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

Da Zhang[1,2]
Larry Barbaras[1]
Matthew R. Palmer[1,2]

Beth Israel Deaconess Medical Center | Harvard Medical School Teaching Hospital

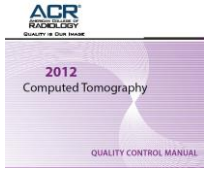
1: Beth Israel Deaconess Medical Center, Radiology
2: Harvard Medical School

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/2
- Requirements from the JC (effective Jul/2015)



http://time.com/3601047/medical-scans-hidden-dangers/
http://www.forbes.com/sites/jamesconca/2014/12/03/do-not-fear-the-radiation-in-medical-scans/

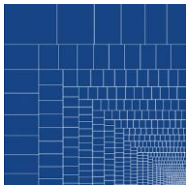
JC Requirements

- 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.
  - Bench mark dose levels with external references.
  - Set up dose thresholds specific to individual exam types.
- These are non-trivial tasks!

4

Challenges: Highly Fragmented Data

- An extreme example from a workhorse GE CT750
  - 18 mo. of data, 16587 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



Beth Israel Deaconess Medical Center | Harvard Medical School Teaching Hospital

5

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, CAFU, CA F/U
  - Above 300 lbs, > 300 lbs, 300+ lbs, above 300
  - Without contrast, I-, C-, NON-CON, W/O
  - Thorax vs. Chest, etc.

Beth Israel Deaconess Medical Center | Harvard Medical School Teaching Hospital

6

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

Beth Israel Deaconess Medical Center | Harvard Medical School Teaching Hospital

7

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 light-weight 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

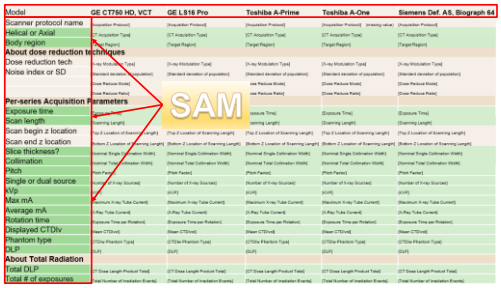
8

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

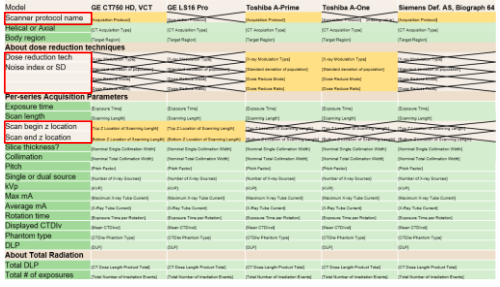
9

CT Dose Info in RDSR from Multiple Vendor/Models



10

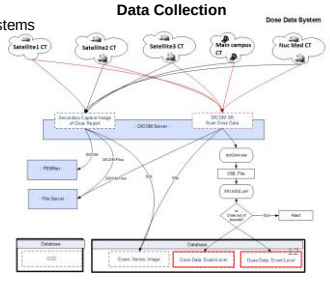
CT Dose Info Not Globally Available



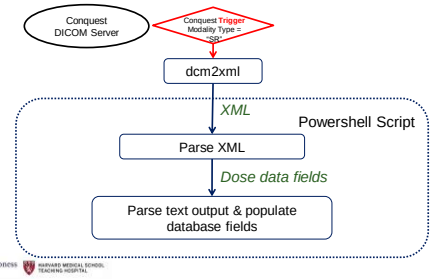
11

Steps of Implementation: Data Collection

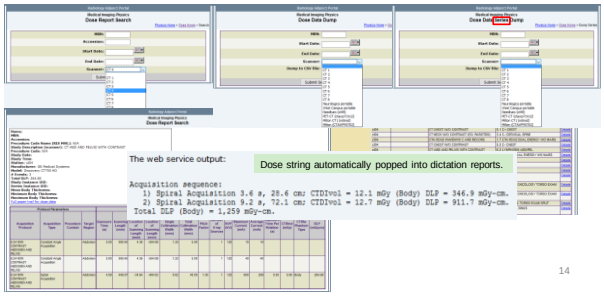
- Planning
- Implementation of major sub-systems
  - Data collection
  - Data mining
- Exam level vs. Event level
  - Event level: about individual scans
    - E.g., kVp, mAs, per-series CTDIv
  - Exam level: about the entire exam
    - E.g., total DLP, total mAs, etc.



Data Collection: Processing of RDSR



HTML-Based User Interface of the Data Collection System

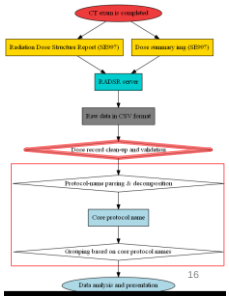


Need for Automated Data Selection

- 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



16

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)

SAM

17

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

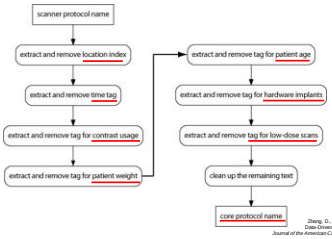
Beth Israel Deaconess Medical Center | HARVARD MEDICAL SCHOOL TEACHING HOSPITAL

18

Zhang, D., Savage, C.A., Li, X. and Liu, B., 2015. Data-Driven CT Protocol Review and Management. Journal of the American College of Radiology, 12(3), pp.267-272.

Protocol Decomposition

- Protocol names can be decomposed automatically:



Zheng, D., Savage, C.A., Li, X. and Liu, B., 2015. Data-Driven CT Protocol Review and Management. *Journal of the American College of Radiology*, *12*(3), pp.267-272.

Protocol Decomposing Results

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

	Scanner protocol name	Core protocol name	Contrast tag	Weight tag	Hardware tag	Low Dose tag	Loc Index	Time tag	# of Exams
2	5.19 CHEST ROUTINE 135 - 200 LBS (~/>/+ 6/2011 LOW DOSE*	CHEST ROUTINE	(~/>/+	135 - 200 LBS		LOW DOSE	5.19	Jun-11	656
3	5.22 CHEST_ROUTINE_UNDER_200LBS 6/2011 *559	CHEST ROUTINE GSI		UNDER_200LBS			5.22	Feb-16	237
4	5.20 NON_CON_CHEST_ROUTINE_135_200LBS 6/2011	CHEST ROUTINE	NON_CON	135_200LBS			5.2	Jun-11	166
5	5.21 CHEST_ROUTINE_UNDER_200LBS 6/2011 *559	CHEST ROUTINE GSI		UNDER_200LBS			5.21	Aug-13	165
6	5.2 PE PROTOCOL 135-300lbs 1/2010 LOW DOSE	PE PROTOCOL		135-300lbs		LOW DOSE	5.2	Mar-10	166
7	5.18 CHEST ROUTINE UNDER 135 LBS (~/>/+ 6/2011 LOW DOSE*	CHEST ROUTINE	(~/>/+	UNDER 135 LBS		LOW DOSE	5.18	Jun-11	156
8	5.2 PE_PROTOCOL_136_200_LBS 12/2013 GSI	PE PROTOCOL GSI		136_200_LBS			5.2	Dec-13	153
9	5.22 CHEST_ROUTINE_UNDER_200LBS 6/2011 *559	CHEST ROUTINE GSI		UNDER_200LBS			5.22	Aug-13	149
10	5.20 CHEST ROUTINE OVER 200 LBS (~/>/+ 6/2011 LOW DOSE*	CHEST ROUTINE	(~/>/+	OVER 200 LBS		LOW DOSE	5.2	Jun-11	144
11	5.39 DIFFUSE LUNG DISEASE 135 - 200 LBS 12/2008 LOW DOSE	DIFFUSE LUNG DISEASE		135 - 200 LBS		LOW DOSE	5.39	Dec-12	127
12	5.40 PULMONARY NODULE F/U 136-200 LBS 6/2012 LOW DOSE	PULMONARY NODULE F/U		136-200 LBS		LOW DOSE	5.40	Jun-12	104
13	5.22 CHEST_ROUTINE_UNDER_200LBS 12/2013 *559	CHEST ROUTINE GSI		UNDER_200LBS			5.22	Dec-13	99
14	5.2 PE_PROTOCOL_136_200_LBS 7/2013 GSI	PE PROTOCOL GSI		136_200_LBS			5.2	Jul-13	98
15	5.19 NON_CON_CHEST_ROUTINE_UNDER_135LBS 6/2011	CHEST ROUTINE	NON_CON	UNDER_135LBS			5.19	Jun-11	86
16	5.18 NON_CON_CHEST_ROUTINE_135_200LBS 6/2011	CHEST ROUTINE	NON_CON	135_200LBS			5.18	Jun-11	77
17	6.16 CCPO 1 CHEST/ABP UNDER 200lbs 11/2011	CCPO 1 CHEST/ABP		UNDER 200lbs			6.16	Nov-11	63
18	5.21 NON_CON_CHEST_ROUTINE_OVER_200LBS 6/2011	CHEST ROUTINE	NON_CON	OVER_200LBS			5.21	Jun-11	63
19	5.3 PE PROTOCOL >200lbs 2/2010 LOW DOSE	PE PROTOCOL		>200lbs		LOW DOSE	5.3	Mar-10	63
20	5.38 DIFFUSE LUNG DISEASE UNDER 135 LBS 12/2008 LOW DOSE	DIFFUSE LUNG DISEASE		UNDER 135 LBS		LOW DOSE	5.38	Dec-08	62
21	5.1 PE PROTOCOL <135lbs 2/2010 LOW DOSE	PE PROTOCOL		<135lbs		LOW DOSE	5.1	Mar-10	61
22	5.8 FEMUR/ LOWER LEG + HARDWARE + 200 lbs 12/2012	FEMUR/ LOWER LEG		+ 200 lbs	HARDWARE		5.8	Dec-12	24

Zheng, D., Savage, C.A., Li, X. and Liu, B., 2015. Data-Driven CT Protocol Review and Management. *Journal of the American College of Radiology*, *12*(3), pp.267-272.

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

NumPy matplotlib python pandas

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), …

```
para_set = {"output_path": output_path_general + "\\ABP",
            "general_protocol_name_override": "ADD/PEL",
            "prot_inc_exc_regexp_pair_list": [(r'(?i).*(ADD.*PEL|AP).*',
                                               r'torso;chest;biopsy;run;ctv;ctac;ctu;research'),
                                              (r'ADD.*PEL', r'torso;chest;biopsy;run;ctv;ctac;ctu;research'),
                                              (r'ab.*p', r'torso;chest;biopsy;run;ctv;ctac;ctu;research'),
                                              # mesenteric bowel is an abp protocol
                                              (r'mesenteric;bowel', r'research'),
                                              (r'CTE', r'\bPE\b;research')],
```

Beth Israel Deaconess Medical Center | Harvard Medical School Teaching Hospital

22

Some Results: Interactive Exploration of Dose Data

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

```
dcc = DoseCrossCompare(target_path_general,
                       naming_scheme='scan_protocol_name',
                       prot_inc_exc_regexp_pair_list=[(r'chest;pain', r'')],
                       CT_list=['CT-1',
                                'CT-2',
                                'CT-3',
                                'CT-4',
                                'CT-5',
                                'CT-6',
                                'CT-7',
                                'CT-8',
                                'Needham-CT'],
                       study_date_from_str="2016-01-01",
                       study_date_to_str="2016-06-30",
                       bool_verbose=False)
dcc.display_dose_summary('all')
```

```
Needham-CT
Core protocol name:  ('NON GATED CHEST PAIN (PE)', 'NON GA
WITH ABD/PELVIS (PE/GENERAL)', 'PE/DISSECTION (CHEST PAI
Number of exams:  254
CTDIv:  10.6  mGy
CTDIv STD:  5.5  mGy
DLP:  445.6  mGy*cm
DLP STD:  296.7  mGy

-----------------------------------------
 Dose info for the all CTs together
-----------------------------------------
Core prot names from all scanners:
CTA NON GATED CHEST PAIN
CTA NON GATED CHEST PAINGA/P
NON GATED CHEST PAIN
NON GATED CHEST PAIN (PE)
NON GATED CHEST PAIN W ADD PEL
NON GATED CHEST PAIN WITH ADD/PEL
NON GATED CHEST PAIN WITH AP
NON GATED CHEST PAIN(PE/DISSECTION)
NON-GATED CHEST PAIN WITH ABD/PELVIS (PE/CTE)
NON-GATED CHEST PAIN WITH ABD/PELVIS (PE/GENERAL)
PE/DISSECTION (CHEST PAIN)

Number of exams:  1532
CTDIv:  12.3  mGy
CTDIv STD:  5.5  mGy
DLP:  472.3  mGy*cm
DLP STD:  336.1  mGy
```

23

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

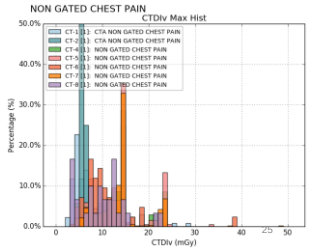
```
dcc = DoseCrossCompare(target_path_general,
                       naming_scheme='scan_protocol_name',
                       prot_inc_exc_regexp_pair_list=[(r'chest;pain', r'ab;a.?p'),],
                       CT_list=['CT-1',
                                'CT-2',
                                'CT-3',
                                'CT-4',
                                'CT-5',
                                'CT-6',
                                'CT-7',
                                'CT-8',
                                'Needham-CT'],
                       study_date_from_str="2016-01-01",
                       study_date_to_str="2016-06-30",
                       bool_verbose=False)
dcc.display_dose_summary('all')
```

```
CT-8
Core protocol name:  ['NON GATED CHEST PAIN']
Number of exams:  30
CTDIv:  10.0  mGy
CTDIv STD:  4.0  mGy
DLP:  324.3  mGy*cm
DLP STD:  164.1  mGy

Needham-CT
Core protocol name:  ('NON GATED CHEST PAIN (PE)', 'NON GA
Number of exams:  229
CTDIv:  10.6  mGy
CTDIv STD:  5.7  mGy
DLP:  418.0  mGy*cm
DLP STD:  281.8  mGy

-----------------------------------------
 Dose info for the all CTs together
-----------------------------------------
Core prot names from all scanners:
CTA NON GATED CHEST PAIN
NON GATED CHEST PAIN
NON GATED CHEST PAIN (PE)
NON GATED CHEST PAIN(PE/DISSECTION)
PE/DISSECTION (CHEST PAIN)

Number of exams:  1371
CTDIv:  12.2  mGy
CTDIv STD:  5.6  mGy
DLP:  484.5  mGy*cm
DLP STD:  248.1  mGy
```

24

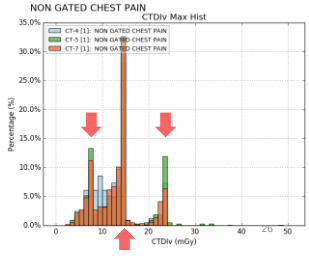
Graphical Exploration of Dose Data: Histogram

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



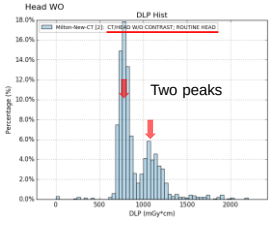
Graphical Exploration of Dose Data: Histogram

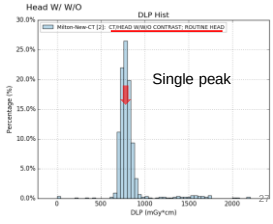
- Similar distribution observed from 3 GE 750 scanners
- Size-specific protocols: 3 BMI groups, 3 peaks



Graphical Exploration of Dose Data: Histogram

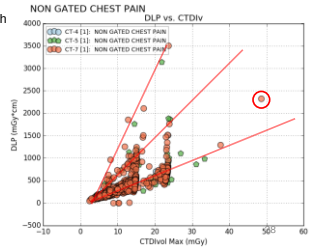
- Histogram also helps avoid errors in protocol classification
  - Left figure: "CT/Head W/O CONTRAST" and "ROUTINE HEAD" were grouped together
  - "ROUTINE HEAD" turned out to be a two-phase scan; it should be grouped with W/WO





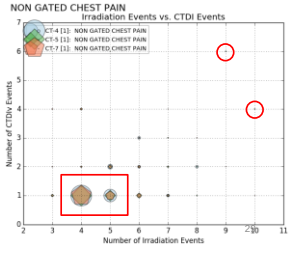
Graphical Exploration of Dose Data: Scatter Plot

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



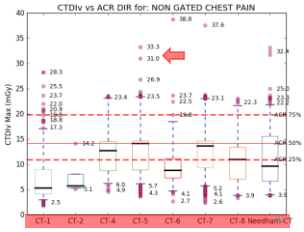
Graphical Exploration of Dose Data: Scatter Plot

- 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



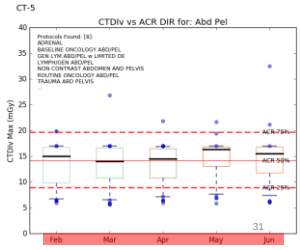
Graphical Exploration of Dose Data: Box Plot

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



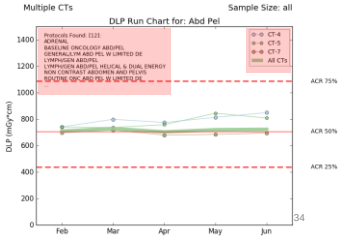
Graphical Exploration of Dose Data: Box Plot

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



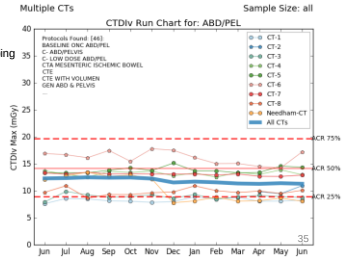
Graphical Exploration of Dose Data: Run Chart

- 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

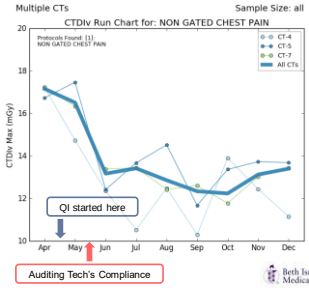


Graphical Exploration of Dose Data: Run Chart

- When combining data from many scanners
  - High level summary
  - 9 scanners, 46 ABP protocols
    - Work with lead tech to verify grouping
  - Very busy figure

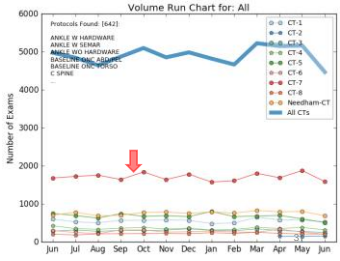


Use Run Chart to Show the Effect of a CT QA Project



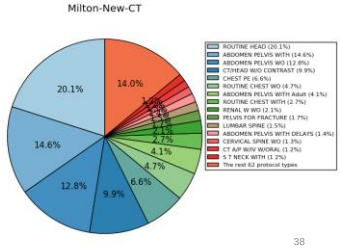
Graphical Exploration of Dose Data: Run Chart of Volume

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



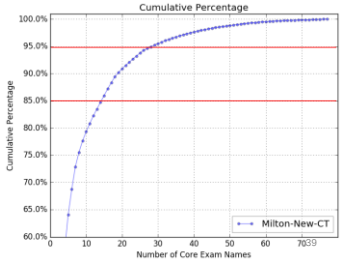
Graphical Exploration of Dose Data: Pie Chart

- Pie charts for evaluating the complexity of the protocol space



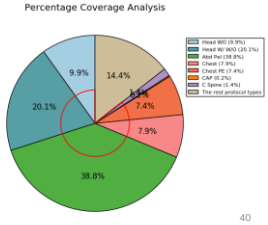
From Pie Chart to Cumulative Percentage

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



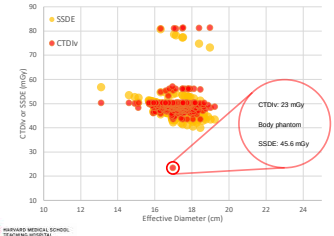
Graphical Exploration of Dose Data: Pie Chart

- 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!