Data Driven CT Dose Monitoring and Protocol Management

Da Zhang, PhD, DABR

Beth Israel Deaconess Medical Center | HARVARD MEDICAL SCHOOL TEACHING HOSPITAL

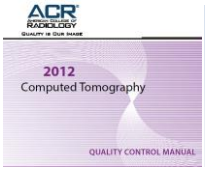
1

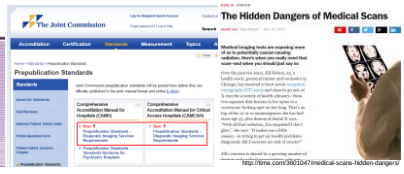
TABLE OF CONTENTS

2

Dose Concerns and Regulatory Requirements

- The rise of concerns of medically induced radiation
- Requirements from ACR accreditation and JC
- Dose monitoring and protocol review/revision/management





The Rise of Data Mining and Data Analytics

- "Data is the new electricity." -- Satya Nadella, CEO of Microsoft
- "Turn data into (actionable) insights." – data evangelists

Image from: http://anyfit.biz/enrbi-general

Advantages of Radiology in the Age of Data

- We are early adopters of informatics standardization in healthcare

HL7

DICOM
Digital Imaging and Communications in Medicine

RSNA
Radiological Society of North America

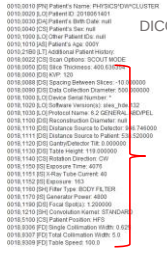
J Am Med Inform Assoc. 2018 Jul; 25(7): 885–893.
Published online 2018 May 29. doi: 10.1093/jamia/ocy053
PMCID: PMC6016707
PMID: 29850823

The LOINC RSNA radiology playbook - a unified terminology for radiology procedures

NRDR DIR Dose Index Registry
AMERICAN COLLEGE OF RADIOLOGY

Advantages of Radiology in the Age of Data

- This means we have some infrastructure ready to use



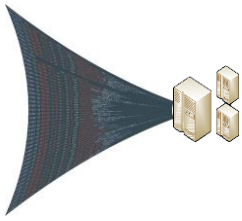
DICOM

Image from: www.radiologybusiness.com

TABLE OF CONTENTS

7

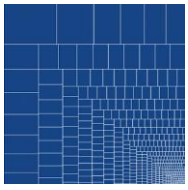
Challenge 1: Diversity of vendor, model, & vintage



8

Challenge 2: Highly Fragmented Data

- An example from our workhorse CT750 (in ED)
  - 1 yr. of data, 18569 exams
  - 118 protocols in exam records
- From 5 GE CTs in my institution
  - 1 yr. of data, 48046 exams
  - 365 protocols in exam records
- Valid reasons of so many protocols
  - Fine-tuned & individualized protocols
  - Mixture of old and new data in the exam records
- Tremendous amount of data to analyze
  - How to cover them effectively?



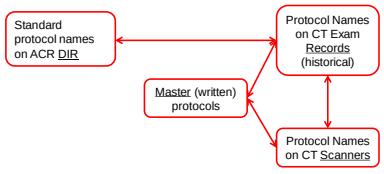
9

Challenge 3: The Clash of Names and Name Mapping

- 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.

Name Mapping

- There exists several sets of names concerning patient dose and CT protocols

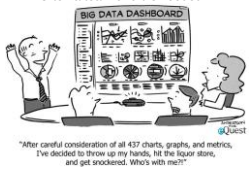


Tower of Babel

The Tower of Babel by Pieter Bruegel the Elder (1563), from wikipedia.org

Challenge 5: General Concerns

- Clinical users can be bombarded with overloaded information
- Data does not magically turn into information, knowledge, and insights
  - It takes effort to identify complex patterns or cause-effect relationships from data
  - Often a team effort is needed



"After careful consideration of all 437 charts, graphs, and metrics, I've decided to throw up my hands, hit the liquor store, and get snockered. Who's with me?!"

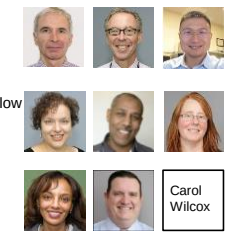
"Let's shrink Big Data into Small Data ... and hope it magically becomes Great Data."

Left image: curtesy of datafloq.com
Right image: curtesy of smartinsights.com

TABLE OF CONTENTS

Technical Development and Teamwork

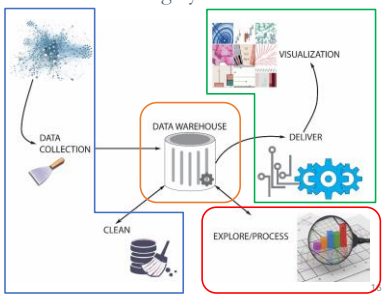
- A technical solution developed in-house
  - Geeks with enthusiasm and skill set
  - Desire of high-level flexibility from the system
  - Open source software resources
- Teamwork and integration into clinical workflow
  - Very supportive department
  - Teamwork of
    - Modality and division leadership
    - CT leadership
    - Physics
    - Informatics
    - Community hospital leadership



Carol Wilcox

General Components of A Data Mining System

- Data collection and data cleaning
- Data warehousing
- Data exploration
- Data visualization



The Data Collection Part

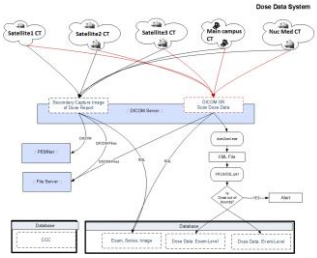
80% of a data scientist's valuable time is spent simply finding, cleansing, and organizing data. -- IBM

17

https://www.ibm.com/blogs/bluemix/2017/08/ibm-data-catalog-data-scientists-productivity/

Data Collection Subsystem

- We don't need to scrape every bit from the internet -- info is spoon-fed to us



18

Radiation Dose Structure Report – A Blessing

- Radiation Dose Structured Report (RDSR) became widely available, thanks to the MITA XR29 initiative
  - RDSR can be trivially converted to xml for data collection
  - CTDIv & DLP always present for CT

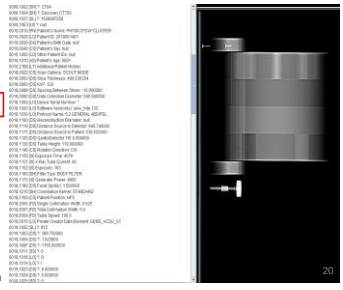
19

Useful Information Imbedded in DICOM Data

- Rich info in DICOM attributes
- Useful info not always reported
  - e.g. Dose modulation parameters
  - e.g. Patient BMI
  - e.g. Water equivalent diameters

SAM

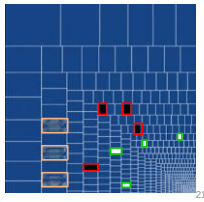
dicom.nema.org/MEDICAL/dicom/current/output/chtml/part03/sect_C.8.2.html



20

The Data Cleaning Part

- Direct results of data collection are generally dirty
- For example, the dose data can be polluted by
  - Invalid data, e.g., non-patient exams, incomplete exams
  - Duplicated records (e.g., produced by splitter)
  - Missing data



21

A Flexible and Smart Data Mining Subsystem

- With the wealth of data flooding in, how to smartly select the right data to answer clinical questions is the key.
- The capability to generate various data slicing is needed.

Implement A Flexible Data Slicing Algorithm

- Build flexible data selection criteria using regular expressions for including and excluding desired patterns
  - (inclusion_regex, exclusion_regex), (INC, EXC), (INC, EXC), …
- Script-driven, OOP-based modular system
  - To explore data interactively
  - To prepare plots for presentation in a batch-processing style

```
para_set = {"output_path": output_path_general + "\\ABP",
            "general_protocol_name_override": "ABD/PEL",
            "prot_inc_exc_regexp_pair_list": ((r'LYM.*GEN.*(ABD.*PEL|AP).*',
                                               r'torso;chest;biopsy;run;ctv;ctacctu;research'),
                                              (r'ABD.*PEL', r'torso;chest;biopsy;run;ctv;ctacctu;research'),
                                              (r'ab.*p', r'torso;chest;biopsy;run;ctv;ctai;ctacctu;research'),
                                              # mesenteric bowel is an abp protocol
                                              (r'mesenteric;bowel', r'research'),
                                              (r'CTE', r'\bPE\b;research')),
```


Data Exploration with Flexible Searching Criteria

- 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 G
WITH ABD/PELVIS (PE/GENERAL)', 'PE/DISSECTION (CHEST PAI
Number of exams:  254
CTDIv:  10.9  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 PAIN&A/P
NON GATED CHEST PAIN
NON GATED CHEST PAIN (PE)
NON GATED CHEST PAIN W ABD PEL
NON GATED CHEST PAIN WITH ABD/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:  330.1  mGy
```

Data Exploration with Flexible Searching Criteria

- 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'abp: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:  18.8  mGy
CTDIv STD:  4.8  mGy
DLP:  324.3  mGy*cm
DLP STD:  164.1  mGy


Needham-CT
Core protocol name:  ['NON GATED CHEST PAIN (PE)', 'NON G
Number of exams:  229
CTDIv:  18.6  mGy
CTDIv STD:  5.7  mGy
DLP:  418.0  mGy*cm
DLP STD:  201.8  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:  404.5  mGy*cm
DLP STD:  248.1  mGy
```

25

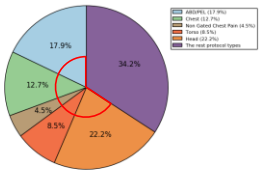
TABLE OF CONTENTS

26

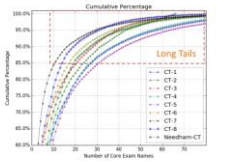
Before 2016-Q4: Review of Bread & Butter Protocols

- Review of only Bread & Butter Protocol
  - The problem is: How to cover the rest long-tail exam types
  - The purpose of CTQA is not merely meeting regulatory requirements, but to find issues

Percentage Coverage Analysis



How many protocols do we need to review to cover, e.g., 85% of the volume?

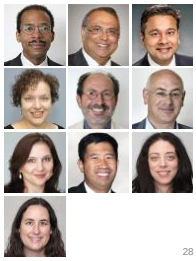


27

Since 2016-Q4: Section-by-Section Review

- Section-by-Section Rotational Review with the support of different sections
  - 2016-10: Neuro
  - 2017-02: MSK
  - 2017-06: Abdominal and Chest
  - 2017-09: Abdominal
  - 2018-01: Procedural CT
  - 2018-04: Cardiovascular CT
  - 2018-07: Participation with ACR DIR
  - 2018-10: Chest
  - 2019-01: Neuro
  - 2019-04: Abdominal

  One rotation
- Section leaderships are heavily involved



28

Section-by-Section Review

201901: Neuro

- Major Protocols
  - C+ Head non-DE: 12369
  - C-Spine: 1559
  - Head and Neck CTA: 1380
  - Sinus: 1252
  - C+ Neck: 1010
  - Facial: 659
  - L-Spine: 629
  - Head CTA non-DE (2 series): 409
  - Head/neck perfusion: 372
  - T-Spine: 297
- Minor Protocols
  - DE Head single scan: 152
  - C- Neck: 130
  - C+ Head non-DE: 112
  - Head CTA and CTV: 60
  - DE Head CTA: 22
  - C-/C+ Head non-DE: 112
  - C-/C+ Head DE: 20
  - Head/brain perfusion: 61

201904: ABP/Torso

- Major protocols:
  - ABP single scan: 8227
  - CAP 1-scan: 2471
  - CAP 1-scan TRAUMA: 805
  - CAP 2-scan: 2232
  - CTU: 1092
  - ER CTU: 1096
  - CAP multi-scan: 583
  - CTA ABP 2-scan: 784
- Minor protocols:
  - CTA CAP 1-scan: 435
  - CTA CAP 2-scan: 512
  - CTA ABP multi-scan: 482
  - ABD Only 2-scan: 122
  - ABD Only multi-scan: 562
  - ABP 2 scan: 325
  - ABP multi-scan: 85

201701: MSK

- General exam types, exam #, percentage (Year 2016)
  - MSK-Pelvis (337, 24.8%)
  - Knee (168, 12.4%)
  - Knee Fullerton (33, 2.4%)
  - Ankle (132, 9.7%)
  - Shoulder (123, 9%)
  - Tibia (110, 8.1%)
  - Femur (97, 7.3%)
  - Foot (87, 6.4%)
  - Hip (79, 5.8%)
  - Wrist (78, 5.7%)
  - Total coverage: 89.2% of the entire MSK cohort (1358 MSK exams)

201810: Chest

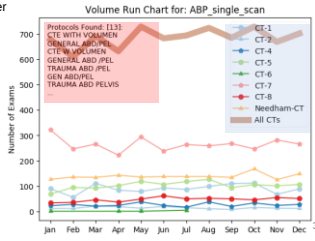
- Major Protocols
  - C- Chest: 4654
  - C+ Chest: 2170
  - Chest Lung Screening: 1217
  - Non-Gated Chest Pain: 2976
- Minor Protocols
  - Chest CTA/CTV: 109
  - Chest Super Dimensional: 178
  - Chest Biopsy: 257
  - T-Spine: 309

201804: Cardiac

29

Graphical Exploration: Volume Run Chart

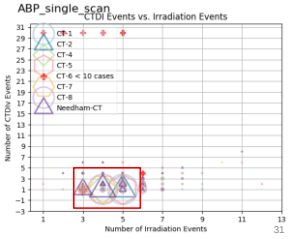
- Volume run-chart can be used to show the changes over time, across scanners
  - ABD/PEL exam from multiple scanner
  - Combined result from 13 protocols



30

Graphical Exploration: Scatter Plot of Radiation Event

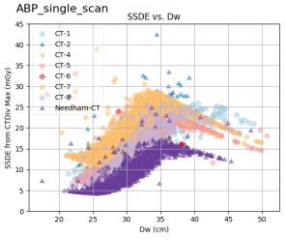
- Scatter plot of scan events gives clue about the # of scans and radiation events in an exam
  - E.g., ABP Single-phasic exams
    - 1 true CT scan
    - 3-5 total irradiation events
      - CT scan
      - Scout views
      - Monitoring phases
  - Note: marker size ∝ # of occurrence



Graphical Exploration: CTDIv and SSDE vs. Patient Size

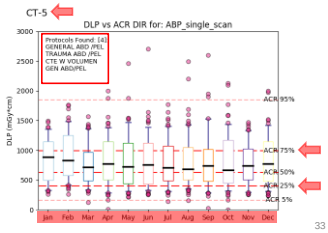
- With scatter plot of dose vs patient water equivalent diameter, we can visually and intuitively find performance inconsistency across scanners
  - Single phase ABP exam
  - Note the dark purple data points
    - Same model as CT8
    - Much lower dose settings
    - Impact on image quality
  - SSDE: Normalize CTDIv according to
    - Patient effective diameter
    - Water equivalent diameter
    - Water equivalent area

SAM



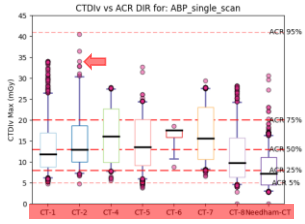
Graphical Exploration: Box Plot Over Time

- Boxplots can also be used to show the changes over time
  - Single-scan ABD/PEL exam from 1 scanner (CT5)
  - Combined result from 4 protocols



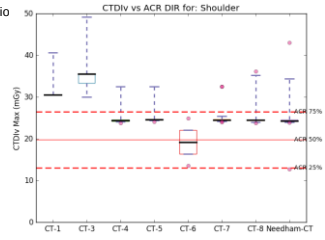
Graphical Exploration: Box Plot Across CT

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



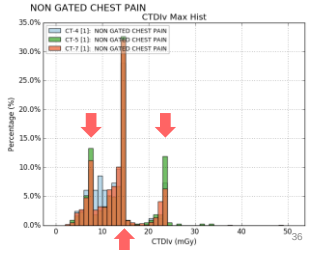
Real Issue Found: MSK Shoulder Protocols With Fixed mA

- Boxplots also help to find issues in the practice
  - Fixed technique vs dose modulatio
  - Higher dose than benchmark

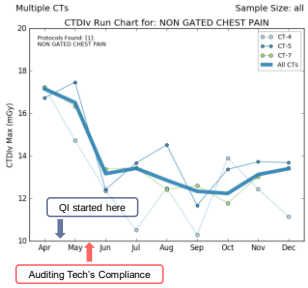


Graphical Exploration: Histogram of a BMI-based protocol

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



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



A Protocol Error Found from Histogram

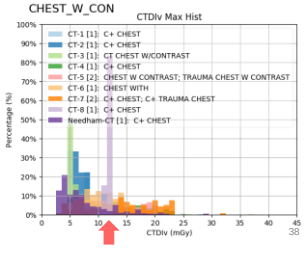
- Single Scan Chest
- Identified issue from data
  - CT8 C+ Chest uses a fixed technique
- Confirmed later by Bob Butler

From: Butler,Robert D. (BIDMC - Radiology)
Sent: Monday, October 29, 2018 3:28 PM
To: Zhang,Da (BIDMC - Radiology); Wilcox,Carol
Subject: RE: CTQA meeting of BIDMC and Needham -Follow-up Action Item

Hi Da and Carole,

I hope all is well. I wanted to follow up with both of you on C+ Chest dose issues at Chestnut Hill. For some reason or another, our protocol was set to a fixed 300 mA. No idea why. We adjusted the technique to match the protocol. I will look into it further and look to see when this started. But we should be all set moving forward.

Thanks,
Bob

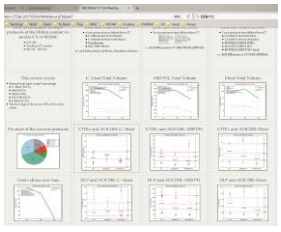


Script-Driven Data Analytics Report Generation

- Semi-automatic Generation of Reports for CT Quality Assurance meetings

Text based script

HTML based slides



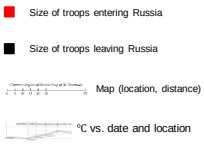
Summary

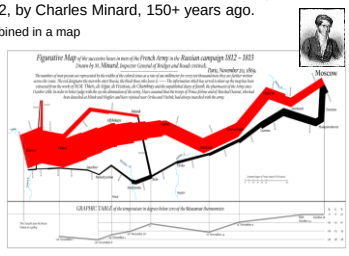
- CT dose monitoring and protocol control are crucial components of the dynamic QA programs of modern imaging environments.
- A data mining pipeline can be formed to handle data collection, cleansing, warehousing, exploration, and visualization of these valuable operational and dosimetry data, as well as to answer various clinically relevant questions.
- Collaborations among different clinical teams and careful plans of protocol review sessions are critical for the effective integration into the clinical workflow and comprehensive coverage of CT protocols.

40

Outlook: From Data to Insights

- A famous data visualization masterpiece is the visualization of Napoleon's Russian campaign of 1812, by Charles Minard, 150+ years ago.
  - Multi-dimensional data combined in a map





41

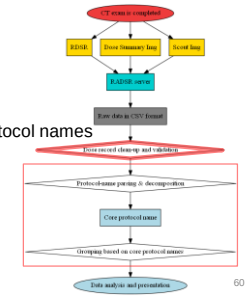
Image from: http://https://en.wikipedia.org/wiki/French_invasion_of_Russia

Thank you for your attention!

- I'd love to hear your thoughts
  - Email: dzhang8@bidmc.harvard.edu

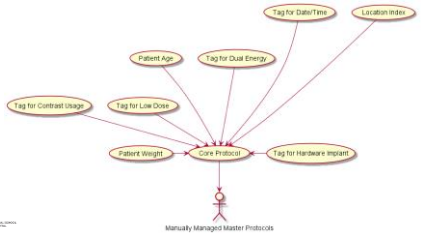
Data Mining: Cleansing, Validation, and Classification

- Data cleansing and validation against
  - Non-patients CT scans
  - Duplicated records
  - Invalid/incomplete data
- Parse, decompose, and normalize protocol names
  - To solve the “clash of names”
  - To build classifiers for the dose data
  - To group right data for presentation



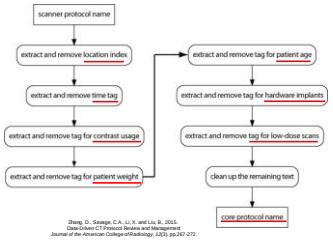
Protocol Name Decomposition

- Observation: many descriptive phases/fields are added to the core part to form a protocol name



Protocol Name Decomposition

- Protocol names can be decomposed automatically:



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 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 - 288 LBS (>3V-6/2011 LOW DOSE*	CHEST ROUTINE	CON	135 - 300 LBS		LOW DOSE	5.19	Jun-11	616
3	5.22 CHEST_ROUTINE_UNDER_300LBS (CON) *559	CHEST ROUTINE 559		UNDER_300LBS			5.22	Feb-13	157
4	5.20 NON_CON_CHEST_ROUTINE_135_288LBS 6/2011	CHEST ROUTINE	NON_CON	135_300LBS			5.2	Jun-11	266
5	5.21 CHEST_ROUTINE_UNDER_300LBS 6/2012 *559	CHEST ROUTINE 559		UNDER_300LBS			5.21	Aug-12	165
6	5.2 PE PROTOCOL 135-300lbs 3/2010 LOW DOSE	PE PROTOCOL		135-300lbs		LOW DOSE	5.2	Mar-10	160
7	5.18 CHEST ROUTINE UNDER 135 LBS (>3V- 6/2011 LOW DOSE*	CHEST ROUTINE	>3V	UNDER 135 LBS		LOW DOSE	5.18	Jun-11	156
8	5.2 PE_PROTOCOL_136_285_LBS 12/2013 559	PE PROTOCOL 559		136_285_LBS			5.2	Dec-13	153
9	5.22 CHEST_ROUTINE_UNDER_300LBS 6/2013 *559	CHEST ROUTINE 559		UNDER_300LBS			5.22	Aug-13	149
10	5.20 CHEST ROUTINE OVER 288 LBS (>3V-6/2011 LOW DOSE*	CHEST ROUTINE	>3V	OVER 288 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-08	127
12	5.40 PULMONARY NODULE F/U 136-288 LBS 6/2012 LOW DOSE	PULMONARY NODULE F/U		136-288 LBS		LOW DOSE	5.40	Jun-12	104
13	5.22 CHEST_ROUTINE_UNDER_300LBS 12/2013 *559	CHEST ROUTINE 559		UNDER_300LBS			5.22	Dec-13	99
14	5.2 PE_PROTOCOL_136_300_LBS 7/2013 559	PE PROTOCOL 559		136_300_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_300LBS 6/2011	CHEST ROUTINE	NON_CON	135_300LBS			5.18	Jun-11	77
17	8.16 CCPO 1 CHEST/ABP UNDER 200lbs 11/2011	CCPO 1 CHEST/ABP		UNDER 200lbs			8.16	Nov-11	63
18	5.21 NON_CON_CHEST_ROUTINE_OVER_300LBS 6/2011	CHEST ROUTINE	NON_CON	OVER_300LBS			5.21	Jun-11	63
19	5.3 PE PROTOCOL >288lbs 3/2010 LOW DOSE	PE PROTOCOL		>288lbs		LOW DOSE	5.3	Mar-10	62
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 3/2010 LOW DOSE	PE PROTOCOL		<135lbs		LOW DOSE	5.1	Mar-10	61
22	9.8 FEMUR/ LOWER LEG = HARDWARE < 200 lbs 12/2012	FEMUR/ LOWER LEG		< 200 lbs	HARDWARE		9.8	Dec-12	24

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.