Precompetitive Data Sharing
and Collaboration to Develop
PD-1/PD-L1 Combinations

National Cancer Policy Forum

17 July 2018
Disclosures

I am a salaried employee and shareholder of Foundation Medicine
1. Background & Context: Biomarkers & Real World Data (RWD)
2. A case study in bridging data silos to unlock the potential of RWD
3. Potential impact of RWD on immunotherapy research and development
Biomarkers & Real World Data (RWD)
Biomarkers in Cancer Immunotherapy

Increasing importance
Growing complexity
Emerging challenges

1. Complexity
   No single biomarker today completely predicts who will or won’t respond to immunotherapy

2. Heterogeneity
   Each study looks at a single therapy or regimen, specific endpoints, select biomarkers and thresholds

3. Generalizability
   Clinical trials study selected populations – how do these results apply to the range of patients treated in the “real world”
Real World Data is emerging as a key enabler

Numerous efforts are actively underway to bridge data silos (genomic + clinical data) and institutional silos (across institutions and across practice settings) to begin to apply real world data to important problems in oncology.
Use of Real World Evidence
Beginning to understand opportunities to use real world endpoints derived from the EHR

Establishing a Framework to Evaluate Real-World Endpoints

July 10, 2018
Washington, DC
The Clinicogenomic Database (CGDB)

A collaboration between a genomics lab (Foundation Medicine) and an EMR/data company (Flatiron Health) to bridge real world data silos
2000+ samples sent each week to Foundation Medicine for analysis from 100s of centers across the country.

Sequencing is performed on a highly validated\(^1\) platform, identifying all classes of mutations in >300 genes.

Reports are sent to ordering doctors with genomic findings and therapies and trials to consider.

FoundationCORE™ contains >200,000 de-identified results, including genomic alterations, tumor mutational burden, microsatellite instability (MSI), and, for many patients, PD-L1 results.

\(^1\) Frampton et al., Nature Biotech 2013
Flatiron: EHR data integrated at the source

2M
Active Patients

2,500
Clinicians

265
Cancer Clinics

800
Unique Sites of Care
Linking genomic and clinical datasets

Flatiron

- Identified Patient-level Data
- Cleaned Demographic Inputs
- SFTP
- Token Engine
- Escrow Seed
- Token List
- Common Patient Token Subset
- De-identified Clinical Data Subset
- De-identified, Linked Clinico-Genomic Dataset
- Segregated Secure Environment
- Patient IDs hashed again

Third party linking agency

- Identified Patient-level Data
- Cleaned Demographic Inputs
- SFTP
- Token Engine
- Escrow Seed
- Token List
- Common Patient Token Subset
- De-identified Genomic Data Set
- De-identified, Linked Clinico-Genomic Dataset
- Segregated Secure Environment

FMI

- Identified Patient-level Data
- Cleaned Demographic Inputs
- SFTP
- Token Engine
- Escrow Seed
- Token List
- Common Patient Token Subset
- De-identified Genomic Data Set
- De-identified, Linked Clinico-Genomic Dataset
- Segregated Secure Environment

This entire process is repeated quarterly at FH-FMI for longitudinal refresh of the data (e.g., new chart abstraction).

Final linked clinico-genomic dataset cannot be re-identified or linked to other FH or FMI datasets (unique patient ID).

Red - Identified Data
Blue - De-identified Data
Yellow - De-identified Linked Data
Green - Third party linking agency’s technology
Development and Validation of a Real-World Clinico-Genomic Database

Gaurav Singal1, Peter Grant Miller2, Vinayeta Agarwala3, Jie He4, Anvia Gossiaux5, Gerald LU, Shannon Frank2, David Broucek6, Bynen Bowyer7, Thomas Carol8, Ezra Reve9, Keith Sadi-Rathkopf10, Ivan Ivanov10, Alex Parker11, Ameri Guirgis, Garrett Michael Frumpton11, Ann Jaskiw11, Dana Fauchois12, Nathan Coleman Nastlis12, Amy Picker Abraham7, Vincent A. Miller1,11

Foundation Medicine, Inc., Cambridge, MA, and Flatiron Health, New York, NY

BACKGROUND

• Genomic findings have diagnostic, prognostic, and predictive utility in clinical oncology.
• Population studies have been limited by reliance on large, expensive, or institutional cohort
 • Integrating electronic health records (EHR) and genomic data collected as part of routine clinical practice may overcome these hurdles.

RESULTS

CLINICAL CHARACTERISTICS

Table 1. Clinical characteristics of patients in the clinico-genomic database. The distribution of features such as median age, scoring history, and histology are consistent with prior studies.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age</td>
<td>66.00</td>
</tr>
<tr>
<td>Cancer</td>
<td>45.00</td>
</tr>
<tr>
<td>Survival</td>
<td>26.00</td>
</tr>
<tr>
<td>Sex</td>
<td>Male</td>
</tr>
<tr>
<td>Race</td>
<td>Black</td>
</tr>
<tr>
<td>Number</td>
<td>212</td>
</tr>
</tbody>
</table>

GENOMIC CHARACTERISTICS

Figure 5. Genomic characteristics of the NSCLC tumors in the clinico-genomics database are largely consistent with prior studies in large populations, including the TCGA. As expected, the presence of a driver mutation (EGFR, ALK, ROS1, MET, BRI1, RET, or ERBB2) was associated with younger age, female gender, and non-smoking status.

CONCLUSIONS

• We have built a de-identified, HIPAA-compliant, real-world clinico-genomic database by linking longitudinal clinical data and high-resolution genomic information. The dataset consists of 212 NSCLC cases, more than 30,000 total cases, and is both growing and updated as an ongoing basis.
• The clinico-genomic database shares similar genomic and clinical characteristics as NSCLC-  
  • Future studies will explore these patterns, which are expected to be common across datasets.

FUTURE DIRECTIONS

GROWTH OF THE CLINICO-GENOMIC DATABASE

Figure 7. Future trends and growth of the clinico-genomic database by disease (June 2017). The DB950 serves 30 tumor types and continues to grow and receive updates on a quarterly basis.

FUTURE APPLICATIONS

• General Needs
  Population for whom current treatments do not exist
  Therapies for whom appropriate population needs to be better defined
• Trial Design
  Characterizing the natural history of a biomarker-defined population for clinical trials
  Integration of NGS testing into trials to further biomarker and drug discovery
• Targeted Therapy
  Predicting genomic lesions for drug development
  Better understanding of mechanisms of resistance to current therapies
• Precision Medicine
  Integrating tumor mutation burden into predictive and prognostic algorithms
  Defining genomic subpopulations with differential sensitivity to checkpoint blockade
  Related approaches to combining targeted therapy with checkpoint blockade

TESTING AND THERAPEUTIC RESPONSE PREDICTION

Figure 6. Using the DB950 to understand and predict response to therapy. The presence of an EGFR mutation and presence of resistance (defined as both survival and clinical response to treatment) were used to predict response to EGFR-targeted therapy when appropriate was associated with increased OS and higher tumor burden associated with increased resistance on the TKI 1085 treatment. Updated.
Potential Impact of RWD

RWD may help address some of the existing challenges with cancer immunotherapy biomarker research.
Opportunity 1: Generalizability

Clinical trial results, especially for genomic populations, may benefit from RWD for generalizability.

We need more answers about immunotherapy for the elderly

By Ankur Parikh / June 22, 2018

Half of cancer patients are over the age of 65, while only 17% of patients in clinical trials are.

Studies of rare biomarkers, especially when histology agnostic, often include few (or no) patients in some histologies.

Challenge 1: Generalizability
Real world data can help us evaluate non-clinical trial populations

In the Flatiron-Foundation Medicine Clinicogenomic Database of ~4000 patients with non-small cell lung cancer:

- Did Not Receive Immunotherapy (n=2778, 68%)
- Received Immunotherapy (n=1286, 32%)
- >65yo (n=698, 54%)
- <65yo (n=588, 46%)
Use of Cancer Immunotherapies in the Real-World in the Setting of Microsatellite Instability

Agnieszka Agarwala MD PhD, Aneela Gioia PhD, Gaurav Singal MD, Claire O’Donnell, Gemali Li PhD, Ken Kao, David Bourque, Dima Feuchtbaum, Garrett Alpha Cobbo PhD, Allen Nunnelly, Sarah Kowalcz, Stacy Subbukolu PhD, Garrett Frampton PhD, Shijai AI MD PhD, Amy Abernethy MD PhD
1 Flatiron Health, New York, NY
2 Foundation Medicine, Cambridge, MA
3 Stanford University School of Medicine, Stanford, CA

Background
In May 2017, the NCCN issued its first dosing algorithm approval.1 Frequent microsatellite instability-high (MSI-H) or microsatellite-repair deficient (MMR-D) tumors are

Real-world data on the frequency of MSI/MMR across tumor types and the impact on treatment selection are scant, especially outside of colorectal cancer.1,2

Methods
This retrospective study utilized the EHR-derived Flatiron Foundation Medicine
linked Clinical-Genealogy Database. The study cohort included all patients diagnosed in the Flatiron Health Network (108 community and academic oncology practices, representing 606 distinct sites of care across the US) between January 2011 and May 2017, who were identified as MSI-H over the next generation tumor sequencing as part of routine clinical care. The clinical and genomic data were linked at the patient level via a third-party linking agency in a HIPAA-compliant process. The resulting database was used to identify

- Clinical data were curated from electronic health records, through normalization and standardization of structured and unstructured data elements (e.g., demographics, medication administration).
- Tumor type was determined by pathologist review of specimens submitted to Foundation Medicine.
- Genomic data were obtained from Foundation Medicine next-generation tumor sequencing across the coding regions of ≥100 cancer-related genes. Data included mutation level data, as well as spectrum-level MSI status and tumor mutation burden (TMB-M).

TMB-M was calculated using the following equation:

TMB-M = (Number of detected somatic coding mutations) / (Total exon length) 

Each microsatellite locus had its own estimate of TMB-M.

The next-generation sequencing-based "MSI score" was translated into colorectal MSI-H for MSI-H, MSI ambiguous, or microsatellite-stable (MSS) by uncrewished clustering of specimens for which MSI status was previously assessed via gold standard methods (e.g., H-J.

Figure 1: Frequency of MSI-H across tumor types.

Table 1: Comparison of MSI stability and high TMB mutation burden in patients with common tumor types.

<table>
<thead>
<tr>
<th>Tumor Type</th>
<th>Breast Cancer</th>
<th>Colorectal</th>
<th>Non-Small Cell Lung Cancer</th>
<th>Other tumor types combined</th>
<th>Unknown primary</th>
</tr>
</thead>
<tbody>
<tr>
<td># MSI-H (%)</td>
<td>4.2 (2.2%)</td>
<td>7.2 (2.3%)</td>
<td>8.8 (2.0%)</td>
<td>12.0 (1.7%)</td>
<td>10.7 (1.0%)</td>
</tr>
<tr>
<td># MSI Stable (%)</td>
<td>124 (4.3%)</td>
<td>140 (3.5%)</td>
<td>233 (2.3%)</td>
<td>323 (3.1%)</td>
<td>274 (2.6%)</td>
</tr>
<tr>
<td># MSI Ambigious (%)</td>
<td>10.5 (2.5%)</td>
<td>22 (0.5%)</td>
<td>26 (0.5%)</td>
<td>58 (1.0%)</td>
<td>67 (1.0%)</td>
</tr>
<tr>
<td># MSI Unknown (%)</td>
<td>440 (2.5%)</td>
<td>409 (2.4%)</td>
<td>462 (2.1%)</td>
<td>267 (2.3%)</td>
<td>242 (2.1%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Conclusions
In May 2017, the NCCN issued its first dosing algorithm approval. The rate of MSI-H in cancers of unknown primary is higher than many other cancer types. MEK/MAPK assessment (including by NGS) may be particularly important in this clinical setting in order to guide treatment strategies. Early data suggests that the rate at which MSI-H patients are being treated with checkpoint inhibitors is increasing over time. Further exploration of the real-world effectiveness of immune checkpoint inhibitors in the MSI-H population is still needed.

Because different immunotherapy regimens are under study, more patients receiving checkpoint inhibitors do not have MSI-H tumors. As such, exploration of additional biomarkers for immunotherapy response is critical. Tumor mutational burden, which may be associated with expression patterns, is an example of a biomarker that is being explored as an additional predictor of clinical benefit from immune checkpoint inhibitors.

Limitations
Before September 2017, MSI status was calculated retrospectively from tumor sequencing data, and was not, in most cases, reported to be treating physicians in order to guide treatment decisions.

In some cases (e.g., in colorectal cancer patients), physicians may have separately ordered MSI testing when another diagnostic laboratory is available to order the MSI status for their patients.

As with any retrospective real-world data set, this cross-section cohort may not be a representative sample of all cancer patients.

References

This work was supported by ASCO, the ASCO Foundation, and Flatiron Health. The authors declare no conflict of interest.
Opportunity 2: Heterogeneity

Each study studies individual patient populations, therapies, and biomarkers

CHECKMATE-226

Therapies: Nivolumab + Ipilimumab
Biomarkers: PD-L1 (multiple thresholds), Tumor Mutational Burden (10 mut/Mb)

KEYNOTE-189

Therapies: Pembrolizumab + Chemotherapy
Biomarker: PD-L1 (multiple thresholds)

Opportunity 3: Complexity
RWD can help tackle the complexity of immunotherapy response and resistance

Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma
Jesse M. Zaretsky, B.S., Angel Garcia-Diaz, Ph.D., Daniel S. Shin, M.D.,
Helena Escuin-Ordinas, Ph.D., et al.

Observations from 4 patients who developed resistance to immunotherapy in melanoma at a single academic medical center, who ultimately underwent advanced sequencing


STK11/LKB1 Mutations and PD-1 Inhibitor Resistance in KRAS-Mutant Lung Adenocarcinoma

Combining RWD from multiple academic medical centers led to the discovery of a new resistance mechanism for immunotherapy

Skoulidis F et al. Cancer Discovery 2018
Concluding Thoughts

- Multiple real world data sets are emerging: across academic medical centers, consortia, and industry-driven initiatives.
- These datasets can complement and extend data generated through clinical trials – helping generalize findings, evaluate multiple therapies with common biomarkers, and explore the complexity of evolving disease.
- Additional applications could include better trial design and assessment of post-approval efficacy, especially of histology agnostic therapies.

Open Questions

- Can real world data be used to accelerate development of therapies for niche patient populations by serving as a real world control arm?
- Can real world data be used to facilitate discovery of new targets and combinations, perhaps by engaging patients in research?
- How will complex biomarkers ultimately be translated into clinical care and decision making, and what role will real world data and technology play?
Thank you