



Published in final edited form as:

Acad Med. 2016 March ; 91(3): 433–440. doi:10.1097/ACM.0000000000001028.

The Changing Nature of Scientific Sharing and Withholding in Academic Life Sciences Research: Trends From National Surveys in 2000 and 2013

Darren E. Zinner, PhD,

Associate professor, Schneider Institutes for Health Policy, Heller School for Social Policy and Management, Brandeis University, Waltham, Massachusetts

Genevieve Pham-Kanter, PhD, and

Assistant professor, Department of Health Management and Policy, Drexel University School of Public Health, Philadelphia, Pennsylvania, and a research fellow, Edmond J. Safra Center for Ethics, Harvard University, Cambridge, Massachusetts

Eric G. Campbell, PhD

Professor of medicine, Mongan Institute for Health Policy, Massachusetts General Hospital, and Harvard Medical School, Boston, Massachusetts

Abstract

Purpose—Since 2000, federal funders and many journals have established policies requiring more open sharing of data and materials post-publication, primarily through online supplements and third-party repositories. This study examined changes in sharing and withholding practices among academic life scientists, particularly geneticists, between 2000 and 2013.

Method—In 2000 and 2013, the authors surveyed separate samples of 3,000 academic life scientists at the 100 U.S. universities receiving the most NIH funding. Respondents were asked to estimate the number of requests for information, data, and materials they made to and received from other academic researchers in the past three years. They were also asked about potential consequences of sharing and withholding.

Results—Response rates were 63.9% (1,849/2,893) in 2000 and 40.8% (1,165/2,853) in 2013. Proportions of faculty in 2000 and 2013 who received, denied, made, or were denied at least one request were not statistically different. However, the total volume of requests received from or made to other scientists dropped substantially (19.4 received in 2000 vs. 10.8 in 2013, $P < .001$; 8.4 made in 2000 vs. 6.6 in 2013, $P < .001$). Faculty in 2013 also made an average of 8.4 requests

Correspondence should be addressed to Darren Zinner, Brandeis University, 415 South Street, MS035, Waltham, MA 02453; telephone: (781) 736-3971; dzinner@brandeis.edu.

Other disclosures: None reported.

Ethical approval: Both the 2000 and 2013 surveys were approved by the institutional review board at the Massachusetts General Hospital.

Previous presentations: Portions of this research were presented at the Cold Spring Harbor Laboratory International Biology of Genomes Conference; May 2015; Cold Spring Harbor, New York.

Supplemental digital content for this article is available at [LWW INSERT LINK TO SDAPP1 FILE] and [LWW INSERT LINK TO SDAPP2 FILE].

to third-party repositories. Researchers in 2013 were less likely to report that sharing resulted in new research or new collaborations.

Conclusions—The results show a dramatic shift in sharing mechanisms, moving away from a peer-to-peer sharing model toward one based on central repositories. This may increase efficiency, but collaborations may suffer if personal communication among scientists is deemphasized.

Open disclosure of experimental methods and results is required to achieve optimal progress in science. This scientific sharing allows researchers to publicly vet and build upon others' results and to avoid duplicative work. Most public research funders require the open dissemination of research results and materials as a condition of grants. However, the personal rewards of science—individual recognition and career advancement—are predicated upon the number and quality of the individual scientist's scientific publications and patents, which require periods of secrecy to secure priority.¹ These two incentive structures conflict when a scientist is asked to share specific information, data, and materials mentioned in a journal article.

In a national survey conducted in 2000, we documented the level of sharing and withholding of materials, techniques, and data post-publication among U.S. life sciences researchers broadly, with a focus on geneticists in particular.² We found that the reasons geneticists cited most frequently for withholding requested information, data, or materials were not wanting to expend the effort to produce them (80%) and needing to protect their own ability (53%) or the ability of a graduate student, postdoctoral fellow, or junior faculty member (64%) to publish.

Since 2000, researchers in the life sciences have witnessed a rapid uptake in policies and recommendations to encourage scientific sharing.³⁻⁵ For example, recent National Institutes of Health (NIH) policies specify that data sharing plans must be submitted in grant applications and that the data, materials, and tools developed from NIH-funded research must be made readily available to the research community.⁶ Most top journals require that data and detailed methodological information be made available online.^{7,8} Also, third-party repositories have been created to aggregate data and biomaterials, making sharing easier and more efficient.⁹ Meanwhile, social and technological trends have created easier pathways for scientific collaboration, including an increased emphasis on team science, the use of online forums for scientists (e.g., ResearchGate), and the ubiquity of internet connectivity.

To empirically document how data sharing and withholding have changed in the life sciences as a result of these new scientific sharing initiatives, we repeated our 2000 survey in 2013 to estimate the prevalence and magnitude of sharing requests made, honored, and denied. We studied the most commonly funded types of academic biomedical research in order to be able to generalize nationally. We also focused on certain subsets of researchers-- faculty in clinical departments, in basic life science departments, and in genetics departments or programs--to identify important trends both in the formal process of and cultural attitudes toward scientific sharing across the life sciences research community.

Method

Sample selection

In 2000 and 2013, we selected separate samples of 3,000 university faculty members to survey using an identical, four-step process.

First, we selected the 100 U.S. higher education organizations that received the most NIH extramural research support, based on the latest list of NIH awards by organization¹⁰ (i.e., fiscal years 1998 and 2010 for the 2000 and 2013 surveys, respectively). We combined universities with their medical schools if they were listed separately by the NIH; a handful of organizations did not have a corresponding medical school (e.g., California Institute of Technology, Massachusetts Institute of Technology). Using the same data source, we aggregated NIH awards by recipient department to identify the types of departments that received the most NIH funding. We separated these departments into clinical and basic life science groups and selected the top five departments in each group. The clinical departments were internal medicine, neurology, pathology, pediatrics, and psychiatry; the basic life science departments were biochemistry, biology, microbiology/immunology/virology, pharmacology, and physiology. Our 2000 survey documented that both sharing and withholding were more pronounced among geneticists than other life scientists, which may be attributable to the greater scientific competitiveness in the field and the potential commercial value of genetic products.² Consequently, in 2013 as in 2000, we included genetics departments and programs in their own subgroup.

Second, for each institution, we randomly selected one of the five clinical departments and one of the five basic life science departments. We also selected all genetics departments and programs (i.e., if the institution had two or more genetics programs, we selected all of them).

Third, we collected the names of all research faculty within these selected departments using departmental Web sites. We maintained three separate strata of faculty: faculty in clinical departments, faculty in basic life science departments, and faculty in genetics departments/programs. To avoid the inclusion of lab managers, research assistants, or hospital staff members not truly functioning as faculty researchers, individuals in the clinical stratum were excluded if they had not published at least one research article listed in the U.S. National Library of Medicine's MEDLINE database within the last three years. We randomly ordered faculty names by stratum. We retrieved addresses and telephone numbers from departmental and university Web sites for the faculty selected for the sample.

Finally, we added a fourth stratum of faculty: principal investigators listed by the NIH as being directly funded by the Human Genome Project (HGP) and/or the National Human Genome Research Institute (NHGRI) in the previous five years.¹⁰ These individuals included researchers from the institutions listed above (duplicates removed from the other three strata), other academic institutions, and independent research centers. We similarly extracted names and addresses for these faculty from departmental and institutional Web sites.

For each survey, a stratified sample of 3,000 faculty members was aggregated and sampling weights were calculated by the probability of inclusion in each of the four strata to better generalize to the entire population of U.S. life sciences researchers.¹¹ Responses from faculty in clinical and basic life science departments carried a larger weight than responses from faculty in genetics departments/programs because there are many more clinical and basic life sciences departments and these departments include larger numbers of faculty compared with genetics departments/programs. The 2000 sample included all 219 grantees of the NHGRI and all 1,547 faculty in genetics programs/departments; the remainder of the sample (n = 1,234) was randomly selected so that half came from basic life science departments (n = 617) and half from clinical departments (n = 617). The 2013 sample included all 483 grantees from the NHGRI and 1,317 faculty in genetics programs/departments; the remainder of the sample (n = 1,200) was randomly selected so that half came from clinical departments (n = 600) and half from basic life science departments (n = 600).

Survey design and administration

The design of the 2000 survey instrument was informed by two focus group discussions and 20 semi-structured interviews with knowledgeable biomedical researchers, as well as discussions with colleagues and reviews of the literature. The 2013 survey instrument included questions identical to those used in 2000 and 21 new items. Of the new questions, three were included in this study's analysis. These three questions pertain to the role of journal policies on the use of online supplements and third-party repositories in scientific sharing—important developments identified from the two focus groups and 10 semi-structured interviews conducted to inform the design of the 2013 survey. (The 2013 survey instrument is available as Supplemental Digital Appendix 1 at [LWW INSERT LINK to SDAPP1 FILE].)

In each study, the faculty in the sample were sent a letter, a fact sheet describing the study, a survey instrument with a postage-paid return envelope (with no identifying information), and postage-paid confirmation postcard (with a respondent identification number). Faculty were asked to complete the survey and the postcard and mail them back separately. This process enabled tracking of nonrespondents via the postcards while ensuring respondents' complete anonymity, because the survey instrument had no unique identifying information. Approximately six weeks later, nonrespondents were mailed a reminder letter and an additional survey with \$10 cash. Persistent nonrespondents were contacted by telephone approximately three weeks after that and encouraged to participate.

The 2000 survey was administered by the Center for Survey Research (CSR) of the University of Massachusetts Boston between March and July 2000. The 2013 survey was administered by Harris International between January and June 2013. Surveys were returned to CSR (2000) or Harris International (2013), who compiled the data. Both surveys were approved by the institutional review board at the Massachusetts General Hospital.

Measures and variables

Self-identification of geneticists—Respondents identified themselves as geneticists by responding “yes” to the following question:

Do you consider yourself a genetics researcher? *By genetics researcher we mean someone whose research involves any of the following: (1) identification of genomes, genes, or gene products in any organism; (2) study of the structure, function, or regulation of genes or genomes; (3) comparison of genes and genomes between species or populations.*

Measuring sharing and withholding—The surveys used multiple measures of data sharing and withholding. First, we asked scientists how many times in the last three years they had *made a request* to another academic scientist to provide information, data, or materials concerning published research. We then asked those who had made such requests to estimate the number of their requests that were denied or were honored only after a significant delay.

Second, we asked scientists how many times in the last three years they had *received requests* from other academic scientists for information, data, or materials concerning their published research. We then asked those who had received such requests to estimate the number that they denied or honored only after a significant delay.

Consequences of sharing and withholding—The surveys had three batteries of questions regarding the potential consequences of sharing and withholding. The first battery asked whether, as a result of another academic scientist’s failure to share information, data, or materials, the respondent had ever had a publication significantly delayed; been unable to confirm others’ published research; abandoned a promising line of research; stopped collaborating with another academic scientist; appealed to a funding agency, journal, or professional association; refused to share their own information, data, or materials with that person or group; or delayed sharing with that person or group.

The second battery asked about the positive and negative consequences of sharing, such as whether the respondent had formed collaborations that led to grants, formed collaborations that led to publications, opened a new line of research, or performed research that would otherwise not have been possible. This set of questions also asked whether the respondent had been “scooped” by another scientist; compromised the ability of a graduate student, post-doctoral fellow, or junior faculty member to publish; or been unable to benefit commercially from their results.

The third battery addressed respondents’ perceptions of the effects of withholding, asking how data withholding among academic scientists affects the progress of science in their field, the level of communication in their field, the education of students and postdoctoral fellows, the progress of their research, the quality of their relationships with other academic scientists, and their satisfaction with their professional career. The response categories were “enhances greatly,” “enhances somewhat,” “no effect,” “detracts somewhat,” and “detracts greatly.”

Policies to promote sharing—In the 2013 survey, we posed a new battery of questions to measure the prevalence and influence of journal policies in promoting scientific sharing. Faculty were asked whether, in the past three years, they had been required by a journal to submit a detailed description of their methods as an online supplement, to make data available as an online supplement, to submit biomaterials (e.g., tissues, reagents, organisms) to a third-party repository, or to submit data to a third-party repository. For ease of reporting, we combined responses for the first two items (e.g., answered “yes” to providing methods or data in an online supplement) and the last two items (e.g., answered “yes” to submitting biomaterials or data to a third-party repository).

Analysis

The data were analyzed using standard statistical techniques using Stata SE 13.1 for Windows (Stata Corp, College Station, Texas). Except for raw counts of survey responses (see Table 1 and the top row of Table 2 [LWW: THESE ARE NOT THE TABLE 1 AND 2 CALLOUTS; CALLOUTS APPEAR IN RESULTS SECTION]), all data were weighted to adjust for differential nonresponse and probability of selection within sampling strata. All tests were two-sided and tested at the $P = .05$ level. For univariate statistics, means and standard deviations (SDs) were calculated. Bivariate analyses of unweighted differences in proportions and means were tested using chi-square tests.

Multivariate logistic regression analyses were used to determine whether the probability of a life scientist making or receiving a request was different between the two survey periods. These regressions controlled for gender, whether scientists were trained in the United States, whether they engaged in human subjects research, whether they had research funding from industry, and whether they had engaged in commercial activities (including having applied for patents, patents issued, patents licensed, a product under regulatory review, a product on the market, or a start-up company). Regressions also controlled for the number of peer-reviewed articles published by the responding scientist in the last 3 years, split into low (0–5 publications), medium (6–15 publications), and high (> 16 publications) categories. These variables were selected according to the results of our previous research into the causes of data withholding in the life sciences² and unpublished data from personal interviews and focus groups.

For multivariate analyses that examined the likelihood of a life scientist denying a request (i.e., another scientist denying a request made by a respondent or a respondent refusing a request made by another scientist), in addition to the variables described above we included variables representing the volume of requests received in the last three years.

Results

For each survey, 3,000 faculty were invited to participate. In the process of administering the surveys, 107 scientists in the 2000 sample and 147 scientists in the 2013 sample were deemed ineligible because they had died or retired or were out of the country, on sabbatical, not located at the sampled institution, or lacking faculty appointments. In 2000, 1,849 of the remaining 2,893 scientists responded, yielding an overall response rate of 63.9%. In 2013, 1,165 of the remaining 2,853 faculty responded, yielding a response rate of 40.8%.

In 2000, through the 256 follow-up telephone interviews conducted as part of that survey, we found nonrespondents were significantly more likely to be full professors and significantly less likely to be geneticists than respondents. In 2013, telephone interviews with nonrespondents were not conducted, but based on the postcards returned ($n = 979$) and information gleaned from departmental Web sites, response rates were significantly lower among faculty holding MD degrees versus those with PhD degrees and among faculty in clinical departments versus those in basic life science departments.

Characteristics of respondents

Table 1 provides the characteristics of the 1,849 academic life scientists who responded to the 2000 survey and the 1,165 academic life scientists who responded to the 2013 survey. The 2013 survey respondents reflected the changing demographics of science: Compared with respondents in 2000, they were more likely to be female (24.4% in 2000 vs. 29.4% in 2013, $P = .003$), to have trained outside the United States (12.7% in 2000 vs. 19.4% in 2013, $P < .001$), and to conduct research on human subjects (33.0% in 2000 vs. 38.3% in 2013, $P < .003$). Conversely, they were less likely to self-identify as a geneticist (67.1% in 2000 vs. 52.8% in 2013, $P < .001$) and to be engaged in commercial activities (33.3% in 2000 vs. 27.0% in 2013, $P < .001$).

Prevalence of sharing and withholding

Across the two time periods, there were no significant differences in the proportion of faculty who had made requests to other academic scientists in the past three years for information, data, or materials concerning published findings: In 2000, 66.7% of respondents reported making at least one request compared with 65.6% in 2013 ($P = .65$). Similarly, there were no statistical differences in the proportions of faculty who reported receiving (76.0% in 2000 vs. 73.4% in 2013, $P = .26$) or denying (10.1% in 2000 vs. 8.7% in 2013, $P = .40$) at least one request, or the percentage of requests denied (2.1% in 2000 vs. 2.0% in 2013, $P = .11$).

The proportions of faculty making and receiving requests were also stable across subgroups of researchers across the two surveys. Geneticists (84.4% in 2000 vs. 81.2% in 2013, $P = .22$) and faculty in basic life science departments (77.0% in 2000 vs. 77.6% in 2013, $P = .85$) were more likely to make requests of another scientist than were faculty in clinical departments (49.7% in 2000 vs. 51.4% in 2013, $P = .70$). Likewise, geneticists (91.7% in 2000 vs. 89.5% in 2013, $P = .24$) and faculty in basic life science departments (84.9% in 2000 vs. 81.3% in 2013, $P = .23$) were more likely to receive requests compared with faculty in clinical departments (61.7% in 2000 vs. 62.4% in 2013, $P = .86$).

From 2000 to 2013, the proportion of respondents who made at least one request of another academic scientist decreased for nearly all types of requests for additional information not included in the publication—lab techniques (54.0% in 2000 vs. 39.1% in 2013, $P < .001$), pertinent findings (33.7% in 2000 vs. 24.0% in 2013, $P < .001$), phenotypic information (15.6% vs. 13.4%, $P = .11$), and genetic sequences (17.0% in 2000 vs. 10.1% in 2013, $P < .001$)—as well as for requests for biomaterials including probes, cell lines, tissues, reagents,

and organisms that were mentioned in the publication (58.5% in 2000 vs. 44.6% in 2013, $P < .001$).

Volume of requests

The volume of requests life scientists received for information, data, or materials after publication of a finding significantly declined from 2000 to 2013 (Figure 1). The average number of requests received by respondents in the last three years (among those who received at least one request) was 19.4 (SD = 1.2) in 2000 compared with 10.8 (0.6) in 2013 ($P < .001$). Declines in the average number (SD) of requests received were also seen across the sample subgroups: faculty in clinical departments (10.3 [1.1] in 2000 vs. 6.7 [0.8] in 2013, $P = .007$), faculty in basic life science departments (18.5 [2.7] in 2000 vs. 12.6 [1.4] in 2013, $P = .05$), and geneticists (26.8 [1.9] in 2000 vs. 13.2 [1.0] in 2013, $P < .001$).

Likewise, the average number (SD) of requests made by respondents to other academic scientists in the last three years (among those who made at least one request) decreased by nearly 25%, from 8.4 (0.4) requests in 2000 to 6.6 (0.3) requests in 2013, $P < .001$. Declines in the volume of requests made were witnessed across the sample subgroups: faculty in clinical departments (30.2% decrease, $P = .03$), faculty in basic life science departments (15.4% decrease, $P = .16$), and geneticists (21.6% decrease, $P < .01$). Details by sample subgroup are presented in Figure 2.

Characteristics of researchers making, receiving, and refusing requests

We conducted four separate multivariate analyses to determine the characteristics of respondents who, over the past three years, (1) made at least one request to another academic scientist, (2) were denied at least one request, (3) received at least one request from another academic scientist, and (4) refused at least one request for information, data, or materials. (Full regression statistics are available in Supplemental Digital Appendix 2 at [LWW INSERT LINK TO SDAPP2 FILE].)

Controlling for all listed factors, respondents in 2013 were half as likely as those in 2000 to experience a denial of a request (odds ratio [OR], 0.51; 95% confidence interval [CI], 0.39–0.66). Overall, among all respondents (2000 and 2013), there was no statistical difference by gender or by domestic versus foreign training in the likelihood of making or receiving a request. Geneticists were significantly more likely than other life scientists to make requests (OR, 3.42; 95% CI, 2.68–4.36) and to receive requests (OR, 4.61; 95% CI, 3.46–6.14). Researchers engaged in human subjects research were half as likely as those not engaged in such research to make requests (OR, 0.46; 95% CI, 0.36–0.59) and to receive requests (OR, 0.50; 95% CI, 0.387–0.66).

In 2000, industry support and engagement in commercial activities (e.g., patenting or licensing an innovation) were significant predictors of secrecy and data withholding.² In 2013, researchers with industry support were roughly twice as likely as those without industry support to make (OR, 2.34; 95% CI, 1.82–3.01) and to receive (OR, 1.94; 95% CI, 1.47–2.56) at least one request, but they were not more likely to experience a denial of a request or to refuse a request. However, as in 2000, scientists engaged in commercial

activities in 2013 were still more likely to refuse a request (OR, 1.63; CI, 1.13–2.35) or to be denied a request (OR, 1.42; CI, 1.09–1.85) compared with scientists not engaged in commercial activities.

Prevalence of online supplement and third-party repository requirements

Since our 2000 survey, research findings, biomaterials, and especially genetic data have been increasingly aggregated into online and independent repositories to create more efficient mechanisms to combine and share information.¹² In 2013, in response to our new questions about online supplements and data repositories, 44.2% of life scientists reported they were required by a journal to submit a detailed description of scientific methods or data as an online supplement and 24.8% reported they were required by a journal to submit biomaterials or data to a third-party repository. This pattern was more pronounced for geneticists: 58.5% were required to submit additional materials online and 47.1% made biomaterials or data available to a third-party repository. In comparison, 36.4% of faculty in clinical departments and 50.7% of faculty in basic life science departments were required to submit additional materials online, and 12.3% of faculty in clinical departments and 30.9% of faculty in basic life science departments made data or materials available to a third-party repository.

Repositories also served as a source for scientific sharing. Over the previous three years, 29.5% of 2013 respondents had initiated at least one request with a third-party repository for information, data, or materials concerning published research. They made an average of 8.4 (SD = 1.2) requests to repositories (Figure 1). Details by subgroups are presented in Figure 2. Faculty reported that 2.4% of repository requests were denied outright, 5.6% were honored only after significant delays, and 3.8% were abandoned due to protracted delays or difficulties (Figure 1).

Cumulatively, 10.8% of researchers experienced at least one denial of a repository request; 23.6% experienced a significant delay; and 5.9% believed the request was honored in a way that was misleading or inaccurate. In addition, 17.9% of respondents reported abandoning at least one request to a third-party repository due to protracted delays or difficulties. Despite these problems, 39.0% of respondents (and 61.5% of geneticists) in 2013 reported that repositories have helped the progress of their research.

When peer-to-peer and third-party requests in 2013 were viewed cumulatively, the total volume of requests had substantially increased (Figure 2). This total volume was especially high among geneticists: 43.3% of geneticists made at least one request to a third-party repository, averaging 11.1 (SD = 2.1) requests in the last three years.

Consequences and perceived effects of sharing and withholding

Both the 2000 and 2013 surveys asked about the consequences of sharing and withholding (Table 2). Across both time periods, a similar proportion of faculty reported being “scooped” by another scientist (28.1% in 2000 vs. 25.7% in 2013, $P = .24$) or that sharing compromised the ability of a junior member of the team to publish (10.1% in 2000 vs. 9.4% in 2013, $P = .55$).

Compared with respondents in 2000, faculty in 2013 were significantly less likely to report that sharing resulted in collaborations that led to new grants (39.5% in 2000 vs. 32.1% in 2013, $P = .001$) or new publications (66.9% in 2000 vs. 45.5% in 2013, $P < .001$). Moreover, faculty in 2013 believed that sharing was less helpful toward innovation, with a lower proportion of faculty than in 2000 indicating that sharing opened a new line of research (37.5% in 2000 vs. 27.3% in 2013, $P < .001$) or allowed for research that would have otherwise not been possible (43.6% in 2000 vs. 28.9% in 2013, $P < .001$).

While geneticists and faculty in basic life science departments indicated higher overall levels of collaboration and research opportunities stemming from sharing, compared with faculty in clinical departments, approximately 20–30% fewer respondents in all three subgroups reported positive benefits in 2013 than in 2000 ($P < .05$ for all comparisons). Scientists in 2013 appeared less likely than scientists in 2000 to “punish” other academics who refused to share: they were less likely to indicate they had stopped collaborating with them (22.0% in 2000 vs. 17.2% in 2013, $P = .01$), delayed sharing with them (13.2% in 2000 vs. 7.5% in 2013, $P = .001$), or refused to share with them again (9.8% in 2000 vs. 4.6% in 2013, $P < .001$).

However, while one-third to one-half of life scientists in 2013 perceived that withholding had negative effects on their work, the effect of withholding was less impactful overall than in 2000, with lower proportions reporting that it detracted from the level of communication in their field, the quality of their relationships with other academic scientists, the education of students and postdoctoral fellows, and the progress of their research and of science in their field overall ($P < .05$ for all comparisons).

Prevalence of other forms of withholding

While the absolute number of request denials reported was lower in 2013 than in 2000, the 2013 survey identified other areas in which information was withheld to safeguard researchers' future interests. In 2013, 24.0% of respondents reported they had intentionally excluded pertinent information from a manuscript submitted for publication in order to protect their scientific lead, and 39.5% admitted they had excluded pertinent information from a presentation of published work at a national conference or meeting. In addition, some respondents reported they had intentionally delayed the publication of their results by more than six months to honor an agreement with a collaborator (9.6%); protect the priority of a graduate student, postdoctoral fellow, or junior faculty member (6.1%); protect their own scientific lead (5.1%); allow time for a patent application (3.7%); or to meet the requirements of an industry (3.1%) or nonindustry (1.3%) sponsor.

Discussion

This study is the longest running cross-sectional study of data sharing and withholding in the life sciences. It was intended to examine the impact of recent developments in scientific sharing that have largely been driven by the requirements of scientific funders and scientific journals.

The results of this study show that the mechanisms of scientific sharing in the academic life sciences have changed dramatically from 2000 to 2013. Peer-to-peer sharing, as measured by the average number of requests a scientist makes or receives from others, dropped significantly. This decrease can likely be attributed to the increasing use of third-party data repositories, journal online supplemental materials, and other institution-based mechanisms for storing and distributing scientific resources, which were in their infancy in 2000.

Overall, these changes in sharing mechanisms have likely resulted in more sharing in science. In 2000, academic life scientists averaged 8.4 requests made for information, data, or materials in the previous three years. In 2013, academic life scientists averaged 15.0 requests in the previous three years (6.6 to other academic scientists plus 8.4 to third-party repositories; see Figure 1). Furthermore, most of these requests were honored: From 2000 to 2013, the incidence of denials remained the same. So while the prevalence of withholding has remained stable, the volume of honored requests has increased significantly when all sharing pathways are considered. These trends are more pronounced among geneticists, whose average of total requests made over the three previous years was 10.4 in 2000 but was 19.6 in 2013 when requests to third-party repositories were included (see Figure 2).

Therefore, from a policy perspective, what are the implications of scientific sharing becoming more institutionalized? First, it is likely that use of institution-based mechanisms has streamlined the methods by which additional data, tools, and materials can be obtained, making the sharing process more efficient. With one submission to a third-party repository, for example, a scientist can effectively share a scientific resource with multiple requesters and thereby eliminate the top barrier to sharing: the effort required to comply with a request.²

However, the increase in efficiency inherent in third-party repositories appears to have come with a cost. Life scientists in 2013 were significantly less likely than those in 2000 to report that sharing of scientific materials had resulted in forming collaborations that led to grants or scientific papers, opening a new line of research, or conducting research that would not otherwise have been possible. This decrease may indicate that important aspects of scientific collaboration may be lost when the process is depersonalized; not all information can be readily transferred by intermediaries. The sharing of tacit information, like research techniques, and creative “brainstorming” interactions are often best conducted through peer-to-peer dialogue. Thus, the reduction in personal communication among scientists—as occurs via requests between individual researchers—may hamper scientific innovation and reduce the outcomes the new sharing mechanisms were intended to promote.¹³ These implications are especially salient for young researchers and the ways in which they will share information and form collaborations. Their mentors will need to work with them to establish new mechanisms for training and new pathways for collaboration. Future research should explore the potential differences in the outcomes of scientific sharing based on the mechanism by which the sharing took place.

Although our results suggest that there is more overall sharing in the academic life sciences in 2013 compared with 2000, data withholding remains a common problem. In 2013, a significant number of scientists reported intentionally leaving pertinent information out of

journal articles (25%) and presentations of published work at conferences (49%) to protect their scientific lead. These findings should raise concerns as journals and conferences are still the primary mechanisms of communication in science. Without full and open communication of methods, techniques, and data, researchers will be unable to replicate published biomedical findings, which undermines one of the basic tenets for the self-corrective nature of science.^{14,15}

Overall, this study provides an overview of how certain policy changes have affected scientific sharing in the academic life sciences research enterprise in the United States. Our findings should, however, be viewed with an understanding of the limitations inherent in survey research. Specifically, the drop in response rates between the 2000 and 2013 surveys is notable. While several studies have documented a general decline in responses to surveys,¹⁶ our 2013 survey was fielded at the time of the 2013 U.S. government budget sequester, which may have further suppressed participation. In addition, while the characteristics of respondents were similar across both surveys (see Table 1), we cannot rule out the possibility that non-collaborative researchers would be less likely to respond.

Additionally, because we relied on self-reporting, our estimate of the percentage of life science researchers who withheld information or denied requests from other academic scientists likely constitutes a lower bound estimate of the proportion who actually participate in this behavior, since respondents are often reluctant to admit engaging in behavior that may be perceived as less than desirable.¹⁷ Further, because our sample included faculty at only research-intensive institutions, our results may not be applicable to faculty at institutions receiving less extramural research support. Finally, other forms of secrecy, such as refusals to publicly present research findings and not discussing research with others, also affect the progress of science and should be factored into future research and policy formulation.

As a whole, this study's findings suggest that changes in policies regarding and the mechanisms of scientific sharing since 2000 have created a more open, productive atmosphere where scientific data and tools can be exchanged more easily, especially in genetics. Policymakers must continue to balance the efficiencies gained from online supplements and central repositories with the potential collaborations lost when peer-to-peer communication is deemphasized.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding/Support: Both the 2000 and 2013 studies were funded by the National Human Genome Research Institute of the National Institutes of Health, grant numbers 5R01HG01789 and 5R01HG006281, respectively.

References

1. Dasgupta P, David P. Towards a new economics of science. *Res Pol.* 1994; 23:487–532.

2. Campbell EG, Clarridge BR, Gokhale M, et al. Data withholding in academic genetics: Evidence from a national survey. *JAMA*. 2002; 287(4):473–480. [PubMed: 11798369]
3. National Research Council. *Sharing Publication-Related Data and Materials: Responsibilities of Authorship in the Life Sciences*. Washington, DC: The National Academies Press; 2003.
4. Ross JS, Krumholz HM. Ushering in a new era of open science through data sharing: The wall must come down. *JAMA*. 2013; 309(13):1355–1356. [PubMed: 23508736]
5. Mello MM, Francer JK, Wilenzick M, et al. Preparing for responsible sharing of clinical trial data. *N Engl J Med*. 2013; 369(17):1651–1658. [PubMed: 24144394]
6. National Institutes of Health. [Accessed August 13, 2015] NIH grants policy statement (Availability of research results). <http://grants.nih.gov/grants/sharing.htm>. Published October 2013
7. Alsheikh-Ali AA, Qureshi W, Al-Mallah MA, Ioannidis JPA. Public availability of published research data in high-impact journals. *PLoS One*. 2011; 6(9):e24357. [PubMed: 21915316]
8. Vines TH, Andrew RL, Bock DG, et al. Mandated data archiving greatly improves access to research data. *FASEB J*. 2013; 27(4):1304–1308. [PubMed: 23288929]
9. Ramos EM, Din-Lovinescu C, Bookman EB, et al. A mechanism for controlled access to GWAS data: Experience of the GAIN Data Access Committee. *Am J Hum Gen*. 2013; 92(4):479–488.
10. National Institutes of Health. [Accessed October 10, 2014] NIH awards by location & organization. Research Portfolio Online Reporting Tools. <http://report.nih.gov/award/index.cfm> Published October 6, 2014
11. Kish, L. *Survey Sampling*. New York, NY: John Wiley & Sons; 1965.
12. Thessen AE, Patterson DJ. Data issues in the life sciences. *ZooKeys*. 2011; 150:15–51. [PubMed: 22207805]
13. Fischer BA, Zigmond MA. The essential nature of sharing in science. *Sci Eng Ethics*. 2010; 16(4): 783–799. [PubMed: 21108019]
14. Merton, RK. *The Sociology of Science: Theoretical and Empirical Investigations*. Storer, N., editor. Chicago, Ill: University of Chicago Press; 1973.
15. Ioannidis JPA. Why most published research findings are false. *PLOS One*. 2005; 2(8):e124.
16. Afrken CL, Balon R. Declining participation in research studies. *Psychotherapy and Psychosomatics*. 2011; 80:325–328. [PubMed: 21829043]
17. Fowler, FJ. *Survey Research Methods*. Newbury Park, Calif: Sage; 1993.

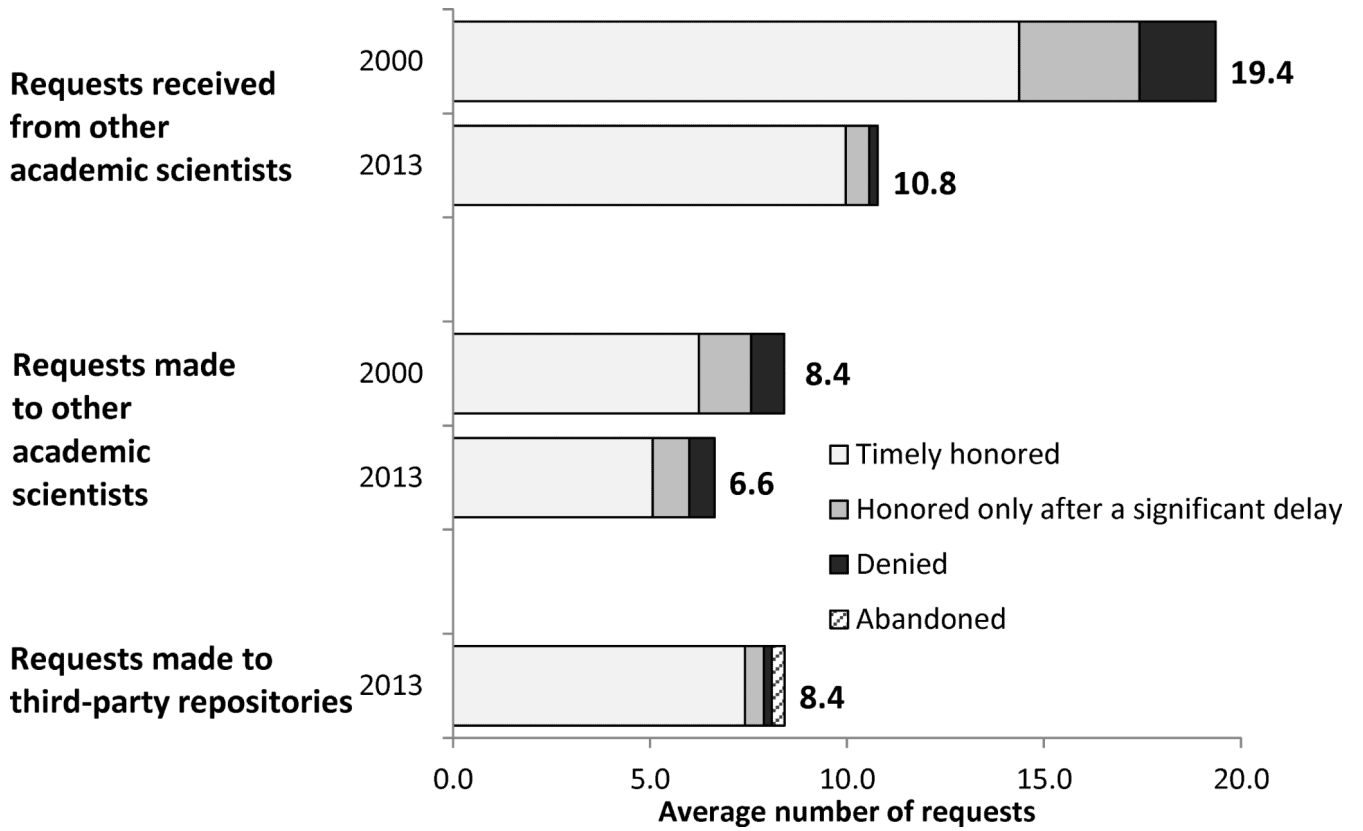


Figure 1. Average number of requests made or received by academic life scientists (who made or received at least one request) in the past three years for data, materials, or information related to published research, 2000 vs. 2013 national surveys. Data are weighted by survey strata.

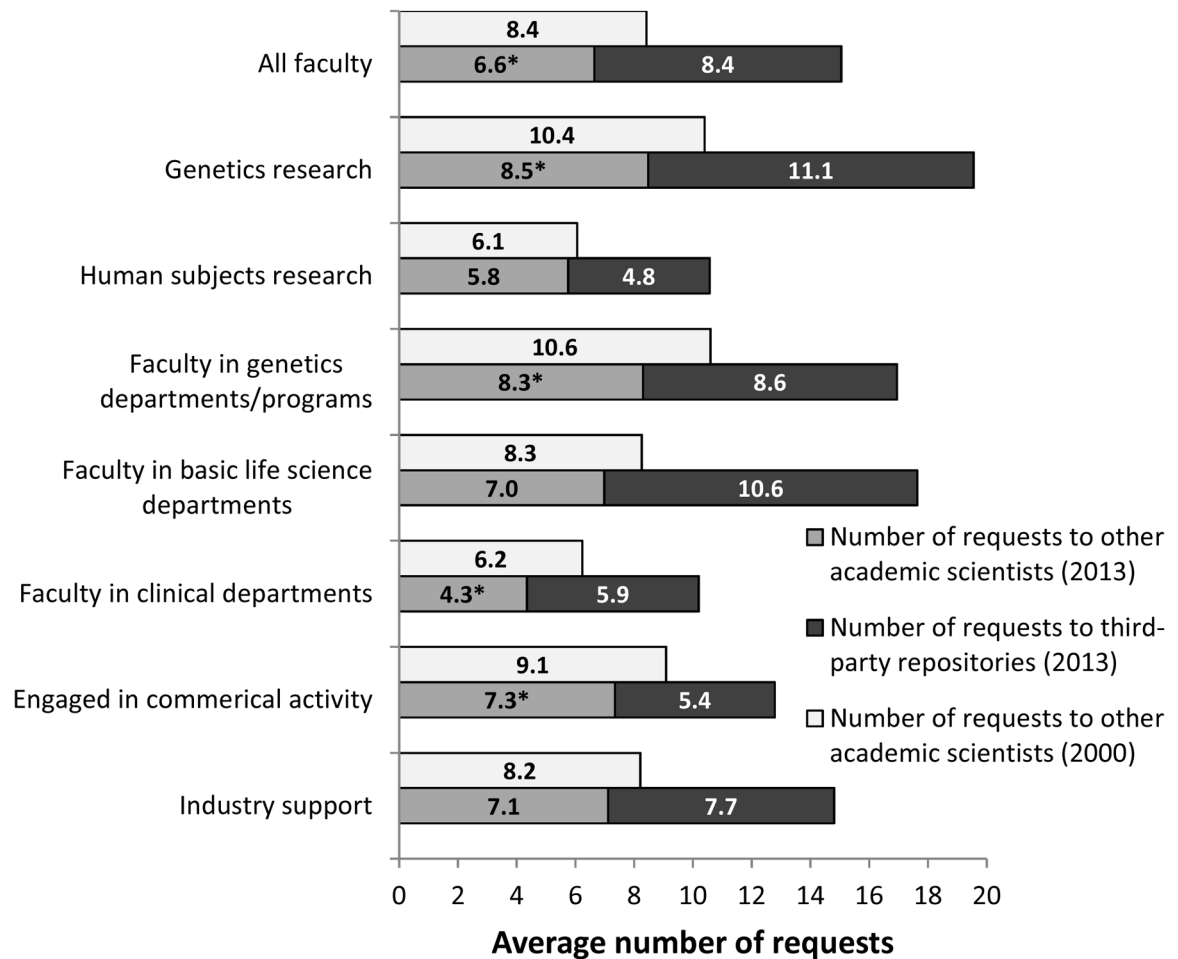


Figure 2. Average number of requests made by academic life scientists (who made at least one request) in the past three years to other academic scientists (2000 vs. 2013) and to data repositories (2013) for data, materials, or information related to published research, by researcher characteristic (see Table 1). Data from 2000 and 2013 national surveys of academic life scientists, weighted by survey strata. * $P < .05$ for difference in number of requests to other academic scientists: 2000 vs. 2013.

Characteristics of Academic Life Scientists Responding to National Surveys on Scientific Sharing and Withholding, 2000 and 2013^a

Table 1

Characteristic	All respondents		2000 respondents		2013 respondents		P value ^c
	No.	% ^b	No.	% ^b	No.	% ^b	
All	3,014	100	1,849	100	1,165	100	
Sex							
Male	2,183	72.4	1,374	74.3	809	69.4	.003
Female	794	26.3	452	24.4	342	29.4	
Academic rank							
Professor	1,483	49.2	916	49.5	567	48.7	
Associate professor	799	26.5	500	27.0	299	25.7	< .001
Assistant professor	641	21.3	404	21.8	237	20.3	
Other/unknown	97	3.2	29	1.6	68	5.8	
Trained in the United States							
Yes	2,518	83.5	1,588	85.9	930	79.8	< .001
No	461	15.3	235	12.7	226	19.4	
Genetics research							
Yes	1,855	61.5	1,240	67.1	615	52.8	< .001
No	1,140	37.8	600	32.4	540	46.4	
Human subjects research							
Yes	1,057	35.1	611	33.0	446	38.3	.003
No	1,939	64.3	1,228	66.4	711	61.0	
Industry support							
Yes	1,964	65.2	1,216	65.8	748	64.2	.19
No	1,008	33.4	599	32.4	409	35.1	
Engaged in commercial activity							
Yes	930	30.9	616	33.3	314	27.0	< .001
No	2,036	67.6	1,192	64.5	844	72.4	
Publications in last 3 years							

Characteristic	All respondents		2000 respondents		2013 respondents		P value ^c
	No.	% ^b	No.	% ^b	No.	% ^b	
Low (0–5)	947	31.4	610	33.0	337	28.9	
Medium (6–15)	946	31.4	618	33.4	328	28.2	.002
High (16)	1,009	33.5	589	31.9	420	36.1	

^a All values reflect raw counts of survey responses and are not weighted by sampling strata.

^b Percentages may not sum to 100% depending on proportion of respondents who left the question blank.

^c P values are based on chi-squared tests of independence comparing categorical variables across the two time periods.

Table 2

Consequences and Perceived Effects of Scientific Sharing and Withholding Among Academic Life Scientists, 2000 and 2013 National Surveys

Item	2000 respondents ^a	2013 respondents ^a	P value
No. of respondents (unweighted)	1,849	1,165	
Negative consequences of another academic scientist's failure to share, % answering "yes"			
Stopped collaborating with another academic scientist	22.0	17.2	.01
Been unable to confirm others' published results	21.1	24.2	.11
Abandoned a promising line of research	17.1	16.6	.80
Had a publication significantly delayed	17.2	15.1	.22
Delayed sharing with that person or group	13.2	7.5	.001
Refused to share your information, data or materials with that person or group	9.8	4.6	< .001
Negative consequences of sharing with another academic scientist, % answering "yes"			
Been "scooped" by another scientist	28.1	25.7	.24
Compromised the ability of a graduate student, post- doctoral fellow, or junior faculty member to publish	10.1	9.4	.55
Been unable to benefit commercially from your results	5.9	2.6	< .001
Positive consequences of sharing with another academic scientist, % answering "yes"			
Formed collaborations that led to publications	66.9	45.5	< .001
Performed research that would otherwise not have been possible	43.6	28.9	< .001
Formed collaborations that led to grants	39.5	32.1	.001
Opened a new line of research	37.5	27.3	< .001
Perceived effects of withholding, % answering "detracts somewhat" or "detracts greatly"			
The progress of science in your field	64.5	56.7	.001
The level of communication in your field	67.0	54.9	< .001
The quality of your relationships with other academic scientists	52.4	44.0	.001
The progress of your research	47.9	40.4	.002
The education of students and post-doctoral fellows	47.6	38.8	< .001
Your satisfaction with your professional career	38.2	42.9	.04

^aTotal survey respondent numbers are unweighted; all other data are weighted by strata based on the probability of being selected for inclusion in the sample. The number of responses for any individual cell in the table may vary slightly, depending on the small proportion of respondents who left the question blank.