Big Data : New and Emerging Big Data Strategies in Oncology

Bringing Value from Big Data Analytics into Clinical Practice

Charles Mayo, Ph.D.
University of Michigan

AAPM Annual Meeting
TU-B-605-08:50-9:10
Room:605

MICHIGAN MEDICINE

Disclosure

Work is supported in part by a grant from Varian Medical Systems

MICHIGAN MEDICINE

It takes a village to make a Big Data Analytics Resource System

Randy Ten Haken, PhD
Marc Kessler, PhD
Dan McShan, PhD
Issam El Naqa, PhD
Jean Moran, PhD
Martha Matuszak, PhD
Scott Hadley,PhD
James Balter, PhD
Yue Cao, PhD
Dale Litzenberg, PhD

Avi Eisbruch, MD
Jim Hayman, MD
Shruti Jolly, MD
Dawn Owen, MD
Reshma Jagsi, MD
Michelle Mierzwa, MD
Ted Lawrence, MD PhD

Grant Weyburn
Carlos Anderson, PhD
Xioaping Chen
John Yao, PhD
Lynn Holevinski

Sue Merkel
Sherry Machnak
Denise Goodman
Dawn Johnson
Latifa Bazzi
Grace Sun
Alex Hayes
Parth Janni

MICHIGAN MEDICINE

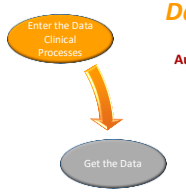
Data – Value Cycle

Manual

Extract, Transform, Load

Minimal upstream coordination required

Limited to relatively to small numbers of patients (10s-100s)



Data – Value Cycle

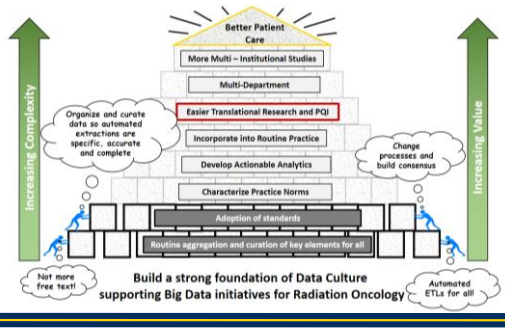
Automated Electronic

Extract, Transform, Load

Large numbers of patients (>> 1000s)

Lots of upstream coordination required

- access
- standardization
- people
- resources
- technical skills



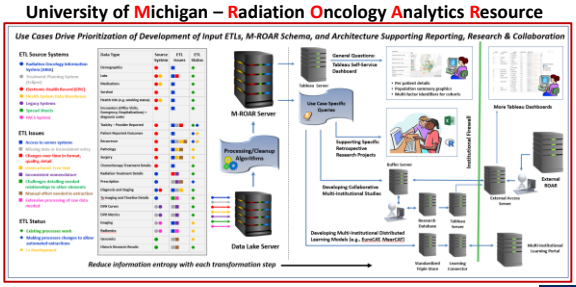
So you want to build a Big Data Analytics Resource System ?

Most of your effort is going to be in building and improving ETL processes

- It takes a multi-disciplinary community that wants to make it real
- Invest in people with diverse skill sets
- Need commitment from leadership

MICHIGAN MEDICINE

University of Michigan – Radiation Oncology Analytics Resource



MICHIGAN MEDICINE

Be proactive on the Ethics of Access

- Acknowledge that data are people and can do harm
- Recognize that privacy is more than a binary value
- Develop a code of conduct for your organization, research community, or industry

Zook, Barocas, Pasquale, et al. Ten simple rules for responsible big data research. PLOS Computational Biology | https://doi.org/10.1371/journal.pcbi.1005399 March 30, 2017

MICHIGAN MEDICINE

Culture Shift: Standardize entry of Diagnosis and Staging -> Volume, Value

Data in the Electronic Medical Record

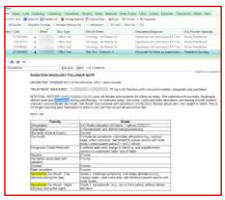
Encounter Notes in EMR (EPIC) → M-ROAR

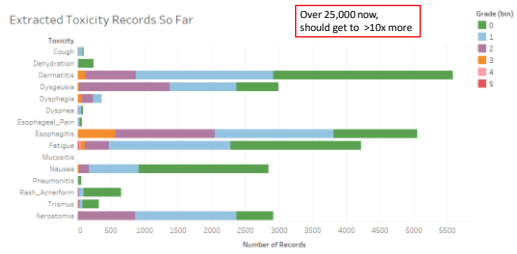
- Huge **volume** of text data available
- M-ROAR access (ie **velocity**) is fast (seconds)
- Potentially really **valuable** source of information
- The problem is **variability** …
  the solution is standardization

EMR Access + standardization -> Volume, Value

- *Automate harvesting regular data entered into notes in EMR*
- *Presentation standardizations improve harvest-ability*

MiChart (EPIC) → ETL → M-ROAR





We can do even better…

Culture Shift: Build templates in EMR with standardized schema for **key data** elements

- ***Standardized schema*** *designed to function with regular expressions*
- *Physician selection from drop down lists of standard values - >* ***Fast, Easy, Accurate***

Disease Site Information

|> Treatment Area= Primary <|
|> Status= No Evidence of Disease <|

|> KPS = 80 (range 0 - 100) <|

Toxicities

|> Dyspnea=0 (none) <|
|> Cough=1 (symptomatic, non-narcotic medication only indicated) | Attribution=possibly related <|

|> Dysphagia=0 (none) | Attribution=none <|
|> Chest Wall Pain=0 (none) <|
|> Fatigue=1 (mildfatigue over baseline) | Attribution=possibly related <|

Requests for data tend to fall into three categories

Type of query	Typical value
Practice Quality Improvement (PQI)	Evidenced based approach to improving clinical processes and patient care
Translational Research	Provide data sets needed for publications and grants
Administrative Support	Ease access to data needed by front office e.g. Certificate of Need, Regulatory Groups, Institutional evaluation

Look at practical examples of using this resource to provide value

…. and what we've learned along the way

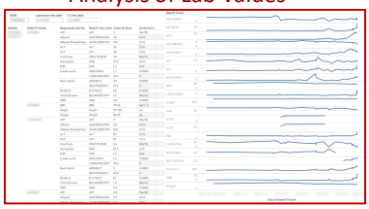
Tableau Dashboards providing end user self-service

SRS and SBRT Utilization Analytics

How can I look at how our SRS/SBRT program is evolving ?

Value Categories
- Administrative Support
- Practice Quality Improvement

Analysis of Lab Values



How can I look at trends in labs for a patient or look at labs for a set of patients?

EMR -> MROAR
- Batch processing possible
- Reduce work
- Integrate with treatment data

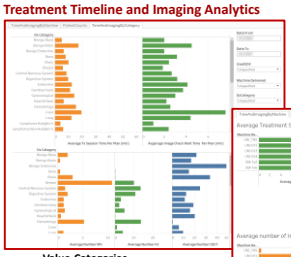
Value Categories
- Practice Quality Improvement
- Translational Research

Patient Cohort Identification

How can I find a list of patients treated in a particular way?

Value Categories
- Practice Quality Improvement
- Translational Research
- Administrative Support

Treatment Timeline and Imaging Analytics



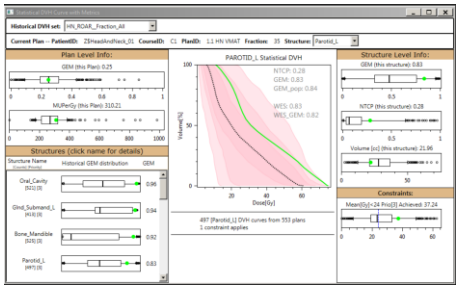
How long does it take to treat our patients and what imaging do we use?

Value Categories
- Practice Quality Improvement
- Translational Research
- Administrative Support

Can we use our historical DVH data when we are examining new treatment plans?

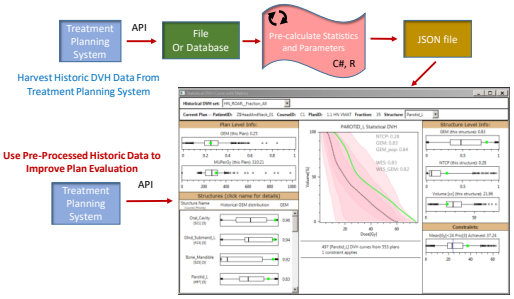
- Statistical DVH Dashboard
- Disease Site DVH Metric Summaries
- Practical Statistical Metrics
  - Generalized Evaluation Metric
  - Weighted Experience Score
  - Difficulty Ranking Score
  - Experienced based priorities

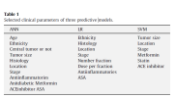
Statistical DVH Dashboard – *Plan Summary Panel*

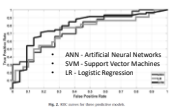


Application is run from treatment planning system.

Uses visual and statistical metrics to compare this plan to historical plans

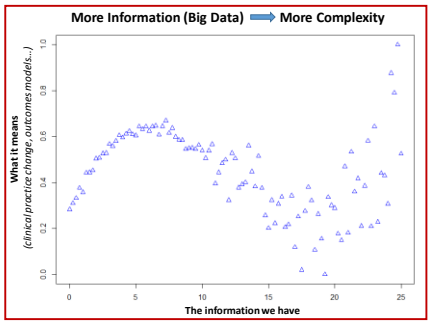






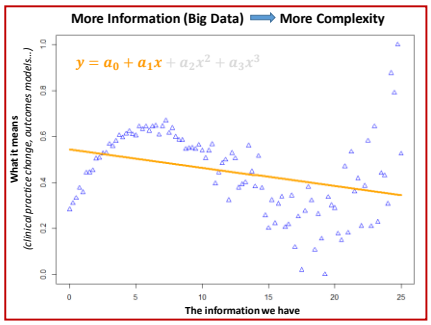
Support for Machine Learning

- ML algorithms are data hungry: models and validation
- Need realistic representations of clinical distributions move from 10's to 1000's of patients
- Foundation for resolving differences between models



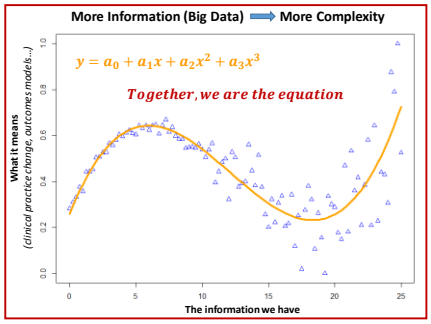
Community Science

- Variability in data and processes undermines reaching the real potential of Big Data



Community Science

- Variability in data and processes undermines reaching the real potential of Big Data
- No one institution can be the solution for these issues



Community Science

- Variability in data and processes undermines reaching the real potential of Big Data
- No one institution can be the solution for these issues
- Viable solutions from community of individuals solving issues in their institutions then collaborating on shared solutions

https://www.flickr.com/photos/86369280@N00/albums/72157684027833246

MICHIGAN MEDICINE

Summary

Analytics from Big Data fit readily into Clinical practice supporting
- Translational Research
- Practice Quality Improvement
- Administrative Support

Effort needed to build a Data Culture
- Clinical practice changes
- Support for access, extraction and curation

Community Science
- Development and publication of common standards
- Multi-institutional data sets

MICHIGAN MEDICINE