ORIGINAL CONTRIBUTION

Association of *BRCA1* and *BRCA2* Mutations With Survival, Chemotherapy Sensitivity, and Gene Mutator Phenotype in Patients With Ovarian Cancer

Da Yang, PhD
Sofia Khan, PhD
Yan Sun, MD, PhD
Kenneth Hess, PhD
Ilya Shmulevich, PhD
Anil K. Sood, MD
Wei Zhang, PhD

Context Attempts to determine the clinical significance of *BRCA1/2* mutations in ovarian cancer have produced conflicting results.

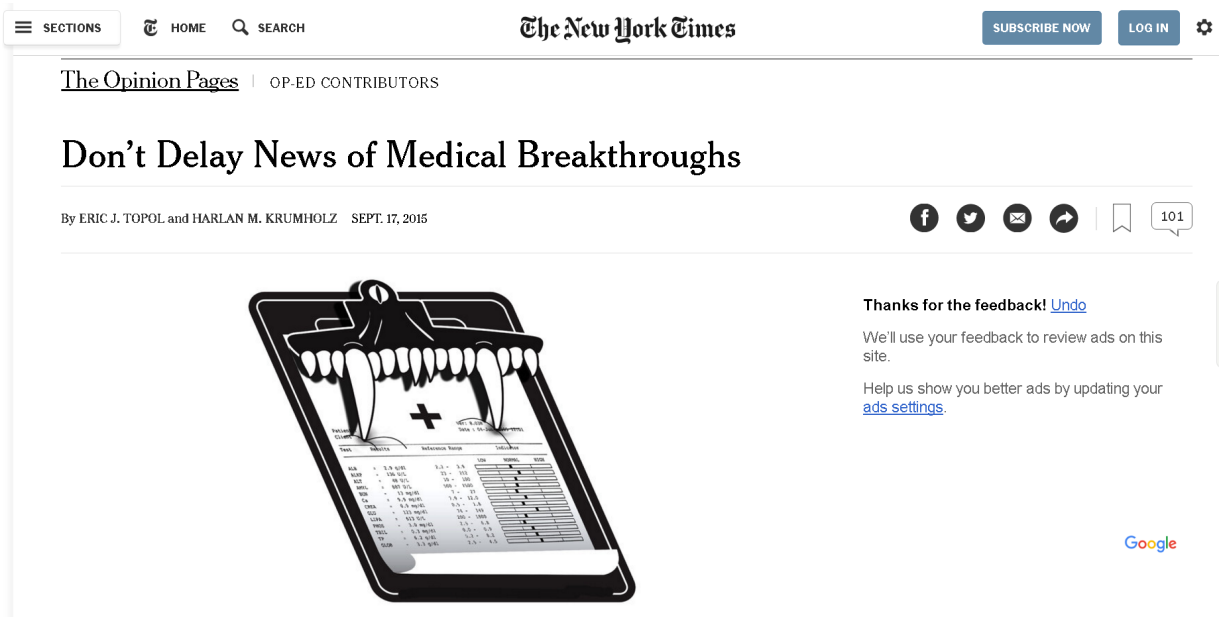
Objective To determine the relationships between *BRCA1/2* deficiency (ie, mutation and promoter hypermethylation) and overall survival (OS), progression-free survival (PFS), chemotherapy response, and whole-exome mutation rate in ovarian cancer.

Design, Setting, and Patients Observational study of multidimensional genomics and clinical data on 316 high-grade serous ovarian cancer cases that were made public between 2009 and 2010 via The Cancer Genome Atlas project.

Main Outcome Measures OS and PFS rates (primary outcomes) and chemotherapy response (secondary outcome).

INCREASED SURVEILLANCE OF *BRCA1/2* germ line mutation carriers is a gen-

Yang, D., et al. (2011). "Association of BRCA1 and BRCA2 mutations with survival, chemotherapy sensitivity, and gene mutator phenotype in patients with ovarian cancer." JAMA 306(14): 1557-1565.



Medicine needs to change its approach to releasing new, important information. Throughout science we are seeing more rapid modes of communication. The traditional approach was not to publish until everything was finalized and ready to be chiseled in stone. But these sorts of delays are unnecessary with the Internet.

http://www.nytimes.com/2015/09/18/opinion/dont-delay-news-of-medical-breakthroughs.html?_r=0

EDITORIAL

A SPRINT to the Finish

Jeffrey M. Drazen, M.D., Stephen Morrissey, Ph.D., Edward W. Campion, M.D., and John A. Jarcho, M.D.

We were therefore surprised by the call from Topol and Krumholz for immediately "placing the data on the NIH website."[2] We believe that it is critical to give the investigators, on behalf of the study participants, who invested years of their lives in the study, the opportunity to see what led the sponsor to stop the trial and then the opportunity to distill a clinical message from it. There are cogent reasons to follow this approach

Drazen, J. M., et al. "A SPRINT to the Finish." New England Journal of Medicine 2015;373:2174-2175

"Research Parasite"

The NEW ENGLAND JOURNAL of MEDICINE

EDITORIALS

Data Sharing

Dan L. Longo, M.D., and Jeffrey M. Drazen, M.D.

The aerial view of the concept of data sharing is beautiful. What could be better than having high-quality information carefully reexamined for the possibility that new nuggets of useful data are lying there, previously unseen? The potential for leveraging existing results for even more benefit pays appropriate increased tribute to the patients who put themselves at risk to generate the data. The moral imperative to honor their collective sacrifice is the trump card that takes this trick.

However, many of us who have actually conducted clinical research, managed clinical studies and data collection and analysis, and curated data sets have concerns about the details. The first concern is that someone not involved in the generation and collection of the data may not understand the choices made in defining the parameters. Special problems arise if data are to be combined from independent studies and considered comparable. How heterogeneous were the study populations? Were the eligibility criteria the same? Can it be assumed that the differences in study populations, data collection and analysis, and treatments, both protocol-specified and unspecified, can be ignored?

A second concern held by some is that a new class of research person will emerge — people who had nothing to do with the design and execution of the study but use another group's data for their own ends, possibly stealing from the research productivity planned by the data gatherers, or even use the data to try to disprove what the original investigators had posited. There is concern among some front-line researchers that the system will be taken over by what some researchers have characterized as "research parasites."

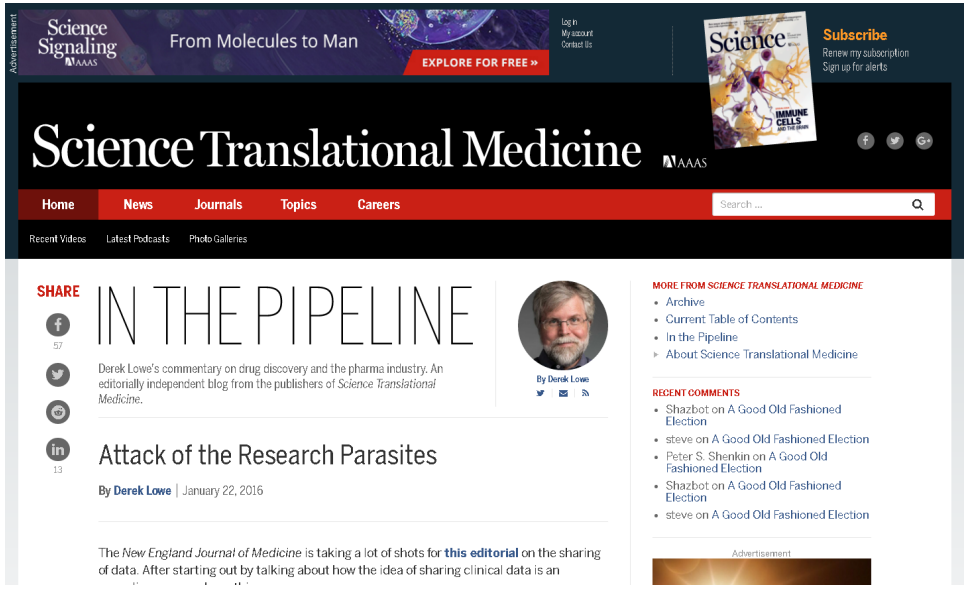
This issue of the *Journal* offers a product of data sharing that is exactly the opposite. The new investigators arrived on the scene with their own ideas and worked symbiotically, rather than parasitically, with the investigators holding the data, moving the field forward in a way that neither group could have done on its own. In this case, Dalerba and colleagues[1] had a hypothesis that colon cancers arising from more primitive colon epithelial precursors might be more aggressive tumors at greater risk of relapse and might be more likely to benefit from adjuvant treatment. They found a gene whose expression appeared to correlate with the expression of genes that characterize more mature colon cancers on gene-expression arrays and whose product was reliably measurable in resected colon cancer specimens by immunohistochemistry. To assess the clinical value of this potential biomarker, they needed a sufficiently large group of patients whose archived tissues could be used to assess biomarker expression and who had been treated in relatively homogeneous way.

They proposed a collaboration with the National Surgical Adjuvant Breast and Bowel Project (NSABP) cooperative group, a research consortium funded by the National Cancer Institute that has conducted seminal research in the treatment of breast and bowel cancer for the past 50 years. The NSABP provided access to tissue and to clinical trial results on an individual patient basis. This symbiotic collaboration found that a small proportion (4%) of colon cancers did not express the biomarker and that the survival of patients with those tumors was poorer than that of patients whose tumors expressed the biomarker. Furthermore, when the effect of adjuvant chemotherapy was assessed, nearly all

Longo, D. L. and J. M. Drazen (2016). "Data Sharing." New England Journal of Medicine 374(3): 276-277.

A second concern held by some is that a new class of research person will emerge — people who had nothing to do with the design and execution of the study but use another group's data for their own ends, possibly stealing from the research productivity planned by the data gatherers, or even use the data to try to disprove what the original investigators had posited. There is concern among some front-line researchers that the system will be taken over by what some researchers have characterized as "research parasites."

Longo, D. L. and J. M. Drazen (2016). "Data Sharing." New England Journal of Medicine 374(3): 276-277.



It's that second paragraph that really sets people off, and it is unfortunately worded. Science actually advances on this sort of thing – calling people who use or build on previous data sets "research parasites" is actually fairly silly. ..... We stand on each other's shoulders in this business; that's how science works. It's not the scientists that worry me here.

http://blogs.sciencemag.org/pipeline/archives/2016/01/22/attack-of-the-research-parasites

Forbes LOGIN

YOUR READING LIST

Data Scientists = Research Parasites?

+20 comments in the last 24 hours

Brace Yourself: How You Will Pay To Rescue ObamaCare

Active on Facebook

Is Roman Reigns Heading Back To The Main Event Following WWE SummerSlam Feud With Rusev?

+1 comments in the last hour

Western Sanctions Slam Into Russia, With No Relief In Sight

Morgan Stanley Voice: Transition Time for Boomer Businesses

+3 comments in the last 24 hours

Immigration Minister Should Restore The Canadian Experience Class

JAN 21, 2016 @ 10:54 PM 11,132 VIEWS

Data Scientists = Research Parasites?

David Shaywitz, CONTRIBUTOR

I write about entrepreneurial innovation in medicine. FULL BIO

Opinions expressed by Forbes Contributors are their own.

This Thursday, Dec. 29, 2011 photo shows the entrance to the editorial offices of the New England Journal of Medicine in Boston. (AP Photo/Michael Dwyer)

http://www.forbes.com/sites/davidshaywitz/2016/01/21/data-scientists-research-parasites/#3c54d7563d1c

My Twitter TWTR -0.11% feed exploded in response to this editorial:

Michael Eisen, geneticist at UC-Berkeley: "One of the most shockingly anti-science things ever written."

Sek Kathiresan, cardiologist/geneticist at MGH/Broad: "Shocking to see disparaging term 'research parasite' to describe use of data often created w/ public funds."

Michael Hoffman, computational genomicist at the Princess Margaret Cancer Centre, Toronto: "Fear that others may use data 'to try to disprove what the original investigators had posited' is a dangerous misunderstanding of science."

http://www.forbes.com/sites/davidshaywitz/2016/01/21/data-scientists-research-parasites/#3c54d7563d1c

EDITORIAL

#IAmAResearchParasite

In the midst of steady progress in policies for data sharing, a recent editorial expressed a contrarian view.* The authors described the concern of some scientists about the rise of an underclass of "research parasites" who exploit data sets that are collected and curated by others. Even worse, these parasites might use such data to try to disprove the conclusions posited in the data's original source studies. The editorial raised the points of how anyone not involved in the original study could use the data without misrepresenting it, and the danger of perhaps arriving at erroneous conclusions. The editorial advised instead that data sharing be implemented by involving the authors of the original study as coauthors in follow-up research. The research community immediately took to Twitter under the hashtag #IAmAResearchParasite to voice opposition to the editorial.

Much of what we know about the large-scale features of this planet is apparent thanks to widespread data-sharing practices and the early establishment of data banks in the geosciences. Aspects such as determining the shape of the ocean floor, ocean chemistry, the internal structure of Earth's deep interior, the physics and chemistry of the atmosphere, and many other topics could not have been ascertained from a single investigator's field program. One meta-analysis I published on the South Pacific benefited from observations of my own and those of others, including the 18th-century British explorer Captain James Cook. Involving Cook as a coauthor on my paper was clearly not an option, any more than it would have been feasible or desirable to include the dozens of others, living or dead, who had contributed to the data repository. Many fields, including the biomedical sciences, are now benefiting from meta-analyses of data to better understand the big picture.

Effective data sharing is not trivial or inexpensive to implement, and it takes more than community acceptance of the practice. Agencies supporting research in oceanography have long funded data and sample repositories and have encouraged data and sample deposition by making new awards contingent on compliance. Repositories are *instrumental* in setting formats for data, so much so that standard programs and apps accept and output data in the standard format. The marine community supports data professionals who are responsible for the quality control of data collected on ships and from other major observing programs.

"There are costs... for re-collecting data for new uses."

Often overlooked is the importance of community-established metadata, so that those not involved in the original research will know what the data mean. As an example, in an oceanographic temperature data base, the community had to agree on what T(0) meant. Was it temperature at atmospheric pressure? Temperature at the sea surface?

Communities must discourage low-quality data collection. A well-attended poster presentation at one prominent scientific meeting some years ago compared the crossover errors (misfits) of non–time-dependent measurements (such as depth soundings) from ships' tracks where they intersected in the world's oceans. Any discrepancy at a crossing could be attributed to poor data quality control on either ship, but with thousands of crossings, institutions with systematically more misfits than others stand out. The results did not escape the attention of the funding agencies that support ship time.

There are costs to implementing data reuse, but there are also costs for irreproducible research and for re-collecting data for new uses. And no amount of funding can reconstruct lost ephemeral or time-dependent phenomena for which the data were not well curated. No more excuses: Let's step up to data sharing.

– **Marcia McNutt**

Marcia McNutt is Editor-in-Chief, Science *Journals*

PHOTOS: (INSET) ISHMEDIA/ALAMY; ITOP) STACEY PENTLAND PHOTOGRAPHY

*D. L. Longo, J. M. Drazen, *N. Engl. J. Med.* **374**, 276 (2016).

10.1126/science.aaf4701

Published by AAAS

McNutt, M. (2016). "#IAmAResearchParasite." Science 351(6277): 1005-1005.

In the midst of steady progress in policies for data sharing, a recent editorial expressed a contrarian view.* The authors described the concern of some scientists about the rise of an underclass of "research parasites" who exploit data sets that are collected and curated by others. Even worse, these parasites might use such data to try to disprove the conclusions posited in the data's original source studies. The editorial raised the points of how anyone not involved in the original study could use the data without misrepresenting

McNutt, M. (2016). "#IAmAResearchParasite." Science 351(6277): 1005-1005.

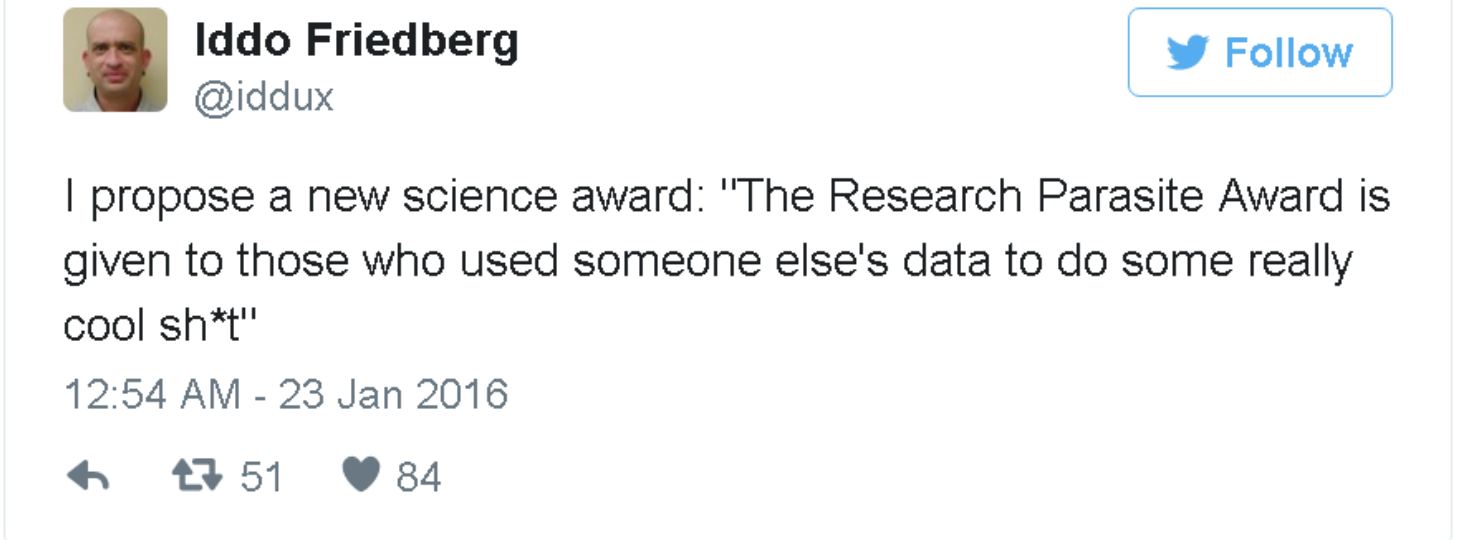
gram. One meta-analysis I published on the South Pacific benefited from observations of my own and those of others, including the 18th-century British explorer Captain James Cook. Involving Cook as a coauthor on my paper was clearly not an option, any more than it would have been feasible or desirable to include the dozens of others, living or dead, who had contributed to the data repository. Many fields, including the biomedical sciences, are now benefiting from meta-analyses of data to better understand the big picture.

Effective data sharing is not trivial or inexpensive but with thousands of crossings, institutions with systematically more misfits than others stand out. The results did not escape the attention of the funding agencies that support ship time.

There are costs to implementing data reuse, but there are also costs for irreproducible research and for recollecting data for new uses. And no amount of funding can reconstruct lost ephemeral or time-dependent phenomena for which the data were not well curated. No more excuses: Let's step up to data sharing.

– **Marcia McNutt**

McNutt, M. (2016). "#IAmAResearchParasite." Science 351(6277): 1005-1005.



http://www.greenelab.com/parasite-award/

PSB Awards for rigorous secondary data analysis (The "Parasites")

APPLICATION PROCESS

For either award, submit an application by **October 14, 2016 at 5PM HST** (Hawaii Standard Time) to **parasite.award@gmail.com**. An application requires:

- A nomination letter describing how each selected paper meets the criteria for the award. Self nominations are encouraged, and all nominees must be aware that they have been nominated.
- Junior Parasite (aka the sporozoite): a PDF of one paper published after peer review on which the application will be judged.
- Sustained Parasitism (aka the merozoite): PDFs of three papers published after peer review on which the application will be judged.

http://www.greenelab.com/parasite-award/

The NEW ENGLAND JOURNAL of MEDICINE

EDITORIAL

Data Sharing and the *Journal*

Jeffrey M. Drazen, M.D.

We want to clarify, given recent concern about our policy, that the *Journal* is committed to data sharing in the setting of clinical trials. As stated in the Institute of Medicine report from the committee[1] on which I served and the recent editorial by the International Committee of Medical Journal Editors (ICMJE),[2] we believe there is a moral obligation to the people who volunteer to participate in these trials to ensure that their data are widely and responsibly used. *Journal* policy will therefore follow that outlined in the ICMJE editorial and the IOM report: when appropriate systems are in place, we will require a commitment from authors to make available the data that underlie the reported results of their work within 6 months after we publish them.

In the process of formulating our policy, we spoke to clinical trialists around the world. Many were concerned that data sharing would require them to commit scarce resources with little direct benefit. Some of them spoke pejoratively in describing data scientists who analyze the data of others.[3] To make data sharing successful, it is important to acknowledge and air those concerns.[3] In our view, however, researchers who analyze data collected by others can substantially improve human health.

We need your help to move medicine forward and improve patient care. Our enemy is disease. By working in collaboration, as we have suggested,[3] biologists, data scientists, and clinical trialists can advance the art, and everyone will gain. Clinical trial data are some of the highest quality data in medicine. They should be used responsibly and extensively to help alleviate suffering. We believe that we will all benefit most if this is done collaboratively, but the *Journal*'s data sharing policy will apply in all settings.

Disclosure forms provided by the author are available with the full text of this article at NEJM.org.

This article was published on January 25, 2016, at NEJM.org.

1. Committee on Strategies for Responsible Sharing of Clinical Trial Data. Sharing clinical trial data: maximizing benefits minimizing risk. Washington, DC: National Academies Press, 2015.
2. Taichman DB, Backus J, Baethge C, et al. Sharing clinical trial data — a proposal from the International Committee of Medical Journal Editors. N Engl J Med 2016;374:384-6.
3. Longo DL, Drazen JM. Data sharing. N Engl J Med 2016;374:276-7.

DOI: 10.1056/NEJMe1601087

Drazen, J. M. (2016). "Data Sharing and the Journal." New England Journal of Medicine 374(19): e24.

Participants in clinical research volunteer in order to support the development of scientific knowledge and help future patients. Inherent in their commitment is the belief that research will lead to new insights that will be disseminated. As clinical researchers, we fully support the concept of data sharing as fundamental to achieving this goal.

Cardiovascular, T. A. R. O. C. f. C. E. o. S. S. (2016). "Sharing Data from Cardiovascular Clinical Trials — A Proposal." New England Journal of Medicine 375(5): 407-409.

The NEW ENGLAND JOURNAL of MEDICINE

Perspective

AUGUST 4, 2016

Strengthening Research through Data Sharing

Elizabeth Warren, J.D.

Data sharing has incredible potential to strengthen academic research, the practice of medicine, and the integrity of the clinical trial system. Some benefits are obvious: when

to the data underlying trial results can provide an avenue for independent confirmation of results and further analyses of the data set, raising the bar for academic

Warren, E. (2016). "Strengthening Research through Data Sharing." New England Journal of Medicine 375(5): 401-403.

Back to presenation