectly validate the data used.[49,68,105] Finally, 2 records from the UK Biobank reported a substantial increase in loss-of-function variants observed when the number of individuals with genetic data in the biobank increased.[50,106] This is particularly significant for the translation of knowledge into clinical practice because loss-of-function variants can be clinically relevant.

Appendix Table 2 found at https://doi.org/10.1016/j.jval.2023.02.017 expands on those records from literature reviews, meta-analyses, and descriptions of biobank networks that reported some gene- and variant-related discovery.

Clinical Outcomes and Outputs

Development of guidelines

Support for guideline development was reported in 4 records as an output of biobanks. In 1 case, research from the biobank on a specific oral therapy was widely adopted and recommended in consensus guidelines for treatment of cytomegalovirus disease.[69] For 2 other records, use of the biobank data led to a more standardized approach for the assessment of muscle strength among children[32] and contributed to formalizing a first of its kind audit for a surgical technique.[89] Finally, a literature review on rare diseases[19] concluded that, by facilitating multicenter collaboration, discussions stemming from research conducted using the biobank led to expert discussion about treatment protocols and best practices.

Figure 3. Indicators and records reported.

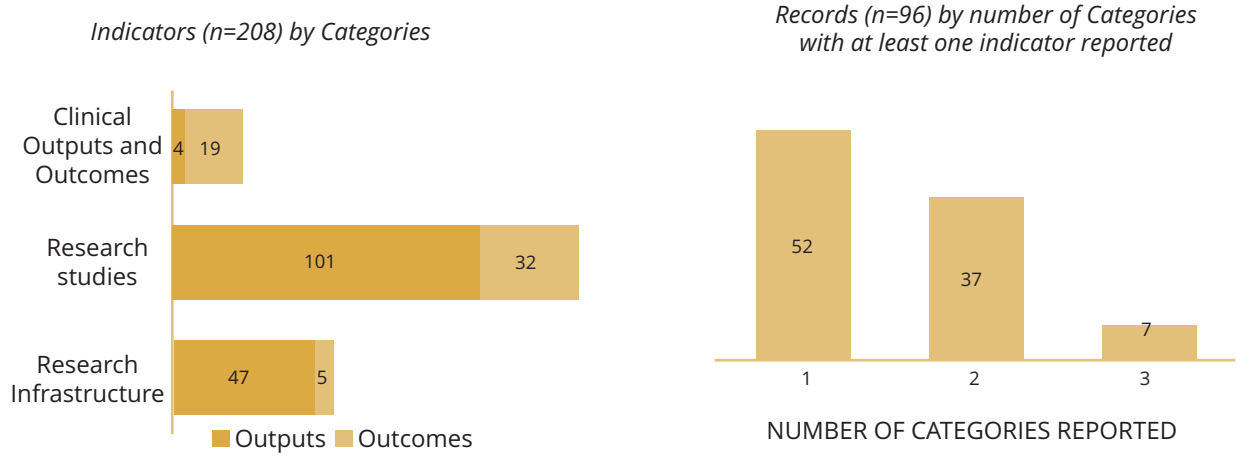


Table 2. Number of records (and associated references) by indicator and study type.

Categories	Indicators	Reviews, meta-analysis, and descriptions of biobank networks (n = 10)		Description of individual biobanks and research center/projects (n = 86)	
		Number of records	References	Number of records	References
Research infrastructure	Grants	2	20,26	16	27-42
	Collaborations	2	19,43	27	27,28,30–34,37,39,40,42,44–59
	Increased representation	1	60	4	27,44,61,62
Research studies	Publications	7	19,20,26,43,63–65	47	27–31,33–36,38–41,44–48,54,55,57–59,62,66-88
	Projects	5	19,20,26,43,64	33	27–29,32,35,39,40,42,48,51,53,55–57,62, 66,69,70,74,75,77,81,83–85,87,89–95
	Presentations			5	27,37,39,55,59
	Patents			4	29,36,66,70
	Gene- and variant-related discoveries*	7	26,43,60,63,96–98	22	31,46,49,50,55,61,68,70,76,95,99–110
Clinical outputs and outcomes	Developed guidelines	1	19	3	32,69,89
	Improved diagnosis and treatment support	5	19,20,26,43,111	14	27,36,70,91,104,109,112–119

*Note that given the broad scope of this indicator one record can report >1 gene- and variant-related discovery. In particular, 29 records reported a total of 32 discoveries, defined as novel findings related to the role of a specific gene, loci, or variant in human disease.

Improved clinical diagnosis and treatment support

Nineteen records reported an impact in terms of obtaining a genetic diagnosis, which in some records resulted in opportunities for management, whether in the form of genetic counseling,[26] referrals to additional specialists,[27] or pharmacological intervention.[112,113] For example, for participants in the Precision Oncology for Young People Collaborative Program[114] and in the 100 000 Genomes Project for rare diseases,[63] treatment or medical management recommendations were provided to 82% and 25% of participants, respectively. Results also reflected how the use of data from biobanks to support treatment decisions and diagnoses is increasing to include cascade testing for relatives and to evaluate population-wide screening programs in unselected populations.[115] Notably, as a result of data stored in biobanks, diagnoses could be achieved retrospectively, from further information being stored and reanalyzed leading to new diagnoses in individuals who originally had uninformative genetic testing.[19,120]

Discussion

The evidence summarized by our review suggests that biobanks containing genetic information and genomic data initiatives have succeeded in expanding the scientific knowledge base and affecting clinical practice and improving care. We created a taxonomy to structure our findings that broadens the scope of biobank benefits beyond outputs (a direct product of the biobank activities) to also include results with impacts on the knowledge base and clinical practice (the outcomes). Furthermore, indicators are divided into 3 main categories based on their contribution to the knowledge translation process: research infrastructure, research studies, and clinical output and outcomes. We found most evidence (63% of studies) was focused on research study-related outputs and outcomes (see Fig. 2), particularly the number of publications and gene-disorder association discoveries, with far fewer reporting about clinical impact or practice changes. Notably, findings on research studies and infrastructure outcomes, including increased representation and a wide range of gene- and variant-related discoveries, highlight the unrealized value in data sharing and integration through biobanks.

Consistent with our results, a previous systematic review on biobanks for rare diseases[19] found predominance on research activity outputs. Additionally, a review of high impact publications from the UK Biobank found no immediate clinically relevant improvements in risk prediction, screening, or treatment arose from the biobank.[121] Although we also found no clinical outputs or outcomes reported among records corresponding to the UK Biobank, we did find some indicators within the clinical category (Fig. 2) reported by other biobanks, mostly related to genetic diagnosis, but also treatment and guideline development.

We also found that only 6 records (excluding biobank websites) stated that one of their core objectives was to summarize results from biobank use (whether at an output or an outcome level), with the rest only reporting indicators incidentally and with great disparities in reporting standards. In addition, biobanks websites, the resource most accessible to all stakeholders, mostly reported on easily quantifiable outputs such as publications, grants, and collaborations, rather than outcomes resulting from biobank activities (eg, new discoveries, improvements in patient care). The variability of reported indicators, besides a lack of intentional collection and reporting, could also be driven by failing to apply a comprehensive value framework. Although some biobank evaluation literature[7,16,17] and work on biobank/bioresource impact factors[18,122,123] exists, the most comprehensive model was only published recently (2020)[7] and thus has not been widely applied. Indeed, among all 96 records included in this review, we found no intentional use of existing value frameworks for biobanks, with one exception,[20] which likely contributed to the low number of records (n = 7) containing at least one indicator in all 3 categories.

Additionally, how biobanks expand the knowledge base and affect clinical care should be complemented with a close analysis of internal metrics of biobank operation. These include successful interoperability between biobank-related systems such as

registries and electronic medical records for detailed and accurate phenotypic information, as well as quality of sample collection, processing and analysis, and cost and human resources. Although not the focus of this review, we should acknowledge that these factors could represent barriers in successful knowledge translation and should be investigated in further studies.

Results from this study should be considered in light of its limitations. First, given the rapid nature of the review and its associated constraints depending on chosen steps,[124] we may have excluded relevant records reporting outputs/outcomes. Nevertheless, considering the rigor of our review process and inclusion of a variety of sources, it is likely that excluded studies (if any) did not affect the trends and main results found by the review. Likewise, by including abstracts and not conducting critical appraisals, the quality of the included evidence likely varies. Second, overlap among the indicators can exist and that might have not been acknowledged by the original source. For example, gene-disorder discoveries that also resulted in publications could be counted more than once within the research studies category. Similarly, given that most records only reported output/outcome indicators incidentally, collection was based on wording use by the authors without the reviewer conducting data extraction making additional derivations or assumptions. Consequently, some indicators might have not been reported though in fact collected by the biobank. Third, by not including individual gene- and variant-related discoveries, the number of such discoveries is likely understated for biobanks that have not published a review or summary of the findings. Finally, and most importantly, we issue a word of caution on interpreting the link between greater data access through biobanks and resulting outputs/outcomes. In most records included in this review, no alternative scenario to compare with the use of the biobanks was provided, and in the case of research centers or projects with an associated biobank, the link becomes less clear due to the lack of comprehensive reporting of processes and resources. We should note this is not an issue exclusive to this study but one that should be discussed further across the literature on the value of biobanks.

By reviewing and summarizing the existing evidence on biobanks outputs and outcomes, this study offers 3 main recommendations for biobank users to improve data collection, reporting, and impact evaluation. First, we recommend prospectively and intentionally collecting a comprehensive set of outcomes, outputs, and internal metrics that reflects the impact of biobank-supported research results on research and clinical practice and demonstrate value to all potential stakeholders. Second, we advise to report collected indicators in a standardized and detailed manner (ie, summary statistics, timelines, progress status when it comes to projects supported by the biobank, sources and nature of grants and collaborations) to facilitate comparability across time and different biobanks and business models. Third, we recommend informing these efforts in collecting comprehensive data on biobank impact and standardized reporting by leveraging value frameworks built on the existing conceptual and applied biobank literature. Moreover, biobanks containing genetic information should aim to develop a distinct supplemental value framework comprising genetics-specific indicators. The present review begins this work by including indicators such as gene- and variant-related discoveries, increased genetic representation, clinical genetic diagnosis, changes in treatment, and management in its framework. Overall, a useful tool in systematically planning the collection of relevant indicators is the performance measurement framework proposed within the results-based management approach,[15] which can be used to assess biobanks beyond outputs to also focus on comprehensive impacts. Any new value framework for genetics biobanks could also incorporate elements of existing learning health systems frameworks[125] given that sharing and storing information through biobanks allows for reanalysis, new diagnoses, and potential advancements in other outcomes as new information is added in each "learning cycle."

Conclusions

Although privacy and ethical issues associated with greater data access and sharing through biobanks have been well documented, the full range of their impacts on health research, clinical practice, and health outcomes have not been systematically documented, which undermines our ability to conduct value assessment of data repositories. Results from this review highlight that, although there is emerging evidence on the various benefits stemming from use of biobanks containing genetic information, greater attention should be paid to the systematic collection and reporting of indicators, which will require intentional application of value frameworks specifically designed for biobanks containing genetic information.

Supplemental Materials

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2023.02.017.

Article and Author Information

Accepted for Publication: February 24, 2023

Published Online: April 19, 2023

doi: https://doi.org/10.1016/j.jval.2023.02.017

Author Affiliations: Collaboration for Outcomes Research and Evaluation (CORE), Faculty of Pharmaceutical Sciences, The University of British Columbia, Vancouver, British Columbia, Canada (Llorian, Kopac, Waliji, Borle, Dragojlovic, Lynd); Department of Medical Genetics, The University of British Columbia, Vancouver, British Columbia, Canada (Elliott); Centre for Health Evaluation and Outcome Sciences (CHÉOS), St. Paul's Hospital, Vancouver, BC, Canada (Lynd).

Correspondence: Larry D. Lynd, PhD, Collaboration for Outcomes Research and Evaluation (CORE), Faculty of Pharmaceutical Sciences, The University of British Columbia, Vancouver Campus, 2405 Wesbrook Mall, Vancouver, British Columbia, Canada V6T 1Z3. Email: larry.lynd@ubc.ca

Author Contributions: *Concept and design*: Rodriguez Llorian, Kopac, Borle, Dragojlovic, Elliott, Lynd
Acquisition of data: Rodriguez Llorian, Kopac, Waliji, Borle, Dragojlovic
Analysis and interpretation of data: Rodriguez Llorian, Kopac, Waliji, Borle, Dragojlovic, Elliott, Lynd
Drafting of the articlet: Rodriguez Llorian, Kopac
Critical revision of the article for important intellectual content: Rodriguez Llorian, Kopac, Waliji, Borle, Dragojlovic, Elliott, Lynd
Obtaining funding: Dragojlovic, Elliott, Lynd
Administrative, technical, or logistic support: Rodriguez Llorian
Supervision: Lynd

Conflict of Interest Disclosure: The authors reported no conflicts of interest.

Funding/Support: Financial support for this study was provided by a Canadian Institutes of Health Research Genomics and Precision Health Top-Up Grant (funding reference number GPT-174782) provided as part of the GenCOUNSEL project. GenCOUNSEL was funded through the Large Scale Applied Research Project Genome Canada competition with cofunding from: Canadian Institute for Health Research, Genome BC, Genome Quebec, Provincial Health Services Authority, BC Children's Hospital Foundation, and BC Women's Hospital Foundation.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the article; and decision to submit the article for publication.

Acknowledgment: We are grateful to Dr Alivia Dey, PMP, for project management. The GenCOUNSEL Study is led by Alison M. Elliott, Jehannine Austin, Bartha Knoppers, and Larry D. Lynd with Project Manager Alivia Dey and includes the following coinvestigators: Shelin Adam, Nick Bansback, Patricia Birch, Lorne Clarke, Nick Dragojlovic, Jan Friedman, Debby Lambert, Daryl Pullman, Alice Virani, Wyeth Wasserman, and Ma'n Zawati.

REFERENCES

1. Banks MA. Sizing up big data. *Nat Med*. 2020;26(1):5–6.
2. Hemingway H, Asselbergs FW, Danesh J, et al. Big data from electronic health records for early and late translational cardiovascular research: challenges and potential. *Eur Heart J*. 2018;39(16):1481–1495.
3. Global healthcare data-storage market by type (flash & solid-state storage & magnetic storage), by storage system (storage area network, direct-attached storage & network-attached storage),by delivery (remote, hybrid and on-premise), by end-user (pharmaceutical & biotechnology companies, hospitals & clinics, research Centers, and others), by region (North America, Europe, the Asia-Pacific, the Middle East & Africa, and Latin America); trend analysis, competitive market share & forecast, 2021-2027. BlueWeave Consulting. https://www.blueweaveconsulting.com/report/global-healthcare-data-storage-market-size-industry-trends-forecast-report-bwc20152. Accessed August 10, 2022.
4. Annaratone L, De Palma G, Bonizzi G, et al. Basic principles of biobanking: from biological samples to precision medicine for patients. *Virchows Arch*. 2021;479(2):233–246.
5. Shaw DM, Elger BS, Colledge F. What is a biobank? Differing definitions among biobank stakeholders. *Clin Genet*. 2014;85(3):223–227.
6. Watson PH, Barnes RO. A proposed schema for classifying human research biobanks. *Biopreserv Biobank*. 2011;9(4):327–333.
7. Rush A, Catchpoole DR, Ling R, Searles A, Watson PH, Byrne JA. Improving academic biobank value and sustainability through an outputs focus. *Value Health*. 2020;23(8):1072–1078.
8. Khoury MJ, Armstrong GL, Bunnell RE, Cyril J, Iademarco MF. The intersection of genomics and big data with public health: opportunities for precision public health. *PLoS Med*. 2020;17(10):1–14.
9. Zawati MH, Knoppers B, Thorogood A, Population Biobanking and International Collaboration. *Pathobiology*. 2014;81(5-6):276–285.
10. Budimir D, Polašek O, Marušić A, et al. Ethical aspects of human biobanks: a systematic review. *Croat Med J*. 2011;52(3):262–279.
11. Porter ME. What is value in health care? *N Engl J Med*. 2010;363(26):2477–2481.
12. Richards JE, Bane E, Fullerton SM, Ludman EJ, Jarvik G. Allocation of resources to communication of research result summaries: biobank participant perspectives. *J Empirical Res Hum Res Ethics*. 2016;11(4):364–369.
13. Hyams T, Bowen DJ, Condit C, et al. Views of cohort study participants about returning research results in the context of precision medicine. *Public Health Genomics*. 2016;19(5):269–275.
14. Laurie G, Mallia P, Frenkel DA, et al. Managing access to biobanks: how can we reconcile individual privacy and public interests in genetic research? *Med Law Int*. 2010;10(4):315–337.
15. Results-based management. Government of Canada. https://www.international.gc.ca/world-monde/funding-financement/results_based_management-gestion_axee_resultats.aspx?lang=eng. Accessed August 15, 2022.
16. Rush A, Matzke L, Cooper S, Gedye C, Byrne JA, Watson PH. Research perspective on utilizing and valuing tumor biobanks. *Biopreserv Biobanking*. 2019;17(3):219–229.
17. Hofman V, Ilie M, Long E, et al. Measuring the contribution of tumor biobanks to research in oncology: surrogate indicators and bibliographic output. *Biopreserv Biobanking*. 2013;11(4):235–244.
18. Mabile L, Dalgleish R, Thorisson GA, et al. Quantifying the use of bioresources for promoting their sharing in scientific research. *GigaSci*. 2013;2(1):7.
19. Garcia M, Downs J, Russell A, Wang W. Impact of biobanks on research outcomes in rare diseases: a systematic review. *Orphanet J Rare Dis*. 2018;13(1):202.
20. Vora T, Thacker N. Impacts of a biobank: bridging the gap in translational cancer medicine. *Indian J Med Paediatr Oncol*. 2015;36(01):17–23.
21. Garritty C, Gartlehner G, Nussbaumer-Streit B, et al. offers evidence-informed guidance to conduct rapid reviews. *J Clin Epidemiol*. 2021;130:13–22.
22. Hamel C, Michaud A, Thuku M, et al. Defining rapid reviews: a systematic scoping review and thematic analysis of definitions and defining characteristics of rapid reviews. *J Clin Epidemiol*. 2021;129:74–85.
23. Tricco AC, Antony J, Zarin W, et al. A scoping review of rapid review methods. *BMC Med*. 2015;13(1):224.
24. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ*. 2020;368:l6890.
25. Murtagh MJ, Demir I, Harris JR, Burton PR. Realizing the promise of population biobanks: a new model for translation. *Hum Genet*. 2011;130(3):333–345.
26. Filocamo M, Baldo C, Goldwurm S, et al. Telethon Network of Genetic biobanks: a key service for diagnosis and research on rare diseases. *Orphanet J Rare Dis*. 2013;8(1):129.
27. Qatar Biobank. https://www.qatarbiobank.org.qa/. Accessed April 27, 2022.
28. European Genome-Phenome Archive (EGA). EGA Consortium. https://ega-archive.org/. Accessed June 10, 2022.
29. UK Biobank Limited. https://www.ukbiobank.ac.uk/. Accessed May 9, 2022.
30. American Association for Cancer Research (AACR) project GENIE. American Association for Cancer Research. https://www.aacr.org/professionals/research/aacr-project-genie/. Accessed June 10, 2022.
31. Conroy M, Sellors J, Effingham M, et al. The advantages of UK Biobank's open-access strategy for health research. *J Intern Med*. 2019;286(4):389–397.
32. Martin N, Krol P, Smith S, et al. A national registry for juvenile dermatomyositis and other paediatric idiopathic inflammatory myopathies: 10 years' experience; the Juvenile dermatomyositis National (UK and Ireland) Cohort biomarker Study and Repository for Idiopathic Inflammatory Myopathies. *Rheumatology*. 2011;50(1):137–145.
33. 23andMe Inc. https://www.23andme.com/en-ca/. Accessed April 21, 2022.
34. Alzheimer's Disease Sequencing Project. University of Pennsylvania, Perelman School of Medicine. https://www.niagads.org/adsp/. Accessed June 10, 2022.
35. Consortium of Investigators of Modifiers of BRCA 1/2 (CIMBA). University of Cambridge. https://cimba.ccge.medschl.cam.ac.uk/. Accessed May 9, 2022.
36. Kohli M. The role of "liquid Biopsy" repositories in cancer care in low to middle-income countries. *J Glob Oncol*. 2018;4(suppl 2):244s–244s.
37. RD-Connect Project. RD-Connect. https://rd-connect.eu/. Accessed May 9, 2022.
38. Tan S. The center for prostate disease research biospecimen bank and multicenter national database: integrating biospecimen and clinical data for precision medicine. In: *Biopreserv Biobanking*. 2020;18(3).
39. The China Kadoorie Biobank (CKB). University of Oxford. https://www.ckbiobank.org/. Accessed June 10, 2022.
40. Tree-house childhood cancer initiative. The Regents of the University of California. https://treehousegenomics.ucsc.edu/. Accessed May 3, 2022.
41. UK10K Project. Wellcome Sanger Institute. https://www.sanger.ac.uk/collaboration/uk10k/. Accessed May 9, 2022.
42. DNA biobank (BioVU). VICTR. https://victr.vumc.org/what-is-biovu/. Accessed June 10, 2022.
43. Lochmuller H, Badowska DM, Thompson R, et al. RD-Connect, NeurOmics and EURenOmics: collaborative European initiative for rare diseases. *Eur J Hum Genet*. 2018;26(6):778–785.
44. Lin JC, Chen LK, Hsiao WWW, Fan CT, Ko ML. Next chapter of the Taiwan Biobank: sustainability and perspectives. *Biopreserv Biobanking*. 2019;17(2):189–197.
45. DatabasE of genomiC varIation and phenotype in humans using Ensembl resources (DECIPHER). Wellcome Sanger Institute. https://www.deciphergenomics.org/about/overview. Accessed June 10, 2022.
46. The Electronic Medical Records and Genomics (eMERGE) network. eMERGE. https://emerge-network.org/. Accessed May 9, 2022.
47. The genetics of Type 2 diabetes consortium (GoT2D). National Institutes of Health. https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_idphs000840.v1.p1. Accessed May 9, 2022.
48. Cancer Cell Line Encyclopedia (CCLE) project. Cancer Cell Line Encyclopedia. https://sites.broadinstitute.org/ccle/. Accessed June 10, 2022.
49. Carey DJ, Fetterolf SN, Davis FD, et al. The Geisinger MyCode community health initiative: an electronic health record–linked biobank for precision medicine research. *Genet Med*. 2016;18(9):906–913.
50. Szustakowski JD, Balasubramanian S, Kvikstad E, et al. Advancing human genetics research and drug discovery through exome sequencing of the UK Biobank. *Nat Genet*. 2021;53(7):942–948.
51. Schmidt AH, Sauter J, Baier DM, et al. Immunogenetics in stem cell donor registry work: the DKMS example (Part 2). *Int J Immunogenet*. 2020;47(2):139–148.
52. Chen WC, Singh E, Muchengeti M, et al. Johannesburg Cancer Study (JCS): contribution to knowledge and opportunities arising from 20 years of data collection in an African setting. *Cancer Epidemiol*. 2020;65:101701.
53. Furlong S. Macquarie University neurodegenerative disease biobank. *Biopreserv Biobanking*. 2020;18(3).
54. Washetine K, Heeke S, Bonnetaud C, et al. Establishing a dedicated lung cancer biobank at the University Center hospital of Nice (France). Why and how? *Cancers*. 2018;10(7):220.
55. Saudi Human Genome Program. https://shgp.kacst.edu.sa/index.en.html#achievements. Accessed April 27, 2022.
56. The Scottish Genomes Partnership. https://www.scottishgenomespartnership.org. Accessed May 9, 2022.

57. Malecki K, Nikodemova M, Schultz AA, et al. The Survey of the Health of Wisconsin (SHOW) Program: an infrastructure for Advancing Population Health Sciences. *Front Public Health*. 2022;10:464.
58. Hallberg P, Yue QY, Eliasson E, Melhus H, Ås J, Wadelius M. SWEDEGENE—a Swedish nation-wide DNA sample collection for pharmacogenomic studies of serious adverse drug reactions. *Pharmacogenomics J*. 2020;20(4):579–585.
59. The Clinical Genome Resource (Clingen). ClinGen. https://clinicalgenome.org/. Accessed May 3, 2022.
60. Partanen JJ, Häppölä P, Zhou W, et al. Leveraging global multi-ancestry meta-analysis in the study of idiopathic pulmonary fibrosis genetics. *Cell Genom*. 2022;2(10):100181.
61. Wei CY, Yang JH, Yeh EC, et al. Genetic profiles of 103,106 individuals in the Taiwan biobank provide insights into the health and history of Han Chinese. *npj Genom Med*. 2021;6(1):10.
62. Tohoku medical Megabank organization. Tohoku University. https://www.megabank.tohoku.ac.jp/english/. Accessed April 21, 2022.
63. Glynn P, Greenland P. Contributions of the UK Biobank high impact papers in the era of precision medicine. *Eur J Epidemiol*. 2020;35(1):5–10.
64. Zhou L, Catchpoole D. Spanning the genomics era: the vital role of a single institution biorepository for childhood cancer research over a decade. *Transl Pediatr*. 2015;4(2):14.
65. Coppola L, Cianflone A, Grimaldi AM, et al. Biobanking in health care: evolution and future directions. *J Transl Med*. 2019;17(1):172.
66. Lee. Utilization performance of bioresources distributed by the national biobank of Korea. In: *Biopreserv Biobanking*. 2018;16.
67. The database of Genotypes and Phenotypes (dbGaP). National Institutes of Health. https://www.ncbi.nlm.nih.gov/gap/. Accessed June 10, 2022.
68. Zawistowski M, Fritsche LG, Pandit A, et al. The Michigan Genomics initiative: a biobank linking genotypes and electronic clinical records in Michigan medicine patients. *Cell Genom*. 2023;3(2):100257.
69. Åsberg A, Humar A, Rollag H, et al. Lessons learned from a randomized study of oral valganciclovir versus parenteral ganciclovir treatment of Cytomegalovirus disease in solid organ transplant recipients: the VICTOR Trial. *Clin Infect Dis*. 2016;62(9):1154–1160.
70. Greenwald P. Abstract SS01-01: view from the NCI: NCI's Division of Cancer Prevention. *Cancer Prev Res (Phila)*. 2010;3(suppl 12):SS01-01.
71. Psychiatric Genomics Consortium (PGC). https://pgc.unc.edu/. Accessed May 3, 2022.
72. van Rooij IALM, van der Zanden LFM, Bongers EMHF, et al. AGORA, a data- and biobank for birth defects and childhood cancer. *Birth Defects Res A*. 2016;106(8):675–684.
73. deCODE Genetics. https://www.decode.com/. Accessed May 9, 2022.
74. Genomic data commons (GDC). National Institutes of Health. https://gdc.cancer.gov/access-data. Accessed April 25, 2022.
75. Castro VM, Gainer V, Wattanasin N, et al. The Mass General Brigham biobank Portal: an i2b2-based data repository linking disparate and high-dimensional patient data to support multimodal analytics. *J Am Med Inform Assoc*. 2022;29(4):643–651.
76. Matchmaker Exchange. https://www.matchmakerexchange.org/. Accessed April 27, 2022.
77. Million Veteran Program. US Department of Veteran Affairs. https://www.mvp.va.gov/pwa/. Accessed May 9, 2022.
78. MSSNG. https://research.mss.ng/. Accessed May 9, 2022.
79. openSNP. https://opensnp.org/. Accessed June 10, 2022.
80. Buhmeida A. Biobanking in Saudi Ariabia: the CEGMR biobank experience. *Biopreserv Biobanking*. 2015;13(3).
81. The All of Us Research Program. National Institutes of Health. https://www.researchallofus.org/. Accessed June 10, 2022.
82. Glover J, Krailo M, Tello T, et al. A summary of the osteosarcoma banking efforts: a report from the Children's Oncology Group and the QuadW Foundation: summary of the Osteosarcoma Banking Efforts. *Pediatr Blood Cancer*. 2015;62(3):450–455.
83. Senthil G, Dutka T, Bingaman L, Lehner T. Genomic resources for the study of neuropsychiatric disorders. *Mol Psychiatry*. 2017;22(12):1659–1663.
84. Rosales XQ, Thompson JLP, Haas R, et al. The North American mitochondrial disease registry. *J Transl Genet Genomics*. 2020;4(2):81–90.
85. Link biobank for health discovery. Boston Children's Hospital, Precision. http://biobank.childrens.harvard.edu/. Accessed June 10, 2022.
86. Tiwari S. A prostate cancer biorepository meeting international collaborative demands. *Biopreserv Biobanking*. 2015;13(3).
87. The Undiagnosed Diseases Network (UDN). https://undiagnosed.hms.harvard.edu/. Accessed June 10, 2022.
88. Caleyachetty R, Littlejohns T, Lacey B, et al. United Kingdom biobank (UK Biobank): JACC focus Seminar 6/8. *J Am Coll Cardiol*. 2021;78(1):56–65.
89. Busheri L, Dixit S, Nare S, et al. Breast cancer biobank from a single institutional cohort in an urban setting in India: tumor characteristics and survival outcomes. *Cancer Treat Res Commun*. 2021;28.
90. Mistry J. The national Children's Cancer and Leukaemia Group tissue bank ~ facilitating a comprehensive range of paediatric oncology research. 2016;63(suppl 3).
91. Aquilante CL, Kao DP, Trinkley KE, et al. Clinical implementation of pharmacogenomics via a health system-wide research biobank: the University of Colorado experience. *Pharmacogenomics*. 2020;21(6):375–386.
92. International Cancer Genome Consortium (ICGC). https://dcc.icgc.org/. Accessed May 9, 2022.
93. Olson JE, Ryu E, Hathcock MA, et al. Characteristics and utilisation of the Mayo Clinic biobank, a clinic-based prospective collection in the USA: cohort profile. *BMJ Open*. 2019;9(11):e032707.
94. Paul S. *McCain Genito-Urinary biobank (MGB) at the university health network*. Toronto, Ontario: Biopreserv Biobanking; 2020;18.
95. Turner CF, Pan H, Silk GW, et al. The NIDDK central repository at 8 years—ambition, revision, use and impact. *Database (Oxf)*. 2011;2011:bar043.
96. Sakaue S, Kanai M, Karjalainen J, et al. Trans-biobank analysis with 676,000 individuals elucidates the association of polygenic risk scores of complex traits with human lifespan. *Nat Med*. 2020;26(4):542–548.
97. Zhou W, Kanai M, Wu KHH, et al. Global Biobank Meta-analysis Initiative: Powering Genetic Discovery Across Human Diseases. *Cell Genom*. 2022;2(10):100192.
98. Perola M. FinnGen-study at its midway: already identifying dozens of novel variants for tens of diseases and traits. Eur J Hum Genet;28(suppl 1):1-140.
99. Ko YA, Hayek S, Sandesara P, Samman Tahhan A, Quyyumi A. Cohort profile: the Emory cardiovascular biobank (EmCAB). *BMJ Open*. 2017;7(12):1–7.
100. Keller MA, Gordon ES, Stack CB, et al. Coriell Personalized Medicine Collaborative®: a prospective study of the utility of personalized medicine. *Pers Med*. 2010;7(3):301–317.
101. Shah SN, Gammal RS, Amato MG, et al. Clinical utility of pharmacogenomic data collected by a health-system biobank to predict and prevent adverse drug events. *Drug Saf*. 2021;44(5):601–607.
102. Wu Y, Byrne EM, Zheng Z, et al. Genome-wide association study of medication-use and associated disease in the UK Biobank. *Nat Commun*. 2019;10(1):1891.
103. McInnes G, Lavertu A, Sangkuhl K, Klein TE, Whirl-Carrillo M, Altman RB. Pharmacogenetics at scale: an analysis of the UK Biobank. *Clin Pharmacol Ther*. 2021;109(6):1528–1537.
104. Erikstrup C, Pedersen O, Ullum H. Current opinion in donor health-personalised donor care. Paper presented at: 29mth Regional Congress of ISBT; June 22-26, 2019; Basel, Switzerland. https://onlinelibrary.wiley.com/doi/10.1111/vox.12792. Accessed June 10, 2022.
105. Dumitrescu L. On the utility of biobanks linked to electronic medical records in genome-wide association studies. *HUGO J*. 2010;4:126–127.
106. Van Hout CV, Tachmazidou I, Backman JD, et al. Exome sequencing and characterization of 49,960 individuals in the UK Biobank. *Nature*. 2020;586(7831):749–756.
107. Van Hout CV, Tachmazidou I, Backman JD, et al. Whole exome sequencing and characterization of coding variation in 49, 960 individuals in the UK Biobank. *Eur J Hum Genet*. 2019;27(S2):1043–1173.
108. Park J, Lucas AM, Zhang X, et al. Exome-wide evaluation of rare coding variants using electronic health records identifies new gene–phenotype associations. *Nat Med*. 2021;27(1):66–72.
109. The 100,000 genomes project pilot investigators. 100,000 genomes pilot on rare-disease diagnosis in health care—preliminary report. *N Engl J Med*. 2021;385(20):1868–1880.
110. Hung RJ. MS12.04 The International Lung Cancer Consortium (ILCCO), an international study to identify risk factors for lung cancer development. *J Thorac Oncol*. 2019;14(10 suppl):S180.
111. Traschütz A, Reich S, Adarmes AD, et al. The ARCA registry: a collaborative global platform for advancing trial readiness in autosomal recessive cerebellar ataxias. *Front Neurol*. 2021;12:677551.
112. Alver M, Palover M, Saar A, et al. Recall by genotype and cascade screening for familial hypercholesterolemia in a population-based biobank from Estonia. *Genet Med*. 2019;21(5):1173–1180.
113. Collen L. Clinical phenotypes, molecular analysis, and outcomes in very early onset inflammatory bowel disease. *Gastroenterology*. 2021;160(6):S-718.
114. Grover A, Terry Fox P. Abstract 5413: Terry Fox PRecision Oncology for Young peopLE (PROFYLE): a Canadian precision medicine program for children, adolescents and young adults with hard-to-treat cancer. *Cancer Res*. 2020;80(suppl 16):5413.
115. Buchanan AH, Manickam K, Meyer MN, et al. Early cancer diagnoses through BRCA1/2 screening of unselected adult biobank participants. *Genet Med*. 2018;20(5):554–558.
116. Reynolds T, Jack K, Kemeny M, et al. Abstract A090: community partnerships expand access to tumor molecular profiling. *Cancer Epidemiol Biomarkers Prev*. 2020;29(6 suppl 2):A090.
117. Palover M. Biobank-based genotype-first model for return of actionable findings: follow-up and extended impact by cascade screening of BRCA1/2 carriers. *Eur J Hum Genet*. 2020;28(suppl 1):693–694.
118. Leitsalu L, Alavere H, Tammesoo ML, Leego E, Metspalu A. Linking a population biobank with national health registries—the Estonian experience. *J Pers Med*. 2015;5(2):96–106.

119. East KM, Kelley WV, Cannon A, et al. A state-based approach to genomics for rare disease and population screening. *Genet Med*. 2021;23(4):777–781.
120. Jarvik GP, Evans JP. Mastering genomic terminology. *Genet Med*. 2017;19(5):491–492.
121. Smedley D, Smith KR, Martin A, et al. 100,000 genomes pilot on rare-disease diagnosis in health care—preliminary report. *N Engl J Med*. 2021;385(20):1868–1880.
122. Cambon-Thomsen A. Assessing the impact of biobanks. *Nat Genet*. 2003;34(1):25–26.
123. Cambon-Thomsen A, Thorisson GA, Andrieu S, BRIF Workshop Group. The role of a bioresource research impact factor as an incentive to share human bioresources. *Nat Genet*. 2011;43(6):503–504.
124. Hamel C, Michaud A, Thuku M, et al. Few evaluative studies exist examining rapid review methodology across stages of conduct: a systematic scoping review. *J Clin Epidemiol*. 2020;126:131–140.
125. Menear M, Blanchette MA, Demers-Payette O, Roy D. A framework for value-creating learning health systems. *Health Res Policy Syst*. 2019;17(1):79.