Informatics and Information in Radiation Oncology:

OncoSpace

John W. Wong, Ph.D.
Todd McNutt, Ph.D.
Department of Radiation Oncology

Supported in Part by Elekta-Impac
“Today, most business ---- down to the smallest corner grocery store have better information about their sales and inventories than even affluent medical practices have about their patients. ……………”

--- Michael Bloomberg

to the Academy National Health Policy Conference, 05/07
NIH Roadmap 3: Re-engineering the Clinical Research Enterprise

  - “One of the greatest inefficiencies of the current model of clinical research in our country is the lack of a sustaining infrastructure (which includes shared resources, common data standards, and effective use of information technology among researchers), as well as the lack of a convenient forum to share best practices and learn from one another’s mistakes and successes.”
Cancer Bio-informatics Grid (caBIG)

- NCI: cancer Bioinformatics Grid (caBIG) provides infrastructure support
  - clinical trials management systems
  - integrative cancer research
  - tissue banks and pathology
  - Image workspace

- Not directed to address specific research or clinical questions
Data “Loss” at the Institutional Level

- Data we are capturing
  - Labs, Images, Treatment Plans
- Data we are sending away
  - Patients in protocols
- Data we are storing
  - Disparate databases
- Data (experience) we are not capturing
  - Discarded treatment plans (and decision making process)

- Information and knowledge are Not captured systematically
- Not utilized efficiently to impact research and patient care

OCI_50th 2008, JWW
Challenges of data longevity and re-use

• RTOG
  – Formed 1968, funded since 1971
  – Activated 300 trials
    • 40 on-going
    • 60,000 patients enrolled
  – Established QA, credentialing process for RTP and dosimetry
  – Centralized data repository; lacks secondary research
  – No measure of impact on community practice
Multiple Informatics Initiatives at JHU

- Johns Hopkins University Health Systems
  - Committee for Health Informatics
  - *Johns Hopkins Medical Image Archive (JHMIA)*
  - *I4M: Integration of Imaging, Information and Intervention in Medicine*
  - Clinical Trial Groups
  - Industrial collaborations
    - Microsoft (Almaga -- Healthcare Informatics)
    - IBM (Computational Medicine)
    - Harris Corporation (Multi-disciplinary clinic)
    - .............
The JHMIA Program– (Radiology)

A single archive where all medical images and other non-textual data (and associated reports, etc.) from across a Healthcare Enterprise are stored

- Clinical (300 TB now --- 700 TB in 2 years)
- Research
- Waveform
- Genomic (planned)
- Proteomic (planned)

• Medical Image Archive ➔ Medical Data Archive?
## JHMIA

**An Enterprise Image Archive**

<table>
<thead>
<tr>
<th>Current Participants</th>
<th>Committed</th>
</tr>
</thead>
<tbody>
<tr>
<td>JHH Radiology</td>
<td>BMC Vascular</td>
</tr>
<tr>
<td>JHH Vascular Surgery</td>
<td>BMC Adult Echo Cardio</td>
</tr>
<tr>
<td>JHH Peds Cardiology</td>
<td>Endoscopy (GI)</td>
</tr>
<tr>
<td><strong>JHH Rad Oncology</strong></td>
<td></td>
</tr>
<tr>
<td>JHH Adult Echo Cardio</td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td></td>
</tr>
<tr>
<td>Cardiology</td>
<td></td>
</tr>
<tr>
<td>BMC Radiology</td>
<td></td>
</tr>
<tr>
<td>Ophthalmology</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Potential</th>
</tr>
</thead>
<tbody>
<tr>
<td>OB/GYN</td>
</tr>
<tr>
<td>Pathology</td>
</tr>
<tr>
<td>Howard County General Hospital</td>
</tr>
</tbody>
</table>
JHMIA and I^4M

Analytic Database(s): Query and Security

Analytic and Change Tools: Extraction of Information

Decision Support: Data-mining Statistical Modeling

Web-service

 Goal: To improve both medical research and patient care

OCI_50\textsuperscript{th} 2008, JWW
The enterprise model: JHMIA and I^4M

<table>
<thead>
<tr>
<th>Robotic Surgery</th>
<th>Radiation Oncology</th>
<th>Ophthalmology</th>
</tr>
</thead>
</table>

Infra-structure: JHMIA, Radiology, TeraMedica,

- Analytic Database
- Shape and Change Tools
- Data-mining
- Decision Support

Challenges to implement across multi-disciplines:
- Data Standards
- Workflow, Procedure, and Management Differences
  - Different intervention time-scale
OncoSpace: Closed Loop Adaptive Radiation Therapy

Patient

Data – Information – Intervention – Response

Real Time Image Guided Intervention

Treatment Re-optimization; Early Treatment Assessment, ...

Population: New protocol, New dose level, New standards

OncoSpace Infrastructure
**OncoSpace:**
Radiation Oncology as the $I^4M$ test-bed

<table>
<thead>
<tr>
<th>Robotic Surgery</th>
<th>Radiation Oncology</th>
<th>Ophthalmology</th>
</tr>
</thead>
</table>

**Infra-structure:** JHMIA, Radiology, TeraMedica,

- Analytic Database
- Shape and Change Tools
- Decision Support
- Data-mining

**OncoSpace**
Extending the OncoSpace Model: Sharing Research and Clinical Care

JHU

I4M Infra-structure

Institute n

Institute 1

JHU

Genomics

Ophthalmology

Radiation Oncology
OncoSpace

• OncoSpace is a new research infra-structure based on Radiation Oncology as a “use-case” model

• Bioengineering Research Partnership
  – multi-disciplinary: Radiation Oncology, Radiology, Physics and Astronomy, Computer Science and Biostatistics
  – multi-institutional: Hopkins, clinical partner sites
  – IMPAC
Distributed Research Model

Current Trial Practice

- Data Collection
- Patient Tx
- Follow up
- Literature Search
- Journal Publication

Hypothetical Future Practice

- Data Collection
- Patient Tx
- Follow up
- Treatment Protocol
- Literature Search
- Journal Publications
- Data Analysis and Integrity Checks
- Publication of Data to DB’s

Increased potential for data reuse

OCI_50th 2008, JWW
Data Delivery in Cooperative Research: Hitting the Wall

FTP and GREP are not adequate

- You can GREP 1 MB in a second
- You can GREP 1 GB in a minute
- You can GREP 1 TB in 2 days
- You can GREP 1 PB in 3 years
- You can FTP 1 MB in 1 sec
- You can FTP 1 GB / min (~1 $/GB)
- 2 days;1K$ / 3 years and 1M$

- 50 MB local DICOM transfer takes 1 min
- 100 patients x 10 (3D) scans = 5 - 10 TB
- A factor of 10 improvement in access speed cannot offset the growth in data and complexity

- Rethink databases’ function
  - following the CS community
OncoSpace: Adapting the SkyServer Approach

SDSS is a collaborative effort to map 25% of the sky
SkyServer publishes data from the SDSS
>> 100’s of new discoveries in astrophysics
Increased scale and scope for research

Shared resources
  – Methodology
  – Software
  – Expertise
  – Experience
New opportunities
  – Analysis
  – Visualization
  – User experience

Skyserver.sdss.org

Alex Szalay PhD - JHU
Jim Gray PhD - Microsoft
OncoSpace: Adapting the SkyServer Approach

- **Active Databases**
  - There is too much data to move around, *take the analysis to the data!*
  - Do all data manipulations at database
    - *Build custom procedures and functions in the database*
  - Established Web-service for broad access
    - Query across multiple databases
- Automatic parallelism guaranteed
### Database Consideration

**Operational** vs **Analytical**

<table>
<thead>
<tr>
<th>Operational</th>
<th>Analytical</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Workflow management</td>
<td>• Decision Support</td>
</tr>
<tr>
<td>• Patient records and archival</td>
<td>• On-line Analytical Processing</td>
</tr>
</tbody>
</table>

**Data Extraction & Filtering**

- How do I organize my data? **Typically Hierarchical**
- DICOM RT
- OO principles

- How might I analyze my data? **Star Schema**
- Design to support analysis...
- Fast query

**OIS, OCIS, EPR**

- TPS, PACS

**OncoSpace**
OncoSpace: Work in progress (McNutt)

- **Technologies**
  - SQL Server 2005
  - Ruby on Rails

**OncoSpace**

- **Membership**
- **Vocabulary**
- **Data Preparation**
- **Analytical Database**

OCI_50th 2008, JWW
This image appears to be a database diagram, likely representing a schema for a medical or radiation therapy database. The diagram includes various tables and relationships, such as:

- Patient
- Lesions
- Primary Tumors
- Pathology Features
- Lab Values
- Medical Histories
- Family Histories
- Social Histories
- Radiation Summaries
- Dose Grids
- Radiotherapy Sessions
- Prescribed Drugs
- Surgical Procedures
- Clinical Events
- Outcomes

Each table and relationship is connected with lines and arrows, indicating the flow of data or entities within the system. The diagram is complex and would require a detailed understanding of the database schema to fully interpret its contents.
Hopkins OncoSpace

Clinicians

View/Analyze Data

Researchers

View/Analyze Data

Bio-Statisticians

View/Analyze Data

Tools

MS Web Services

Security

Active Database

Project 3

Project 2

Project 1

PACs

MIS

IMPAC/RTP

Labs

OCI_50th 2008, JWW
OncoSpace: 4 Projects

1. Integration of clinical workflow with data collection to populate OncoSpace.
2. Optimize database architecture for secured distributed web-access
3. Tools for query, analysis, and navigation of OncoSpace to derive information from various classes of questions
4. Bio-statistic research and development to support data mining and ensure valid decision making from the OncoSpace Systems and Nested Experiments.
Influence of Shape
OVH maps the shape of OAR to a volume-distance plane through target **expanding** and **shrinking**.

**Cumulative Overlap Volume Histogram (COVH)**

- **Volume**
- **Distance**
  - Shrinkage
  - Expansion
  - -2mm
  - -1mm
  - 0
  - 1mm
  - 2mm
Sphere: radius 7cm
Rectangle: 3.7*3.7*12.1cm
Spatial resolution: 0.1*0.1*0.1cm
Image size: 291*291*291pix
Parotid: \( V(30\text{Gy}) < 50\% \) of volume

Dose corresponding 50% of volume:

\[ 32.5\text{Gy} = T[0.4685\text{cm}, 4.411\text{cm}, 5.189\text{cm}] \]
Treatment plan evaluation (parotids)

Re-plan 22: R parotid (31Gy)

Dose corresponding 50% of volume:
- : 0—25Gy
- : 25—30Gy
- : 30—35Gy
- : 35—40Gy
- : 40—45Gy
- : 45—50Gy
- : 50—55Gy
- : >60Gy
- : >60Gy
Right parotid: 50% volume 31Gy → 24.4Gy

Re-plan: Original plan

Brainstem  Cord

Left parotid

: Re-plan

: Original plan
OncoSpace: Physician’s Tool for Personalized Medicine

- Enable natural recall of past experience’s with patients
  – display of data that match physician’s way of thinking
- Allow other physicians to share the experience
- Allow other centers to contribute and use OncoSpace to broaden the stored experience
- caBIG compliance to insure data reuse and sharing
OncoSpace: as a Collaborative Research Model

- All queries should be IRB approved
- RTOG as a service to legitimize query
  - Data is live for re-use
  - Demonstrate Patient Care Improvement
Challenges

• OncoSpace Design
• Statistical Research
• Quality Assurance
• HIPPA and Security
OncoSpace: as Physician’s Tool for Personalized Medicine

• Enable natural recall of past experience’s with patients – display of data that match physician’s way of thinking
• Allow other physicians to share the experience
• Allow other centers to contribute and use OncoSpace to broaden the stored experience
• caBIG compliance to insure data reuse and sharing
OncoSpace: as a Collaborative Research Model

- All queries should be IRB approved
- RTOG as a service to legitimize query
  - Data is live for re-use
  - Demonstrate Patient Care Improvement
Challenges

• OncoSpace Design
• Statistical Research
• Quality Assurance
• HIPPA and Security
Future Role of Physics in Radiation Therapy

• Technological focus
  – Dose escalation; proton?
  – How high can we deliver?
  – How high do we need?

• Bridging discoveries to RT
  – Room to de-escalate

• Information and Informatics
  – Improve effectiveness and efficiency of research
  – Disseminate knowledge for care
Future Role of Physics in Radiation Therapy

• Technological focus
  – Dose escalation; proton?
  – How high can we deliver?
  – How high do we need?

• Bridging discoveries to RT
  – Room to de-escalate

• Information and Informatics
  – Improve effectiveness and efficiency of research
  – Disseminate knowledge for care