

PERSPECTIVE OPEN

The Open Translational Science in Schizophrenia (OPTICS) project: an open-science project bringing together Janssen clinical trial and NIMH data

Marsha A. Wilcox¹, Adam J. Savitz¹, Anjené M. Addington², Gary S. Gray³, Eva C. Guinan⁴, John W. Jackson^{5,6}, Thomas Lehner⁷, Sharon-Lise Normand⁸, Hardeep Ranu³, Geetha Senthil⁷, Jake Spertus⁸, Linda Valeri^{9,10} and Joseph S. Ross^{11,12,13}

Clinical trial data are the gold standard for evaluating pharmaceutical safety and efficacy. There is an ethical and scientific imperative for transparency and data sharing to confirm published results and generate new knowledge. The Open Translational Science in Schizophrenia (OPTICS) Project was an open-science initiative aggregating Janssen clinical trial and NIH/NIMH data from real-world studies and trials in schizophrenia. The project aims were to show the value of using shared data to examine: therapeutic safety and efficacy; disease etiologies and course; and methods development. The success of project investigators was due to collaboration from project applications through analyses, with support from the Harvard Catalyst. Project work was independent of Janssen; all intellectual property was dedicated to the public. Efforts such as this are necessary to gain deeper insights into the biology of disease, foster collaboration, and to achieve the goal of developing better treatments, reducing the overall public health burden of devastating brain diseases.

npj Schizophrenia (2018)4:14 ; doi:10.1038/s41537-018-0055-7

INTRODUCTION

Data sharing

Clinical trial data are the gold standard for evaluating pharmaceutical safety and efficacy. Until recently, these data have been sequestered within companies, sponsors of much of the clinical research conducted on pharmaceutical products. There is an ethical and scientific imperative for transparency and sharing of these data to confirm published results and generate new knowledge.^{1–4} Transparency has become the new standard in pharmaceutical and medical device science.^{5,6} Recent initiatives have highlighted the need for transparency and sharing, among these are The Institute of Medicine's report, "Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risks" and guidance from international organizations and industry.

The Open Translational Science in Schizophrenia (OPTICS) Project, an overview

Janssen and a group of leading research organizations launched the OPTICS Project in 2015. This industry-academic-government collaboration aggregated Janssen clinical trial and federally-funded data about schizophrenia. Researchers from diverse disciplines collaborated in an open-science environment addressing essential questions about schizophrenia, therapeutic safety

and efficacy, and methods development. All IP generated has been dedicated to the public and is free for all to use. In this effort, data about the disorder and clinical trials of therapies were made available to researchers in one place. Our goal was to advance scientific knowledge about the disease, foster collaboration, and create new models for conducting research. Figure 1 depicts this process.

The Yale Open Data Access (YODA) project

The YODA Project was founded in 2011 to promote data sharing among the scientific community and develop a platform that could be used as a means of responsible data sharing.⁷ Through partnerships with Johnson & Johnson and Medtronic, Inc., the YODA Project established policies to make shared data available to investigators, including de-identified participant-level trial data, comprehensive trial reports, and supportive documentation and meta-data. Data are currently available for more than 250 clinical trials of different psychiatric conditions, as well as non-psychiatric conditions, to facilitate research that may advance science or lead to improvements in health and health care delivery. Janssen data for the OPTICS Project were made available through the YODA Project.

¹Janssen Pharmaceutical Research and Development, Titusville, NJ, USA; ²Genomics Research Branch, National Institute of Mental Health in Bethesda, Bethesda, MD, USA; ³Harvard Catalyst/The Harvard Clinical and Translational Science Center, Harvard Medical School, Boston, MA, USA; ⁴Department of Radiation Oncology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA, USA; ⁵Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; ⁶Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA; ⁷Office of Genomics Research Coordination, National Institute of Mental Health, Bethesda, USA; ⁸Department of Health Care Policy (Biostatistics), Harvard Medical School, Boston, MA, USA; ⁹Department of Psychiatry, Harvard Medical School, Boston, MA, USA; ¹⁰Laboratory for Psychiatric Biostatistics, McLean Hospital, Belmont, MA, USA; ¹¹Section of General Internal Medicine and the National Clinician Scholars Program, Department of Medicine, Yale University School of Medicine, New Haven, CT, USA; ¹²Department of Health Policy and Management, Yale University School of Public Health, New Haven, CT, USA and ¹³The Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, CT, USA

Correspondence: Marsha A. Wilcox (MWilcox@its.jnj.com)

Received: 5 October 2017 Revised: 3 May 2018 Accepted: 15 May 2018

Published online: 27 June 2018

National Institute of Mental Health (NIMH) perspective

NIMH has long been a strong proponent of endeavors, such as the OPTICS Project, that encourage researchers to leverage shared resources for broader scientific discovery. The [NIMH Repository and Genomics Resource](#), established under the 1989 NIMH Human Genetics Initiative, currently shares clinical data and samples from over 200,000 subjects spanning many psychiatric disorders and populations. Additionally, NRGR and [Database of Genotypes and Phenotypes](#) (dbGaP) together provide access to genome-wide data from over 87,000 subjects from psychiatric genetics studies. NIMH further provides access to a variety of clinical and data through the recently established [NIMH Data Archives](#). NIMH is eager to foster collaborations with industry partners that will leverage clinical trial data (e.g., genomic screens) that is not broadly available to enhance the secondary analyses done using our publicly available datasets. Such collaborations could provide many opportunities for discovery and may expedite identification and prioritization of potential molecular targets for therapeutic development for psychiatric disorders.

Harvard Catalyst (HC) perspective

The [Harvard Catalyst](#)⁸ and OPTICS Project collaboration began with the “[ReSourcing Big Data](#)⁹” symposium on collaborative data reuse/sharing. One outcome of the symposium was a joint pilot grant program to support analysis of OPTICS data. Grant awardees were selected on a competitive basis from applicants throughout Harvard Medical School and affiliates and supported for 12 months. The HC-OPTICS pilot grant program was successful as measured by the insights produced, the quality of the research and analyses completed, the number of analytic tools developed, publications generated, and the potential for subsequent funding. These outcomes are described in detail elsewhere in this issue. Moreover, several observations from this pilot grant project may be relevant to future data reuse efforts (Table 1).

The HC-OPTICS pilot project demonstrated the potential for highly productive reuse of existing clinical trials data sets.

Lessons learned

This pilot was designed to inform future efforts of this kind. The most important lesson learned was the importance of a collaborative environment; on-going dialogue with other investigators and the data holders contributed to successfully completing analyses and producing manuscripts.

From a pragmatic perspective, it is important that one never underestimate the time required to gain access to data, which includes IRB approval and especially institutional signing of Data Use Agreements, as well as the time required for external investigators to gain facility with the shared data. To address the former, Data Use Agreements should be standardized and pursued as soon as possible. To address the latter, data sharing platforms should ensure that data descriptions, dictionaries, protocols, and statistical analysis plans are available early, preferably during the application period, and that the computing environment is up-to-date, including statistical software. Finally, we would recommend that external investigators reproduce sample characteristics and primary results prior to pursuing their independent research analyses to identify possible discrepancies or data misunderstandings as soon as possible. All parties should anticipate frequent communications (e.g., questions about the data, the statistical software, the sponsors analyses), which foster collaboration.

The following manuscripts were the result of work pursued within OPTICS Project:

“[Risk of weight gain for specific antipsychotic drugs: A Bayesian network meta-analysis of individual participant level clinical trial data](#)” by Jacob Spertus, Marcella Horvitz-Lennon, Haley Abing, and Sharon-Lise Normand.

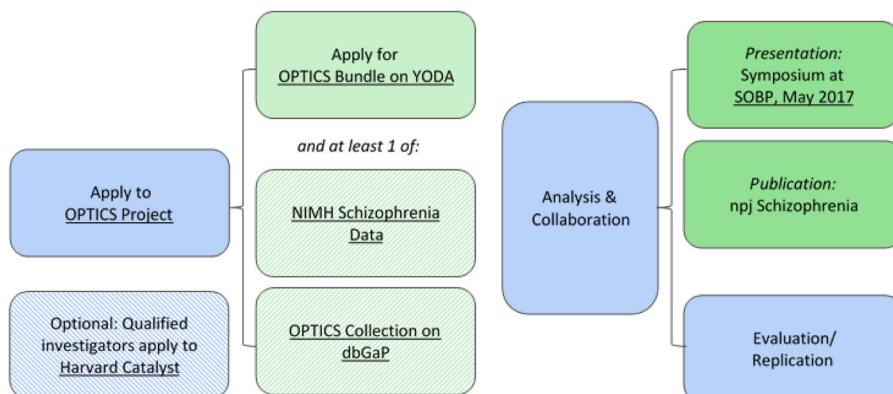


Fig. 1 The OPTICS Project Process

Table 1. Data reuse success factors

Success factor	Observation
Education	Provided good description of opportunity, including face-to-face Q&A Demonstrated access to and use of data set
Data set	Should be complete, consistent, well curated, and readily accessible Lack of familiarity with format or non-standard format adds challenge
Communication	Frequent, both between teams and data set curators and among teams Designed to foster cross-team collaboration, learning, and problem solving
Data Use Agreements	Completed agreement is a limiting factor. Timing and content varies among institutions

"The role of PANSS symptoms and adverse events in explaining the effects of Paliperidone on social functioning: a causal mediation analysis approach" by Xue Zou, Yiwen Zhu, John W. Jackson, Andrea Bellavia, Garrett Fitzmaurice, Franca Centorrino, and Linda Valeri.

CONCLUSIONS

The OPTICS Project was a collaboration in which both observational and interventional data about schizophrenia were made available to researchers in an open-science effort. The goal was to provide a forum for translational science in schizophrenia research. Aggregating and sharing these data enabled researchers to address questions about the disease, therapies, and analytic methods in ways not before possible. While the value of sharing clinical trial data on its own is significant, this project also provided a unique opportunity for collaborative partnerships.

The OPTICS Project is representative of a new paradigm in scientific research which we hope will lead to more collaboration and data sharing among industry and the broader research community. Such cooperative efforts are necessary to gain deeper insights into the biology of disease, to achieve the goal of developing better treatments and reducing the overall public health burden of devastating brain diseases.

Data availability

Janssen clinical trials available at the YODA Project: <http://yoda.yale.edu/multiple-ncts-optics-trial-bundle>
 NIMH: https://nimhgenetics.org/available_data/schizophrenia/
 dbGaP: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/collection.cgi?study_id=phs000887.v1.p1

ACKNOWLEDGEMENTS

The authors would like to acknowledge members of the OPTICS Scientific Advisory Board for their guidance in every aspect of this project, beginning with working through the legal issues through to the submission of the final manuscripts: Jesse A. Berlin, ScD, Senior Vice President, Epidemiology, Johnson & Johnson; Linda Brzustowicz, MD, Principal Investigator and Chair, Department of Genetics, Rutgers University; Jonathan Rabinowitz, PhD, Professor; Bar Ilan University, Israel; Michelle Williams, ScD, Dean, Harvard TH Chan School of Public Health (Drs. Ross, Savitz and Wilcox also served on the board). This project was supported by many individuals at Janssen including: Hal Woodrow for legal advice, Jeff Gardner, Peter Lins, and Stephen Bamford for data curation and management, Mohamad Khashab for advice about health care compliance, and Karla Childers and Sandra Morris for guidance from the Office of the Chief Medical Officer. The authors would also like to acknowledge the assistance of Jessica D. Ritchie, MPH and Emily B. Finn, MPH in background research and manuscript preparation. Both Ms. Ritchie and Mrs. Finn are employees at the Center for Outcomes Research and Evaluation at Yale-New Haven Hospital.

AUTHOR CONTRIBUTIONS

All authors contributed to writing and editing the manuscript; M.A.W. and A.J.S. contributed to the background, the overall project and served on the Scientific Advisory Board; J.S.R. contributed the YODA Perspective and served on the Scientific Advisory Board; A.M.A., T.L., and G.S. provided the NIMH perspective and in the direction of the project; T.L. was also on the Scientific Advisory Board; G.S.G. and E.C.G and H.R. provided the Harvard Catalyst Perspective and participated in conduct of the project; J.W.J., S-L.N., L.V. and J.S. contributed the OPTICS Project investigators' perspective and conducted research for the project.

ADDITIONAL INFORMATION

Competing interests: M.A.W. and A.J.S. are employed by Janssen Research & Development, LLC. In the past 36 months, J.S.R. received research support through Yale University from Johnson & Johnson to develop methods for clinical trial data sharing, from Medtronic, Inc. and the Food and Drug Administration (FDA) to study issues in post-market medical device surveillance, from the Food and Drug Administration (FDA) to establish the Yale-Mayo Clinic Center for Excellence in Regulatory Science and Innovation (CERSI), from the Centers of Medicare and Medicaid Services (CMS) to develop measures for public reporting of hospital and physician quality, from the Blue Cross-Blue Shield Association (BCBSA) to advance pre-market evidence generation for medical products, from the Agency for Healthcare Research and Quality, and from the Laura and John Arnold Foundation to support the Collaboration on Research Integrity and Transparency (CRIT) at Yale. E.C.G. and G.S.G., and H.R., were supported by Harvard Catalyst | The Harvard Clinical and Translational Science Center (National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health Award UL1 TR001102) and financial contributions from Harvard University and its affiliated academic healthcare centers. The content is solely the responsibility of the authors and does not necessarily represent the official views of Harvard Catalyst, Harvard University and its affiliated academic healthcare centers, or the National Institutes of Health. J.W.J., L.V., and S-L.N. and J.S. received support from Harvard Catalyst—The Harvard Clinical and Translational Science Center (National Center for Research Resources and the National Center for Advancing Translational Science Award UL1 TR001102). The authors assume full responsibility for the accuracy and completeness of the ideas presented.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Murugiah, K., Ritchie, J. D., Desai, N. R., Ross, J. S. & Krumholz, H. M. Availability of clinical trial data from industry-sponsored cardiovascular trials. *J. Am. Heart Assoc.* **5**, e003307 (2016).
- Ross, J. S. & Krumholz, H. M. Ushering in a new era of open science through data sharing: The wall must come down. *JAMA* **309**, 1355–1356 (2013).
- Krumholz, H. M. & Peterson, E. D. Open access to clinical trials data. *JAMA* **312**, 1002–1003 (2014).
- Ross, J. S., Lehman, R. & Gross, C. P. The importance of clinical trial data sharing: Toward more open science. *Circ. Cardiovasc Qual. Outcomes* **5**, 238–240 (2012).
- Bauher, H., Golub, R. M. & Fontanarosa, P. B. Data sharing: an ethical and scientific imperative. *JAMA* **315**, 1237–1239 (2016).
- Strom, B., Buyse, M., Hughes, J. & Knoppers, B. M. Data sharing – Is the juice worth the squeeze? *NEM* **375**, 171608–171609 (2016).
- The Yale Open Data Access Project (YODA) (<http://yoda.yale.edu/>) (2018).
- Harvard Catalyst (<https://catalyst.harvard.edu/>) (2018).
- Harvard Catalyst ReSourcing Big Data (<http://catalyst.harvard.edu/programs/reactor/resourcingbigdata.html>) (2018).



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018