

# Perspective on Data-Sharing Requirements for the Necessary Evolution of Drug Development

The Journal of Clinical Pharmacology  
2020, 60(6) 688–690  
© 2020 The Authors. *The Journal of Clinical Pharmacology* published by Wiley Periodicals, Inc. on behalf of American College of Clinical Pharmacology  
DOI: 10.1002/jcph.1607

Jeffrey S. Barrett, PhD, FCP 

## Keywords

data sharing, honest broker, open data, collaboration

At the 2019 annual meeting of the American College of Clinical Pharmacology (Chicago) one of the plenary sessions discussed the expanding roles of data sharing and collaboration. The predominant benefit of data sharing and collaboration is accelerated scientific progress. Advances are clearly valuable to both the pharmaceutical industry and academic researchers in addition to clinicians and the entire health care system, especially when translated into improved patient outcomes, reduced research costs, and decreased time in moving discoveries from the bench to the bedside. Despite the anticipated benefits, sharing research data is still viewed as a work in progress. Likewise, short-term costs to provide data-sharing services are likely to be an issue for some. There were a few obvious take-home messages that resonated from the session. Within the pharmaceutical industry, most real sharing is still occurring in the precompetitive space or targeted in populations in which the financial gains are modest or nonexistent (eg, pediatric oncology, rare diseases, and global health settings). Academic – industry collaborations are broad and can be difficult based on intellectual property considerations and other incentivization issues. Contract research organizations and others in the technology sector will also need to play an important role to bring forward solutions for diverse stakeholders. Meaningful collaboration still requires mutual understanding, and sharing is still problematic for a variety of reasons.

Some positive examples exist, but sustainability is seemingly always in question. Noticeable examples include the National Heart, Lung, and Blood Institute's biological specimen and data repository, Project Data Sphere and the Project Genomics Evidence Neoplasia Information Exchange of the American Association for Cancer Research.<sup>1,2</sup> Although most of these are predominantly academic examples, there are diverse

constituency examples as well. Barriers are typically not an issue of technology limitations.<sup>3</sup> Some of the bottlenecks include the value/overvalue of intellectual property, a lack of resources and expertise, the lack of sharing history or culture of sharing, lack of data-sharing policies and implementation, and the lack of trust for governance around sharing. To be fair the overvalue of intellectual property sentiment is often a reflection of the conservative low-risk legal perspective that typically dictates data-sharing decisions. Clearly, it is in the best interest of the company to protect itself from legitimate intellectual property infringement given the increasingly competitive landscape, but this should not come at the expense of the gains (scientific and financial) that could be made from legitimate collaboration. Some potential solutions have been proposed, but these have mostly been implemented for academic collaborations,<sup>4,5</sup> and industry has been slow to adopt more open data sharing within their organizations, preferring a more traditional Biostat/data management governance model enforced by standard operating procedures focused on protecting clinical data.

Although issues such as data deidentification and the potential for unauthorized reidentification have

---

Bill & Melinda Gates Medical Research Institute, Cambridge, Massachusetts, USA

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Submitted for publication 19 January 2020; accepted 21 February 2020.

## Corresponding Author:

Jeffrey S. Barrett, PhD, FCP, Bill & Melinda Gates Medical Research Institute, 245 Main Street, Cambridge, MA 02139  
Email: Jeff.barrett@gatesmri.org

prevented some from considering sharing and even collaborating, progress has been made in these areas. Analyses of distributed data sets require technical infrastructure and funding to support and maintain the compute environment. It should be possible to achieve meaningful data sharing with embedded research that encourages rather than discourages the growth of a learning health system.<sup>6</sup> Although there is often a mismatch between the explicit motivations, unstated or implicit motivations, and the design of an actual data-sharing policy, this should not dissuade us from pursuing such agreements.<sup>7</sup> As has been observed already, the shift from an aim of changing behavior to changing culture has both subtle and profound implications for policy design and implementation.<sup>8</sup>

One potential solution to the issue of data-sharing governance is an honest broker approach. An honest broker is an entity that keeps sets of private information but distributes parts of those sets to other entities who should not have access to the entire set. The honest broker approach for data sharing offloads the burden of housing identifiable data elements of protected health information (eg, name and address) as well as manage data transfer between clinical and research systems.<sup>3,4,9</sup> The benefit to the sponsor, of course, is the ability to outsource the logging of requests and distribution of data based on contractual rules of engagement tied to data sensitivity. Another advantage is that internal company resources are not used for these activities. Although this is typically a fee-for-service activity for the partner organization, there are some differences from contract to contract, based on the data source and type. Two of the more public examples include the Yale Open Data Access project,<sup>10</sup> <https://yoda.yale.edu/>, and the Duke Clinical Research Institute, <https://dcri.org/our-work/analytics-and-data-science/data-sharing/>,<sup>11</sup> which serve in such a capacity for Johnson & Johnson and Bristol Meyers Squibb pharmaceuticals, respectively. Although these represent a step in the right direction, sharing is still based on low-risk postapproval data only. Likewise, even though honest broker approaches have been engaged for some time, there are still questions about access policies and costs associated with the service as well as potential biases to grantees based on such access.<sup>12</sup> The details of these arrangements, particularly around liability and other legal implications, are also not completely known, and the honest broker as a model is hardly at a stage in which we could consider best practices. Likewise, the honest broker is not the only viable approach, and other third-party solutions involving contract research organizations are also available though not necessarily inexpensive alternatives. A problem with the honest broker approach is that they become a target for hackers and malicious insiders, of course.

Other recommendations for improving sharing within the context of drug development include necessary improvements to the manner and mechanism of internal pharma-sharing solutions. Specifically, these environments need to be more in line with a federated governance model (governance balanced between a central authority and constituent units) and based on information technology (IT) solutions that permit more flexible sharing rules accommodating complex sharing with diverse internal partners, improved and broader data-sharing agreements reflecting the intellectual property considerations of diverse external stakeholders and sharing considerations that can change over time (eg, new partners, change in partners, revised agreements). The generation of an honest broker approach best practices and other commercial solutions that reflect diverse stakeholders and accommodate global data privacy concerns would also help. This does not preclude sponsors from homegrown solutions, of course, although the time and expense including the loss of head count to these activities are not trivial.

Since 2014 the industry has endorsed a commitment to share deidentified individual patient data on request.<sup>13</sup> Two separate studies have confirmed that the extent to which this actually occurs within a reasonable time frame (2 years) is 15% or less.<sup>14,15</sup> Issues identified were highlighted by the lack of data-sharing policies/processes and data-sharing policy conditions that exclude access on the basis of ongoing follow-up and regulatory activity. Although implementation undoubtedly takes time, it would seem that progress is stalled at the moment. Data sharing is in fact a well-studied problem, and other industries (eg, publishing and financial) have worked through their own growing pains to engage in more meaningful sharing. For the pharmaceutical industry to sustain itself and embrace the innovation and collaboration necessary to thrive in a value-based health care system, it will have to learn to share in a manner it is unaccustomed to and address any IT and legal barriers in addition to adopting the requisite internal policies. Likewise, academic investigators will need to cope with internal intellectual property concerns and embrace potential stewardship in an open manner and do their part to address the requisite guidelines, standard operating procedures, and governance to ensure that they are appropriate sharing partners. This will require a more transparent conversation with all relevant stakeholders in which the benefit:risk to sharing is objectively calibrated.

## References

1. AACR Project GENIE Consortium. AACR project GENIE: powering precision medicine through an international consortium. *Cancer Discov.* 2017;7(8):818-831.

2. Ross JS, Ritchie JD, Finn E, et al. Data sharing through an NIH central database repository: A cross-sectional survey of BioLINCC users. *BMJ Open*. 2016;6(9):e012769.
3. Boyd AD, Saxman PR, Hunscher DA, et al. The University of Michigan Honest Broker: A web-based service for clinical and translational research and Practice. *J Am Med Inform Assoc*. 2009;16(6):784-791.
4. Boyd AD, Hosner C, Hunscher DA, Athey BD, Clauw DJ, Green LA. An 'Honest Broker' mechanism to maintain privacy for patient care and academic medical research. *Int J Med Inform*. 2007;76(5-6):407-411.
5. Jarquin PB. Data Sharing: Creating Agreements In support of community-academic partnerships. 2012 Colorado Clinical and Translational Sciences Institute & Rocky Mountain Prevention Research Center. [http://trailhead.institute/wp-content/uploads/2017/04/tips\\_for\\_creating\\_data\\_sharing\\_agreements\\_for\\_partnerships.pdf](http://trailhead.institute/wp-content/uploads/2017/04/tips_for_creating_data_sharing_agreements_for_partnerships.pdf). Accessed January 2020.
6. Platt R, Ramsberg J. Challenges for sharing data from embedded research. *N Engl J Med*. 2016;374(19):1897.
7. Neylon C. Building a Culture of data sharing: policy design and implementation for research data management in development research. *Res Ideas Outcomes*. 2017;3:e21773.
8. Piwowar HA, Becich MJ, Bilofsky H, Crowley RS, on behalf of the caBIG Data Sharing and Intellectual Capital Workspace. Towards a data sharing culture: recommendations for leadership from academic health centers. *PLoS Med*. 2008;5(9):e183.
9. Choi HJ, Lee MJ, Choi C-M, Lee JH, Shin S-Y, Lyu Y, Park YR, Yoo S. Establishing the role of honest broker: bridging the gap between protecting personal health data and clinical research efficiency. *PeerJ* 2015;3:e1506.
10. Yale Open Data Access Report 2013. <https://yoda.yale.edu/>. Accessed January 2020.
11. Institute of Medicine, Board on Health Sciences Policy, Committee on Strategies for Responsible Sharing of Clinical Trial Data, "Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk," Washington, DC: National Academies Press; 2015:153-155.
12. Ross J, Waldstreicher J, Bamford S, et al. Overview and experience of the YODA Project with clinical trial data sharing after 5 years. *Sci Data* 2018;5:180268.
13. Principles for responsible clinical trial data sharing. 2013. <http://phrma-docs.phrma.org/sites/default/files/pdf/PhRMAPrinciplesForResponsibleClinicalTrialDataSharing.pdf>. Accessed February 2020.
14. Murugiah K, Ritchie JD, Desai NR, Ross JS, Krumholz HM. Availability of clinical trial data from industry-sponsored cardiovascular trials. *J Am Heart Assoc*. 2016;5(4):e003307.
15. Hopkins AM, Rowland A, Sorich MJ. Data sharing from pharmaceutical industry sponsored clinical studies: audit of data availability. *BMC Med*. 2018;16(1):165.