Result Review System for Clinical, Research and Clinical Trial Imaging

DICOM SR, PS, RT, Segmentation and Proprietary Formats

David A. Clunie, MBBS, PixelMed Publishing, LLC and CoreLab Partners, Inc.

Motivation …

Ease of review of results for multiple patients and time points, regardless of the tools used to create annotations and derive quantitative information..

Scenarios include:

- ❑ Clinical practice for individual patients
- ❑ Clinical trials of conditions that change over time
- ❑ Imaging biomarker development
- ❑ Comparison of different techniques
- ❑ Comparison of different tools
- ❑ Background air suppression - for windowing & inversion

No industry consensus on a single universal format for all applications capable of encoding or modeling all scenarios

DICOM has standardized objects to store structured, numeric and coded content, image and 3D coordinates, rendered appearance and application-specific structure sets (such as for RT). Implementation is inconsistent and the amount and form of information varies, and is often not interchangeable between applications despite the same underlying format or template.

These objects include:

- ❑ DICOM Structured Reports (SR)
- ❑ DICOM Segmentations (SEG)
- ❑ DICOM Radiotherapy Structure Sets (RTSS)
- ❑ DICOM Presentation States (PS)

Waiting for consensus, the "perfect" format, or expecting existing tools to change are not viable short-term options.

The Solution …

Leverage existing standards but also support proprietary variants as well as entirely proprietary objects.

Many tool developers and project or collaborative groups invent their own specific formats, binary, text or XML-based, but with a unique and specific schema or information model. Given documentation or a recognizable patterns, one can convert these generically or on a project-specific basis.

E.g., the Lung Image Database Consortium (LIDC) made up their own XML format for encoding "probability maps" of truth estimates based on multiple readers' outlines; one can convert these to DICOM SEG objects (for the map) referenced by DICOM SR objects (for the measurements and context).

The relevant "context" of each measurement or annotation is defined by the use-case, regardless of the encoding.

Each scenario, project or experiment shares common concepts such as patient (or subject), time point (date of study), lesion (recognizable over time), reader (human or machine performing the measurement), etc. Regardless of the encoding, or organization into single or multiple files, these can be extracted into a single, simple tabular model without losing the relationship to images and coordinates from which measurements were made.

Common DICOM SR template as an intermediate form.

Some tools natively produce an oncology-oriented, cumulative, hierarchically organized, reader and time-point based DICOM SR with coordinates or referenced DICOM SEG objects; a simplified form of this is used as the intermediate form when converting other DICOM objects (like RTSS) or proprietary objects (like LIDC) to DICOM SR, using defined codes for context, before tabulating.

Finding values and context to convert can be challenging.

Absent explicit structured values, one may need to extract these from filename or folder hierarchy, or plain text format conventions.

The Table and the Plot …

Tabular presentations are a natural form for representing and navigating multi-dimensional quantitative data.

- ❑ Lowest node in the hierarchy is a row
- ❑ Measurements and observations are columns (siblings)
- ❑ Higher level nodes are replicated for each row (parents)
- ❑ Attributes of higher level nodes are also columns (context)
- ❑ Rows or cells are hyperlinked to rendered images

Plots of numerical and categorical data provide visual insight into the meaning of the information.

- ❑ Easily produced from tabulated values
- ❑ Can be stratified by selected column values
- ❑ Can be annotated with derived statistical information
- ❑ Features (e.g., outliers) can by hyperlinked to images

The Process Flow…

Pattern of storage, organization, structure and encoding of relevant values is identified for each project or data source.

- ❑ Identify information model and objects used for data
- ❑ Derive classes from tools to convert to standard model
- ❑ Tabulate values and context from standard model
- ❑ Derive generic or project-specific statistics
- ❑ Plot relevant generic or project-specific values or statistics

Subject	Reader	Time Point	Lesion	Volume mm3
0001	Reader 1	2010/06/01	1	355
0001	Reader 2	2010/06/01	1	375
0001	Reader 1	2010/06/01	2	3896
0001	Reader 2	2010/06/01	2	4764
0001	Reader 1	2011/07/01	1	471
0001	Reader 2	2011/07/01	1	289
0001	Reader 1	2011/07/01	2	3801
0001	Reader 2	2011/07/01	2	4834