such ACOs cannot be identified on the basis of financial performance because the benchmarks don't represent expected spending under the status quo.[3] In general, it's difficult to quantify a single ACO's response to program incentives — akin to divining a single patient's treatment response in a clinical trial. The distinction between benchmarks and expected spending will grow more obvious as benchmarks increasingly incorporate a regional component; one would not expect ACOs' spending to converge to the regional average under the status quo. As benchmarks become increasingly disconnected from expected spending, performance-based terminations would oust more ACOs that are generating savings.

It remains unknown whether the MSSP will unravel into a large subsidy for providers with already low spending, but the evidence raises concerns about strategies for accelerating savings by holding ACOs to a specified pace. Efforts to improve the MSSP should build on, rather than jeopardize, its early success. Instead of setting a faster pace for high-spending ACOs, CMS could set stronger incentives to elicit a faster, albeit uncertain and varied, pace in ways that wouldn't compromise participation. Benchmarks could be based on ACOs' historical baseline spending and increased annually at a desirable rate, but without periodic "rebasing" to newly achieved levels.[5] Without rebasing or aggressive regionalization of benchmarks, downside risk could be incorporated gradually. If sufficient convergence in spending can be fostered among a region's ACOs, regional benchmarks could be considered. In addition, lower-spending ACOs could be given higher shared-savings rates to support their more challenging task of cutting waste when there's less to cut, instead of subsidies in the form of regionally adjusted benchmarks.

MSSP incentives have been weak. Calls for greater savings and the instinct to lay a pathway are understandable. Relying entirely on incentives, without holding providers to an expected savings pace, may be less satisfying to regulators because it requires a leap of faith. But we need to know how ACOs would respond to stronger incentives, and the MSSP's voluntary nature constrains reform options. If we give ACOs a pace to beat instead of a reason to set their own, we may never know what could have been.

Disclosure forms provided by the authors are available at NEJM.org.

From the Department of Health Care Policy, Harvard Medical School (J.M.M., B.E.L., M.E.C.), the Division of General Internal Medicine and Primary Care, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School (J.M.M.), the Division of General Internal Medicine and Primary Care, Department of Medicine, Beth Israel Deaconess Medical Center (B.E.L.), the Department of Otolaryngology–Head and Neck Surgery, Massachusetts Eye and Ear Infirmary (V.K.R.), and Harvard Business School (V.K.R.) — all in Boston.

1. McWilliams JM, Hatfield LA, Landon BE, Hamed P, Chernew ME. Medicare spending after 3 years of the Medicare Shared Savings Program. N Engl J Med 2018;379:1139-49.
2. Verma S. Quality Payment Program (QPP) year 1 performance results. Centers for Medicare & Medicaid Services blog. November 8, 2018 (https://www.cms.gov/blog/quality-payment-program-qpp-year-1-performance-results).
3. McWilliams JM, Landon BE, Chernew ME. Letter to Seema Verma, administrator, Centers for Medicare & Medicaid Services, RE: [CMS–1701–P] Medicare program; Medicare Shared Savings Program; accountable care organizations — pathways to success; proposed rule. September 18, 2018 (https://hmrlab.hcp.med.harvard.edu/files/hmrtest2/files/september_18_2018_comment_letter_re_cms-1701-p_medicare_program_mssp.pdf).
4. Kanter GP, Pauly MV. Coordination of care or conflict of interest? Exempting ACOs from the Stark Law. N Engl J Med 2019;380:410-1.
5. Douven R, McGuire TG, McWilliams JM. Avoiding unintended incentives in ACO payment models. Health Aff (Millwood) 2015;34:143-9.

DOI: 10.1056/NEJMp1900537

Innovation in Genomic Data Sharing at the NIH

Bruce M. Psaty, M.D., Ph.D., Stephen S. Rich, Ph.D., and Eric Boerwinkle, Ph.D.

In October 2018, the National Institutes of Health (NIH) released draft provisions for data sharing from federally funded research (osp.od.nih.gov/wp-content/uploads/Data_Sharing_Policy_Proposed_Provisions.pdf). For research involving human participants, the recommendations highlight a key tension: "Sharing scientific data derived from human specimens warrants special attention and should be done as broadly as possible, consistent with the consent of the individual participants." Making publicly funded data widely available not only increases its value but also enhances reproducibility of findings.

The NIH position on widespread data sharing is not new.

The National Center for Biotechnology Information established the database of Genotypes and Phenotypes (dbGaP) in 2007 as a repository for individual-level data used in gene-mapping research.[1] An investigator-initiated data access request for specific data sets, with review and approval by an NIH data access committee and local institutional attestations, is required for obtaining data. This process ensures that the proposed research is consistent with the participants' informed consent and protects their privacy. Over time, funding for genotype-array studies was contingent on the study investigators' depositing phenotype and genotype data into dbGaP. Even though some studies with extensive phenotype data may not have had informed-consent documents that specifically addressed the depositing of individual-level data into a repository such as dbGaP for widespread sharing, institutional review boards nonetheless usually approved this data-sharing mechanism.

Technological advances that enabled the performance of genomewide association studies transformed the conduct of common-disease epidemiology. Originally, investigators conducting NIH-funded cohort studies reported study-specific findings with little effort at replication. The large number of statistical tests required for a genomewide association study posed a special challenge, and few studies were sufficiently large to provide adequate statistical power for detecting small-to-modest but biologically meaningful effect sizes. The need for large sample sizes and the importance of replication in genetic discovery were powerful incentives for scientific collaboration.

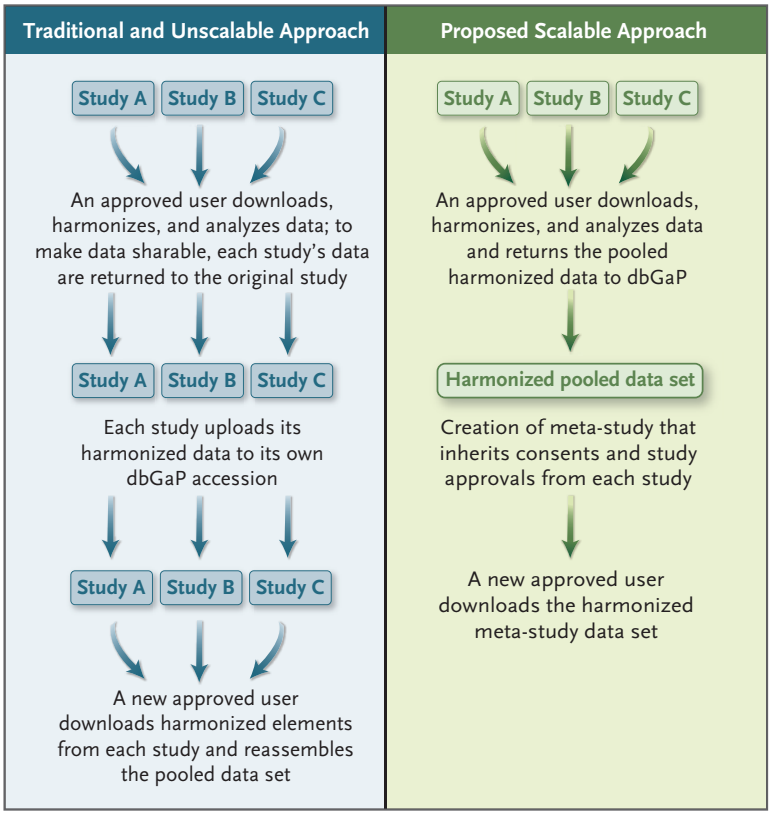
Collaborators had two general options: create a single pooled data set or use meta-analysis. For the creation of a pooled data set in a multistudy collaboration, the participant-level data can be obtained either from the individual studies or from dbGaP. The data-distribution agreements of both the individual studies and dbGaP, however, prohibit the transfer of those pooled data to investigators at other sites. Thus, the considerable work of creating harmonized phenotype data cannot be shared. To create a pooled data set available to all contributing investigators in a multistudy collaboration, the study-specific data-distribution agreements would require bilateral approvals between each pair of studies and their institutions. For 3 studies, 6 agreements would be required; for 10 studies, 90 agreements would be required; and so forth.

To avoid the administrative complexities of sharing individual-level data among studies, a meta-analysis approach using summary statistics from each study became the preferred approach.[2] Hundreds of genomewide association studies have not only identified thousands of variants associated with a variety of phenotypes but have also revealed new disease mechanisms and therapeutic targets. A recent meta-analytic publication includes data on more than 1 million participants.[3] A key, but hidden, component of the prospectively planned meta-analysis is the harmonization of phenotype data across studies without ever formally bringing the individual-level data together.

Technological advances that made genome sequencing affordable have generated new pressure to create pooled data sets. Genome sequencing identifies a large number of rare variants, and efficient use of these extensive genetic data sets favors a pooled central analysis rather than meta-analysis. Low-frequency variant data also sparked the development of new statistical methods, including burden tests that aggregate test statistics for rare variants across genomic regions.[4] For all but the simplest analyses of genome-sequence data, pooling of individual-level data becomes far more efficient than meta-analysis.

Large, NIH-supported, whole-genome-sequencing projects — the Trans-Omics for Precision Medicine (TOPMed) program of the National Heart, Lung, and Blood Institute (NHLBI), the National Human Genome Research Institute Genome Sequencing Program (GSP), the National Institute on Aging Alzheimer's Disease Sequencing Project, and the National Cancer Institute Cancer Genome Atlas — have highlighted the need for efficient methods of data pooling. The TOPMed plan calls for sequencing of about 140,000 whole genomes from more than 80 studies focused on heart, lung, blood, and sleep disorders, and the GSP supports mendelian and complex-disease exome and whole-genome sequencing in a similar number of participants.

In TOPMed, for instance, all study data are deposited in and retrieved from dbGaP. One of the responsibilities of the data coordinating center is phenotype-data harmonization. These harmonized data, however, cannot be returned directly to dbGaP as a pooled data set for general access; rather, the contract-funded data-coordi-



Approaches to the Sharing of Genomic Data.

nating center separates the data by study and returns each study's data to its own accession. For investigator-initiated pooling efforts, the process is even more complicated: the pooled data from each study would be returned to the individual study, and that study's investigators would upload their part of the pooled data to their study's own accession in dbGaP (see left flowchart). It is possible for a new user to submit dbGaP data requests, obtain the harmonized data from each study, and reassemble the harmonized pooled data set that was originally created, but the reassembled data set cannot be shared with others. The value created by the considerable investigator effort in data harmonization is squandered in complying with regulations established in an earlier era under an outdated model of a simple data repository. The development of a new approach, for instance, one with the ability to upload the harmonized pooled data set as a meta-study that inherits the appropriate consents and study approvals, would facilitate data sharing (see right flowchart).

Working with NIH project officers, several investigators from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium developed a consortium agreement that requires only one agreement per study to pool data.[5] Phenotype data are harmonized at the local level and then uploaded to a single site and pooled. Instead of distributing copies of the data sets to analysts around the world, a cloud-based Analysis Commons brings together genotype and phenotype data from participating studies in a platform that is accessible to multiple authorized analysts. This framework addresses many of the challenges of multicenter analyses of sequence data, including robust security protocols, data-sharing mechanisms, phenotype harmonization, integrated multi-omics analyses, annotation, and computational flexibility. Nevertheless, access by nonstudy investigators remains a challenge.

Part of the NIH Data Commons, the NHLBI Data STAGE (Data Storage, Toolspace, Access, and analytics for biG-data Empowerment, www.nhlbi.nih.gov/science/data-storage-toolspace-access-and-analytics-big-data-empowerment-data-stage) is a cloud-based platform designed for "innovative computing solutions that meet the needs of the NHBLI and our research community." TOPMed is one of the flagship data sets. The long-term goal is to leverage "emerging opportunities in data science to open new frontiers in heart, lung, blood, and sleep (HLBS) research." Data STAGE will rely on dbGaP permissions to control access to individual-level data.

Major advances leveraging these large-scale genomic and phenotype data require not only contemporary analytics based on deep learning and artificial intelligence, but also administrative and regulatory innovation. Without revisions or novel approaches, the use of dbGaP as a mechanism for providing access to data in Data STAGE will consume investigator time, increase the burden on dbGaP personnel, and

hinder progress in large-scale collaborative efforts. Within large institute-based efforts, a consortium agreement, similar to the one used in the Analysis Commons, that treats all study investigators as part of one multicenter project would accelerate discovery and translation. New mechanisms could be developed for sharing data (primary, harmonized, and manipulated) with outside, or nonstudy, investigators. In addition, it would be helpful for data sets from large-scale collaborations to be archived and available, since they can facilitate the reproducibility of findings and serve as a testing ground for novel methods. Access to these archived data sets can still be controlled and participants' consent respected.

Administrative innovations in data sharing to promote big-data science will not emerge on their own. The NIH can devise a new set of data-access policies and regulations that would be fit for the purpose and appropriate for current and future forms of biomedical data. To promote maximum use of publicly funded research data, the NIH can mandate intelligent and innovative data-sharing methods, similar to those that made research funding of genomewide association studies contingent on an agreement to submit data to dbGaP.

Disclosure forms provided by the authors are available at NEJM.org.

From the Cardiovascular Health Research Unit, Departments of Medicine and Epidemiology, University of Washington, Seattle (B.M.P.); the Center for Public Health Genomics, University of Virginia, Charlottesville (S.S.R.); and the School of Public Health, University of Texas at Houston, Houston (E.B.).

1. Mailman MD, Feolo M, Jin Y, et al. The NCBI dbGaP database of genotypes and phenotypes. Nat Genet 2007;39:1181-6.
2. Psaty BM, O'Donnell CJ, Gudnason V, et al. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: design of prospective meta-analyses of genome-wide association studies from 5 cohorts. Circ Cardiovasc Genet 2009;2:73-80.
3. Evangelou E, Warren HR, Mosen-Ansorena D, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. Nat Genet 2018;50:1412-25.
4. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet 2011;89:82-93.
5. Brody JA, Morrison AC, Bis JC, et al. Analysis commons, a team approach to discovery in a big-data environment for genetic epidemiology. Nat Genet 2017;49:1560-3.

DOI: 10.1056/NEJMp1902363