Informatics to enable routine personalized cancer therapy [U01 CA180964]

Elmer V. Bernstam, MD
UT School of Biomedical Informatics

Funda Meric-Bernstam, MD
UT MD Anderson Cancer Center
Problem

From: Xiong, Wen
Sent: Thursday, June 12, 2014 10:13 AM
To: Meric-Bernstam, Funda
Cc: Piha-Paul, Sarina Anne
Subject: mutation report

Dr. Meric:

Please review the patient for BKM120. She has PIK3CA mutation.

Name: [redacted]
MRN: [redacted]
Protocol: 2013-0682
Diagnostics: Fallopian tube/Ovarian CA

Thanks,

Wendy Xiong,
Clinical Studies Coordinator.
Investigational Cancer Therapeutics.
The University of Texas M. D. Anderson Cancer Center
1400 Holcombe Blvd, Houston TX 77030
Phone: 713-745-3029
Pager: 713-745-3348
Fax: 713-792-3535
wxiong1@mdanderson.org

5 page report attached
Overview

- Biomedical Literature (MEDLINE)
- Clinical Practice Guidelines
- Clinical Trial Database
- Biomedical Literature (MEDLINE)
- Institutional Clinical Data
- Patient Clinical Profile
- Patient Molecular Profile
- Analysis
- Analysis
- Personalized Decision
- Local Constraints
- Patient Preferences
Real and immediate problem

• Clinical application at MD Anderson
  – The Khalifa Institute for Personalized Cancer Therapy (IPCT)
    • “Clearinghouse” protocol
      – >4000 patients with molecular profiles
    • Human-curated “gene sheets”
      – Gold standards for evaluation of automated methods
  – Department of Investigational Cancer Therapeutics
    • Phase I program
    • Primarily “trial-driven”
      – Problem is: given patient → assign to appropriate clinical trial
      – ~800 ongoing therapeutic trials at MD Anderson alone

• Plans discussed for UT System clinical trial recruitment
Approach

• Adapt and apply existing informatics techniques (SBMI)
• Evaluate against human-curated gold standards (IPCT/ICT)
• Apply to clinical trial matching
  – Outcome = recruitment
Current projects and progress

• Information curation for decision support
  – Gene knowledge sheets
  – Trials and drugs
  – Patient reports
  – Actionable gene panels

• Automated identification of molecular effects of drugs (AIMED)
  – Done, published
  – Working on improvements (e.g., ranking, statistical NLP)

• Creating a pipeline for literature curation
  – Ongoing
Decision Support via Knowledgebase Development and Molecular Tumor Board

Alterations: Copy number, Fusions, SNVs

Report functional consequences of each alteration (clinical data, functional genomics, functional predictions)

Functional Alteration in Driver Gene?

Relevant targeting drugs (direct and indirect)

Report evidence for using each drug in the context of altered gene/disease/molecular subtype

Actionable?

Retrieve clinical trials using the relevant drugs

Prioritization of mutations/targets
Identifying optimal approved or investigational treatment
Current Genomic Selection Strategies for Targeted Trials

97 unique trials with 53 unique genomic alterations and 62 drugs
Gene Knowledge Sheets

• Created “knowledge sheets” on 27 relevant genes
  – E.g., PIK3CA = 11 pages
• For each gene, detailed mutations
  – found to disrupt the protein structure
  – their therapeutic implications
  – targeted therapies
  – Trials
    • That select patients with this aberration
    • Of therapies targeting this aberration
• Open access (with registration) at http://personalizedcancertherapy.org
  – 95,469 page views during 41,562 visits by 22,728 unique users in the first year
  – Dropped after registration required
Trials and Drugs

• Curated over 1500 clinical trials
  – all MD Anderson trials + others
  – details of selection criteria in context of molecular evidence

• Curated over 600 drugs
  – all drugs used in MD Anderson clinical trials + others
  – providing genes targeted by drugs
Personalized Cancer Therapy

1. Molecular Profiling
2. Prognostic Markers
   - Markers predictive of drug sensitivity/resistance
   - Markers predictive of adverse events

Personalizedcancertherapy.org --- April 2014
@kennamshaw on twitter for announcement
Genomics Knowledgebase

- Gene and variant level information
- Mutations/copy number changes
- Frequency of alterations
  - in TCGA
  - In Cosmic
  - In MD Anderson IPCT experience

- Therapeutic Implications

- Drugs targeting alterations

- Clinical Trials
  - Genotype-specific Trials
  - Genotype-relevant trials
Other Progress

• Hot spot testing on over 6000 pts
• Targeted exome sequencing /annotated almost 2000 samples
  – Unusual responder program
  – Clearinghouse testing
• Optimizing annotation
  – Functional? Driver? CANDriver
  – COSMIC and TCGA frequencies
  – Therapeutic Implications>
  – Level 1 or 2 of evidence—clinical data
  – Preclinical data
    • Functional genomics
• Germline calls/return of incidental results
  – First large scale return of incidental results initiative (Meric-Bernstam et al Ann Onc 2015)
• Tools for point care annotation of variants
• JAMIA special issue on precision medicine informatics
  – Near record submissions (within 2-3 for any special issue)
  – 17 papers accepted, coming out in print July 2016
  – Cancer was a common application area
Personalized Results Reporting

• Leveraging in-house knowledge base → annotations for single patients
• Report that includes
  – Targetable aberrations
  – Targeted therapies
  – Summary context information (narrative)
  – Related genomic alterations
  – FDA approved therapy options
  – MD Anderson clinical trials
  – Clinical trials at other institutions
Actionable Gene Panel

- Gene recognition tool
- Clinical trial pipeline
- In-house drug database
  - Clinical “actionability” of genes in CLIA-certified gene panels - panel agnostic
    - 3-tier system
- Help clinicians identify mutations with available therapeutic options
  - Future design of panels with only actionable genes
Leveraging the literature to identify drugs and targets

Trevor Cohen MBBS, PhD (Co-I)
Associate Professor, UT SBMI
AIMED pipeline

Question
“Find drugs that target genes”

Normalization

Semantic Query
Drug-gene relation finder

Post-filtering
Drug lists, concepts lists, entity types, frequency

Knowledge Source
Medline + Clinical trials

SemRep

Knowledgebase
Semantic predications + Co-occurrence

Initial Result Set
Entities + Supporting documents

Answer List
Entities + Supporting documents
### Performance

- **“False positive” results**
  - ~26% actually true positive (i.e., should be on gold standard)
  - ~61% “should be considered” (i.e., not gold standard, but should be reviewed)

<table>
<thead>
<tr>
<th>Marketed</th>
<th>Yes</th>
<th>Pred.</th>
<th>&gt;4</th>
<th>INHIBITS, INTERACTS WITH, COEXISTS WITH</th>
<th>phsu, orch</th>
<th>2,251</th>
<th>80</th>
<th>0.69</th>
<th>0.3</th>
<th>0.42</th>
<th>0.55</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Yes</td>
<td>Pred.</td>
<td>-</td>
<td>INHIBITS, INTERACTS WITH, COEXISTS WITH</td>
<td>-</td>
<td>299</td>
<td>61</td>
<td>0.35</td>
<td>0.3</td>
<td>0.32</td>
<td>0.34</td>
</tr>
<tr>
<td>2</td>
<td>Yes</td>
<td>CoOccSen</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4,723</td>
<td>205</td>
<td>0.5</td>
<td>0.17</td>
<td>0.25</td>
<td>0.36</td>
</tr>
<tr>
<td>1 / 2</td>
<td>Yes</td>
<td>CoOccSen</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3,875</td>
<td>40</td>
<td>0.29</td>
<td>0.18</td>
<td>0.22</td>
<td>0.26</td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
<td>CoOccDoc</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1,609</td>
<td>129</td>
<td>0.25</td>
<td>0.19</td>
<td>0.22</td>
<td>0.24</td>
</tr>
<tr>
<td>All Phases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12,757</td>
<td>515</td>
<td>0.39</td>
<td>0.21</td>
<td>0.27</td>
<td>0.33</td>
</tr>
</tbody>
</table>
Ongoing work

• Ranking
  – Heuristics
  – “Learning to rank”

• Develop annotation pipeline
  – Drugs $\rightarrow$ genes
    • Drugs
    • Papers

• Two tasks
  – New gene
  – Update gene
    • Avoid repeat review of papers/drugs
Lessons learned

• Hard task
  – No real “gold standard”
    • Constantly changing
    • Inter-observer variability
  – Her2 intracellular domain protein = drug

• 100% recall is not what users actually want
  – Ranking is important

• Have to get workflow right
Publications and products

Thank you

Questions?