DICOM Image quality and harmonization in MIDRC

Paul Kinahan, PhD
University of Washington
Technology Development Projects (TDPs)

TDP3: Developing and implementing quality assurance and evaluation procedures for usage across the MIDRC

• Spans from data ingestion through data access and dissemination
• Includes data quality metrics, data provenance, data processing audits paths, and task-directed data distribution
• Procedures are intended to help enable research and translation with MIDRC data
• Developing harmonized data definitions and labeling methods, hosting of data science challenges, and benchmarking of algorithm performance
TDP 3 Activities to develop and implement MIDRC quality assurance and evaluation procedures

TDP3a: Develop COVID phantoms (digital and physical)

TDP3b: Conduct measurement of CT & CXR image quality & harmonization techniques for image data ingestion

TDP3c: Develop benchmarking methods for technology assessment & clinical tasks in COVID-19 research and translation

TDP3d: Develop task-based distribution methods
TDP3a: Development of digital and physical imaging phantoms (John Boone)

- Evaluation of existing general (ACR) and specific (Corgi) phantoms
- Use of automated image analysis tools
- Developing a digital phantom to assess impact of de-identification methods
Analysis of Image Quality

Corgi Report

Input Image

TOSHIBA Aquilion Precision
Study Date: Jul 31, 2020
Study Time: 12:25 PM
Axial FOV: 37 cm
Longitudinal FOV: 31 cm
Tube Voltage: 120 kV
Tube Current: 400 mA
CTDIvol: 14.5 mGy

Uniformity

Mean value: 0.2 AU
Cupping artifact: 11.9%
Uniformity map

Spatial Resolution

\( \text{ Spatial Resolution } \)

- \( d_{50} \text{ axial: 0.45 mm} \)
- \( d_{50} \text{ sag: 0.77 mm} \)
- \( f_{50} \text{ axial: 0.73 mm}^{-1} \)
- \( f_{50} \text{ axial: 0.06 mm}^{-1} \)

Noise

Mean noise: 20.8 AU

Cone-Beam Artifact

\( \text{ Z max at -2.2 cm: 55%} \)
\( \text{ Z max at -7.8 cm: 32%} \)

Corgi Modules

Module count: 5
Phantom x,y offset: 0.5 cm
Phantom z tilt: 0.1°
TDP3b: Assessment of image quality on ingestion into MIDRC

Two main components:

1. Assessment of the images themselves where there are manual and semi-automated approaches available to search for and assess known structures and artifacts (e.g. from implants)

2. Use of the image meta-data that is stored in the DICOM image
   • An exploratory component is to combine parameters derived from the images with the meta-data to general image quality metrics

Efforts on evaluating image data quality early in the ingestion process are driving the efforts of the Data Quality and Harmonization SC
Explorer data model for cohort selection
## Cohort selection characteristics

The following list includes the variables (properties) at the IMAGE SERIES (aka Exam) level. Users will be able to search on these data for all IMAGE SERIES. These data are defined as, "information related to a CR/DX/CT imaging series to which one or more computed radiographs/digital radiographs (x-rays)/a stack of CT slices belong."

Note: While we understand most of this information is provided within each Image File, we want to limit the amount of searching done at that level as this could cause performance issues. Ideally, we want to eliminate anything that is completely or heavily redundant for each image. In other words, these properties are going to be "rolled up" from the image files and some of these properties are arrays because more than one value might be relevant for a single series.

<table>
<thead>
<tr>
<th>Node</th>
<th>Property</th>
<th>Description</th>
<th>Data type</th>
<th>MRI</th>
<th>CT</th>
<th>CR</th>
<th>DX</th>
<th>Include as filter</th>
<th>Needs lookup table</th>
<th>Required</th>
<th>DICOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>series</td>
<td>A general description of or additional information about the imaging series.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>series_uid</td>
<td>The identification code or scan number of the imaging series.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>series_acquisition_type</td>
<td>Sequence type.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>series_contrast_bolus_agent_number</td>
<td>Contrast bolus agent number.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>series_convolution_kernel</td>
<td>Convolution kernel.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>series_detector_type</td>
<td>Detector type.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>series_exposure_modulation_type</td>
<td>Exposure Modulation Type.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>series_image_type</td>
<td>Image type.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>series_lossy_image_compression</td>
<td>Lossy image compression.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>series_magnetic_field_strength</td>
<td>Magnetic field strength.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>series_manufacturer</td>
<td>Manufacturer.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>series_manufacturer_model_name</td>
<td>Manufacturer's Model Name.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>series_modality</td>
<td>Modality.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>series_pixel_spacing</td>
<td>Pixel Spacing.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>series_scanning_sequence</td>
<td>Scanning Sequence.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>series_slice_thickness</td>
<td>Slice Thickness.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>series_spacing_between_slices</td>
<td>Spacing Between Slices.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>series_spacing_resolution</td>
<td>Spatial Resolution.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>series_view_position</td>
<td>View Position.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>series_data_format</td>
<td>Format of the data files.</td>
<td>string</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>node</td>
<td>property</td>
<td>raw_value (submitted by contributors)</td>
<td>raw_count</td>
<td>proposed mapped midrc_value (RadLex Playbook)</td>
<td>Coding Scheme</td>
<td>Code</td>
<td>difference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------------------------</td>
<td>---------------------------------------</td>
<td>-----------</td>
<td>---------------------------------------------</td>
<td>---------------</td>
<td>-----------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>XR CHEST 1 VIEW AP</td>
<td>477</td>
<td>XR CHEST 1V (RadLex Playbook)</td>
<td></td>
<td>RIPI2502</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>CXR</td>
<td>432</td>
<td>XR CHEST (non-RIP)</td>
<td></td>
<td>RIPI2502</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>XR CHEST AP PORTABLE</td>
<td>204</td>
<td>XR CHEST 1V</td>
<td></td>
<td>RIPI2502</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>CHEST 1V</td>
<td>76</td>
<td>XR CHEST 1V</td>
<td></td>
<td>RIPI2502</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>CT CHEST</td>
<td>58</td>
<td>CT Chest</td>
<td>LOINC</td>
<td>24627-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>THORAX PE</td>
<td>34</td>
<td>CT CHEST W EMBOLUS</td>
<td></td>
<td>RIPI147</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>XR CHEST AP ONLY</td>
<td>19</td>
<td>XR CHEST 1V</td>
<td></td>
<td>RIPI2502</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>CT PE CHEST</td>
<td>17</td>
<td>CT CHEST W EMBOLUS</td>
<td></td>
<td>RIPI147</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>THORAX/ABDOMEN/PELVIS + CONT</td>
<td>14</td>
<td>CT CHESTAB/PELV W</td>
<td></td>
<td>RIPI2503</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>XR CHEST 2 VIEWS PA AND LATERAL</td>
<td>13</td>
<td>XR CHEST 2V</td>
<td></td>
<td>RIPI2503</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>CT CHEST PULMONARY EMBOLISM (CTPE)</td>
<td>13</td>
<td>CT CHEST W EMBOLUS</td>
<td></td>
<td>RIPI147</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>Non-contrast Chest CT</td>
<td>11</td>
<td>CT CHEST WO</td>
<td></td>
<td>RIPI16</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>Low-close Chest CT</td>
<td>7</td>
<td>CT CHEST WO SCREEN</td>
<td></td>
<td>RIPI6002</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>THORAX PE + LOW DOSE</td>
<td>5</td>
<td>CT CHEST W EMBOLUS</td>
<td></td>
<td>RIPI147</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>CT CHEST ABDOMEN PELVIS W</td>
<td>5</td>
<td>CT CHESTAB/PELV W</td>
<td></td>
<td>RIPI2503</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>XR CHEST PA AND LATERAL</td>
<td>5</td>
<td>CT CHEST 2V</td>
<td></td>
<td>RIPI2503</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>THORAX + CONT</td>
<td>3</td>
<td>CT CHEST W</td>
<td></td>
<td>RIPI18</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>CT ANGIOGRAM CHEST</td>
<td>3</td>
<td>CT CHEST ANGIO W/W/O</td>
<td></td>
<td>RIPI6</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>XR ACUTE ABDOMINAL SERIES W PA CHEST</td>
<td>2</td>
<td>XR ABD SUPINE/ERECT+CHEST</td>
<td></td>
<td>RIPI6044</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>XR CHEST PA ONLY</td>
<td>2</td>
<td>XR CHEST 1V PA</td>
<td></td>
<td>RIPI6035</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>CARDIAC CORONARY CTA</td>
<td>1</td>
<td>1 CT HEART CORONARY ANGIO W</td>
<td></td>
<td>RIPI6003</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>CT CHEST HIGH RESOLUTION</td>
<td>1</td>
<td>1 CT CHEST W LUNG PARENCHYMIA</td>
<td></td>
<td>RIPI6001</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>CT PE</td>
<td>1</td>
<td>1 CT CHEST EMBOLUS</td>
<td></td>
<td>RIPI147</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>imaging_study</td>
<td>study_description</td>
<td>CT CHEST WO SCREEN</td>
<td>1</td>
<td>1 CT CHEST W O SCREEN</td>
<td></td>
<td>RIPI6002</td>
<td>TRUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TDP 3b Data Quality Activities

The MIDRC will use data management methods harmonized across all participating organizations at three critical stages:

1. intake, including curation, de-identification, abstraction, and quality assessment
2. annotation and labelling of imaging and other data using semi-automated approaches
3. distributed access and query methods
Schematic of the DICOM image data quality SOP

MIDRC Data Flow: Deidentification, curation, annotation, other manipulations

Legend:
- DICOM object
- Service
- Other data
Cohort selection characteristics

Current selection:
- Cases: 588
- Studies: 1,238
- Series: 1,590
- Instances: 54

Manufacturer (series-level):
- Agfa: 33.8%
- Philips: 55.4%
- Others: 6.7%

Reconstruction Diameter:
- 360.0: 36.0%
- 360.0000: 13.4%
- Others: 6.7%

Manufacturer model name (series-level):
- Agfa: 33.8%
- CR 95: 55.4%
- Brilliance 64: Others: 6.7%

Modalities (series-level):
- CR: 45.3%
- DX: 32.8%
- CT: 21.9%

KVP:
- 94: 23.4%
- 949: 23.4%
- 1828: 6.7%
- Others: 6.7%

Distance to Detector:
- 0: 64.5%
- 1: 64.5%
- 2.2: 35.5%
- Others: 6.7%

Spiral Pitch Factor:
- 1.2: 50.1%
- 1.5: 9.7%
- Others: 6.7%

CTDvol:
- 0.68: 45.9%
- 1.2: 38.6%
- Others: 6.7%
Added DICOM Image Quality Metrics – Under Consideration

- mobile_xray_unit (assuming we can figure out a way to do this)
- missing_slices/images (can we do this automatically?)
- image_artifacts (similarly can we do this automatically?)
- image_modified (e.g. altering burned-in information)
- incorrect_header_information (e.g. wrong body_part_examined NOT SURE (won't we remove these cases?)
- MIDRC_acceptable_image_quality (assuming we can figure out how to do this automatically on a large scale)
- phantom_image
- MIDRC_high_noise_image (assuming we can figure out a way to do this automatically)
- MIDRC_low_resolution_image (assuming we can figure out a way to do this automatically)
- study_description_mapped
- manufacturer_mapped
- body_part_examined_mapped
- annotations_provided
- deidentification_profile_used
- CT slice thickness categories (thin, medium, thick)
- CT dose categories (ultra low dose (CTD\text{vol}<1mGy), low dose (1mGy< CTD\text{vol}< 3mGy), modest dose (3 mGy<CTD\text{vol}< 10 mGy), diagnostic dose (10mGy< CTD\text{vol}).
Synergies between AAPM TDP 3 and CRPs 9-12

TDP 3a Develop COVID phantoms (Boone)
• Physical and digital phantoms to assess quality
TDP 3b Measure and Harmonize image quality (Kinahan)
• Assessment of images and meta data
TDP 3c Benchmarking methods (McNitt-Gray)
• Clarification paper on COVID-19 task-based performance metrics
• Online decision tree for performance metrics; benchmarking
TDP 3d Task-based distribution methods (Myers)
• Task-based distribution methods for Challenges and industry
• Initiated MIDRC Bias & diversity working group
• Assessing diversity in incoming data
• Developing sequestering algorithm

CRP 9 AI for detection and diagnosis (Armato)
• Grand Challenges
• Development / testing of AI diagnostic methods
CRP 10 Visualization & Explainability of AI (Giger)
• "Lessons learned”; AI vs. AI med imaging vs. AI COVID-19 imaging
• “Explainability & Interpretability” paper
• Development / testing of AI prognostic methods
CRP 11 Image-based biomarkers for radiogenomics (Chen)
• Enclaves for integrating and interoperability of multi-modality (multi-omics) data sources
• Containerizing techniques (how to) as MIDRC connects with other COVID-19 registries
CRP 12 Data quality, provenance, and harmonization (Fedorov)
• Introduction of active tracking methods w DICOM