# Challenges, Opportunities and Gaps in Informatics and Cancer Research

## IOM, National Cancer Policy Forum

## February 2012

**Lawrence N. Shulman, MD**
Dana-Farber Cancer Institute
Brigham and Women's Hospital
Harvard Medical School
Boston, Massachusetts, USA

# Disclosures

- No financial conflicts of interest
- American Society of Clinical Oncology (ASCO)
  - Member of HIT Workgroup
  - Chair of Quality of Care Committee
- Commission on Cancer
  - Vice-chair, Quality Integration Committee
- Certification Commission for Health Information Technology (CCHIT)
  - Member – Oncology workgroup
  - Co-chair – Clinical Research workgroup

*Some (me included) believe we are at an "inflection" point in cancer research, and have the opportunity for a truly transformative approach that will accelerate not only our understanding of the biology of cancer, but also our development of new, more effective therapies*

*But to do that we need to take advantage of technologic opportunities*

# Some Random Thoughts on Research Opportunities

- Rapid advances in technology are resulting in a dramatic increase in the output of genomic and molecular data related to cancer biology

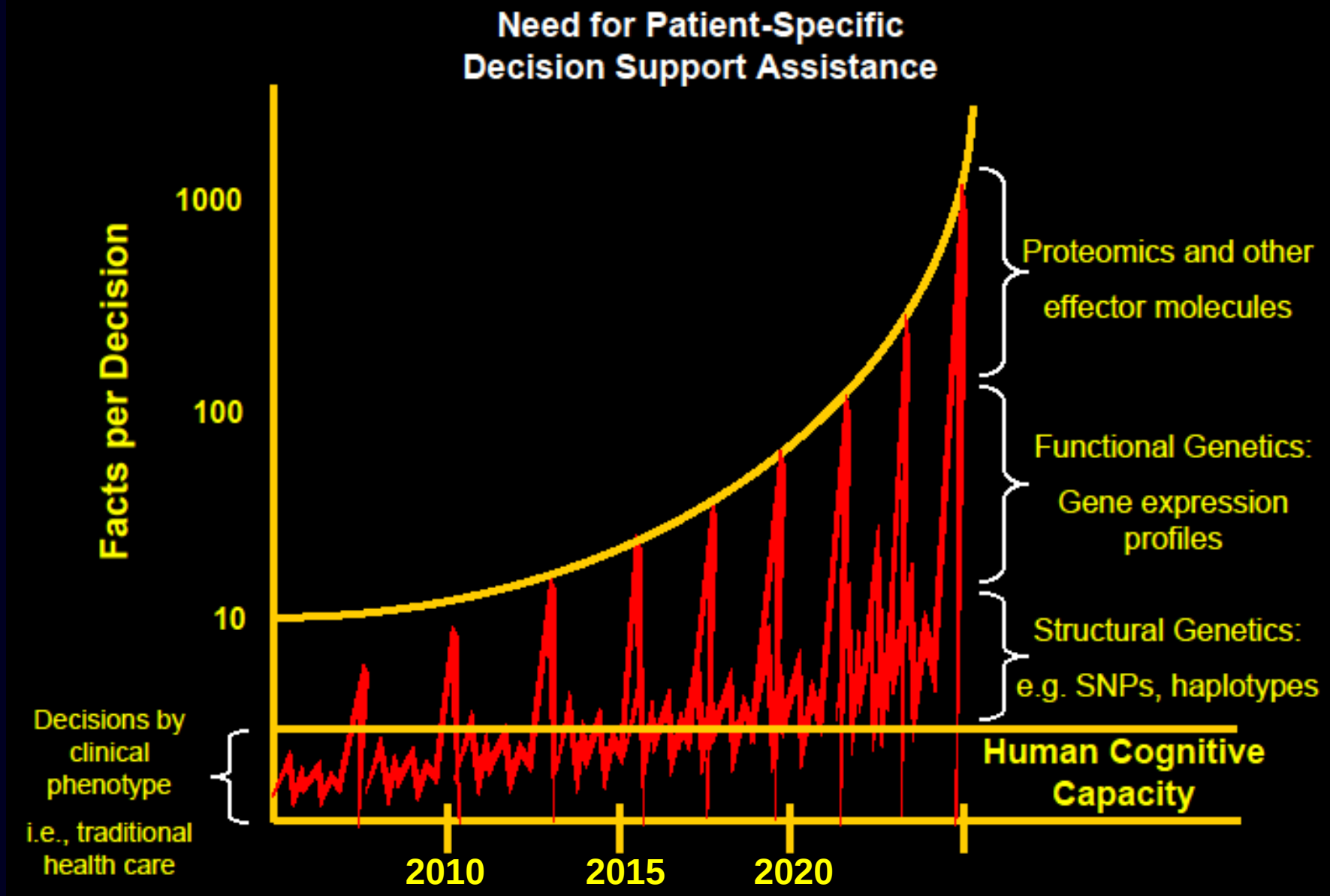
- Traditionally these investigations have been focused on the inherent biology of a cancer cell – sometimes somewhat in isolation of the clinical setting

# Some Random Thoughts on Research Opportunities

- These emerging data can inform our understanding of:
  - Basic cancer biology – driver mutations, etc (somatic)
  - Cancer epidemiology (somatic and germline)
  - Cancer behavior – growth rate and metastases, etc (somatic)
  - Response to specific therapies (somatic and germline)
  - Toxicity to specific therapies (germline)
  - Optimal care for patient or cohorts of patients (survival and toxicity)
  - And others….this is an incomplete list

# Some Random Thoughts on Research Opportunities

- The sheer volume of emerging data is more than any of us can keep track of
- …or make appropriate or interesting connections between




From Carolyn Compton, NCI

# Some Random Thoughts on Research Opportunities

- Research data – particularly genomic and molecular data – will be much more powerful when inter-connected with clinical data
- For research data to be optimally useful, it must be structured and in a database
- For clinical data to be useful it must contain certain critical elements, and it must be structured and in a database
- IT infrastructure is necessary to
  - Be HIPAA compliant
  - Enable data exchange (data liquidity)
  - Mine combined data

# Clinical Medicine and Clinical Data

- At the same time, clinical medicine is becoming increasingly more complex – the number of new data points required to make clinical decisions, and the number of new drugs available
- Slowly but surely we are moving towards electronic health records (EHRs) to facilitate practice, supply decision support, etc
- The adoption of electronic health records is a tremendous opportunity to supply structured clinical data – but the current state of EHRs does not optimally support this
  - Many critical data elements missing
  - Data elements that are included are often not in structured format

# What is the Essential Clinical Data Set that is Needed?

- Patient demographics
- Tumor type and anatomic and non-anatomic staging
- Treatment plan, treatment intent, and treatment received
- Tumor response
- Toxicity
- Patient reported outcomes
- Disease-free progression, and overall survival
- Others???

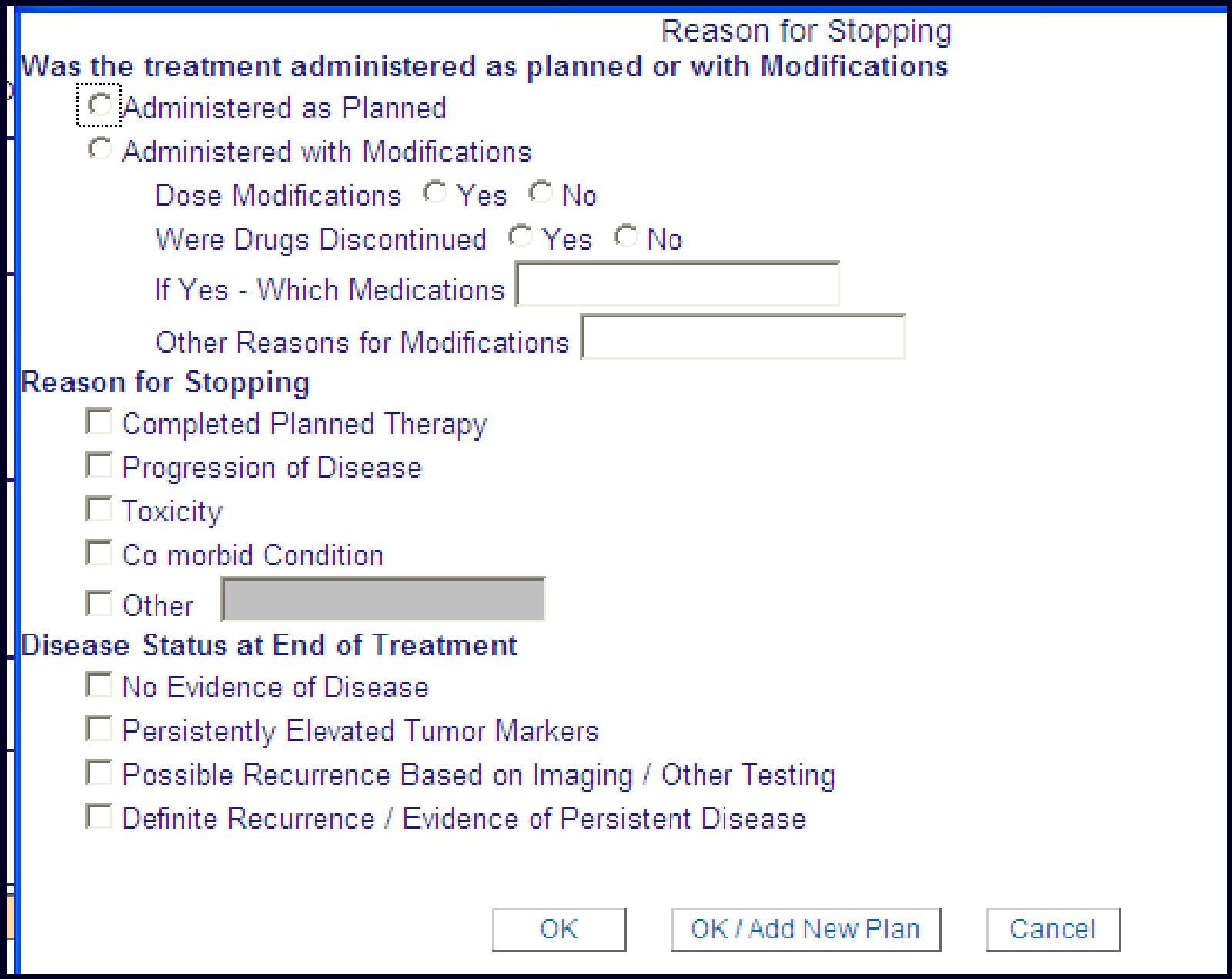
Tumor Type: Breast

| Tumor Description | |
|---|---|
| Dx Date: | 07/01/2008 |
| T: | T2 |
| N: | pN3c |
| M: | M0 |
| Group Stage: | IIIC |
| Prov. Stage: | IIIC |
| Site: | Left |
| ER: | Positive |
| PR: | Positive |
| HER2/neu IHC: | 0 |
| Inv Histology: | Lobular |
| Inv Hist Grade: | Well Diff (grade 1) |
| Sites of Recurrence/Metastasis: | |
| 03/18/2010 Comments: | Intra Abdominal peritoneal |

− Treatment Plan #1 DFCI (Fin

| Regimen | | |
|---|---|---|
| Goal of Treatment / Indication | Palliative / Non-curative; | |
| *Diagnosis* | *Regimen* | *Chemo Med* |
| Breast | PACLITAXEL 80 | PACLITAXEL (TAXOL ) (80 mg/m2) progression |
| **Additional Information** | | |
| ECOG Performance Status | 0=Fully active, able to carry on all pre-disease perfo | |

# Module to be completed when stopping a treatment course



***The most successful programs will be those that inter-connect research and clinical activities and data, in an organized and efficient manner, with as broad a database as possible.***

# SPARKS

## (Synergistic Patient And Research Knowledge Systems)

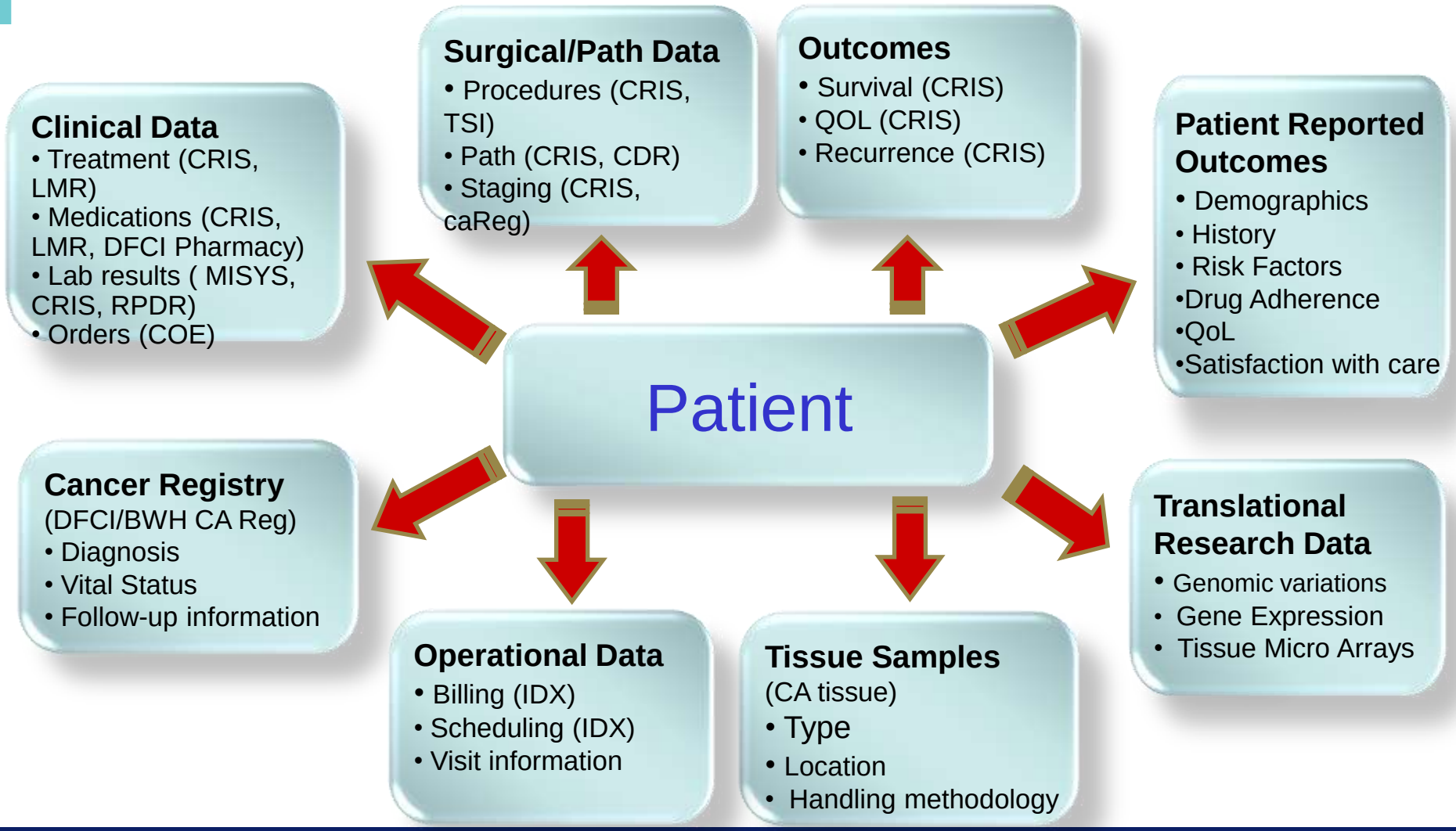
Vision: To provide cutting edge and collaborative Institutional Informatics Framework to **accelerate scientific discoveries** and their **translation into clinical practice** to enable early diagnosis, **personalized** treatment, cure and prevention of cancer and related diseases

Objective:

Implement policies, standards, systems and tools that facilitate collection, integration, mining, analysis and interpretation of biomedical data to accelerate scientific discoveries and their translation into personalized medicine and clinical practice.

# Long Term Goal: Establish An Integrated Patient-Centric Clinical genomic Data Model & Systems for Enabling Translational Research & Personalized Medicine

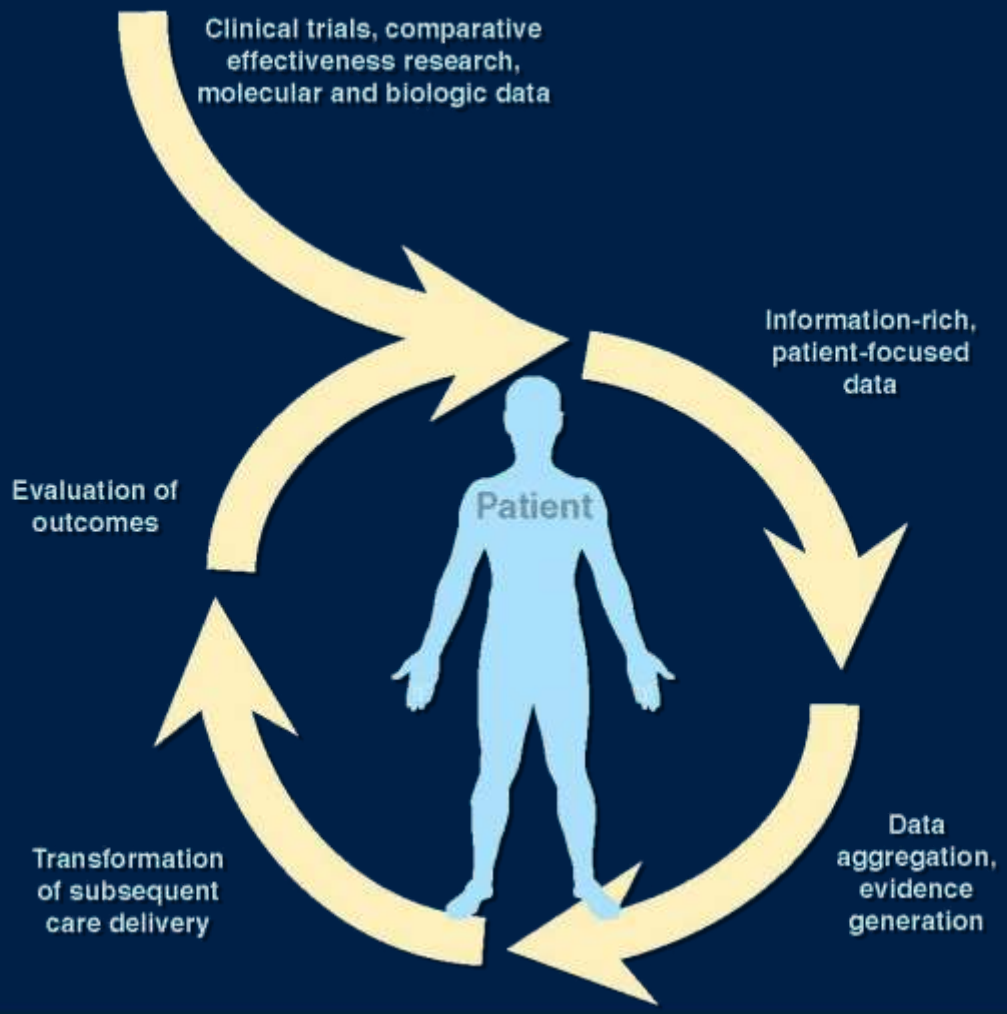
DANA-FARBER CANCER INSTITUTE



# IOM ‘Rapid Learning System for Cancer Care’

“In this framework, routinely collected real-time clinical data drive the process of scientific discovery, which becomes a natural outgrowth of patient care”

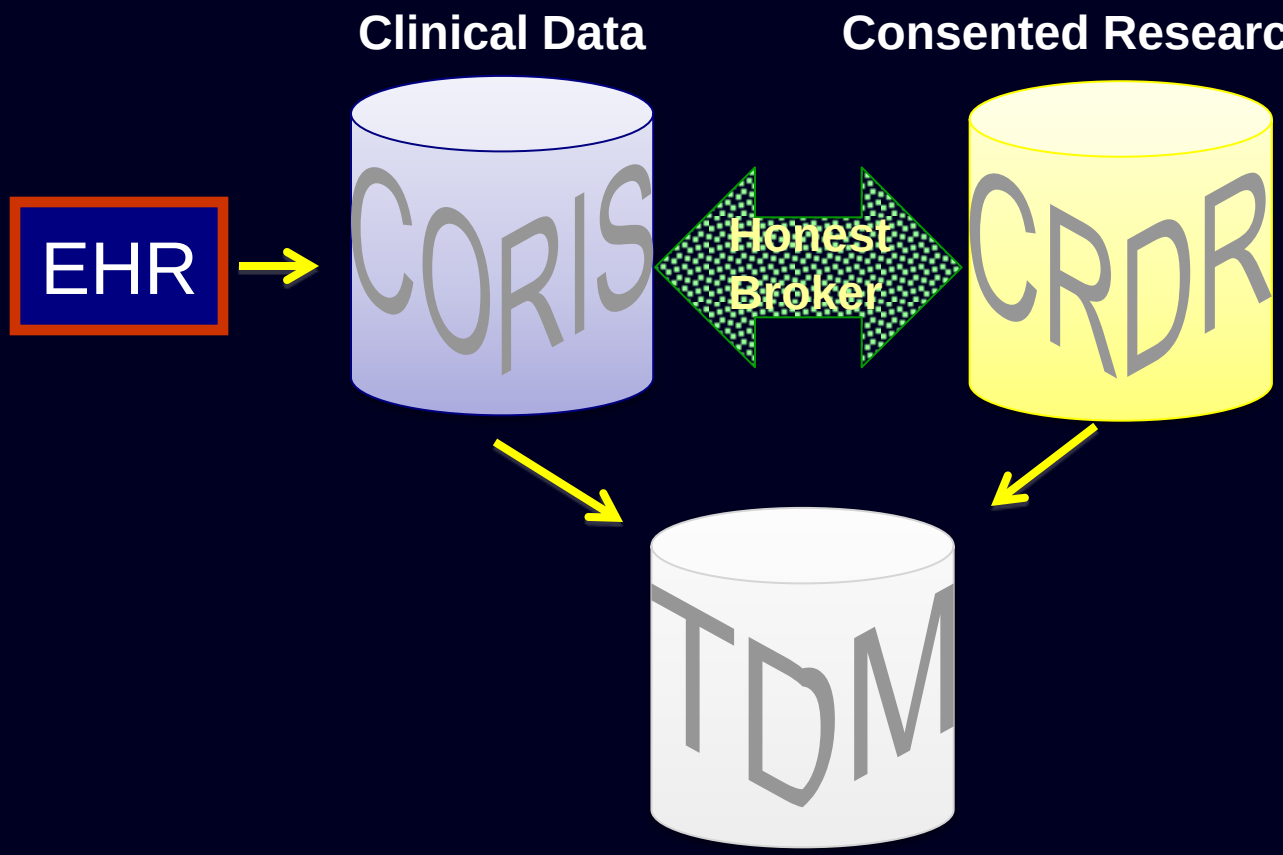
Abernethy et al, Rapid-Learning System for Cancer Care, JCO 2010



# The Life-Cycle of a Gene Mutation

- Basic research laboratory discovery
  - Unclear clinical significance
- Association with clinical syndrome
- Translational research ties mutation to disease prognosis, or response to particular therapy (HER2 or EGFR kinase mutation) – research-based clinical investigation
- Clinical significance is validated and testing becomes standard CLIA certified

# Combined Clinical/Molecular Queries




**“Transient Data Mart”**
**Houses data to be used in combined query**
**Destroyed after query is complete**

# Some Types of Queries – Cohort Studies

- **Tell me, in aggregate, the number of patients who have:**
  - **mutation X in their tumors**
  - **their diagnoses**

  **de-identified data – exploratory investigation performed by "honest broker" – looking for an actionable mutation in multiple tumor types**

  - **HER2 amplification in breast, gastric and salivary gland tumors and response to trastuzumab**
  - **Bcr-abl, and c-kit in CML, GIST tumors and response to imatinib**
  - **Covered by "umbrella" protocol**

# Some Types of Queries – Cohort Studies

- **Tell me the frequency of mutation X in:**
  - **women with ER positive, HER2 positive, metastatic breast ca**
  - **between the ages of 50-65**
  - **who had progressive disease on tamoxifen**

  **de-identified data – exploratory investigation performed by "honest broker"**

  - **Looking for a mutation that influences ER and HER2 targeted therapy (PI3 kinase mutations)**

  - **Covered by "umbrella" protocol**

# Research Protocol with Patient Identifiers

- **Tell me all the actual patients with:**
  - **ER positive, HER2 positive metastatic breast ca**
  - **between the ages of 50-65**
  - **with hormone resistant disease**
  - **who have mutation X**

  **identified data with IRB approval**

  - **Looking for a mutation that influences ER and HER2 targeted therapy (PI3 kinase mutations)**
  - **Individual patient information looking for candidates eligible for a clinical trial**
  - **Covered by specific IRB approved protocol**

# A Plea for Robust Electronic Health Records

- All clinical data codified
- Detailed demographic data
- Detailed tumor characteristics and staging data
- Detailed treatment histories
- Codified treatment responses and treatment resistance development
- Codified genomic and molecular data
- Inter-operable, standard-adherent data

**We are not close to this aspiration**

# A Plea for Robust Research Databases

- Structured genomic and molecular data
- Interoperability – standards
- Linked with relevant data – clinical and research

**We are not close to this aspiration**

***We could do this:***

1. ***At the laboratory level***
2. ***At the cancer center level***
3. ***At a national or international level***
4. ***Can we include both academic centers and community practices?***

***These 3 approaches have different benefits and risks, but one might argue that the transformational opportunities will not be realized without us being able to "pull off" #3 and #4***

***And…….***

- ***Doing #3 and 4 will most benefit scientific discovery, outcomes research, and the individual patient***

- ***But to really take advantage of "Big Data" we need mechanisms that permit actual data exchange***

***The most successful programs will be those that inter-connect research and clinical activities and data, in an organized and efficient manner, with as broad a database as possible.***

**Thank you**