Mining data to develop planning and treatment quality metrics

SEAAPM Annual Meeting
Raleigh, NC
April 23, 2015

Charles Mayo, Ph.D.
Mayo Clinic, Rochester, MN
Collaborators

Robert Foote, MD
Michael Haddock, MD
Scott Stafford, MD
Robert Miller, MD
Yolanda Garces, MD
Nadia Laack, MD
Ivy Petersen, MD
Elizabeth Yan, MD
Thomas Pisansky, MD
Kenneth Oliver, MD
Richard Choo, MD
Sameer Keole, MD
Michele Halyard, MD
Robert Mutter, MD
Brian Davis, MD, PhD
Stephanie Childs, MD
Daniel Ma, MD
Jan Sarkaria, MD
James Martenson, MD
Christopher Hallemeyer, MD
Michele Neben Wittich, MD
Lauren Baker, PMP
Mark Parry
Mike Grinnell
Keith Krupp
Monty Manley
Chris Brent
Geri Pumper
Steve Conners
Kristi Klein, RN
Nancy Heinzelman
Mirek Fatyga, Ph.D.
Jack Cusma, Ph.D.
TJ Whitaker, Ph.D.
Wan Chan, Ph.D.
Research Disclaimer:

Author has an unrelated grant with Varian Medical Systems.

Spelling Disclaimer:
Mining data to develop planning and treatment quality metrics

Most of the work
- Planting Ideas, Cultivating a data focused culture
- Enabling technologies and clinical processes

Big Data
Metrics for Plan Quality
We want to “follow the data” to make meaningful decisions on how to improve treatments to get better outcomes for our patients.

There is significant heterogeneity in treatment parameters that vary among clinics, providers, time, technology, etc. To tease out details on what can be shown to matter, large, detailed, longitudinal datasets are needed.

It should be easier to show what treatment factors correlate with outcomes.
We need to define the metrics that are going to be used to define the quality measurement. These metrics should be clinically useful and should be gettable from automated processes.
Periodic Cleveland Clinic report (~83 pages) provided to physicians to give detailed data on outcomes for all disease sites.

Example of how routine collection of data can be incorporated into clinician discussions of evidence based medicine and used as a baseline to better inform patients.

Requires a commitment to “follow the data”

Courtesy of John Suh, MD
Knowledge Based Clinical Practice Improvement System (aka KPI or Kπ)

Mayo RadOnc System to routinely gather and analyze outcomes data for all patients

Vision - Routine aggregation and analysis of data for all patients to inform practice on the effect of treatment choices on outcomes and direct improvements

Outcomes DBs for other specialties e.g. ENT or Breast Surgery

Rad Onc Outcomes DB

Automated outcomes DB pattern analysis applications (e.g. Data Mole)

Automated Reporting Applications

Multi-Institutional Registries (e.g. ASTRO NROR, RadioGenomics Consortium)

Treatment and Follow-up data (Plan and Prescription Parameters, DVH, Survival, PRO, QOL, Toxicity, etc.) for all patients and all modalities (Photons, Protons, Brachytherapy)

Specify optimal treatment planning parameters

Physician evaluation of treatment options

The basis of knowledge is information
Data Dictionary of KPI Dev Data Elements as of 10/10/2014

Enhanced Demographics: Name, ClinicID, DOB, DOD (i.e. survival), County, State, Country, Postal Code, Race, Ethnicity, Religion, Marital Status, Gender, E-Mail

Diagnosis and Staging: Date of Entry, Basis, T,N,M + G,H,N,P,R,S, etc, OverallStage, StagingSystem, Laterality, ICD9 (ICD10), ICD0, Ranking, Primary Site, DistantMets, Recurrence, ICD9 of Primary, Diagnosis Date, Diagnosis Method

Toxicity: Date, Grading System, Grade, Cause, Certainty

Patient Reported Outcomes: Date of PRO, Templates, Questions, Answers

Labs (Current): Height, Weight, BMI, Neutrophils, Platelet Count, Lymphocytes, Hemoglobin, Leukocytes, PSA (~3,000,000 rows of data ≥ 1/1/2011)

Treatment Course: CourseID, Course Start Date (based on treatment records), Course End Date (based on treatment records)

Treatment Rx: Number of course fractions, number of treatments per course fraction, dose to each target volume (i.e. all the data in the Rx in Planning Templates)

Treatment Delivery Details for Each Plan: Facility, Machine, NFractions Treated, NFractions Planned, Total Dose Delivered, Total Dose Planned, Number of Beams, Plan Name, Plan DicomUID, TotalMU, TotalBeamOnTime, TotalTreatmentDeliveryTime, TotalTreatmentSessionTime, IsProton, IsBrachy, IsSBRT, IsBreathHold, UsedStaticIMRT, UsedHybridIMRT, UsedVMATIMRT, UsedHybridVMAT, UsedWedges, UsedNonCoplanarBeams, UsedHalfBeamX, UsedHalfBeamY, UsedIGRT, UsedCBCT, UsedX06, UsedX06FFF, UsedX10, UsedX10FFF, UsedX18, UsedX18FFF, UsedE06, UsedE09, UsedE12, UsedE16, UsedE20, CouchVrt_Mean, CouchVrt_Stdev, CouchLng_Mean, CouchLng_Stdev, CouchLat_Mean, CouchLat_Stdev, CouchRotation_Mean, CouchRotation_Stdev, CouchPitch_Mean, CouchPitch_Stdev, CouchRoll_Mean, CouchRoll_Stdev, (Other Brachy and proton specific items to be added)

Treatment DVH Curves: Structure, Volume[cc], Max[Gy], Min[Gy], Mean[Gy], Median[Gy], Stdev[Gy], DVH curve (Percent Volume vs Absolute Dose as point pairs)

Treatment DVH Metrics: Values stored for course composite plans (e.g. 1st course + Boost), as in planning templates. This enables rapid identification of patient groups according to metrics that are most relevant to disease site groups, (i.e. find all esophagus patients with lung_total:V20Gy[%] > 20) The treatment DVH curves can be used to pull other values, but the searches based on DVH curves are slower than searches on DVH metrics.

Treatment Details Specific to Disease Site: e.g. Breast target details questionnaire, CU Androgen questionnaire, Head and Neck HPV status

Recurrence Status: HadLocalRelapse, Date of Local Relapse, HadRegionalLymphNodeRecurrence, Date of Regional LymphNodeRecurrence, HadDistant Relapse, Date of Distant Relapse, Site of Distant Relapse, Cause of Death (Cancer-Local, Cancer-Regional, Cancer-Distant, Treatment, Other, Unknown)
How will our clinic be able to gather “Big Data”?

- Technology is a much smaller step than culture changes needed for implementation: consensus (inter and intra institutional), process, changes in work duties, QA

- Can do a lot with existing treatment planning and radiation oncology information systems

- Think through what data elements you want /need in the long run, how they are related and then develop a strategy of small, manageable steps.
How to get there?

**Technology**
- Software/database systems for aggregating information
- Software systems for analytics
- Integration with other systems

**Culture**
- Need to shift thinking about data related to treating our patients.
- Thinking about the data not just for treatment of the patient before us, but for systematic aggregation to help all the patients yet to come.
- Implication is accepting limitations in options, standardizations
- Potentially more work to quantify data – “free text” is hard to use

Baby Steps – a lot of them

To move a group you have to help them believe in the vision.  
As you create working examples that show it is real and doable, then they will lead the way.

Pick working examples that can positively impact work flow in clinic and add value to current practice

Identify and tackle the “enabling” steps one by one. This positions you to grow your effort.

A few options here
- DIY – Use in house staff with expertise or train
- Use consultants to help build
- Purchase from current vendor (ROIS, TPS)
- Purchase from 3rd party vendor

This… only you can do

Assume you have the technology, what do you have to change about your practice to enable the technology to get the data?
- Consensus in your practice
- Standardize practice
- Change who does what
Standardization and nomenclatures are needed to enable the automation needed to handle Big Data

“A foolish consistency is the hobgoblin of little minds, adored by little statesmen and philosophers and divines.”

Ralph Waldo Emerson - Self Reliance

“Mind? I have no mind. I’m a computer.”

and database programmers

A SQL query walks up to two tables in a restaurant and asks: “Mind if I join you?”

Manual effort is the enemy, free text is its cousin

Cost, FTE, Inconsistency, Speed
Standardize Diagnosis and Staging

Assume you are going to use standard codes (ICD-9/10, ICD-0, TNM+, histology, etc) to filter your searches to find the patients of interest to you.

- What information is input for every patient?
- How do you handle metastatic sites? e.g. is metastatic prostate 198.5 with a secondary of 185?
- Is the diagnosis and staging linked to the course for easy, computer lookup?
- Are you inputting courses and diagnosis from outside institutions so that you have a complete record?
- Have you put in place a QA process/peer review to be sure you are not going to have doubts about the data when you look back?

This... only you can do

Assume you have the technology, what do you have to change about your practice to enable the technology to get the data?

- Consensus in your practice
- Standardize practice
- Change who does what
Once you standardize diagnosis and staging, you are in position to query your existing ROIS database to get a wealth of information on your practice. The tech is relatively easy.

Once you standardize your inputs, you can pull large data sets from your existing ROIS system to answer many clinical questions:

- Find me all the patients treated for metastatic prostate cancer with SBRT at facility X in the last year.
- Find me all the breast patients treated with IMRT in the last 3 years for whom we used a couch kick.
- Find me all the left sided stage IA lung cancer patients for whom we used an 18MV beam.
- Calculate statistics on the range of couch variation for each of our treatment sites so that we can use it to set our couch tolerances.
- Calculate the last year’s work load for machine X so that we can use it in the shielding calculation for our replacement machine.
- Find me all the breast patients we treated with breath hold.
- Find me all the SBRT liver plans that I treated last year using VMAT.
How can big data fit into making our patient’s more safe?

- Insert a statistical layer for consistency check of parameters with historical probability distributions for the parameters.
- Different doesn’t mean wrong (not different doesn’t mean right) but it does highlight attention for a closer look.

http://patientsafetyed.duhs.duke.edu/module_e/swiss_cheese.html
Using information from treatment records to define “expected” probability distributions

Retrospective statistics could be used in an automated plan check program to highlight sections for special attention

**How many MU’s is unusual?**

<table>
<thead>
<tr>
<th>Tissue Type</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBRT Liver VMAT</td>
<td>24</td>
</tr>
<tr>
<td>SBRT Spine VMAT</td>
<td>203</td>
</tr>
<tr>
<td>SBRT Lung VMAT</td>
<td>34</td>
</tr>
<tr>
<td>SBRT Lung FfF</td>
<td>30</td>
</tr>
<tr>
<td>Head and Neck VMAT</td>
<td>122</td>
</tr>
<tr>
<td>Esophagus Hybrid VMAT</td>
<td>15</td>
</tr>
<tr>
<td>Esophagus VMAT</td>
<td>36</td>
</tr>
<tr>
<td>Esophagus FfF</td>
<td>149</td>
</tr>
<tr>
<td>Breast Hybrid VMAT</td>
<td>34</td>
</tr>
<tr>
<td>Breast VMAT</td>
<td>46</td>
</tr>
<tr>
<td>Breast FfF</td>
<td>1136</td>
</tr>
<tr>
<td>Lung Hybrid VMAT</td>
<td>22</td>
</tr>
<tr>
<td>Lung VMAT</td>
<td>191</td>
</tr>
<tr>
<td>Lung IMRT</td>
<td>2</td>
</tr>
<tr>
<td>Lung FfF</td>
<td>360</td>
</tr>
<tr>
<td>Prostate VMAT</td>
<td>501</td>
</tr>
<tr>
<td>Prostate IMRT</td>
<td>39</td>
</tr>
</tbody>
</table>

**Can we predict couch height by patient weight?**

**How should we set our tolerance tables base on our experience?**
Build consensus with physician disease site groups define standard DVH metrics and objectives to use for all patient treatment plans ~ 18 months

• Supports physician lead initiative to develop and define standards of practice for treatment plans.

• Replace free text word documents with standardized tabular templates

• Critical point in dialog for building consensus is distinction between agreement on what metrics we measure vs. the constraint value and priority

  Agree on what to measure for all

  lung_total  V20Gy[%] < 25%  Priority = 1

  Enable per patient change from default of constraint/priority

• While defining vanilla (standard), must take an approach that allows for chocolate (per patient changes)
Standardize structure and DVH Nomenclatures along with Rx and DVH metrics measured

Normal tissue naming schema is left to right: general to specific with laterality at the end. Character string length, use of capitals, spaces, etc are guided by vended systems used in the clinic (simulator, planning system, information system, etc) constrain format

For targets (PTV, CTV, GTV, ITV) take an approach that allows a standard name plus an alias in the database e.g. ptv_high = PTV7200

Using both a standard name and an alias, means when pulling data from the database we can identify the volume getting the highest dose for any plan or treatment site (ptv_high) independent of the specific name used in the plan (ptv6300).
Define a DVH nomenclature schema that fully defines all parts of the curve and can be expanded upon to accommodate other DVH derived metrics as they evolve.

endpoint name(calculation parameters)[output units]

For points on DVH curve, the nomenclature
- accommodates all combinations of relative and absolute, dose and volume
- defines units of output result value
- distinguishes between high and low dose fractions of the structure volume
- works with regular expression operators for automated data processing

Example of use for radiobiological metrics: V35EQ2Gy(4)[%]
Several groups are coordinating efforts to address nomenclature for radiation oncology.

NRG Oncology

BRIEF REPORT AND OPINION

Radiation Therapy Digital Data Submission Process for National Clinical Trials Network

Jialu Yu, PhD,* William Straube, MS,† Charles Mayo, PhD,‡ Tawfik Gladduhi, PhD,* Walter Bosch, DSc,∥ Kenneth Ulin, PhD,* Stephen F. Kry, PhD,* James Galvin, DSc,* and Ying Xiao, PhD*

*IROC/NRG/Jefferson Medical College, Philadelphia, Pennsylvania; †IROC St. Louis, Department of Radiation Oncology, Washington University School of Medicine, St. Louis, Missouri; ‡Department of Radiation Oncology, Mayo Clinic, Rochester, Minnesota; ¶IROC Rhode Island, Department of Radiation Oncology, Rhode Island Hospital, Providence, Rhode Island; and §§IROC Houston, Department of Radiation Physics, University of Texas MD Anderson Cancer Center, Houston, Texas

Received May 6, 2014, and in revised form May 28, 2014. Accepted for publication May 30, 2014.

AAPM Task Group No. 263 - Standardizing Nomenclature for Radiation Therapy

Members represent multiplicity of stakeholders – institutions, vendors, national regions and international, academic/non-academic, physicians, physics, AAPM/ASTRO

Left Optic Nerve[12]: Lt Optic Nerve, OPTIC Nerve, OPTIC_Nerve, OPTIC nerve, OPTICNERV, OPTICNERV_L, OptICnerve_L, LOPIC, OpticNerve_L, Left Optic Nerve

Left Lung[12]: Lt Lung, Lung L, LUNG L, Lung L, LLUNG, Llung

Both Lungs[12]: Lungs, LUNGS, LUNG TOTAL, Lung total, combined Lung,肺癌, LUNG, LUNGs, Lung Bilat, Lung Both

8th cranial nerve[7]: CN VIII(5), CN_vIII(2)

Right External Iliac Artery[2]: A Iliac R, a Iliac e_r
What do you do when your nomenclature differs from the nomenclature for a TRIAD submission of DICOM files?

Write a script using Evil Dicom (thanks Rex Cardan, UAB).
Iterative Process

Building consensus on the IT design and function.

**Free text Word**
- Physician driven

**Standardized formatted Word**
- Physician + Physicist driven

**Stand alone application that demonstrates automation and software driven templates**
- Physicist + Physician driven

**Production application that uses database**
- IT driven with multidisciplinary committee: physicians, dosimetrists, therapists, physicists
Application becomes our standard prescription. Also serves as documentation tool for image setup, notes, IMRT justification, etc.

Physician groups define consensus for DVH metrics for all treatment sites; what to measure and default values for constraints and prioritizations.
Now generate a report as part of routine care that compares desired vs achieved DVH metrics for each patient. Use this in plan check to highlight areas for special attenuation.

Save DVH metrics in database to mine results later

We constructed this system at a time when the vended system was very limited. Now more built in and scripting.

Vended systems (ROIS/TPS) are maturing rapidly to enable standardization of nomenclature, prescriptions and reporting

- Built in modules
- Scripting capabilities

You likely can use existing tools in your system to aggregate DVH metrics or use scripting APIs to create them.

The most important step is to standardize on what to measure

Then you are in position to begin learning from the statistics on your own experience
Using the data to improve our practice – gains for research and quality improvement

Sets the stage for constructing plan check software that uses recent retrospective data on distributions of values for DVH metrics, for highlighting values for a new plan that should get extra attention.
Pool data among institutions to define what is normal

Explore variations in treatment techniques and effects on DVH parameters

Together these efforts position us to evaluate individual plans in the context of the history of previous plans.
Median, box 50% CI, whiskers 95% CI
Median, box 50% CI, whiskers 95% CI
Considering the pooled data prompted, ideas about how to improve.

Systematically gathering the data enabled demonstrating the improvement.
### Completing the loop

Use data on what was achieved in DVH metrics as basis to set new constraints to use as defaults for future plans in the planning templates.

<table>
<thead>
<tr>
<th>Suggested Mean Doses (range)</th>
<th>Mean heart dose (Gy)</th>
<th>Ipsilateral lung V20(%)</th>
<th>Total Lung V20 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tangential chest wall/breast</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Right</strong></td>
<td>4 → 0.5 (0.2 – 1.6)</td>
<td>15 → 10.4 (2.7 – 18.6)</td>
<td>10 → 5.8 (1.6 – 11.1)</td>
</tr>
<tr>
<td><strong>Left</strong></td>
<td>4 → 1.1 (0.3 – 5.1)</td>
<td>15 → 8.9 (0 – 19.2)</td>
<td>10 → 4.4 (0.2 – 9.4)</td>
</tr>
<tr>
<td><strong>Chest wall/breast + SCV</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Right</strong></td>
<td>4 → 1.9 (0.5 – 8.0)</td>
<td>25 → 24.0 (14.3 – 36.3)</td>
<td>10 → 15.0 (7.9 – 28.3)</td>
</tr>
<tr>
<td><strong>Left</strong></td>
<td>4 → 3.2 (0.7 – 9.1)</td>
<td>25 → 23.9 (9.9 – 36.7)</td>
<td>10 → 10.9 (4.7 – 15.9)</td>
</tr>
<tr>
<td><strong>Chest wall/breast + SCV + IM</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Right</strong></td>
<td>4 → 2.1 (0.5 – 5.4)</td>
<td>25 → 32.1 (9.8 – 44.2)</td>
<td>10 → 18.2 (5.7 – 29.3)</td>
</tr>
<tr>
<td><strong>Left</strong></td>
<td>4 → 4.0 (1.4 – 9.2)</td>
<td>25 → 27.7 (8.9 – 43.4)</td>
<td>10 → 13.1 (7.7 – 20.2)</td>
</tr>
</tbody>
</table>
With routine collection of DVH metrics comes functionality for data pooling and one element for meaningful plan quality metrics that inform practice and address Affordable Care Act.
Using benchmarks to compare different practices
Example Rectum:V65Gy[%] for 4 groups
What about correlations of DVH metrics to toxicities?
You can use your existing radiation oncology outcomes information system to query to pull toxicities.

Remember to get the connection to diagnosis and staging right.
Diarrhea: Black: Grade 0, Red: Grade 1
When is no data, data and when is it just no data?

Another iteration on changing culture to think about treatment records as like a scientist as well as like a clinician.
Urinary Frequency: Black-Grade 0, Purple- Grade 1, Red-Grade 2
Urinary Frequency: Black-Grade 0, Purple-Grade 1, Red-Grade 2
Summary

Standardize
• Make the input data consistent so that computer systems can automatically extract and reliably process it.
• Build in QA processes on your data so that you will believe it

Extract
• Use the capabilities of your current electronic systems
• The exercise of pulling large data sets from your existing ROIS and TPS systems will improve your understanding of connections and needed consistencies

Extend
• Train or get outside help if you need it
• Coordinate with other groups interested in data pooling to strengthen your processes and put the data to use
• This will work best if efforts are coordinated among institutions

Demonstrate
• Show use of data from your electronic systems to define your practice norms and demonstrate improvement
• Be prepared to iterate. Changing processes and changing minds takes sustained effort

*Changing culture to think as a scientist as well as a clinician about data usage will require more effort than constructing the technology to use it.*