Development of an In-House CT Dose Monitoring and Management System Using Open-Source Resources – Pearls and Pitfalls

Da Zhang1,2
Larry Barbaras1
Matthew R. Palmer1,2

1: Beth Israel Deaconess Medical Center, Radiology
2: Harvard Medical School

Content of This Talk

- Motivation and design of an in-house CT dose monitoring system
- Steps of implementing such a project
- Pitfalls we went through and lessons we learned

Background: Dose Concerns and Regulatory Requirements

- The rise of concerns of medically induced radiation
  - CT: #1 dose contributor
  - Waves of articles in lay media
- Requirements from ACR accreditation (effective Dec/2013)
- Requirements from the JC (effective Jul/2015)
JC Requirements

- In a brief summary, hospitals need to
  - Review protocols periodically and keep protocols current with input from interpreting radiologist, medical physicist and lead imaging technologist.
  - Benchmark dose levels with external references.
  - Set up dose thresholds specific to individual exam types.

- These are non-trivial tasks!

Challenges: Highly Fragmented Data

- An extreme example from a workhorse GE CT750
  - 18 mo. of data, 16,687 exams
  - 708 protocols in exam records
  - On average 23.4 exams/protocol
  - A disaster to manually analyze the data

- Protocols fine-tuned & individualized for
  - Advanced features, e.g., dual energy CT, MAR
  - Patient size/age
  - Clinical indication: baseline vs. follow up

- Mixture of old and new data in the exam records
- Fleet of scanners from different vendors and models

Challenges: The Clash of Names

- Synonyms, abbreviations, and typos exist in protocol names, e.g.,
  - Abdomen/Pelvis, ABP, Abd/Pel, Abd-Pel
  - Cancer Follow Up, CA FU, CA-FU, CAPU, CA FU
  - Above 300 lbs, > 300 lbs, 300+ lbs, above 300
  - Without contrast, W, C, NON-CON, W/O
  - Thorax vs. Chest, etc.
Opportunities

- Great opportunity to solve these problems and to make innovations
- Radiation Dose Structured Report (RDSR) became widely available, thanks to the MITA XR29 initiative

Motivation and the Planning Phase

- Motivation
  - Geeks with enthusiasm in informatics and desire to demonstrate value
  - High level of desired flexibility of the system
  - Supportive department
  - Open source software resources!
- Top level design: two subsystems
  - A lightweight dose information collection system with simple user interface
  - A flexible and evolving data analysis framework for dose tracking and protocol management
- Figure out what data to collect from PACS and other hospital IT systems

A Good Wealth of Dose Info from PACS

- Four possible sources of CT dose info from PACS
  - Radiation Dose Structured Report (RDSR)
    - Easy to parse
    - Contain info of the entire exam and of individual scan series
  - Dose summary images
  - Scout images
  - Axial images
- Small-footprint data collection
  - RDSR + scout images + dose summary images
  - Several megabytes per exam
CT Dose Info in RDSR from Multiple Vendor/Models

|-------|-------------------|-----------------|-----------------|----------------------|--------------|---------------|-----------------|-----------------|------------------|-------------------|-------------------|-------------------|------------------|------------------|------------------|------------------|------------------|------------------|

CT Dose Info Not Globally Available

Steps of Implementation: Data Collection

- **Planning**
- Implementation of major sub-systems
  - Data collection
  - Data mining
- Exam level vs. Event level
  - Event level: about individual scans
    - E.g., kVp, mAs, per-series CTDIvol
  - Exam level: about the entire exam
    - E.g., total DLP, total mAs, etc.
Data Collection: Processing of RDSR

Conquest DICOM Server

xml

Powershell Script

Parse XML

Dose data fields

Parse text output & populate database fields

HTML-Based User Interface of the Data Collection System

Dose string automatically popped into dictation reports.

Need for Automated Data Selection

- With the wealth of data (5000+ exams per month at my institute), how to smartly select the right data to answer various questions is the key.
Data Mining: Cleansing, Validation, and Classification

- Data cleansing and validation against
  - Non-patients CT scans
  - Duplicated records

- Parse and normalize protocol names
  - To solve the "clash of names"
  - To build classifiers for the dose data
  - To group data for presentation

Candidates of Standardized Imaging Procedure Names

- **Scan protocol names**
  - The protocols that techs can choose on scanners
  - The targets for the protocol review
  - E.g., C-Chest, AAA, LYMPH/GEN ABD/PEL

- **Study description**
  - Orderables or billing code names
  - Pulled by CT from RIS/ordering system
  - Available on all scanners
  - E.g., CT CHEST W/CONTRAST, CT 3D RENDERING W/POST PROCESSING

- ACR-DIR allows both as local exam names to be mapped to standardized exam names (RadLex Playbook or ACR Common)

Data Mining: Protocol Name Normalization

- A protocol name usually contains
  - The "core protocol name", i.e., text representing the essential meaning of the protocol
  - Many descriptive phrases
    - Some can be removed without losing essential information
    - E.g., revision date/time
  - You can decide what other parts to be included in the normalized names
Protocol Decomposition

- Protocol names can be decomposed automatically:

Protocol Decomposing Results

- Result: each long protocol name is decomposed into the core part and various descriptors

Key Points of Implementing the Data Mining Framework

- Make use of established building blocks.
  - Find/extend the right wheels
  - Re-invent the wheels

- Selected building blocks in Python toolchain
  - NumPy: numerical calculation and array support
  - Matplotlib: data presentation
  - Python's built-in regular expression module
  - Pandas: data handling, aggregation, and selection
Key Points of Implementing the Data Mining Framework

- Modular design (OOP) of data selection and presentation
- Script-driven
  - Explore data interactively
  - Prepare for data presentation in batch-processing
- Build flexible data selection criteria using regular expressions for including and excluding desired patterns
  - (inclusion_regex, exclusion_regex), (INC, EXC), (INC, EXC), ...

Some Results: Interactive Exploration of Dose Data

- Example: tentative search for “Chest Pain”
  - Results include protocols covering different body parts
  - I want to exclude the ABP exams in this search

Some Results: Interactive Exploration of Dose Data

- A refined search of “chest pain”
  - Excluded all “ABP” exam types
  - Note the script-driven query and processing
    - Easy to run in batch mode
Graphical Exploration of Dose Data: Histogram

- Histogram: checking consistency of dose behavior across multiple CTs
  - Messy with too many scanners’ data
  - Hint of protocol differences

- Similar distribution observed from 3 GE 750 scanners
- Size-specific protocols: 3 BMI groups, 3 peaks

- Histogram also helps avoid errors in protocol classification
  - Left figure: “CT/Head W/O CONTRAST” and “ROUTINE HEAD” were grouped together
  - “ROUTINE HEAD” turned out to be a two-phase scan; it should be grouped with W/WO
**Graphical Exploration of Dose Data: Scatter Plot**

- Scatter plot of DLP vs. CTDIV gives sense of total scanned length (and repeated scans)
  - Approximately, DLP/CTDIV ~ scan length

![Scatter plot of DLP vs. CTDIV]

**Graphical Exploration of Dose Data: Scatter Plot**

- Scatter plot of scan events and "irradiation" events gives sense of how many radiation events are made
  - E.g., most exams of Non-Gated Chest Pain
    - 1 true CT scan
    - 4-5 total irradiation events
      - CT scan
      - Scout views
      - Monitoring phases
  - Note: marker size ∝ # of occurrence

![Scatter plot of scan events vs. irradiation events]

**Graphical Exploration of Dose Data: Box Plot**

- Boxplots is very useful for comparison across CT scanners
  - Outliers marked when falling < 5th or > 95th percentile
  - Compare against ACR DIR

![Boxplot comparison of DLP vs. CTDIV]
Graphical Exploration of Dose Data: Box Plot

- Boxplots can also be used to show the changes over time
  - ABD/PEL exam from one scanner
  - Combined result from 8 protocols

Graphical Exploration of Dose Data: Run Chart

- Run-charts can also be used to show the changes over time, across scanners
  - ABD/PEL exam from 3 scanner
  - Combined result from 12 protocols

Graphical Exploration of Dose Data: Run Chart

- When combining data from many scanners
  - High level summary
  - 9 scanners, 46 ABP protocols
  - Work with lead tech to verify grouping
  - Very busy figure
Use Run Chart to Show the Effect of a CT QA Project

Graphical Exploration of Dose Data: Run Chart of Volume

- When combining all protocols from many scanners together
  - Trending of volume over time
  - 9 scanners, 642 protocols

Graphical Exploration of Dose Data: Pie Chart

- Pie charts for evaluating the complexity of the protocol space
From Pie Chart to Cumulative Percentage

- Examine how many major protocols can cover 85% and 95% of total volume

Graphical Exploration of Dose Data: Pie Chart

- Pie charts also useful to show the coverage of a protocol review session
  - CTQA review for a satellite facility, percentage of coverage: 86%

Caveats: SSDE Can Automatically Rule Out Some Outliers

- Head CT with extremely low CTDIv (body phantom CTDIv value reported in head exam)
Conclusions/Comments

- With freely available software and some local expertise, a highly flexible and usable dose management system can be configured.

- There are non-trivial challenges in terms of data fragmentation, non-standard lexicon, and inconsistencies in the adoption of RDSR capabilities across vendors and platforms.

- The data-rich review process can be very helpful for CT dose and protocol optimization.

Thank you for your attention!