telling's T2 control chart were better than those of the other method, with a diagnosis of 99.49% and accuracy of 99.51%.

Conclusion: In this paper, instead of classification and clustering methods, for the first time, we used a multivariate control chart with the Hotelling's T2 statistic for the diagnosis of patients suspected of bone marrow metastasis. Then, using some patient samples, the performance of this phase I control chart was evaluated, and the results showed the validity of the proposed method. The validation results revealed that the accuracy and specificity metrics were better for the multivariate Hotelling's T2 control chart than for the fuzzy clustering method.

Keywords: Two-Dimensional Discrete Wavelet Transform, Bone Marrow Metastases; Multivariate Hotelling's T2 Control Chart; Fuzzy Clustering; Feature Extraction

• 10.5812/IRANJRADIOL.99147

Multi-Institutional Medical Imaging Research Data Collection: Challenges of Standardization of Protocols and Header Information to Make an Imaging Biobank

Fattane Pourakpour[1,*]; Bahhare Siahloo[1]; Shaghayegh Karimi[1]; Zeinab Khodakarami[1]; Mansoor Fatehi[1]

[1]*Iranian Brain Mapping Biobank, National Brain Mapping Laboratory, Tehran, Iran*

**Corresponding author: Iranian Brain Mapping Biobank, National Brain Mapping Laboratory, Tehran, Iran. Email: fattane.pourakpour@gmail.com*

Abstract

Background: Iranian brain mapping biobank (IBMB) has three major sources of data to be processed, categorized, and tagged for recurrent use by researchers, including research image acquisition, large scale cohort studies, and routine clinical samples from collaborating institutions. A major limitation of samples coming from routine clinical centers is the potential diversity of data parameters that may prevent to merge databases create biobanks.

Objectives: The study was performed to find out the reliability of a multi-institutional case collection.

Methods: Voluntary case collection was performed from four institutions that signed an agreement with the national brain-mapping laboratory. The centers operated machines from different vendors including Siemens and GE and the scanning protocols were diverse according to the operating technologists. An in-house developed application based on MATLAB 2018b was used to extra DICOM header information from the donated studies. All DICOM headers were imported to a unified database to be analyzed according to the modality, vendor type, and operator protocols.

Results: A total number of 1581 cases were entered into the project with 2414 procedures performed over a six-month period. This collection included 199509 series of images for which all tags were extracted. Except for the modality-specific tags (Table 1), all other tags were found to be uniform regardless of the machine and protocol. The most important tag diversity was seen in the MRI scanning parameters: "Protocol Name", "Scan Options", "Scanning Sequence", "Sequence Name", and "Sequence Variant".

Table 1. List of Modality-specific Tags Unshared Between All Studies

MRI Specific Tags	CT Specific Tags	Radiography Specific Tags
Acquisition Matrix	Convolution kernel	Burned in annotation
Angio Flag	Data collection diameter	Detector activation offset from
dBdt	Distance Source to Detector	Exposure
Echo number	Distance source to patient	Detector active time
Echo time	Exposure	Exposure control mode
Echo train length	Exposure modulation type	Exposure in mAs
Flip angle	Exposure time	Exposure in uAs
Heart rate	Filter type	Exposure time in ms
Imaged nucleus	Focal spot	Exposure time in μs
Imaging frequency	Gantry detector tilt	Field of view shape
In-plane phase encoding direction	Generator power	Focal spot
Inversion time	KVP	Image and fluoroscopy area dose

Laterality	Rotation direction	Product
MR acquisition type	Spatial resolution	Images in acquisition
Number of averages	Table height	KVP
Number of phase encoding steps	X-ray tube current	Pixel intensity relationship
Percent phase field of view		Pixel intensity relationship sign
Percent sampling		Quality control image
Pixel bandwidth		X-ray tube current in mA
Repetition time		X-ray tube current in μA
SAR		
Scanning sequence		
Sequence name		
Sequence variant		
Spacing between slices		
Transmit coil name		
Variable flip angle flag		

Conclusion: DICOM 3.0 standard has an invaluable role in the standardization of information incorporated into the image files as a header, which makes multi-institutional data collection feasible to a large extent. Practice-based data elements need to be unified at the time of acquisition or at the time of importing samples into a biobank to make the dataset homogeneous.