Innovations in Linking Survey and Administrative Data

Project Data Sphere®: Progress & Promise

13-September-2017
Organized to Improve Care for Cancer Patients

- Founded in 2001
- CEOs of diverse companies
  - 38 companies overall
  - 18 biopharmaceutical companies
- Mission
  - Reduce cancer risk
  - Enable early diagnosis
  - Facilitate access to treatment
  - Hasten discovery of new treatments

- Founded in 2004
- Life sciences company executives
- Mission:
  - “Be bold and venturesome”
  - Accomplish together what no single company might consider alone
Project Data Sphere:  
An Open-Access Data-Sharing Cancer Research Platform

- Launched **April 8, 2014**
- Broadly **share, integrate, and analyze** cancer clinical trial data
- Powerful embedded **analytic tools free** to all users. Datasets can also be **downloaded**.
- **Industry and NCI-NCTN phase III trials**
  - At launch, **comparator arm data**
  - **Experimental arm data** available beginning April 2017
- **De-identified patient data, data dictionary, protocol, & CRFs**

What if we could share, integrate, and analyze our collective historical cancer research data in a single location?

<table>
<thead>
<tr>
<th>Patients</th>
<th>Datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>98,551</td>
<td>115</td>
</tr>
</tbody>
</table>
Key Metrics: Utilization & Proof of Principle

● Industry and NCI’s NCTN Commitment
  ✓ ~98,000 patient lives of data from 17 data providers

● Densely used by researchers
  ✓ >1,600 researchers have accessed datasets ~8,200 times

● Prostate Cancer DREAM Challenge (July 2015)
  ✓ >500 participants; 58 teams; 21 countries
  ✓ More participants than first 32 DREAM Challenges

● Proof of Principle: Prostate Cancer
  ✓ Pace-setting publications in top-tier peer-reviewed journals, e.g., Journal of Clinical Oncology, Lancet Oncology, Nature, PLOS One, The Oncologist
Dear Dr. Murphy:

We are issuing this Letter of Support for Project Data Sphere, LLC, to encourage industry sponsors and academic centers to share their de-identified patient-level data from adult and pediatric cancer clinical trials and other research studies with Project Data Sphere.

Signed,

Richard Pazdur, M.D.
Director, Oncology Center of Excellence
Food and Drug Administration
# Proof of Principle: Publications in Top-Tier Journals

<table>
<thead>
<tr>
<th>Publication</th>
<th>Author</th>
<th>Publication Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication Date: August 2017, <em>Journal of Clinical Oncology</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advantages of a Truly Open-Access Data-Sharing Model</td>
<td>Monica Bertagnolli, et al.</td>
<td>March 2017</td>
</tr>
<tr>
<td>Publication Date: March 2017, <em>New England Journal of Medicine</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment of a Prognostic Model, PSA Metrics and Toxicities in Metastatic Castrate Resistant Prostate Cancer using Data from Project Data Sphere.</td>
<td>Anthony Joshua, et al.</td>
<td>February 2017</td>
</tr>
<tr>
<td>Publication Date: February 2017, <em>PLOS One</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publication Date: December 2016, <em>Lancet Oncology</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publication Date: November 2016, <em>Lancet Oncology</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Publication</th>
<th>Author</th>
<th>Publication Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>A Patient-Level Data Meta-Analysis of Standard-of-Care Treatments from Eight Prostate Cancer Clinical Trials.</td>
<td>N. Geifman A. Butte</td>
</tr>
<tr>
<td>10</td>
<td>Comparative Effectiveness of Mitoxantrone plus Prednisone versus Prednisone alone in Metastatic Castrate-resistant Prostate Cancer after Docetaxel Failure.</td>
<td>Angela Green, et al.</td>
</tr>
</tbody>
</table>
Models were developed to predict docetaxel discontinuation.

Prediction models could reduce patient enrollment in clinical trials without the loss of statistical power.

Routinely collected clinical features can be used to identify patients who are likely to discontinue treatment.

A DREAM Challenge to Build Prediction Models for Short-Term Discontinuation of Docetaxel in Metastatic Castration-Resistant Prostate Cancer

INTRODUCTION

Despite decades of research and advances in treatment, the long-term prognosis of metastatic castration-resistant prostate cancer (mCRPC) remains poor. Docetaxel was the first cytotoxic drug to improve survival and quality of life in patients with mCRPC and has remained a standard first-line chemotherapy for the treatment of mCRPC. Although several clinical trials have confirmed the population-level survival and palliative benefits of docetaxel, a significant fraction of patients do not respond to docetaxel, and within approximately 8 months, nearly all patients become resistant to treatment or stop therapy. Of patients who initially experience a response to docetaxel, 70% to 80% will eventually discontinue as a result of toxicity, personal reasons, or poor efficacy.

http://ascopubs.org/doi/full/10.1200/CCI.17.00018
Sounding Board:

Data-sharing of cancer trial data is moving to center stage

- “Truly Open-Access Data-Sharing Model” has advantages over gatekeeper models

- Academia, Industry & Regulatory Authorities encourage and support this data-sharing strategy

Advantages of a Truly Open-Access Data-Sharing Model

Monica M. Bertagnolli, M.D., Oliver Sartor, M.D., Bruce A. Chabner, M.D.,
Mace L. Rothenberg, M.D., Sean Khuzin, M.D., M.P.H., Charles Hugh-Jones, M.D.,
David M. Reese, M.D., and Martin J. Murphy, D.Med.Sc., Ph.D.

Thus, the PDS Prostate Cancer DREAM Challenge confirmed that an open-access model empowers global communities of scientists from diverse backgrounds and promotes crowd-sourced solutions to important clinical problems. This level of engagement is not possible with gatekeeper models.


New Insights Are Impossible Without Data Partners
Data Provided by Industry, Academia, Government and Researchers

*Researcher-curated data
Numbers indicate total number of datasets in-platform

Patient Lives in Thousands

<table>
<thead>
<tr>
<th>Organization</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCTN/NCORP</td>
<td>30</td>
</tr>
<tr>
<td>Sanofi Alliance</td>
<td>19</td>
</tr>
<tr>
<td>AZ</td>
<td>9</td>
</tr>
<tr>
<td>Curated(*)</td>
<td>11</td>
</tr>
<tr>
<td>Amgen Janssen Pfizer</td>
<td>2 6</td>
</tr>
<tr>
<td>Millennium</td>
<td>7 7</td>
</tr>
<tr>
<td>Lilly</td>
<td>5 5</td>
</tr>
<tr>
<td>Celgene EMD Serono</td>
<td>4 2</td>
</tr>
<tr>
<td>ECOG-ACRIN</td>
<td>1 1</td>
</tr>
</tbody>
</table>
| MSKCC Bayer Synta Clovis | 1 1 1
New Insights Are Impossible Without Data Partners
Available Data Will Continue to Expand

[Pie chart showing patient lives by tumor type]

- Breast, 26698 (24 datasets)
- Prostate, 21091 (26 datasets)
- Lung, 7675 (12 datasets)
- Leukemia, 9566 (5 datasets)
- Supportive Care, 2347 (6 datasets)
- Colorectal, 10538 (13 datasets)

[Bar chart showing numbers of datasets by tumor type]

- Liver, 544
- Renal Cell, 385
- Ovarian, 349
- Lymphoma, 340
- Melanoma, 326
- Myelofibrosis, 158
- Brain, 579
- Gastric, 667
- Multiple Myeloma, 674
- Pancreas, 1051
- Bone Mets, 702
- Head & Neck, 1110
New Insights Are Impossible Without Users
Open-Access Data Reaches Around the World

Users by Geographic Location

- USA 60%
- Europe (Other) 16%
- UK 6%
- China* 5%
- Canada 4%
- India 3%
- Other** 6%

*incl. Taiwan
**Other Americas, Asia, and Africa
Improving data usability through cataloging and curation

- Project Data Sphere currently accepts de-identified clinical trial data in all shapes and sizes
- Improved cataloging will be beneficial to researchers
  - Common, searchable, metadata (treatment names, tumor types, sponsors, etc.)
- Curation will add data standardization / normalization, but will be challenging
  - Standardize all data: Costly, and many data fields will never be used.
  - Standardize select data: Considerably less expensive, but will invariably omit valuable data fields
The *Project Data Sphere* Difference

- The *Project Data Sphere* cancer research platform aggregates cancer clinical trial data together in a single location.
- Clinical trial data is frequently regarded as ‘pristine’.
- The vast majority of the data is downloadable for analysis and integration with external data sources.
- Open-access of clinical trial data, as described in the NEJM article, is unique to the *Project Data Sphere* platform, and “truly advantageous”.
- Peer-accepted scientific papers are already in print, driven just from comparator data. Other data-driven investigations are ongoing.
QUESTIONS?

Contact information:
Dave Handelsman
Dave.Handelsman@ProjectDataSphere.org
919-531-2694

www.ProjectDataSphere.org
Register and become a user today