Open-Source Radiation Exposure Extraction Engine (RE3) with Patient-Specific Outlier Detection

Samuel J. Weisenthal^{1,2} · Les Folio¹ · William Kovacs¹ · Ari Seff¹ · Vana Derderian¹ · Ronald M. Summers¹ · Jianhua Yao¹

Published online: 7 December 2015
© Society for Imaging Informatics in Medicine 2015

Abstract We present an open-source, picture archiving and communication system (PACS)-integrated radiation exposure extraction engine (RE3) that provides study-, series-, and slice-specific data for automated monitoring of computed tomography (CT) radiation exposure. RE3 was built using open-source components and seamlessly integrates with the PACS. RE3 calculations of dose length product (DLP) from the Digital imaging and communications in medicine (DICOM) headers showed high agreement ($R^2=0.99$) with the vendor dose pages. For study-specific outlier detection, RE3 constructs robust, automatically updating multivariable regression models to predict DLP in the context of patient gender and age, scan length, water-equivalent diameter (D_w), and scanned body volume (SBV). As proof of concept, the model was trained on 811 CT chest, abdomen + pelvis (CAP) exams and 29 outliers were detected. The continuous variables used in the outlier detection model were scan length ($R^2=0.45$), D_w ($R^2=0.70$), SBV ($R^2=0.80$), and age ($R^2=0.01$). The categorical variables were gender (male average 1182.7 ± 26.3 and female 1047.1 ± 26.9 mGy cm) and pediatric status (pediatric average 710.7 ± 73.6 mGy cm and adult 1134.5 ± 19.3 mGy cm).

Keywords Radiation dose · Quality control · Software design · PACS implementation · PACS integration · Open source · Medical informatics applications · Imaging informatics · Data mining · Data extraction · Computers in medicine · Clinical workflow · Automated measurement · Statistical analysis · Quality assurance

Abbreviations

ACR	American College of Radiology
AE	Application entity
ALARA	As low as reasonably achievable
BMI	Body mass index
CAD	Computer-aided detection/diagnosis
CAP	Chest, abdomen + pelvis
CT	Computed tomography
DCMTK	DICOM toolkit
DIR	Dose index registry
DLP	Dose length product
DRL	Diagnostic reference level
D_w	Water-equivalent diameter
FEM	Female
FOV	Field of view
HU	Hounsfield units
IP	Image processing
IRB	Institutional review board
MC	Monte Carlo
NCI	National Cancer Institute
OCR	Optical character recognition
OS	Overscan
PACS	Picture archiving and communication system
PED	Pediatric
QA	Quality assessment
QC	Quality control
RDSR	Radiation dose structured report

✉ Jianhua Yao
jyao@cc.nih.gov
Samuel J. Weisenthal
samuel_weisenthal@urmc.rochester.edu

¹ National Institutes of Health, Clinical Center, Radiology and Imaging Sciences, Clinical Image Processing Service (CIPS), 10 Center Drive, Bethesda, MD 20892-1182, USA

² University of Rochester Medical Center, 601 Elmwood Ave, Rochester, NY 14642, USA

SBV	Scanned body volume
SD	Standard deviations
SL	Scan length
UID	Unique identification number
DICOM	Digital imaging and communications in medicine

Introduction

CT accounts for almost half of medical procedure-related radiation exposure to Americans each year [1] and the risk of cancer associated with CT radiation has been extensively discussed [2–12]. Imaging modalities are now required to report radiation exposure metrics [13] and some states are requiring hospitals to keep radiation records [14]. However, best practices to ensure that the dose is as low as reasonably achievable (ALARA) are not always observed. There is great variation in dose indices across institutions, presumably because of differing “technique factors,” such as settings for peak kilovoltage (kVp) and pitch [15]. Even when best practices are observed, outliers must still be detected. Ideally, an excessive dose would automatically generate an alert for review by a medical physicist [16]. The American Association of Physicists in Medicine (AAPM) has published reference dose levels [17] for body part and age. A group has also developed automated alerts for high-dose computed tomography (CT) exams based on how they compare to institutional average [18]. Recent experience in San Francisco medical centers is also now available for more general comparisons [19].

There is a movement to employ information technology tools to reduce radiation dose [20]. Institutions that wish to monitor radiation dose systematically have the following choices:

1. Purchase a commercial software package. The functions of a remotely designed software package, however, are often not customized to the special need of an institution. There are a variety of methods within these packages to identify sentinel events, but if they are not optimized to a particular institution, there is no means to tailor them in-house. This is a shortcoming because each institution has a unique combination of vendors, protocols, and scanners so a one-size-fits-all prepackaged option is difficult to implement.
2. Amend an open-source dose monitoring package to their workflow. This is an attractive option as evidenced by the widespread adoption of Radiance [21], a completely open-source dose monitoring software. Besides being free-of-charge, open-source software like Radiance can be modified to the particularities of an institution. For example, if a new scanner that uses non-standard dose pages is introduced, a programmer can amend the source code to track that scanner (this was done at our institution). Easily modifiable open-source software is particularly effective in an environment that is in flux, such as a large institution with multiple scanners and protocols with needs to track and improve dose

practices that could change rapidly over the coming decades due to media and legislative pressures.

3. Write a dose tracking software de novo. This allows for the greatest customization. In this paper, we describe how we developed a radiation exposure extraction engine (RE3). We wanted RE3 to be open-source so that it would be freely available and also because open-source software can be tailored to a specific institution’s needs, as mentioned earlier. Radiology picture archiving and communication system (PACS) are complex, and the constantly evolving scanning and archiving technology makes tracking dose a challenge that is particular to an institution. We developed RE3 with open-source components Perl and digital imaging and communications in medicine (DICOM) tool kit (DCMTK). We wanted a versatile exposure-monitoring tool granular in its controls and data (providing slice-, series-, study-, and patient-specific (e.g., scanned body volume (SBV)) data). Vendors are now populating DICOM headers with a volume CT dose index ($CTDI_{vol}$) that corresponds to the modulated current value at the time of image acquisition. Hence, exposure metrics can now be calculated directly from the image DICOM headers. In anticipation of personalized Monte Carlo organ-specific dose estimates and for more systematic outlier detection, we sought a system that extracted slice-specific radiation metrics from the DICOM header instead of series- or study-specific metrics from the dose page, from which data can also be difficult to extract due to non-standard formatting (Fig. 1). Further, we wished to avoid erroneously overestimating dose because of duplicated dose pages. Previous groups have presented methods to extract DICOM header radiation data [22], but there exists no open-source, fully automated, PACS-integrated system that reads from the DICOM headers and runs in the clinical workflow. Such a system must be able to integrate seamlessly with the PACS, automatically determine the number of acquisitions per exam [23], and account for overscan (OS) [24–26].

In this paper, we describe how we solved each of these problems and then tested our system in our large clinical research center containing multiple scanner manufacturers and models and hundreds of research protocols. As an epidemiological pilot study, we show the average dose length product (DLP) for different research protocols and study descriptions. In addition, we show how our system detects context-specific radiation dose outliers using patient characteristics such as age, gender, water-equivalent diameter (D_w) [27] and SBV.

Methods

Our study received institutional review board (IRB) approval and was compliant with the Health Insurance Portability and Accountability Act. RE3 was built with the open-source

Fig. 1 Snapshots of (a) Toshiba and (b) Siemens dose pages from our PACS. Dose pages only provide study- and series- (not slice-) level data. Variation in dose page format makes OCR-based dose monitoring a challenge. Further, dose pages are sometimes duplicated in the PACS, causing a single study to be erroneously counted multiple times

30-Jul-2014 09:36

a

Ward:
Physician:
Operator:

Total mAs 16473 Total DLP 1222 mGycm

	Scan	kV	mAs f ref.	CTDIvol* mGy	DLP mGycm	TI s	cSL mm
Patient Position F-SP							
Topogram	1	120	35 mA	0.14 L	9	6.8	0.6
DE_CAP	2A	100	230 f 230				
	2B	Sn140	174 f 178	18.06 L	1180	0.33	0.6
New Position F-PR							
Contrast							
ProneTopogram	3	120	35 mA	0.14 L	6	4.8	0.6
HR CHEST	4	120	164 f 200	0.75 L	27	1.0	1.0
Last scan no.	15						

Medium	Type	Iodine Conc. mg/ml	Volume ml	Flow ml/s	CM Ratio
Contrast		0	0	0.0	100%
Saline			0	0.0	

Study Date : 2014/02/20

Dose Display : IEC 3.0

Total DLP(mGy.cm) : (Head): - (Body): 2910.20

Total CTDIvol(mGy) : (Head): - (Body): 175.50

1.Cardiac Sub/Chest Follow - Chen

No.	Protocol	#of scan(s)	kVp	CTDIvol (mGy)	DLP (mGy.cm)	Exposure Time (sec)	Total mAs	Range (mm)	Body Size AP/LA	CTDIvol (Max) (mGy)
1	DualScano	1	120			16.28	813	***	***	
2	DualScano	1	120			16.24	1624	***	***	
3	Volume	1	120	9.10 (Body)	109.10 (Body)	0.28	110	120.0	27.85/30.54	9.10 (Body)
4	S&V	1	120	5.50 (Body)	1.10 (Body)	0.28	19	2.0	54.39/51.05	5.50 (Body)
5	SURE Start	1	120	113.00 (Body)	22.60 (Body)	6.64	332	***	54.39/51.05	113.00 (Body)
6	Volume	1	120	33.60 (Body)	403.40 (Body)	0.72	406	120.0	27.85/30.54	33.60 (Body)
7	Helical	1	120	14.30 (Body)	2374.00 (Body)	14.11	6087	1610.0	19.88/31.34	18.70 (Body)

b

components in Perl and the DICOM toolkit (DCMTK) [28], a set of open-source tools to transfer data to and from the PACS. Using the DCMTK, RE3 is seamlessly connected to the clinical PACS and can be configured to run on the fly or retrospectively for a list of exams or date range. Perl scripts within RE3 extract values from the DICOM headers and interface with an image processing (IP) tool that calculates scanned body volume (SBV) and water-equivalent diameter (D_w) for each slice. RE3 can be configured to output data by patient, study, series, or slice, and uses this data to construct an automatically updating multivariable regression model for patient- and exam-specific outlier detection. Data flow between components of the RE3 system is shown in Fig. 2. The execution of RE3 is controlled by a configuration file that is written in an

extensible markup language (.xml) file. The configuration file has several sections:

1. Paths

Program paths—Paths to engine program and DCMTK executables
Log file paths—Paths to log files and temporary files
IP paths—Paths to IP system that measures SBV and D_w

2. PACS connection

Application entity (AE) titles, IP addresses, ports

Fig. 2 Data flow between RE3, the PACS, image processing (IP), the configuration file (Config.), and the generated report. RE3 reads the configuration file, downloads images from the PACS, extracts dose data, sends the images to the image processing for SBV and D_w determination, and then displays all results in a report

3. Engine Scheduler and Modes

The scheduler tells the engine when to run. The default start and end of day (6 and 12 a.m.) may be modified by the user to optimize RE3 incorporation into workflow. The scheduler switch can be turned off to run the engine continuously. The mode specifies how the engine runs. The possible modes include on the fly (in the workflow in real time), for a time range (date range specified by user), or for a list of exams (accession numbers of exams included in text file specified by user). Further, the engine can be set to retrain the outlier detection model on the fly. Lastly, the housekeeping mode indicates whether the image files should be deleted from the local drive after processing (e.g., to save memory).

4. Study Filters

The filters to select studies to be retrieved and processed by the engine are customizable. Typical filters include study description and image number. For example, to process all CT exams, the filter is set as “CT*”. As mentioned earlier, a list of exams can be processed if the engine is in the list mode. RE3 is open-source and available for download from <https://github.com/NIH-CIPS/RE3>.

Data Flow

Since RE3 must download large amounts of data from the PACS, the data flow scheme (Fig. 3) is important. Using the DCMTK query module FINDSCU, RE3 first queries at the study level,

Fig. 3 Schema of RE3 data flow. Specific queries 1 via FINDSCU at the study (Query₁) and series (Query₂) level discover studies in PACS that are retrieved 2 via MOVESCU to the local drive. Each DICOM image is converted 3 via DCM2TXT to a text file so that 4 data can be extracted and dose metrics calculated per slice. Radiation data can be configured to be output at the study, series, and slice level for 5 analysis and monitoring/outlier detection

finding the exam that matches any study filters set in the configuration file, and then looks for series that correspond to the series filters. It retrieves the DICOM images via the data transfer module MOVESCU to the local drive, where it converts them to text files via DCM2TXT. The appropriate tags are then extracted from each text file and passed through a module that uses an algorithm to calculate radiation metrics and the number of acquisitions. RE3 uses these radiation metrics along with patient-specific age, gender, SBV, and D_w and exam-specific scan length (SL) for outlier detection. RE3 can output radiation data by study (total DLP), series (series DLP), or slice (CTDI_{vol} or mAs) and generates epidemiological reports on a weekly basis.

DICOM Extraction

As shown in Fig. 4, RE3 extracts the tags from the DICOM header at the patient, study, series, and slice level (note: these values are extracted for every image; they are grouped here according to study, series, and image). From this data, it calculates DLP, average and total CTDI_{vol}, average milliampereseconds, and image number. SBV and D_w are determined from the images by a computer-aided detection/diagnosis (CAD) system activated by RE3.

Parameter Calculation

Dose Length Product

Data extracted from DICOM header provides information to compute radiation dose metrics such as dose length product

Fig. 4 Values extracted directly from DICOM header with DICOM tags in parentheses. Fields are grouped by the level at which they remain constant

Patient-level	SingleCollimationWidth (0018,9306)
PatientAge (0010,1010)	TotalCollimationWidth (0018,9307)
PatientSex (0010,0040)	kVp (0018,0060)
PatientID (0010,0020)	ExposureTime (0018,1150)
	ConvolutionKernel (0018,1210)
Study-level	BodyPartExamined (0018,0015)
AccessionNumber (0008,0050)	SliceThickness (0018,0050)
ClinicalTrialProtocolID (0012,0020)	TableSpeed (0018,9309)
ClinicalTrialProtocolName (0012,0021)	SeriesTime (0008,0031)
Manufacturer (0008,0070)	AcquisitionDate (0008,0022)
ManufacturersModelName (0008,1090)	AcquisitionNumber (0020,0012)
StudyDate (0008,0020)	SeriesInstanceUID (0020,000E)
StudyTime (0008,0030)	Slice-level
StudyDescription (0008,1030)	AcquisitionTime (0008,0032)
	AcquisitionDateTime (0008,0020A)
Series-level	SliceLocation (0020,1041)
SeriesNumber (0020,0011)	Exposure (0018,1152)
SeriesDescription (0008,103E)	ImageType (0008,0008)
SpiralPitchFactor (0018,9311)	CTDIvol (0018,9345)

(DLP), a radiation exposure metric that depends on $CTDI_{vol}$, the volume CT dose index (which describes the scanner radiation output), and scan length (distance scanned). DLP is a surrogate [29] for patient radiation dose that is provided on the vendor dose page and is commonly used for radiation monitoring purposes since it represents the total radiation exposure over the length of the scan. While DLP is provided on the vendor dose page, it is not found directly among the DICOM headers, but can be calculated from information therein in the following manner.

The number of images in the series (ImageCount) is determined first. Two metrics are then calculated: scan length (ScanLength), which is calculated according to slice location for contiguously acquired scans Eq. (1) and according to slice thickness for non-contiguously acquired scans Eq. (2). Equation (2) was employed as a backup in the event that Eq. (1) was equal to zero. At our institution, this was only the case for high-resolution CT chest exams. This presents a limitation that is treated in the discussion, and series-level

average $CTDI_{vol}$ ($\overline{CTDI_{vol,series}}$) Eq. (3). $SliceLocation_0$ and $SliceLocation_f$ are the slice locations at the first and last slices.

$$ScanLength_1 = SliceLocation_f - SliceLocation_0 \quad (1)$$

$$ScanLength_2 = SliceThickness \times ImageCount \quad (2)$$

$$\overline{CTDI_{vol,series}} = \frac{\sum CTDI_{vol,slice}}{ImageCount} \quad (3)$$

In addition to the scan length, overscan (OS) needs to be taken into account. OS, a known feature of helical scans, is also referred to as over-ranging since it occurs because the scanner must scan beyond the boundaries of the scan for complete data interpolation [30]. Initial DLP calculations from DICOM headers were found to underestimate the DLP. The missing factor appeared to be small for most scans, but correlated highly with mAs, especially at the front and end of the scan. To account for OS, vendor-specific regression models that depended on the initial and final mAs or on the average mAs (Eq. (4)) were trained on a week of CT data. OS was

determined by manual comparison of our system's calculated DLP with the DLP recorded on the dose page.

$$\overline{\text{mAs}}_{\text{series}} = \frac{\sum \text{mAs}_{\text{slice}}}{\text{ImageCount}}; \quad (4)$$

We generated OS regression models for Siemens Eq. (5), Toshiba Eq. (6), and Philips Eq. (7) scanners.

$$\text{OS}_{\text{Siemens}} = 40.99 + 0.01 \times \text{mAs}_0 + 0.01 \times \text{mAs}_f; \quad (5)$$

$$\text{OS}_{\text{Toshiba}} = 2.93 + 0.56 \times \overline{\text{mAs}}_{\text{series}} \quad (6)$$

$$\text{OS}_{\text{Philips}} = -13.90 + 0.36 \times \text{mAs}_0 + 0.43 \times \text{mAs}_f \quad (7)$$

OS adjustment was not required for discrete scanning protocols such as high-resolution chests.

Now that all the necessary values are acquired, DLP can be calculated via Eq. (8):

$$\text{DLP}_{\text{series}} = \overline{\text{CTDI}}_{\text{vol}} \times \text{ScanLength}_{1 \text{ or } 2} + \text{OS} \quad (8)$$

Series DLPs are summed to give DLP for the entire study (DLP_{study}) Eq. (9):

$$\text{DLP}_{\text{study}} = \sum \text{DLP}_{\text{series}} \quad (9)$$

Body Volume and Water-Equivalent Diameter

Now that DLP is calculated for each scan, we wish to view it in the context of the patient's scanned body volume and composition. To do so, an in-house image processing tool measures scanned body volume (SBV) and water-equivalent diameter (D_w) for each slice. These values are exam- and patient-specific and can be used to generate a patient-specific dose model. To compute SBV, the body area is segmented from the background and scanner table using region growing and active shape models [31]. Lung area is excluded since it is mostly air, which does not absorb radiation. Figure 5 shows an example of computed SBV.

D_w is a commonly used metric that takes into account the area (cm^2) and attenuation (Hounsfield units, HU) of each pixel. It gives the diameter of a body of water that has the same total attenuation as the scanned body. D_w is calculated for each slice Eq. (10):

$$D_{w_{\text{slice}}} = \sqrt{\sum \frac{(\text{pixelHU} + 1000)}{1000}} \times \text{pixel area} \times \frac{4}{\pi} \quad (10)$$

SBV and D_w are then summed at the series-level Eqs. (11, 12), and study-level Eqs. (13, 14):

$$\text{SBV}_{\text{series}} = \sum \text{SBV}_{\text{slice}}; \quad (11)$$

$$D_{w_{\text{series}}} = \sum D_{w_{\text{slice}}}; \quad (12)$$

$$\text{SBV}_{\text{study}} = \sum \text{SBV}_{\text{series}}; \quad (13)$$

$$D_{w_{\text{study}}} = \sum D_{w_{\text{series}}}. \quad (14)$$

Extraneous Series Exclusion

In a CT study, there may be multiple reconstructions of the same acquisition. Hence, there may be series within a study that did not contribute to the total radiation exposure. Further, studies within the PACS may also be duplicated. These duplicated studies may be a subset of or identical to the original study. Here, we call these extraneous. To ensure that one and only one series contributed to the exam DLP for each acquisition, extraneous series and studies should be excluded from DLP calculation. To manage extraneous series, RE3 uses a combination of values available from the DICOM header: acquisition number, acquisition time, protocol, CTDI_{vol} , and slice thickness. Two distinct (by unique identification number (UID)) series S_1 and S_2 are deemed from the same acquisition if one of the following criteria is met:

1. S_1 and S_2 have the same content time unless they are under the dual energy protocol;
2. S_1 and S_2 have the same acquisition number or acquisition time and kVp and protocol are the same.

If S_1 and S_2 are from the same acquisition: if S_1 and S_2 have different DLPs, the series with the lower DLP will be excluded; if S_1 and S_2 have the same DLPs, the series with the smaller scan length will be excluded.

Excessive Radiation Exposure Detection

One important purpose of RE3 is to detect studies with excessive radiation exposure so that clinicians and patients can be alerted. RE3 achieves this by examining each study in the context of patient- and exam-specific variables. In this study, a large dataset from a 2-month period was used as the training set. The independent variables included SBV and D_w , patient age, pediatric status (cutoff 18 years old) and gender, and scan length. The dependent variable is DLP. To visualize the multivariable regression model, we show plots of all predictors against DLP individually.

The model is trained iteratively to exclude outliers in the training set. During each iteration, a regression model is generated and any points whose residuals are more than 3 standard deviations (SD) from the residual mean are removed from the training set in the next iteration. The iterative training process ends when no such points can be excluded. The model is retrained every week to incorporate new data in the model. The model was set to iterate until the training set size was

Fig. 5 Example (a) 3D and (b) 2D SBV segmentation. To compute SBV, the body is segmented from the background and scanner table and lung area is excluded. SBV of each slice is added to give total SBV. SBV is then used by RE3 to detect patient-specific outliers

constant, but the number of iterations can be decreased to decrease sensitivity of the model.

One model is trained for each study type as determined by the study description. Study-dependent models are created to safeguard against the possibility that SBV and D_w calculations might be very different for different body parts. For example, the SBV metric calculated from a CT chest, abdomen + pelvis (CAP) exam might actually be smaller than the SBV metric calculated for a CT heart (since the field of view (FOV) in a heart might be smaller). Currently, a model has been built for

CT CAP, but the engine is programmed to automatically create a new model for each new study type that it encounters. If the engine encounters a study with a previously unseen study description, a new study type-indexed model dataset is created; if the dataset already exists, the new study is simply appended. If the dataset is adequately large (20 in the current system), RE3 creates the regression model and stores the coefficients and outlier cutoff to a file specific to that model. The outlier detection process then searches for the file that corresponds to a study description to find the correct model.

Fig. 6 **a** RE3 output DLP vs. dose page DLP for contiguous exams ($N=826$). RE3 calculates DLP with high fidelity to dose page ($R^2=0.98$) for all exams; lower outliers are the non-contiguously acquired high-resolution CT chests. **b** RE3 output DLP vs. dose page DLP for non-contiguous studies. **c** CT CAP only RE3 DLP calculations vs. dose page. All CT CAP exam metrics from RE3 compared to dose page ($N=580$ series; excluded 1 because of QA/QC, one because it was duplicated, and three because series on dose page were not present in PACS)

Validation Experiments

Dose Page Comparison

Foremost, we wanted to ensure that the metrics calculated by the engine were correct. To validate the engine DLP calculations, we compared series-level RE3 DLP calculations and acquisition determinations to the PACS dose pages for 1 month of CT scans for all scanners, protocols, and study types. Topograms/scouts were not taken into account and neither were scans with fewer than 20 images. A study was marked if it was a duplicate of another study in the PACS or if its dose page listed an extra series not present in the PACS (e.g., patient movement during a scan may have warranted its

redoing and the initial scan may have been recorded to the dose page but never sent to the PACS). The head technologist noted that not sending both exams to PACS was contrary to protocol, but that there seems to be no hard guideline for such a situation.

Epidemiological Pilot

To show RE3's usefulness as an epidemiological tool and since data had already been retrieved for validation purposes, DLP values for the month of February 2014 were plotted against clinical trial protocol name, scanner, and study type.

Fig. 7 DLP plotted against clinical protocol name (a) and study type (b)

Slice-Level Radiation Exposure

RE3 was designed to output not only series- and study-, but also slice-level radiation exposure in an entirely automated fashion (i.e., the user only needs to specify a list of exams or date range and RE3 will output the slice data only for acquisitions in each exam). This feature of RE3 is intended to provide data for large-scale Monte Carlo-associated organ-dose monitoring. To show RE3 slice-specific capabilities, we show an example CT CAP study with the automatic exposure control (AEC) curve and SBV curve plotted over the topogram scan. To show RE3's ability to extract slice data for a large number of exams, we compare SBV to exposure data at the slice level.

Outlier Detection

To show the outlier detection method employed by RE3, we used data from 2 months of CT CAP to iteratively train a regression model. Outliers were detected on this dataset using the final model.

Results

Dose Page Comparison

Eight hundred fifty-nine studies from 1 month were used for dose page validation. The studies were acquired on three scanners from two different vendors. Thirty-four study types were recorded. Manual comparison with the dose page determined that 12 were duplicated studies (i.e., there was another exam with the same images or a subset of the same images; the engine was later amended to take the series with the larger DLP or scan length as described in the section on extraneous series exclusion(1.3)). Hence, there were 847 unique studies performed on 620 unique patients.

From these studies, there were 1030 unique acquisitions that came from 847 studies. Nine of these studies were excluded; six contained series on the dose page that were not present in the PACS, two were classified manually as quality assessment (QA)/quality control (QC) issues (one was an exam that was started on one date and ended on another, another was started on one machine and ended on another), and one was missing a dose page for comparison. In the remaining 838, RE3 correctly determined the number of acquisitions in 837 of 838. The 1/838 in which it did not determine the correct number of acquisitions was correctly flagged by the engine because it contained two series that appeared to be reconstructions of the same data but had very different DLPs (The engine therefore elected not to make a decision between the series but instead to call for manual attention via a flag. The engine was later amended to take the series with the larger DLP or scan

length as described in the section on extraneous series exclusion (1.3)). Figure 6 shows the high agreement ($R^2=0.98$) between RE3-calculated DLP and the dose pages.

Exam DLPs ranged from 49 to 4101 mGy cm. In only contiguously acquired exams (Fig. 6a; 825 of 838, series scan lengths ranging from 1.4 to 1725 cm), $R^2=0.99$. RE3 was able to output exposure metrics at the study-, series-, and slice-level. Twelve of 838 exams were non-contiguously acquired hi-res chests (scan lengths ranging from 3.1 to 36.1 cm) that required a different method (slice thickness $\times$ number of images rather than slice location difference) for determining scan length. When we calculated DLP using the former method, it correlated with the dose page $R^2=0.99$ (Fig. 6b). In only CT CAP (Fig. 6c), RE3 DLP calculations showed high fidelity ($R^2=0.9987$) to dose pages. For CT CAP only, The data were from three scanners by two manufacturers.

Epidemiological Pilot

In Fig. 7, DLP values are plotted against clinical trial protocol name (Fig. 7a) and study type (Fig. 7b). The three most prominent clinical trial protocols (CHEST ABD PELVIS_W ($N=159$), NIH_CAP_W (132), and NIH_CHEST(58)) had DLPs

Fig. 8 **a** Slice-specific modulation of curve showed by RE3 in example patient. RE3 can output slice-specific data for use in personalized Monte Carlo (MC), for example. **b** Slice-specific SBV vs. exposure (mAs). Data was extracted automatically for 73 exams from 9577 images

314.58 ± 24.95 , 1001.96 ± 32.89 , and 235.56 ± 15.75 mGy cm, respectively. The three most prominent study types (CT CAP ($N=415$), CT chest (129), and CT neck (85)) had DLPs 1092.50 ± 24.93 , 252.24 ± 15.78 , and 474.64 ± 14.52 mGy cm, respectively.

Slice-Level Radiation Exposure

RE3 also determines the number of acquisitions for slice-level data, removing the slices that correspond to series that are reconstructions. RE3 was able to automatically extract slice-specific data. An example slice-specific AEC curve is shown in Fig. 8a. Exposure data was automatically extracted for 74 studies as evidenced in Fig. 8b.

Outlier Detection

The contextual outlier detections were performed on all CT CAP scans excluding images from the Toshiba scanner. In 2 months, our department performed 775 CT CAP exams with one to four passes and DLP ranging from 180 to 4101 mGy cm

(775 patients, age 10 to 91, 431 males, 344 females, 753 were >18 years old, and 22 were ≤ 18 years old). Body mass index (BMI) data was available for 743 patients and ranged 18.7 to 57.1 kg/m². In Fig. 9a, scan length is shown against DLP for 775 exams ($R^2=0.45$). In Fig. 9b, D_w is shown against DLP in 774 (for one exam SBV and D_w could not be computed) exams ($R^2=0.70$). In Fig. 9c, SBV is shown against DLP in 774 exams ($R^2=0.80$). In Fig. 9d, average DLP is shown with standard error bars for male and female patients ($N=775$): the male average was 1182.7 ± 26.3 mGy cm and female average was 1047.1 ± 26.9 mGy cm. In Fig. 9e, average DLP is shown with standard error bars for pediatric and adult patients ($N=775$): the pediatric average is 710.7 ± 73.6 mGy cm and the adult average is 1134.5 ± 19.3 mGy cm. In Fig. 9f, age is plotted as a continuous variable ($R^2=0.01$).

For all exams, SBV ($N=774$) (Fig. 10a) is compared to DLP ($R^2=0.80$) and BMI (Fig. 10b) is compared with DLP ($N=743$, $R^2=0.45$). In only single-phase exams, SBV (Fig. 10c) ($N=595$) is compared with DLP ($R^2=0.75$) and BMI (Fig. 10d) ($N=574$) is compared with DLP ($R^2=0.67$).

Fig. 9 Each feature used in the multivariable regression model plotted separately against DLP

Training of Outlier Detection Model

The results of our iteratively trained regression model over 17 iterations are shown in Fig. 11. Since we removed outliers at each iteration, the training data decreased from 811 at 0 iteration to 750 after 8 iterations (not shown) to 725 after 17 iterations. The final regression model for outlier detection was $DLP = -65.1 + -1.03 \times \text{age} + 63.1 \times \text{female (FEM)} - 36.8 \times \text{pediatric (PED)} + 0.03 \times \text{SBV} - 0.09 \times D_w + 3.82 \times \text{SL}$.

Outlier Detection

In Fig. 12, the outliers that would be detected by the regression model are shown. The final regression model is computed based on points (blue) that remained after 17 iterations. The 3 SD residual + residual mean is used as the cutoff to detect outliers. Twenty-nine high outliers in the 2-month period were detected. The plot shows that even patients with an absolute low DLP might have contextually high DLP and therefore be detected by the model. The DLPs of the high outliers ranged from 969.4 to 4101.1 with an average of 2001.8 ± 131.2 mGy cm. Only a simple threshold at roughly 950 mGy cm would have detected all of the outliers detected by our method.

Discussion

RE3 Benefits

The benefits of RE3 can be summarized as follows:

Design Benefits

Our system (1) runs on the fly in the background, seamlessly connected to the PACS and requires no manual intervention to output a list of contextually detected outliers; (2) is open-source; and (3) extracts data directly from the DICOM header, so it does not make OCR-related transcription errors or fail on non-standard or duplicated dose pages. Direct extraction from the DICOM header also allows for slice-specific data for research or Monte Carlo simulations.

Epidemiological Study

For epidemiological or overall exposure assessment purposes, RE3 is a versatile tool because it can extract any relevant field from the DICOM header and output the data at the study-, series-, and slice-level. The data can be output in a spreadsheet, where it can be further analyzed. Since RE3 is open-source, any institution or researcher can add additional tags of

Fig. 10 **a** SBV vs. DLP (all exams). For all exams, SBV ($N=774$) is compared to DLP ($R^2=0.80$). **b** BMI vs. DLP (all exams). For all exams, BMI is compared with DLP ($N=743$) ($R^2=0.45$). BMI fails as a metric especially with multiphase exams, whereas SBV takes into account multiple exams. **c** SBV vs. DLP (only single-phase). In only single-phase exams, SBV ($N=595$) is compared with DLP ($R^2=0.75$). **d** BMI vs. DLP (only single-phase). In only single-phase exams, BMI ($N=574$) is compared with DLP (0.67)

Fig. 11 Our outlier regression model trained on 811 CT CAP exams. RE3 iteratively trains a regression and then removes data points that are over 3 SD beyond the residual mean. For visualization purposes, we only show the model at first ($R^2=0.82$, $N=811$) and final iteration 17 ($R^2=0.94$, $N=725$). Iteration ceases when the dataset contains no outliers above 3 SD from the residual mean

interest. For example, it could help an institution pinpoint the protocol that is responsible for the highest average pediatric radiation exposure, or identify a scanner with an especially aggressive AEC system.

Outlier Detection

Our system enables the dose reduction team by generating sentinel alerts for exams deemed outliers. Instead of using DLP alone, our system uses patient- and exam-specific

Fig. 12 Outliers that would be detected by the final model from Fig. 11. The model depends only on points that remained after 17 iterations. At each iteration, points that were more than 3 SD (314.7 mGy cm) from the residual mean were removed. The final model is $DLP = -65.1 + 1.03 \times \text{age} + 63.1 \times \text{FEM} - 36.8 \times \text{PED} + 0.03 \times \text{SBV} - 0.09 \times D_w + 3.82 \times \text{SL}$. The 3 SD line is dashed. As shown, our method yielding the final CT CAP model would have detected 29 high outliers in the 2-month period (>3 SD from residual mean)

metrics to identify the anomalies relative to past data on patients of similar habitus, gender, and age, and exams with similar scan length and study type. With RE3's context-dependent outlier detection method but not with a simple threshold; for example, a pediatric scan with a dose four times too high but still low compared to adult scans can be detected. RE3 metrics SBV and D_w are also accurate for multiphase exams whereas BMI does not take into account scan length.

Limitations

The authors realize several general limitations. First, the results from our study might not be generalizable. RE3 was tested solely at one institution with mainly research protocols and relatively low volume. The performance and accuracy of RE3 elsewhere (e.g., in a more clinical setting) is not certain. Further, RE3 was tested solely on one PACS and the compatibility of the RE3 PACS interaction scripts or the DCMTK modules with PACS at different institutions is not certain. Second, the RE3 outlier system might not be suitable for all purposes. RE3 detects outliers based on past institutional data, but not based on levels explicitly set by the institution or outside entities. To achieve the latter, an evaluation team would have to manually analyze RE3 output relative to these objective levels. RE3, therefore, is useful for specific outlier detection, but may need to be used in conjunction with another dose monitoring system for more general dose monitoring purposes. Third, RE3 does not read RDSR reports; the ability to read RDSR could be incorporated into RE3 to improve dose calculation accuracy.

The authors also realize several specific limitations. RE3 might miss irradiation events that are not transferred to the PACS. Patients too large to fit in the large FOV may be underestimated and smaller FOV exams, such as cardiac exams, cannot be compared using SBV since they are coned down and not reflective of patient size. RE3 is not useful for scans with only a dose page and no dose data in the DICOM header (e.g., scans by legacy scanners). The regression model to account for overscan requires manual customization for new scanners; this would require effort proportional to the number of unique scanners at a particular institution. Last, Eq. (2), which was used for scans for which Eq. (1) was equal to zero (only high-resolution CT chest scans at our institution), may not be accurate for any scan in which Eq. (1) is zero (e.g., if a scan that is not a high-resolution CT chest has a scan length of zero, but also happens to have a reconstruction interval that is smaller than the slice thickness), and therefore should be used only if the study fits a particular description.

Future Work

RE3 does not include topograms. Although each topogram accounts for roughly the same exposure (an extra factor could

just be added), we would like to implement a method that actually reads and segments the image in anticipation of doing the same for x-rays.

OS did not require more than a week of training data because overscan was such a small addition to DLP and the error in the OS regression was very low. The OS regressions, however, can be retrained on more data to improve accuracy.

Availability of CTDI_{vol} in the DICOM header allows bypass of the dose pages and access to slice-level metrics for more advanced dose monitoring techniques, such as personalized Monte Carlo organ-dose estimation. During development, we collaborated with an National Cancer Institute (NCI) group to use RE3 slice-specific data and DICOM header values for incorporation of automated Monte Carlo organ-dose estimations, and we hope to make this a automated feature of RE3 in the future.

This paper focused on retrieval of the data rather than its storage, but a project is also underway to connect RE3 with a database for storage and reporting.

Further testing in settings outside of our institution would be useful to assess RE3's performance with different PACS and protocols. For this reason, we have made RE3 available online. The outlier detection system could be improved. Although RE3 automatically detects outliers based on past data, the data that it generates still requires manual analysis relative to diagnostic levels. In the future, it will be important to add these objective diagnostic thresholds (e.g., American College of Radiology (ACR), dose reference levels (DRL), or data from the dose index registry (DIR)) to the RE3 reports, or even to the RE3 model itself, so that the dose reduction team can identify outliers not only relative to past scans but also against these diagnostic levels. RE3 needs to be able to read RDSR. Now that many scans have an accompanying RDSR, it is important to amend RE3 to gather dose data from the RDSR, perhaps obviating the need to develop regression models for overscan and therefore reducing RE3 setup time. RE3 could then extract slice-specific doses from the DICOM headers in conjunction with the RDSR values. The backup method for calculating scan length needs to be more targeted. A study-level filter should be added that prohibits the use of Eq. (2) by all but exams with the study description "Hi-Res CT Chest." In the future, the exam-level dose calculation of these exams (and all exams) could be confirmed against RDSR values.

Acknowledgments The authors thank Sue Powell, MS, and Te Chen, MS, for PACS-related advice; head technologist Dennis Johnson and the NIH CT technologist team; and Roberto Moreno, PhD, for his insight.

Compliance with Ethical Standards

Funding This research was funded by the Intramural Research Program of the National Institutes of Health, Clinical Center.

References

- Mayo JR, Aldrich J, Müller NL: Radiation exposure at chest CT: a statement of the Fleischner Society 1. *Radiology* 228(1):15–21, 2003
- Brenner DJ, Hall EJ: Computed tomography—an increasing source of radiation exposure. *N Engl J Med* 357(22):2277–2284, 2007
- Smith-Bindman R, Lipson J, Marcus R, et al: Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. *Arch Intern Med* 169(22):2078–2086, 2009
- de González AB, Darby S: Risk of cancer from diagnostic X-rays: estimates for the UK and 14 other countries. *Lancet* 363(9406):345–351, 2004
- Shuryak I, Hahnfeldt P, Hlatky L, et al: A new view of radiation-induced cancer: integrating short-and long-term processes. Part I: approach. *Radiat Environ Biophys* 48(3):263–274, 2009
- Pearce MS, Salotti JA, Little MP, et al: Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *Lancet* 380(9840):499–505, 2012
- Sodickson A: CT radiation risks coming into clearer focus. *BMJ* 346:f3102, 2013
- Dijkstra H, Groen JM, Bongaerts FA, et al: The cumulative risk of multiple CT exposures using two different methods. *Health Phys* 106(4):475–483, 2014
- Shore RE: Radiation impacts on human health: certain, fuzzy, and unknown. *Health Phys* 106(2):196–205, 2014
- Lin Z: Cancer risk from low dose radiation depends directly on the organ mass in a general model of radiation-induced cancer risk. *Health Phys* 106(4):459–465, 2014
- Sodickson A, Baeyens PF, Andriole KP, et al: Recurrent CT, cumulative radiation exposure, and associated radiation-induced cancer risks from CT of adults 1. *Radiology* 251(1):175–184, 2009
- Brenner DJ, Doll R, Goodhead DT, et al: Cancer risks attributable to low doses of ionizing radiation: assessing what we really know. *Proc Natl Acad Sci* 100(24):13761–13766, 2003
- [Cited 2015 June]. Available from: <http://www.aapm.org/meetings/amos2/pdf/42-12234-94897-404.pdf>
- California Senate Bill 1237, S. 115111, 115112, and 115113 California Health and Safety Code, 2010
- Lukasiewicz A, Bhargavan-Chatfield M, Coombs L, et al: Radiation dose index of renal colic protocol CT studies in the United States: a report from the American College of Radiology National Radiology Data Registry. *Radiology* 271(2):445–451, 2014
- Rehani MM: Challenges in radiation protection of patients for the 21st century. *Am J Roentgenol* 200(4):762–764, 2013
- AAPM. AAPM recommendations regarding notification and alert values for CT scanners: guidelines for use of the NEMA XR 25 CT dose-check standard. [Cited 2015 June]. Available from: http://www.siiim2012.org/abstract_SSC_Chadalavada.shtml
- Seetharam Chadavada M, Anand Sundaram MS, Woojin Kim MD, William W. Boonn MD, Tessa Cook MD: A RADIANCE-based alerting system for CT radiation dose estimates society for imaging informatics in medicine 2012 scientific session. Orlando, 2012
- Smith-Bindman R, Moghadassi M, Wilson N, Nelson TR, Boone JM, Cagnon CH, et al: Radiation doses in consecutive CT examinations from five University of California Medical Centers. *Radiology* 142728, 2015
- Mueller JW, Vining DJ, Jones AK, et al: JOURNAL CLUB: In vivo CT dosimetry during CT colonography. *Am J Roentgenol* 202(4):703–710, 2014
- Cook TS, Zimmerman SL, Steingall SR, et al: Informatics in radiology: RADIANCE: an automated, enterprise-wide solution for

- archiving and reporting CT radiation dose estimates. *Radiographics* 31(7):1833–1846, 2011
22. Sodickson A, Warden GI, Farkas CE, et al: Exposing exposure: automated anatomy-specific CT radiation exposure extraction for quality assurance and radiation monitoring. *Radiology* 264(2):397–405, 2012
23. Talati RK, Dunkin J, Parikh S, et al: Current methods of monitoring radiation exposure from CT. *J Am Coll Radiol* 10(9):702–707, 2013
24. Tsalafoutas I, Metallidis S: A method for calculating the dose length product from CT DICOM images. *Br J Radiol* 84:236–243, 2011
25. Nicholson R, Fetherston S: Primary radiation outside the imaged volume of a multislice helical CT scan. *Br J Radiol* 75(894):518–522, 2002
26. Tsalafoutas IA: The impact of overscan on patient dose with first generation multislice CT scanners. *Phys Med* 27(2):69–74, 2011
27. Medicine AAoPi. Use of water equivalent diameter for calculating patient size and size-specific dose estimates (SSDE) in CT (Task Group 220). College Park, 2014
28. <http://dicom.offis.de/>. DCMTK-DICOM Toolkit, 2013
29. McCollough CH, Leng S, Yu L, et al: CT dose index and patient dose: they are not the same thing. *Radiology* 259(2):311–316, 2011
30. McCollough C, Cody D, Edyvean S, Geise R, Gould B, Keat N, et al: The measurement, reporting, and management of radiation dose in CT. Report of AAPM Task Group. 23:1–28, 2008
31. Sussman DL, Yao J, Summers RM: Automated fat measurement and segmentation with intensity inhomogeneity correction. SPIE Medical Imaging International Society for Optics and Photonics, 2010