A Cancer Knowledge Network

Accelerate cancer R&D
1) By efficiently searching for treatments and cures
2) By learning from every patient both retrospectively and prospectively
3) By shattering the silos to share vetted knowledge
4) By creating a network of oncologists
<table>
<thead>
<tr>
<th>Go</th>
<th>Jeopardy (random factoids)</th>
<th>Cancer</th>
<th>Self-Driving Cars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closed World</td>
<td>Closed World (wikipedia)</td>
<td>Open World</td>
<td>Relatively Open World</td>
</tr>
<tr>
<td>Static rules (for thousands of years)</td>
<td>Static Knowledge (wikipedia)</td>
<td>Highly Dynamic (biology doesn’t change but everything else does)</td>
<td>Static rules (for decades)</td>
</tr>
<tr>
<td>Simulation is trivial</td>
<td>Simulation is unnecessary</td>
<td>Simulation is essentially impossible</td>
<td>Simulation is easy (Grand Theft Auto!)</td>
</tr>
<tr>
<td>Experiments are free</td>
<td>Experiments are unnecessary</td>
<td>Experiments aren’t just extremely expensive, but every experiment kills people or animals in horrible, painful ways!</td>
<td>Experiments are relatively inexpensive, but errors can kill!</td>
</tr>
<tr>
<td>Feedback is instantaneous</td>
<td>Feedback is instantaneous</td>
<td>Feedback can take years and is very noisy</td>
<td>Feedback is instantaneous and near perfect</td>
</tr>
<tr>
<td>Data is essentially unlimited</td>
<td>Data is essentially unlimited</td>
<td>Data is essentially non-existent</td>
<td>Data is plentiful</td>
</tr>
</tbody>
</table>
Cancer Knowledge Network
Thesis

- Clinical trials are inefficient in the molecular world;
- Because the size of the [Marker x Tx Combinations] space is enormous;
- And the number of advanced cancer cases is relatively small;
- And trial participation rates are very low;
- But every patient needs to be treated, even if innovation is necessary;
- But when oncologists innovate little is learned;
- Because case reports are hard to create;
- And journals and guideline updates are slow and require over-the-top evidence. (Oncologists should be able to use even anecdotal evidence, esp. in “off the map” cases, where there is little high quality evidence.)
- So, we are building a Cancer Knowledge Network;
- To permit oncologists to easily capture, and rapidly share case reports;
- And to easily and rapidly aggregate, peer review, and publish case series.
The decision problem: What to do to who, when?
Treatment Validation in the pre-OMIC Era

A handful of phenotypes (lung, breast, etc.)

A handful of treatments (chemo1, chemo2, etc.)

~100 cells
~1 Million patients/year
~10,000 patients/cell

Treatment selection was a statistical diagnosis problem: The diagnosis gave the treatment directly.
Treatment Validation in the OMIC Era

Millions of omic phenotypes (enormous feature vector!)

Thousands of treatments (New drugs + combinations)

10^9 cells! (11^z)
STILL ONLY A MILLION PATIENTS/YEAR!

Approx. ZERO patients/cell!

Treatment selection and planning are now reasoning problems, based upon huge amount of complex knowledge and data.
Precision Oncology Workflow

- **Scans**
- **Biomarkers**
- **Treated cell**
- **In vivo**
- **In vitro**
- **In silico**

**Serum Markers**

**Biopsies**

**Normal skin cell**

**Original cancer cell**

**Panomics**

**Modeling and Targeting**

**Combo Therapies**

**Test**

**Patient**

**Treatment Planning**

**Serum Markers**

**Biomarkers**

**Treated cell**

**In vivo**

**In vitro**

**In silico**
Precision Oncology: Target Treatments to Patients Based on Molecular Disease Models (MDMs)

Clinical History

Dosing and Sequencing Details

Patient Preferences

Patient Education

Treatment History

Trial Practicalities

Shared Decision Making

Financial Considerations

Drug Availability and Affordability

Guidelines

Tumor Availability for Testing

Disease Model Predictions

Precision Oncology: Target Treatments to Patients Based on Molecular Disease Models (MDMs)
Kinds of knowledge unique to practice:

(most conveniently accessible via MTBs)

1. What parts of the literature are **too early**? (too tentative to be useful in practice)
2. What parts of the literature are **outdated**? (no longer believed to be useful in practice)
3. Knowledge **not available at all** in the literature (e.g., ethical, political)
4. What are the **importance relationships** between the various knowledge elements?
5. How knowledge is **interpreted** in context of use?
6. In-context **cross-domain** translations/relations.
7. Assembly and explanation **heuristics**.
   (the “how to” of practical model-based reasoning)
Kinds of knowledge unique to MTBs:

1. Arguments pro and con hypotheses.
   (Published case reports are highly sanitized and are written explicitly for educational purposes.)
2. Rationales for rejected hypotheses.
   (Often these never show up anywhere else!)
   (e.g., as above: too early, outdated, ethical, political, ...)
4. Implicit (unpublishable) knowledge relating to this particular sort of problem solving.
5. Stats on the sorts of cases that typically occur, and on the sorts of difficulties typically arise.
A recommendation with Level of Evidence B or C does not imply that the recommendation is weak. Many important clinical questions addressed in the guidelines do not lend themselves to clinical trials. Even when randomized trials are unavailable, there may be a very clear clinical consensus that a particular test or therapy is useful or effective.

---

**Table 1. Applying Classification of Recommendation and Level of Evidence**

<table>
<thead>
<tr>
<th>Level A</th>
<th>Level B</th>
<th>Level C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple populations evaluated</td>
<td>Limited population evaluated</td>
<td>Very limited evaluated</td>
</tr>
<tr>
<td>Data derived from randomized or meta-analysis</td>
<td>Data derived from single randomized or nonrandomized studies</td>
<td>Only consensus opinion of experts, case studies, or standard of care</td>
</tr>
</tbody>
</table>

**Sizing of Treatment Effect**

- **Class I**: Benefit >> Risk
  - Procedure/Treatment SHOULD be performed/administered
- **Class IIa**: Benefit >> Risk
  - Additional studies with focused objectives needed
  - IT IS REASONABLE to perform procedure/administer treatment
- **Class IIb**: Benefit ≥ Risk
  - Additional studies with broad objectives needed; additional registry data would be helpful
  - Procedure/Treatment MAY BE CONSIDERED
- **Class III** No Benefit or Class III Harm
  - Procedure/Test or Treatment
  - No benefit
  - Not helpful
  - No proven benefit
  - Harm
  - Excess cost or benefit or harmful
  - Harmful to patients

**Estimate of Certainty (Precision) of Treatment Effect**

- **Only expert opinion, single case study, or standard of care**

**Suggested Phrases for Writing Recommendations**

- Should be recommended
- Is indicated
- Is useful/effective/beneficial
- Is probably recommended or indicated
- May/might be considered
- May/might be reasonable
- Usefulness/effectiveness is unknown/unclear/uncertain or not well established
- Is not recommended
- Should not be performed/administered/other
- Is not useful/beneficial/effective
- Potentially harmful
- Causes harm
- Associated with excess morbidity/mortality
- Should not be performed/administered/other

**Comparative effectiveness phrases**

- Treatment/strategy A is recommended/indicated in preference to treatment B
- Treatment/strategy A should be chosen over treatment B
- It is reasonable to choose treatment A over treatment B

---

A recommendation with Level of Evidence B or C does not imply that the recommendation is weak. Many important clinical questions addressed in the guidelines do not lend themselves to clinical trials. Even when randomized trials are unavailable, there may be a very clear clinical consensus that a particular test or therapy is useful or effective.
Get explanations from MTBs. (Why?) Share them and make them computable.
Molecular Tumor Board

Broad-based clinical experience:
- Medical oncologists
- Surgical oncologist
- Radiation oncologists
- Other medical specialists
- Biostatisticians
- Radiologists
- Pathologists
- Clinical geneticists
- Pathway experts
- Pharmacologists
- Oncological nurses
- Reimbursement Admins
- Computational Biologists
Get explanations from MTBs. Share them and make them computable.

**History / Physical Exam:** The patient was diagnosed in 2008. Biopsy revealed this to be a lymph node lesion which was considered to be a lymph node. The second biopsy revealed 2 lymph nodes were positive for melanoma; no lesion was ever identified.

He then went on to receive radiation therapy and completed one year of interferon therapy which was negative for any evidence of disease. He showed some progression of a right axillary lesion.

**Previous Treatment:**
- 2008: Radiation Therapy
- 2008-09: Interferon Therapy

**Aberrations (see attached for details):**
- **NRAS** Q61R
- DNMT3A W297*
- PAX5 V26G
- SPEN E2267fs*

NRAS encodes N-Ras, a member of the Ras family proteins, which are membrane-associated small GTPases that mediate transduction of growth signals. Activation of Ras signaling causes cell growth, differentiation, and survival by activating the Raf/MAPK/ERK, PI3K, and other pathways (Pylayeva-Gupta et al., 2011; 21993244). Q61R is located in the "switch II" region of Ras family members, which plays a key role in regulating the promotion of downstream signaling; mutation at this location disrupts GTP hydrolysis and results in constitutive activation of the mutant protein (Buhrman et al., 2010; 20194776, Scheffzek et al., 1997; 9219684, Prior et al., 2012; 22589270, Pylayeva-Gupta et al., 2011; 21993244). NRAS activating mutations, specifically Q61R, have been shown to promote malignant behavior in cancer cells, as suppression of NRAS Q61R expression inhibits proliferation, invasion, and migration in melanoma cells (Eskandarpour et al., 2009; 18814281). NRAS mutations have been
Findings were discussed at the ITOMIC tumor board. It was mentioned that complete loss of BRCA1 heterozygous loss of CHEK2 might underlie the patient’s excellent response to Cisplatin, and might predict sensitivity to a Poly-ADP ribose polymerase (PARP) inhibitor (Fong et al., N Engl J Med 2009;361:123-134, Tutt et al., Lancet 2010;376:235-44; Somlo, San Antonio Breast Cancer Conference 2013). Since PARP inhibitors are investigational, it was recommended that the patient consider enrollment into a PARP inhibitor trial. The two FGFR2 missense mutations were also discussed. Patient's tumor has FGFR mis-sense mutations similar to S252W. The COSMIC database lists 44 endometrial cancers bearing the FGFR2(S252W) variant, and 4 endometrial cancers with FGFR2(Y375C). Neither mutation has been described previously in breast cancer. Of note a single COSMIC entry describes an endometrial cancer (1105632) bearing FGFR2(S252W;Y375C). Both variants occur within the ligand binding domain, and S252W expands the repertoire of ligands capable of activating FGFR2 (Greulich and Pollock, Trends Mol Med. 2011;17: doi:10.1016/j.molmed.2011.01.012, Dutt Proc Natl Acad Sci U S A 2008 PMID1855176, Yu et al. Proc Natl Acad Sci U S A. 2000 97:14536-41). ...
Subject 001 is a 45-year-old woman with metastatic TNBC*, originally diagnosed in January 2011 with Stage IIIA involvement of the left breast, who had failed multiple prior chemotherapy regimens by the time of study enrollment. Seven core needle biopsies of an enlarged right axillary lymph node, a left chest wall mass, and left pleural effusion.

* TNBC == Triple Negative Breast Cancer, i.e., absence of markers for sensitivity to estrogen, progesterone, and human epidermal growth factor receptors.
Findings were discussed at the ITOMIC tumor board. It was mentioned that complete loss of \textit{BRCA1} heterozygous loss of \textit{CHEK2} might underlie the patient’s excellent response to Cisplatin, and might predict sensitivity to a Poly-ADP ribose polymerase (PARP) inhibitor (Fong et al., \textit{N Engl J Med} 2009;361:123-134, Tutt et al., \textit{Lancet} 2010;376:235-44; Somlo, San Antonio Breast Cancer Conference 2013). Since PARP inhibitors are investigational, it was recommended that the patient consider enrollment into a PARP inhibitor trial. The two \textit{FGFR2} missense mutations were also discussed. Patient's tumor has FGFR mis-sense mutations similar to S252W. The COSMIC database lists 44 endometrial cancers bearing the \textit{FGFR2(S252W)} variant, and 4 endometrial cancers with \textit{FGFR2(Y375C)}. Neither mutation has been described previously in breast cancer. Of note a single COSMIC entry describes an endometrial cancer (1105632) bearing \textit{FGFR2(S252W;Y375C)}. Both variants occur within the ligand binding domain, and S252W expands the repertoire of ligands capable of activating FGFR2 (Greulich and Pollock, Trends Mol Med. 2011;17: doi:10.1016/j.molmed.2011.01.012, Dutt Proc Natl Acad Sci U S A 2008 PMID1855176, Yu et al. Proc Natl Acad Sci U S A. 2000 97:14536-41). ...
Findings were discussed at the ITOMIC tumor board. It was mentioned that complete loss of *BRCA1* heterozygous loss of *CHEK2* might underlie the patient’s excellent response to Cisplatin, and might predict sensitivity to a Poly-ADP ribose polymerase (PARP) inhibitor (Fong et al., *N Engl J Med* 2009;361:123-134, Tutt et al., *Lancet* 2010;376:235-44; Somlo, San Antonio Breast Cancer Conference 2013). Since PARP inhibitors are investigational, it was recommended that the patient consider enrollment into a PARP inhibitor trial. The two *FGFR2* missense mutations were also discussed. Patient's tumor has FGFR mis-sense mutations similar to *S252W*. The COSMIC database lists 44 endometrial cancers bearing the *FGFR2(S252W)* variant, and 4 endometrial cancers with *FGFR2(Y375C)*. Neither mutation has been described previously in breast cancer. Of note a single COSMIC entry describes an endometrial cancer (1105632) bearing *FGFR2(S252W;Y375C)*. Both variants occur within the ligand binding domain, and S252W expands the repertoire of ligands capable of activating FGFR2 (Greulich and Pollock, Trends Mol Med. 2011;17: doi:10.1016/j.molmed.2011.01.012, Dutt *Proc Natl Acad Sci U S A* 2008 PMID1855176, Yu et al. *Proc Natl Acad Sci U S A.* 2000 97:14536-41). ...
Assemble personalized cancer models
Explain the disease process

S001 INTERPRETATION (continued)

... Subject 001’s tumor had a somatic frameshift insertion in *BRCA1* with loss of heterozygosity (LOH), a heterozygous frameshift deletion in *CHEK2*, and two separate *FGFR2* point mutations generating the variant *FGFR2*(S252W;Y375C), also with LOH. ... A cross-platform comparison of variant allele frequencies showed *FGFR2*(S252W;Y375C) levels to be disproportionately higher within the RNA-Seq dataset compared to UW-OncoPlex and WES, indicating that the variant allele is selectively expressed. Members of the ITOMIC Tumor Board predicted that the *BRCA1* and *CHEK2* mutations may confer susceptibility to PARP inhibitors. In considering the FGFR2 mutations members noted that while neither FGFR2 variant had been reported previously in breast cancer, the COSMIC database contained multiple cases of endometrial cancer associated with each variant individually, and a single case of endometrial cancer in which both S252W and Y375C occurred together. Collectively, these results suggested that *FGFR2*(S252W;Y375C) may be a driver of tumor growth.

This example thanks to Tony Blau, University of Washington. Please do not reproduce or pass on without permission.
Create **and test** treatment plans

S001 INTERPRETATION (continued)

...Both native FGFR2 and the p.S252W variant are susceptible to inhibition by the investigational drugs Dovitinib (Konecny GE, et al. Mol Cancer Ther. 2013;12:632-42) and NVP-BGJ398 (Guagnano et al. Cancer Discov. 2:1118-33, 2012 PMID: 23002168), and the approved drug, Ponatinib (Gozgit et al. Cancer Research: April 15, 2011; Volume 71, Issue 8, Supplement 1) **Hypothesize** Dovitinib treatment. Subject 001’s tumor had a somatic frameshift insertion in BRCA1 with loss of heterozygosity (LOH), a heterozygous frameshift deletion in CHEK2, and two separate FGFR2 point mutations generating the variant FGFR2(S252W;Y375C), also with LOH. ... A cross-platform comparison of variant allele frequencies showed FGFR2(S252W;Y375C) levels to be disproportionately higher within the RNA-Seq dataset compared to UW-OncoPlex and WES, indicating that the variant allele is selectively expressed. **Members of the ITOMIC Tumor Board predicted that the BRCA1 and CHEK2 mutations may confer susceptibility to PARP inhibitors. ...**

This example thanks to Tony Blau, University of Washington. Please do not reproduce or pass on without permission.
S001 ACTIONS AND OUTCOMES

The patient had an excellent response to Cisplatin, but could not continue treatment due to hematological toxicity. The ITOMIC Tumor Board recommended that attempts be made to access a PARP inhibitor trial, and that an FGFR2 inhibitor be considered. AbbVie made the PARP inhibitor, Veliparib, available on a compassionate use basis. We therefore submitted a single patient investigational new drug (IND) application for Veliparib, and this was initiated at a dose of 300 mg twice daily with a planned dose escalation after two weeks to 400 mg twice daily. Nausea prevented this planned dose increase. After 10 weeks of Veliparib treatment the patient developed a new skin rash. A skin biopsy demonstrated tumor cells, and the patient underwent 4 separate punch biopsies of the skin. [...] The patient was able to obtain Ponatinib without cost through ARIAD’s patient assistance program. After performing another set of study-related biopsies upon completion of Veliparib, Subject 001 initiated treatment with Ponatinib. Over the ensuing 4 weeks, her skin lesions regressed, however. Serial measurements of CTCs was consistent with a response to Cisplatin, an initial response but eventual progression on Veliparib, and a response to Ponatinib (Supplemental Figure)

This example thanks to Tony Blau, University of Washington. Please do not reproduce or pass on without permission.
A Cancer Knowledge Network

Accelerate cancer R&D
1) By efficiently searching for treatments and cures
2) By learning from every patient both retrospectively and prospectively
3) By shattering the silos to share vetted knowledge
4) By creating a network of oncologists
A Cancer Knowledge Network

Molecular Tumor Board

1. Capture cases and insights flowing through elite tumor boards

2. Rapidly get them to the right physicians at the right time

3. Track outcomes and treatment reasoning to close the learning loop.

4. Operate CRO 3.0: Global Cumulative Treatment Analysis
Stanford Tumor Board

Stanford Local NanoPub Stack

Case ID
Raw
Translation
Translation
Semantic
Coding

Yale Tumor Board

Yale Local NanoPub Stack

Case ID
Raw
Translation
Translation
Semantic
Coding

Public (Global, Archival) e-Publication

Search and Aggregation Tools return results to Tumor Boards

Professional (Global but Closed) NanoPub Stack

Peer Review

Other Tumor Boards (future)

Vetting

Vetting

E-Publication

Tools return results to Tumor Boards

Search and Aggregation

Tools return results to Tumor Boards

Other Tumor Boards (future)
The Knowledge Cycle

Tech Scribes Case Discussion

Semi-Automatic Double De-Identification

Summarization and Insight Extraction

Two levels of curation:
1. Vetting for colleagues
2. Peer Review for ePublication

Semantic Encoding

ePublication

Applications: Case search, Case series, Curbsides, etc.

Molecular Tumor Board

Other MTBs and Professionals
Cancer Knowledge Network Services

- Capture case details and tumor board discussion
- Enable tumor board meeting case management
- Extract and vet insights (Treatment Rationales)
- Find similar cases
- Aggregate and publish case series and collective insights
- Facilitate expert consults
- Translate case details and insights into computable form
- *Validate hypotheses across the network (CRO 3.0)*
Findings were discussed at the ITOMIC tumor board. It was mentioned that [[(Supports[Refs] (AND Ref001 Ref002 Ref003) -> [(Supports["predicts sensitivity to"] Obs005 -> Hyp001) [(AND [(Supports["Underlies"] Obs005 -> Obs003)](Obs005: (AND Obs001 Obs002))) [Obs001: complete loss of BRCA1] and [Obs002: heterozygous loss of CHEK2]] might underlie [Obs003: the patient’s excellent response to Cisplatin]], and [Hyp001: might predict sensitivity to a Poly-ADP ribose polymerase (PARP) inhibitor]] ([Ref001: Fong et al., N Engl J Med 2009;361:123-134], [Ref002: Tutt et al., Lancet 2010;376:235-44]; [Ref003: Somlo, San Antonio Breast Cancer Conference 2013]]). [(Supports Fct001 -> Hyp002) Since [Fct001: PARP inhibitors are investigational], it was recommended that the patient [Hyp002: consider enrollment into a PARP inhibitor trial]]. The two FGFR2 missense mutations were also discussed. [+++001: [Obsxxx: [Patient's tumor has FGFR mis-sense mutations similar to S252W]]] [Dat001: The COSMIC database lists 44 endometrial cancers bearing the FGFR2(S252W) variant], and [Dat002 ["The COSMIC database lists"] 4 endometrial cancers with FGFR2(Y375C)]. Neither mutation has been described previously in breast cancer ...
S001 INTERPRETATION (continued)

From the Cancer Knowledge Network to CRO 3.0
CRO 3.0: The Need for Active Coordination

Millions of omic phenotypes

Thousands of treatments (New drugs + combinations)

$10^9$ cells!
STILL ONLY A MILLION PATIENTS/YEAR!

Approx. ZERO patients/cell!

You must ACTIVELY control the search across the network of MTBs to optimize the treatment of, and value of patients!
Global Cumulative Treatment Analysis (GCTA)

Exploration/Exploitation Tradeoff across all patient encounters