Clinical Trial Transparency: The FDA Should and Can Do More

Amy Kapczynski and Jeanie Kim

In recent years, the scientific community and regulators have increasingly recognized the value of data transparency in clinical research. Adding to this momentum, the Blueprint for Transparency at the FDA calls upon the U.S. Food and Drug Administration to act as a key intermediary for sharing clinical trial data. We strongly support the Blueprint’s recommendations for the FDA to proactively release substantially more pre-market and post-market clinical trial data submitted by companies. The FDA is in a particularly good position to facilitate public disclosure of data. Tasked with comprehensively regulating the drugs, devices, and biologics on the U.S. market, the agency has access to a wealth of clinical trial data and receives more data than its European counterpart.¹ Notably, the recently appointed FDA Commissioner Scott Gottlieb has expressed his support for data transparency, acknowledging that the agency should leverage its regulatory authority to release data in the public’s interest.²

In this brief commentary, we show that the FDA has the legal authority to share much more clinical trial data that it currently does. We also show that the primary existing route for obtaining such data from the FDA — individual requests under the Freedom of Information Act (FOIA) — can be used to obtain important categories of data, particularly summary data and metadata. But FOIA cannot substitute for a comprehensive data sharing system that prioritizes public health. The FOIA approach can take years, often requires litigation, and will be piecemeal and reactive in nature. A more proactive approach to the release of data would allow the FDA to set priorities and to build a platform for data sharing that maximizes the benefits for patients and the research community. Such a platform could also, when appropriate, craft data use agreements that protect patient privacy while promoting research integrity and transparency.

The FDA Should Proactively Share More Clinical Trial Data

The FDA’s core public health mission is to ensure that medicines are safe and effective for their intended uses — a task primarily accomplished by evaluating the rigor and sufficiency of the evidence submitted by companies. By sharing data from clinical studies, the FDA can improve the evidence base that informs patients, providers, and payers, and help protect the integrity of the clinical research enterprise. In 2015, the Institute of Medicine (IoM) released one of the most comprehensive reports in support of sharing data from clinical trials.³ The report defines three categories of clinical trial data: “individual participant data” (i.e., raw data and the analyzable data set); metadata, or ‘data about the data’ (e.g., protocol, statistical analysis plan, and analytic code); and summary-level data (e.g., summary-level results posted on registries, lay summaries, publications, and [clinical study reports (CSRs) submitted for regulatory review]).⁴

As the report emphasizes, transparency for each category offers distinct benefits. By sharing summary

Amy Kapczynski, J.D., M.Phil., M.A., is a Professor of Law at Yale Law School and faculty director of the Collaboration for Research Integrity and Transparency and the Global Health Justice Partnership. She received her A.B. from Princeton University, M. Phil. from Cambridge University, M.A. from Queen Mary and Westfield College at University of London, and J.D. from Yale Law School. Jeanie Kim, J.D., is a Research Scholar at Yale Law School and a Research Fellow with the Collaboration for Research Integrity and Transparency. She received her A.B. from Washington University in St. Louis and J.D. from Northeastern University School of Law.
results, the FDA can “help[ ] protect against publica-
tion bias.”6 Many clinical trial results that are submit-
ted for regulatory review are unpublished or differ sig-
nificantly from the results that are selectively reported
in journals.6 Disseminating summary data ensures
that patients and the research community have access
to all trial results that are relevant for clinical care, and
not just the positive or favorable outcomes.7 Further-
more, the FDA receives detailed summaries in CSRs
that contain far more information than what is found
in publications.8 Making CSRs publicly available
“allows for better understanding of regulatory deci-
sions and facilitates the use of analyzable data set.”9
Clinical trial protocols should also be shared because
they provide context for understanding published
data, as well as any summary data and analyzable data
that are shared.10
The FDA is reportedly the only regulator that rou-
tinely obtains analyzable data sets.11 The sharing of such
data sets, accompanied by study protocols and sum-
mary data, to qualified independent researchers “allows
for reanalysis, meta-analysis, and scientific discovery
through hypothesis generation.”12 Secondary research-
ers can help identify issues that may have been missed
by regulators or understudied or buried by companies.13

The FDA Can Proactively Share More Clinical Trial Data
Despite the potential benefits, the FDA does not pro-
actively share most of the clinical trial data in its pos-
session in any comprehensive fashion. Congress has
set forth baseline disclosure requirements for the FDA
as well as some confidentiality obligations. However,
much of the data submitted to the FDA falls some-
where between these two poles, giving the FDA discre-
tion to determine what should be made available in the
public's interest and what should be kept confidential.
Federal agencies generally have the authority to
release information to the public if the release is not
otherwise forbidden by law. FOIA, enacted in 1966,
embodies this core presumption in favor of transpar-
ency.14 In passing FOIA, Congress evinced “a general
philosophy of full agency disclosure unless informa-
tion is exempted under clearly delineated statutory
language.”15 Moreover, while FOIA provides several
exemptions, including Exemption 4 for “confidential
commercial information” (CCI) and Exemption 6 for
personal privacy,16 courts have found that Congress
did not intend for the exemptions themselves to cre-
ate absolute bars to disclosure.17 Rather, the exempt-
tions are “workable standards” designed to permit
an agency to withhold certain information without
limiting its discretion to disclose that information
either proactively or in response to a FOIA request.18

In general, even if information falls under an exemp-
tion, agencies have the discretion to release it if there
is a compelling public interest in disclosure related to
the agency's activities and if the disclosure is not
barred by another law.19
The FDA has yet to fully exert its discretion to release
much of the summary data, metadata, and individual
participant data (IPD) in its possession. To do so, the
FDA would first have to recognize that many types of
clinical trial data can be shared without genuine risks
to patient privacy. Summary-level trial results, CSRs,
and study protocols typically contain no patient-spe-
cific information or can be easily redacted to remove
patient identifiers.20 Analyzable data sets present
more complex privacy concerns, and some may be
difficult to fully de-identify without rendering them
useless for secondary analyses.21 However, emerging
protocols for de-identification make it possible to suf-
ficiently anonymize certain analyzable IPD so that the
risk of re-identification is very small.22 Agencies have
the discretion to weigh legitimate personal privacy
concerns against the public interest in disclosure and
to share data if the privacy risks are minimal in relation
to a public benefit.23 The FDA is well positioned
to consider the privacy risks and public health benefits
associated with sharing different types of clinical trial
data, and the FDA and the U.S. Department of Health
and Human Services have both signaled interest in
exploring ways to share de-identified IPD from clinical
trials submitted for regulatory review.24
While agencies have less discretion with respect to
CCI because of overlapping nondisclosure laws that
prohibit federal employees from unauthorized release
of commercial or financial data, they are still entitled
to substantial deference to determine the initial thresh-
old question of what constitutes CCI.25 This is a criti-
cal point for the FDA when considering various types
of clinical trial data. Much of the clinical trial data
that researchers need for meta-analyses and second-
ary analyses simply is not CCI, or can be redacted to
address any CCI concerns. The IoM Report describes
how commercially sensitive information that reflects
a company’s business strategies and clinical develop-
ment processes can be separated from analyzable data
that are more objectively collected and tabulated.26 For
example, CSRs, which contain manufacturing formu-
las or clinical trial site information, can be redacted
to address legitimate CCI issues. Courts have also
rejected CCI arguments for certain types of clinical
research data, including postmarket study protocols
and raw safety data, where the claims of competitive
harm are negligible or vague.27
The FDA’s lack of proactive disclosure is particu-
larly problematic where data relevant to drug safety
is concerned, because Congress has expressed in clear terms its intention that the agency disclose data relevant to that question. The Food and Drug Administration and Amendments Act (FDAAA) instructs the FDA to maintain a website that provides patients and providers with better access to safety information about drugs and biologics. The FDA must post the most recent FDA-issued safety alerts, warning letters, links to the trial registry and results, and "other material determined appropriate by the [agency]." To fulfill its obligation to release “other material” pertinent to patient safety, the FDA should routinely release at least data like CSRs, summary results, full protocols, and analyzable datasets that can be de-identified. These data are pertinent to drug safety, and to the balancing of risks against clinical benefits for particular indications, and can help patients, providers, and the research community fully understand the safety profile of drugs and devices.

Congress has also pressed the FDA to be more forthcoming and has urged the agency to incorporate broader transparency policies for the benefit of the public. Congress has done this not only in congressional hearings, but also through specific laws, such as statutory provisions that mandate the release of “action packages” — the FDA’s summaries of all safety and effectiveness data in its possession — for every approved new drug or biologic. These disclosure requirements are intended to address the discrepancies between the comprehensive information that the FDA possesses and the selective information that is publicly available, which in many cases has led to widespread patient harms. In order to further bridge information gaps and increase the value of clinical research data, the FDA should revisit whether various types of data legitimately fall under FOIA exemptions, particularly Exemptions 4 and 6, and affirm the scientific and public health value of data sharing. This is particularly appropriate where, as in the case of data that sheds light on drug safety, Congress has expressed its view that an interest is especially compelling.

**Leveraging FOIA to Obtain Clinical Trial Data from FDA: A Partial Solution**

When faced with requests for particular clinical trial data, the FDA has in fact released many types of clinical trial data, implicitly conceding that such disclosures do not raise commercial confidentiality or personal privacy concerns. We recently used FOIA to seek access to clinical trial data for Gilead’s blockbuster Hepatitis C drugs, sofosbuvir (Sovaldi) and ledipasvir/sofosbuvir (Harvoni). Although it took two years of litigation, the FDA has now released tens of thousands of pages of summary data and metadata, including safety and effectiveness summary-level data, full protocols that include analysis plans and amendments, and CSRs. The agency redacted very little — for example, select information about ingredients and manufacturing information that was commercially confidential, and participant contact information that implicated privacy. Notably, Gilead had intervened in the case early on, thereby presumably consenting to all data disclosures and implicitly admitting that there are few CCI and privacy concerns relevant to these categories of data, and that those that exist can be addressed through simple redactions.

Our suit did not resolve the extent to which FOIA can be used to access IPD and analyzable datasets. The orientation of FOIA — a disclosure law designed for all types of governmental information — may sometimes make it a blunt tool for these purposes. FOIA rests on a philosophy of broad public dissemination and equal
access to information. Once one entity receives information under FOIA, the public as a whole is presumed to have the right to access the same material. Courts have historically found that selective or conditional data disclosure arrangements are not consistent with the purpose of FOIA. However, such arrangements — like data use agreements with confidentiality provisions — may, depending on the circumstances, be the best means of sharing analyzable data to secondary researchers at a reasonable cost, while also protecting patient privacy and commercial interests.

Until the FDA proactively releases data on a routine basis, individual FOIA requests are the only mechanism to obtain data that the agency does not release. Our experience with FOIA shows that the process can be a very powerful tool for obtaining clinical trial data, at least of the summary and metadata variety, but that FOIA also has important limits. First, valid FOIA requests can go unfulfilled without the aid of a lawyer to take the agency to court for its failure to timely respond. Second, even when successful, the process is slow. Requests for clinical trial data are likely to be put in the slower “complex” queue because of the high volume and complexity of the data as well as the need for redactions, and so typically it will take years to resolve. Despite the hundreds of hours of legal assistance, it took us nearly two years to begin receiving data pursuant to our FOIA request. The process can be slow even where a research question is exceptionally urgent, and the FDA grants “expedited processing.” In 2014 and 2015, the FDA completed two requests that were granted expedited processing; it took the agency 693 days and 862 days respectively to finish document production.

FOIA is also better suited to individual requests for specific data than for systematic release of data of scientific and public health importance. Production of data generally occurs piecemeal for practical reasons, and recipients of data may, but are not obliged to, release the data they receive to others. Proactive release of data by the FDA would be preferable to the current reactive approach for many reasons. It would allow the agency to ensure that researchers have equitable access to data. The agency could also — and should — prioritize, releasing first those categories of information that are both important and readily redactable, such as CSRs and protocols for widely prescribed drugs. Proactive release would also allow the agency — possibly with additional appropriations — to create a dedicated and centralized platform, alone or in conjunction with other entities, that would give investigators with legitimate scientific and public health inquiries access to redacted and de-identified datasets, similar to the National Institutes of Health’s database for biomedical and clinical research.

The FDA could also design optimal conditions for data sharing. For an example, where appropriate, the agency could implement data use agreements that prohibit improper uses of shared data or further promote transparency by requiring that results of studies using the data be publicly shared.

By proactively sharing data, the FDA can address the limitations of FOIA and create data sharing policies that promote the health and safety of all Americans. The public interest in data disclosure is more urgent and compelling now than when the FDA first formulated its disclosure policies. With modern advances in data generation and analyses, there is an even greater potential for data sharing to enhance and accelerate medical knowledge. By proactively sharing data, the FDA can better fulfill its responsibilities to patients and public health.

Note
The work was funded in part by the Laura and John Arnold Foundation, through a research grant to Yale University to support the Collaboration for Research Integrity and Transparency. The sponsor had no role in the design or drafting of the article.

Acknowledgments
The authors would like to thank Gregg Gonsalves, Margaret McCarthy, Joseph Ross, Cortelyou Kenney, Joshua Wallach, and Alex Egilman for their comments.

References
2. S. Gottlieb, *Answers to Written Questionnaire from Members of the Senate Committee on Health, Education, Labor, and Pensions* (April 26, 2017): at 58. (“I am a strong proponent of data transparency for patients, physicians, and manufacturers. I have long advocated that the FDA release more information related to its review process ... If confirmed, I will be committed to ... the issue of data transparency and new ways that FDA could potentially make important information more readily available to the public.”).
4. *Id.*, at 7 (emphasis added).
5. *Id.*
8. *Id.*, at 110.
10. *Id.*, at 100, 102-103, 105.
11. Id., at 68-69.
12. Id., at 7.
18. Chrysler Corp., 441 U.S. at 294 (quoting H. R.Rep. No. 1497, 89th Cong., 2d Sess., 2, 5 (1966)); see also Dept. of the Air Force v. Rose, 425 U.S. at 361 (The FOIA exemptions “do not obscure the basic policy that disclosure, not secrecy, is the dominant objective of the Act,” and thus, they must be “narrowly construed”).
19. CNA Fin. Corp. v. Donovan, 830 F.2d 1132, 1134-1135 (D.C. Cir. 1987) (“the agency’s decision to release the data normally will be grounded either in its view that none of the FOIA exemptions applies, or in its belief that release is justified in the exercise of its discretion, even though the data fall within one or more of the statutory exemptions.”); see Jurewiez v. U.S. Dep’t of Agriculture, 741 F.3d 1326, 1332 (D.C. Cir. 2014) (finding that a substantial privacy interest “must be balanced against any public interest in disclosure … [to the extent that] disclosure of the information sought would … let citizens know what their government is up to.”) (quoting U.S. Dept. of Def. v. Fed. Labor Relations Auth., 510 U.S. 487, 497 (1994)). Where CCI is concerned, however, the balancing of public interests may be inappropriate because the category overlaps with other laws that flatly forbid agencies from making unauthorized disclosures of commercial data. CNA Fin. Corp. v. Donovan, 830 F.2d 1132, 1140 (D.C. Cir. 1987) (finding that 18 U.S.C. § 1905 “appears to cover practically any commercial or financial data collected by any federal employee” such that information that falls under Exemption 4 is barred from disclosure unless otherwise authorized). But, the U.S. Court of Appeals for the Seventh Circuit has suggested a different interpretation — that § 1905 was intended to protect a narrower category of information than Exemption 4, thereby preserving some agency discretion to disclose information that falls within Exemption 4. Gen. Elec. Co. v. U.S. Nuclear Regulatory Comm’n, 750 F.2d 1394, 1402 (7th Cir. 1984) (“Exemption 4 is broadly worded, and it is hard to believe that Congress wanted seekers after information to stub their toes on a rather obscure criminal statute almost certainly designed to protect that narrower category of trade secrets … whose disclosure could be devastating to the owners and not just harmful”).
22. IoM Report, supra note 1, at 208-213 (Appendix B) (referring to the de-identification methods provided in the Privacy Rule of the U.S. Health Insurance Portability and Accountability Act (HIPAA) as “a good launching point for examining best practices” for sharing analyzable clinical trial data).
23. U.S. Dept. of the Air Force v. Rose, 425 U.S. at 372 (finding that Exemption 6 requires a balancing of the individual’s right to privacy against the public’s right to disclose under FOIA); Consumers’ Checkbook Ctr. for the Study of Servs. v. U.S. Dept. of Health and Human Servs., 554 F.3d 1046, 1057 (D.C. Cir. 2009) (stating that FOIA’s “presumption favoring disclosure … is at its zenith under Exemption 6”) (quoting Nat’l Ass’n of Home Builders v. Norton, 309 F.3d 26, 37 (D.C. Cir. 2002)); see also Jurewiez v. U.S. Dept of Agriculture, 741 F.3d 1326, 1331-34 (D.C. Cir. 2014) (in reverse-FOIA case, referring to an agency’s decision that that any personal privacy concerns are minimum and outweighed by the public’s interest “in assessing whether the [agency] is fulfilling its statutory mandate” and ‘gaug[ing] the effectiveness of [agency] inspections by comparing data … with publicly available inspection reports’.
24. In 2013, the FDA proposed sharing de-identified analyzable safety and efficacy datasets, acknowledging that such data “have tremendous potential to … provide new opportunities for innovation in medical product development.” “Availability of Masked & De-Identified Non-Summary Safety & Efficacy; Request for Comments,” 78 Federal Register 33421, 33422 (June 3, 2013). More recently, in 2016, the U.S. Department of Health and Human Services has expressed a willingness to explore whether ClinicalTrials.gov can “provide[] the scaffolding on which individual participant data … (the next frontier in transparency) and other trial “meta-data” can be organized in the future,” and the agency “anticipate[s] that ClinicalTrials.gov can be used in the future to catalyze IPD sharing.” “Clinical Trials Registration and Results Submission Final Rule,” 81 Federal Register 64,981, 64,984, 64,991 (Sept. 21, 2016) (codified at 42 C.F.R. Pt. 11).
25. CNA Fin. Corp. v. Donovan, 830 F.2d 1132 (D.C. Cir. 1987) (finding that 18 U.S.C. § 1905 is “co-extensive” with FOIA’s Exemption 4 for CCI, but holding that the agency’s determination that the information at issue is not CCI to be reasonable); see also Jurewiez v. U.S. Dept of Agriculture, 741 F.3d 1326, 1331 (D.C. Cir. 2014) (Exemption 4 “requires a showing of both actual competition and a likelihood of substantial competitive injury … [and the court] will generally defer to the agency’s predictive judgments as to the repercussions of disclosure”) (internal quotations omitted).
26. IoM Report, supra note 1, at 259-60.
28. 21 U.S.C. § 355(r); FDAAA § 915.
30. The FDA does release downloadable analyzable datasets on a quarterly basis containing de-identified synonyms of individual adverse event reports that are collected in the FDA Adverse Event Reporting System (FAERS) database, available at https://www.fda.gov/drugs/guidanceregulatoryinformation/surveillance/adversedrugeffects/ucm082193.htm (last visited November 2, 2017).
31. U.S. Cong. House, Committee on Energy and Commerce, Subcommittee on Oversight and Investigations, Hearing on Publication and Disclosure Issues in Anti-Depressant Pediatric Clinical Trials September 9, 2004, 108th Cong. 2d sess. at 25, 37-35, 41, available at <https://www.gpo.gov/fdsys/pkg/CHRG-108hrrg69049/html/CHRG-108hrrg69049.htm> (last visited November 2, 2017) (Reps. Deutsch, Bass, and DeGette questioning Dr. Janet Woodcock, then acting Dep’t of Health and Human Services has expressed a willingness to explore whether ClinicalTrials.gov can “provide[] the scaffolding on which individual participant data … (the next frontier in transparency) and other trial “meta-data” can be organized in the future,” and the agency “anticipate[s] that ClinicalTrials.gov can be used in the future to catalyze IPD sharing.” “Clinical Trials Registration and Results Submission Final Rule,” 81 Federal Register 64,981, 64,984, 64,991 (Sept. 21, 2016) (codified at 42 C.F.R. Pt. 11).
32. 21 U.S.C. § 355(a)(2); FDAAA § 916.
33. House Committee on Energy and Commerce Committee Hearing, supra note 31; see also Complaint, The People of the


35. 5 U.S.C. § 552(a)(3) (all records requested under FOIA that are not exempt must be made “promptly available to any person”); Dept of Justice v. Reporters Comm. for Freedom of Press, 489 U.S. 749, 771 (1989) (FOIA is “clearly intended ... to give any member of the public as much right to disclosure as one with a special interest”) (quoting NLRB v. Sears, Roe- back & Co., 421 U.S. 132, 149 (1975)).

36. See Swan v. SEC, 96 F.3d 498, 500 (D.C. Cir. 1996) (“Once records are released, nothing in FOIA prevents the requester from disclosing the information to anyone else. The statute contains no provisions requiring confidentiality agreements or similar conditions.”); Maricopa Audubon Soc. v. U.S. For- est Serv., 108 F.3d 1082, 1088–89 (9th Cir. 1997) (holding “that FOIA does not permit selective disclosure of information only to certain parties, and that once the information is disclosed . . . it must also be made available to all members of the public who request it.”)

37. IoM Report, supra note 1, at 13 (“data use agreements are a promising vehicle for reducing ... risks and related disincentives for sharing clinical trial data).