

# Evaluation of Irreversible Compression Ratios for Medical Images Thin Slice CT and Update of Canadian Association of Radiologists (CAR) Guidelines

David Koff · Peter Bak · André Matos · Geoff Norman

Published online: 21 November 2012  
© Society for Imaging Informatics in Medicine 2012

**Abstract** In June 2008, the Canadian Association of Radiologists published its Standards for Irreversible Compression in Digital Diagnostic Imaging within Radiology (Canadian Association of Radiologists 2012). The study suggested that at low levels of compression there was no difference in diagnostic accuracy between uncompressed JPEG and JPEG 2000. There were two exceptions; CT neurological and CT body images resulted in lower rating of image quality (Koff et al., J Digit Imaging 22(6):569–78, 2009). The slice thicknesses used in the previous study were greater than 5 mm. However, other studies (Ringl et al., Radiology 240:869–87, 2006) suggest that thin CT slices might modify image tolerance to irreversible compression. Therefore, a new clinical evaluation using CT slices less than 3 mm was initiated. We examined CT images in four body regions (chest, body, musculoskeletal, and neurological). Twenty-five radiologists from across Canada participated. Each read a total of 70 CTs in his specialty; 10 at each of seven levels of compression (uncompressed, JPEG and JPEG 2000 at low, medium, and high compression (varying by region)). Each reader diagnosed the case, rated his confidence, and compared the compressed to the uncompressed image and rated the degree of degradation. Data were analyzed for sensitivity, specificity, accuracy, confidence, and degradation at three levels and two types of compression as well as the original image. There were no overall differences in sensitivity, specificity, accuracy, or confidence. JPEG images, at all

levels of compression, were rated lower in terms of perceived difference (4.16/5 vs. 4.53/5 for JPEG 2000 and 4.68/5 for uncompressed). However, the rating of perceived difference was not significantly correlated with accuracy. Analysis of individual body regions did not reveal any systematic effects of compression in any region.

**Keywords** Compression · CT

## Introduction

Access to imaging data is becoming more widespread as regions/nations implement Electronic Health Record (EHR) solutions. Expectations for accessibility are changing with physicians requiring access to images dating back many years in near real-time. To satisfy these expectations, EHR projects are driving the deployment of imaging repositories that host data online for the legal retention period and in some instances for the lifetime of the patient.

New technological advancements including multislice CT scanners and functional MRI have dramatically increased the size and number of digital images generated by medical imaging departments. CT scan images represent up to 50 % of a Picture Archiving and Communication System (PACS) storage space; with each CT study consisting of hundreds of slices, easily more than 1,000 images. Storage needs are predicted to further increase when departments are required to archive 3D volume renderings and other complex advanced image reformats of the original images.

Despite the fact that the cost of storage is dropping, the savings are largely surpassed by the increasing volume of data being generated. In addition, the cost of operation remains high, with a mandatory data migration and legal retention period of 7 years for adult patients and 7 years past the age of consent for adolescence. Considering the fact that

---

D. Koff · P. Bak · A. Matos  
Department of Radiology, McMaster University,  
Hamilton, Canada

G. Norman (✉)  
Department of Clinical Epidemiology and Biostatistics,  
McMaster University,  
MDCL 3519, 1280 Main St. W,  
Hamilton, ON L8S 2T1, Canada  
e-mail: norman@mcmaster.ca

Canada annually generates 1.5 PB of radiology imaging data, the potential national metrics are significant [4].

While local area network bandwidth within a hospital is adequate for timely access to imaging data, efficiently moving the data between institutions requires wide area network bandwidth, which has a limited availability at a national level. As such, EHR networks cannot support timely access and distribution of large medical images. Conservative levels of data compression can address the storage and transmission needs by enabling more efficient distribution and optimizing archiving of imaging data.

The goal of this study was to determine levels at which irreversible compression can be confidently used in diagnostic imaging applications for thin slice CT images (less than 3 mm thickness). The compression levels investigated were based on accepted levels published in the Canadian Association of Radiologists (CAR) Standards for Irreversible Compression published in June 2008 after a large-scale pan-Canadian evaluation study [2] which did not include in its scope thin slice CT images less than 5 mm.

It has been demonstrated that thin slices are less tolerant to JPEG 2000 compression than thick slices as they have smaller peak signal-to-noise ratio (PSNR) and more visible artifacts because of more random noise due to less photons contributing to the reconstruction of a voxel [5]. Woo et al. [6] stated that the perceived artifacts, more pronounced in thin-section images, have a blurring effect mainly altering the structures of the solid organs and soft tissues. Yamamoto et al. [7] reported that 2-mm thick high-resolution lung CT images are more vulnerable to compression than 10-mm thick conventional lung CT images on the basis of the comparison of PSNR and subjective rating of the image quality. Siddiqui et al. [8] and Siegel et al. [9] reported that thinner sections are less compressible, based on their observations of the PSNR and a computer-based perceptual metric [10] in five chest CT data sets compressed with the JPEG 2000 algorithm. Thus, there is ample evidence that image compression may have a differential effect on thin slice CT images.

One study has directly examined the effect of digital compression on thin slice CT. Ringl et al. [3] asked their readers, for each image pair, to select the image with the best image quality then rate the difference between the compressed and uncompressed image. Their approach was to select the ideal compression ratio as the one with no visual difference between the original and compressed image. However, the relation between visually detectable image degradation and diagnostic accuracy is unknown, as the authors admit.

In the present study, we explicitly examined this question. We began with a large variety of subtle conditions in diverse body parts. Each lesion was then reproduced under varying compression conditions, then diagnosed by

experienced radiologists to determine if compression would have an impact on diagnostic accuracy. We also asked radiologists to rate the degree of degradation and their confidence.

## Materials and Methods

### Materials

The study considered the effects of compression in four radiological areas in thin slices of CT: body, chest, musculoskeletal (MSK), and neurological. Two hundred eighty retrospective examinations were selected from the Hamilton Health Sciences PACS and used to retrieve 4,200 anonymized and uncompressed images. The images were acquired using Toshiba Aquilion 16 Slice, Toshiba Aquilion 64 Slice (Toshiba Medical Systems Corporation, Toshihigi, Japan), and GE VCT 64 Slice (General Electric Healthcare, Milwaukee, WI, USA). Extreme caution was taken to ensure that all the images were exported in original size noncompressed format.

We retrieved 15 slices for each selected exam. In addition, all retrieved images had a slice thickness up to 3 mm. The body cases had slice thickness between 2 and 3 mm, chest (including high resolution and vascular) cases between 1 and 2 mm, musculoskeletal cases between 1 and 2 mm, and neuro cases between 1 and 3 mm. Reporting was performed by all radiologists on a pair of 2 MegaPixels flat panel Digital Imaging and Communications in Medicine (DICOM) monitors (EIZO NANA Corporation, Ishikawa, Japan).

As much as possible, we selected subtle abnormalities for each radiological area, which could be overlooked if the image was somewhat altered, or over compressed. We also tried to restrict the abnormalities to three to four possibilities, however due to the nature of certain radiological areas, including the numerous types of abnormalities encountered, this was not always possible. Here is a full list of abnormalities per radiological area:

*Body abnormalities:* abscess, collection, cyst, gallstone, hemangioma, liver contusion, liver lesion, mass, perihepatic fluid, scar, and subcutaneous emphysema

*Chest abnormalities:* atelectasis, bronchiectasis, fibrosis, ground glass opacity, mass, nodule, pulmonary embolism, pleural plaque, and tree in bud

*MSK abnormalities:* bone lesion and subtle fracture

*Neuro abnormalities:* aneurysm, arteriovenous malformation, hemorrhage, infarct, ischemia, mass, pneumocephalus, sinus thrombosis, stroke, and subgaleal hematoma

**Table 1** Recommended compression ratio by modality and specialty

	CR/DR	CT	US	MR	NM	MG
Angio		10–15		16–24		
Body	20–30	JPEG 10–15 J2K 10	8–12	16–24	9–11	
Breast			8–12	16–24		15–25
Chest	20–30	10–15				
MSK	JPEG 20–30 J2K 20	10–15	8–12	16–24		
Neuro		JPEG 8–12 J2K 8		16–24		
Pediatrics	20–30	10–15	8–12	16–24	9–11	

Images were randomly assigned one of seven possible compression ratios: uncompressed, JPEG low, JPEG medium, JPEG high, JPEG 2000 low, JPEG 2000 medium, or JPEG 2000 high. Compression ratios were based on values determined in the CAR Standards for Irreversible Compression in Digital Diagnostic Imaging.

We investigated the effects of compression on thin slice CT images in four anatomical areas of interest (body, chest, MSK, and neuro). The evaluation consisted of four independent collections of images or worklists, which have been viewed by a minimum of five readers each.

#### Compression Technique

Ensuring diagnostic quality in compressed images is imperative. Significant advances have been achieved in compression algorithms, moving from JPEG, to wavelet, and now to JPEG 2000. Irreversible compression allows a greater size reduction with no significant visual quality loss or image degradation, and the severity of the degradation is strictly dependent on the compression ratio. JPEG is the most widely accepted compression tool, but it has been shown that the JPEG-2000 algorithm may provide higher compression levels than JPEG at an equivalent or higher image quality [11]. The new features in JPEG 2000 compared to JPEG are [7]:

1. Efficient irreversible and reversible compression within a single unified coding framework

2. Progressive transmission and spatial scalability (thumbnails)
3. Superior image quality; broad range of image types
4. Support for region of interest coding
5. Support for continuous tone and bilevel compression (BW and color)
6. Robustness to bit errors (wireless communication applications)
7. Avoids excessive memory usage

In order to provide a fair assessment of existing compression tools, we tested and compared the two most commonly adopted DICOM compression algorithms: JPEG and JPEG 2000. We used compression ratios for each modality and specialty recommended in the CAR Standards as shown in Table 1. As in our previous study, we used a widely adopted commercial compression package, PICTools JPEG 2000 from Pegasus Imaging (Accusoft Pegasus, Tampa, FL, USA; [12]).

#### Compression Ratios

Based on results from our previous study, we tested three different levels of compression for both JPEG and JPEG 2000 compression algorithms at 3 different compression ratios: 6:1 (low), 9:1 (medium), and 12:1 (high). By virtue of the fact that our study included both JPEG and JPEG 2000, we were able to compare the performance of the two compression algorithms. In addition to compressed images, each set of images also contained uncompressed images for

**Table 2** Results of the study for sensitivity, specificity, accuracy, confidence and perceived difference averaged over all body systems

Variable	Uncompressed	JPEG			JPEG 2000		
		Low	Med	High	Low	Med	High
Sensitivity	0.70	0.60	0.52	0.68	0.59	0.68	0.60
Specificity	0.61	0.46	0.80	0.61	0.65	0.57	0.80
Accuracy	0.70	0.58	0.57	0.66	0.60	0.65	0.61
Confidence	4.10	4.00	4.10	3.88	4.06	4.17	3.96
Perceived difference	4.68	4.22	4.12	4.11	4.67	4.67	4.25

**Table 3** Analysis of sensitivity by body system

Body system	Uncompressed	JPEG	JPEG 2000	<i>P</i> (type)	<i>P</i> (Lvl)	<i>P</i> (type×Lvl)
Chest	0.694	0.584	0.575	ns	ns	<0.001
Body	0.714	0.589	0.667	ns	ns	<0.001
MSK	0.625	0.622	0.644	ns	ns	ns
Neuro	0.771	0.602	0.600	ns	ns	ns
Overall	0.701	0.600	0.621	ns	ns	<0.05

evaluation. Each work list included 70 cases, each one having a stack of images for CT scans, representing a ratio of six compressed for one uncompressed image. The entire set of images was randomized and readers were not shown the same image to evaluate twice. Images were collected from other collaborating hospitals in the Hamilton Health Sciences network to ensure that our findings could be generalized to images generated by different brands of acquisition equipment.

Measurement

Our methodology included two recognized evaluation techniques: a subjective assessment based on comparison of original and compressed images, and an objective assessment of diagnostic accuracy. In the subjective assessment, readers compared original and compressed images and rated the degree of degradation. In the objective assessment, readers were presented with images consisting of a mix (one normal/four abnormal) of normal and abnormal cases.

1. Diagnostic accuracy evaluation: Images consisting of a mix of normal cases and identified pathologies were presented to reviewers for their assessment. Using normal images was important as it (1) allowed us to examine the possibility that artifacts may simulate abnormalities in a normal image and (2) it will control for a possible bias that may have occurred if readers knew that each image contains an abnormality. Images were displayed with a grid overlay, which divided the image into four equal quadrants. The grid could be toggled on and off by the reader during the evaluation. Using the grid and drop-down menus, readers were required to identify the type and location of the lesion.

Radiologists then provided a confidence rating of 0–5 (0 is definite absence of lesion and 5 is definite presence).

The diagnoses provided by radiologists were then used to compute sensitivity (true positive rate), specificity (true negative rate), and overall accuracy.

2. Forced choice evaluation technique: Each compressed image was paired with its original and the observer was asked to compare both images and rate the perceptible difference on a scale from 0 to 5, where 5 represents no visible difference and 0 is unacceptable.

Study Design

Sample of Readers

We enrolled 25 readers across Canada to participate in our study. All were certified by the Royal College of Physicians and Surgeons of Canada or an equivalent specialty body. All had between 5 and 25 years experience. To ensure an adequate sample size, we required a minimum of five readers for each of our reading sessions representing a minimum of 20 radiologists for our four reading sessions. Radiologists were asked to read in a manner consistent with their daily experience. They were blinded to the level and type of compression in each image.

The data from each radiologist in each condition (body area (body, chest, MSK, neuro)), compression type (JPEG, J2000), compression ratio (low, med, and high) was first used to compute an overall sensitivity (true positive/total positive), specificity (true negative/total negative), and accuracy (true (positive+negative)/total). Mean difference between compressed and uncompressed images (six-point

**Table 4** Analysis of specificity by body system

Body system	Uncompressed	JPEG	JPEG 2000	<i>P</i> (type)	<i>P</i> (Lvl)	<i>P</i> (type×Lvl)
Chest	0.762	0.806	0.889	ns	ns	ns
Body	0.571	0.548	0.492	ns	ns	<0.001
MSK	0.700	0.578	0.627	ns	ns	ns
Neuro	0.417	0.450	0.684	ns	<0.05	<0.05
Overall	0.612	0.608	0.673	ns	<0.10	ns

**Table 5** Analysis of accuracy by body system

Body system	Uncompressed	JPEG	JPEG 2000	<i>P</i> (type)	<i>P</i> (Lvl)	<i>P</i> (type×Lvl)
Chest	0.714	0.638	0.619	ns	ns	<0.05
Body	0.686	0.557	0.643	0.05	ns	<0.05
MSK	0.640	0.627	0.600	ns	ns	ns
Neuro	0.717	0.573	0.633	ns	ns	ns
Overall	0.689	0.599	0.624	ns	ns	ns

scale, where 5=no difference, 0=large difference) and mean confidence in diagnosis (six-point scale, where 5=100 % confidence, 0=no confidence) were computed.

### Analysis

The data were then subjected to an analysis of variance. First, an overall analysis of variance (ANOVA) was conducted across all body parts. There were three factors in this analysis—body region (four levels) compression (0=uncompressed, 1=JPEG, 2=J2000) and level (1=low comp, 2=med comp, 3=high comp (only for type=1, 2)). Overall correlations between accuracy, confidence, and perceived difference were calculated. The ANOVA was then repeated for each body region. Separate ANOVAs were performed for sensitivity, specificity, accuracy, confidence, and perceived difference.

### Results

The data on sensitivity, specificity, accuracy, confidence, and perceived difference, averaged across body systems, is shown in Table 2. Results are displayed in histogram form for each of the five dependent variables, with *p* values noted.

#### Sensitivity

Mean value for the uncompressed images was 0.70, for JPEG was 0.60, for JPEG 2000 was 0.62. There was no significant difference by Type ( $p=0.566$ ), or by level of compression ( $p=0.629$ ). The type by level interaction was significant ( $p=0.037$ ) reflecting that JPEG was low on

medium compression, but JPEG 2000 was high. Because there was no overall significant difference by type or level and no consistent trend, we conclude that compression had no significant effect on sensitivity.

#### Specificity

Mean value for the uncompressed images was 0.61, for JPEG was 0.61, for JPEG 2000 was 0.67. There was no significant difference by type ( $p=0.467$ ) or by level of compression ( $p=0.072$ ) and no interaction ( $p=0.761$ ). Because there was no overall significant difference by type or level and no consistent trend, conclude that compression had no significant effect on specificity.

#### Accuracy

Mean value for the uncompressed images was 0.69, for JPEG was 0.60, for JPEG 2000 was 0.62. There was no significant difference by type ( $p=0.478$ ) or by level of compression ( $p=0.541$ ) and no interaction ( $p=0.566$ ). Because there was no overall significant difference by type or level and no consistent trend, we conclude that compression had no significant effect on accuracy, despite the apparent reduction overall.

#### Confidence

Mean value for the uncompressed images was 4.1, for JPEG was 4.03, for JPEG 2000 was 4.12. There was no significant difference by type ( $p=0.430$ ) or by level of compression ( $p=0.250$ ) and no interaction ( $p=0.143$ ). Because there was no overall significant difference by type or level and no consistent trend, we conclude that compression had no

**Table 6** Analysis of confidence by body system

Body system	Uncompressed	JPEG	JPEG 2000	<i>P</i> (type)	<i>P</i> (Lvl)	<i>P</i> (type×Lvl)
Chest	4.12	3.69	4.07	<0.05	ns	ns
Body	4.44	4.26	4.59	ns	ns	ns
MSK	3.58	3.86	3.94	ns	ns	ns
Neuro	4.23	4.11	3.81	<0.05	ns	ns
Overall	4.09	4.03	4.12	ns	ns	ns

**Table 7** Analysis of perceived difference by body system

Body system	Uncompressed	JPEG	JPEG 2000	<i>P</i> (type)	<i>P</i> (Lvl)	<i>P</i> (type×Lvl)
Chest	4.12	3.69	4.07	<0.05	ns	ns
Body	4.44	4.26	4.59	ns	ns	ns
MSK	3.58	3.86	3.94	ns	ns	ns
Neuro	4.23	4.11	3.81	<0.05	ns	ns
Overall	4.09	4.03	4.12	ns	ns	ns

significant effect on confidence, despite the apparent reduction overall.

### Perceived Difference

Mean value for the uncompressed images was 4.68, for JPEG was 4.16, for JPEG 2000 was 4.53. There was a significant difference by type ( $p=0.05$ ), but not by level of compression ( $p=0.185$ ) and no interaction ( $p=0.453$ ). JPEG images were perceived as significantly worse than J2000 or uncompressed images.

We then analyzed the correlation between confidence rating, perceived difference and accuracy. If image compression is resulting in significant loss of clarity (as measured by perceived difference) this should be reflected in lower accuracy and lower confidence. In fact, the correlations among the measures were very low. Correlation between the perceived difference and accuracy was  $-0.125$  (n.s.). Similarly, the correlation between perceived difference and accuracy was  $-0.055$ . The relation between confidence and perceived difference was  $0.263$  ( $p=0.001$ ). Thus while image degradation was possible detectable (a greater perceived difference for JPEG, but not for JPEG 2000 images), this did not affect confidence or diagnostic accuracy.

Finally, we repeated the analysis within body system, as shown in Tables 3, 4, 5, 6, and 7. To control for type I error rate, we used a Bonferroni correction on the  $p$  value. Since there were a total of 75 computed  $p$  values, we used an alpha of  $0.05/75=0.0007$ . Under these circumstances, although 12 of the 75  $p$  values were less than 0.05, none achieved a significance level of 0.0007. Moreover, on inspection, there was no consistent pattern except that, as noted in the overall analysis, JPEG was rated lower on perceived difference for all four regions, and was significant and relatively large for chest (3.58 vs 4.53 for JPEG 2000 and 4.67 for uncompressed).

### Conclusions

Although some level of degradation was visually detectable on JPEG images, the overall analysis showed that the

compression (level/type) had no affect on subjective assessment or diagnostic accuracy for the anatomical areas investigated.

For diagnostic accuracy, there was no significant difference by type or by level of compression. Because there was no overall significant difference by type or level and no consistent trend, we conclude that compression had no significant effect on accuracy, despite the apparent reduction overall. Similarly, although the highest level of compression appeared to show decreased confidence for both JPEG and JPEG 2000, this was not statistically significant.

Assessment of perceived difference showed a significant difference by type (uncompressed, JPEG, and JPEG 2000) but not by level of compression (6:1, 9:1, and 12:1). However, visually, there appeared to be an approximate downward trend with increased compression. JPEG images were perceived as significantly worse than JPEG-2000 or uncompressed images.

As a general recommendation, we suggest that at the levels of irreversible compression tested, irreversible JPEG and irreversible JPEG 2000 on thin slices (less than 3 mm) CT images, are both appropriate to use for medical images.

### References

1. Canadian Association of Radiologists. Canada Health Infoway. Available at: [http://www.car.ca/Files/Standard\\_Lossy\\_Compression\\_EN.pdf](http://www.car.ca/Files/Standard_Lossy_Compression_EN.pdf). Accessed April 24, 2012
2. Koff D, Bak P, Brownrigg P, Hosseinzadeh D, Khademi A, Kiss A, Lepanto L, Michalak T, Shulman H, and Volkening A: Pan-Canadian Evaluation of Irreversible Compression Ratios ("Lossy"-Compression) for Development of National Guidelines - J Digit Imaging. 2009 Dec;22(6):569–78. Epub 2008 Oct 18
3. Ringl H, Schemthaler RE, Bankier AA, Weber M, Prokop M, Herold CJ, Schaefer-Prokop C: JPEG2000 compression of thin-section ct images of the lung: effect of compression ratio on image quality. Radiology 240:869–87, 2006
4. Bak P: Will the use of irreversible compression become a standard of practice? SCAR News 18:10, 2006
5. McNitt-Gray MF. AAPM/RSNA physics tutorial for residents: topics in CT. Radiation dose in CT, RadioGraphics 2002; 22:1541–1553
6. Woo HS, Kim KJ, Kim TJ, Hahn S, Kim B, Lee KH. JPEG 2000 compression of abdominal CT: difference in tolerance between thin- and thick-section images AJR Am J Roentgenol. 2007;189:535–41

7. Yamamoto S, Johkoh T, Mihara N, et al. Evaluation of compressed lung CT image quality using quantitative analysis. *Radiat Med* 2001; 19:321–329
8. Siddiqui KM, Siegel EL, Reiner BI, Johnson JP, Crave O, Nadar M.: Improved image compression at various slice thickness for multi-slice CT using 3D JPEG 2000 (part 2) in comparison with conventional 2D compression. (abstr) Society for Computer Applications in Radiology Scientific Abstracts. Leesburg, VA: Society for Computer Applications in Radiology, 87–88, 2004
9. Siegel EL, Siddiqui KM, Johnson JP, et al.: Compression of multi-slice CT: 2D vs. 3D JPEG2000 and effects of slice thickness. *SPIE Proceedings*. Bellingham, WA: SPIE, 162–170, 2005
10. Lubin J. The use of psychophysical data and models in the analysis of display system performance. In: Watson AB, ed. *Digital images and human vision*. Cambridge, MA: MIT Press, 1993:163–178
11. Driggers, RG: *The Encyclopedia of optical engineering*, volume 12, New York: Dekker, 2003
12. Accusoft. Pegasus Imaging. Available at: <http://www.accusoft.com/picmedx.htm>. Accessed: April 24, 2012