Image Compression and Chest Radiograph Interpretation: Image Perception Comparison Between Uncompressed Chest Radiographs and Chest Radiographs Stored Using 10:1 JPEG Compression

Douglas P. Beall, Phillip D. Shelton, Thomas V. Kinsey, Maria C. Horton, Brian J. Fortman, Steffen Achenbach, Vadim Smirnoff, Daniel L. Courneya, Bill Carpenter, and John T. Gironda

We have assessed the effect of 10:1 lossy (JPEG) compression on six board-certified radiologists' ability to detect three commonly seen abnormalities on chest radiographs. The study radiographs included 150 chest radiographs with one of four diagnoses: normal (n = 101), pulmonary nodule (n = 19), interstitial lung disease (n = 19), and pneumothorax (n = 11). Before compression, these images were printed on laser film and interpreted in a blinded fashion by six radiologists. Following an 8-week interval, the images were reinterpreted on an image display workstation after undergoing 10:1 lossy compression. The results for the compressed images were compared with those of the uncompressed images using receiver operating characteristic (ROC) analyses. For five of six readers, the diagnostic accuracy was higher for the uncompressed images than for the compressed images, but the difference was not significant ($P > .1111$). Combined readings for the uncompressed images were also more accurate when compared with the compressed images, but this difference was also not significant ($P = .1430$). The sensitivity, specificity, and accuracy values were 81.5%, 89.2%, and 86.7% for the compressed images, respectively, as compared with 78.9%, 94.5%, and 89.3% for the uncompressed images. There was no correlation between the readers' accuracy and their experience with soft-copy interpretation; the extent of radiographic interpretation experience had no correlation with overall interpretation accuracy. In conclusion, five of six radiologists had a higher diagnostic accuracy when interpreting uncompressed chest radiographs versus the same images modified by 10:1 lossy compression, but this difference was not statistically significant.

Copyright © 2000 by W.B. Saunders Company

THE EMERGING picture archiving and communication systems (PACS) and digital imaging in general have introduced both opportunities and challenges. Medical images are now routinely compressed, stored into databases, and transmitted via local area networks. Efficient information management and rapid information transfer are critical components to consider when developing a functional PACS or an effective teleradiology system.

The Department of Defense (DOD) deployed the Medical Diagnostic Imaging Support (MDIS) system in eight medical centers and 14 teleradiology/limited PACS sites. The long-term storage archive

technique used routinely at those sites is JPEG 10:1 lossy compression. The images are interpreted and are accessible as uncompressed data before long-term storage, where the images are compressed and stored on optical discs.

MATERIALS AND METHODS

PACS

The PACS deployed at the Wright Patterson Air Force Base Medical Center in Dayton, OH was one of the first military MDIS systems installed by Lockheed/Martin in 1994. From June 1994 to March 1996, all chest x-ray images were processed by Fuji 7000 computed radiography (CR) readers (Fuji Medical Systems, Stamford, CT) and the uncompressed radiographs were printed on 10- × 14-inch Fuji film. The same images were imported into the PACS (Vantage Windows PACS, Lockheed Martin, Hoffman Estates, IL), stored on 14-inch Kodak optical discs (Eastman Kodak, Rochester, NY), and maintained in a Kodak optical disc jukebox.

Case Selection

Cases for the study were selected from the radiographic library by one of the authors (P.D.S.). There were 222 posteroanterior (PA) chest radiographs collected at Wright Patterson Air Force Base over a 21-month period. These radiographs were then reviewed by five radiologists certified by the American Board of Radiology. During the initial review, the radiologists were given the expected diagnosis and were asked to score the images in three separate categories: agree or disagree with the expected diagnosis, rate degree of conspicuousness from mild to obvious, and rate image quality from poor to good. Each radiologist scored the images separately, and images that did not have complete concurrence between all reviewers as to the diagnosis were excluded from the study. All images that any one of the reviewers rated as poor quality were also excluded from

From the Department of Radiology, Sheppard Air Force Base, 82nd Medical Group, Wichita Falls, TX; Department of Radiology, John Hopkins Hospital, Baltimore, MD; Joint Imaging Technology Project Office, Ft Detrick, MN; and Department of Diagnostic Radiology, Philips University Marburg, Marburg, Germany.

Address reprint requests to Douglas P. Beall, MD, Department of Radiology, 82 MDSS,SGSS, 149 Hart St, Suite 5, Wichita Falls, TX 73611-5300. E-mail: douglas.beall@sheppard.af.mil.

Copyright © 2000 by W.B. Saunders Company

0897-1889/00/1302-1007\$10.00/0

doi:10.1053/jdim.2000.6820

the study. After the initial review and exclusion of various images due to these criteria, 49 cases remained that demonstrated pathology. These 49 images included 19 cases of interstitial lung disease, 19 cases of solitary pulmonary nodule, and 11 pneumothoraces. There were also 101 cases (meeting analogous imaging requirements) that qualified as normal chest radiographs. These were added to the 49 cases demonstrating pathology for a total of 150 cases in the set. The images were then numbered and randomly ordered.

Interpretation

The radiologists read uncompressed images printed on 10- × 14-inch Fuji film placed on a light board. The compressed images were interpreted using PACS with high-resolution monitors (2K × 1.6K pixels with a luminescence of 70 foot/lambert). The readers were allowed to use all image interrogation capabilities of the PACS.

Each case was viewed only once per reader in its uncompressed state and once after compression. The first reading was of the uncompressed images and all readings were accomplished in a single sitting that took from 2 to 4 hours to complete. The second reading involved interpreting compressed soft-copy images and was done 8 weeks after the uncompressed image interpretation session. The readers were instructed that each case had only one diagnosis and there were four diagnostic possibilities: normal, pneumothorax, pulmonary nodule, and interstitial lung disease. Interpreters assigned confidence ratings of low, moderate, and high to each of their findings. The ratings, together with the positive or negative reading, corresponded to six categories for the purpose of receiver operating characteristic (ROC) analysis. The readers were asked to give a diagnosis (when appropriate) and to score each case from 0 to 5, with the following categories represented: category 0 (negative), category 1 (probably negative), category 2 (possibly negative), category 3 (possibly positive), category 4 (probably positive), and category 5 (positive). In scoring the interpretations, location was taken into consideration.

Data Management and Analysis

Data collected during the diagnostic sessions were initially entered by hand on paper scoring sheets and then transferred into a computer database program. Each interpretation was scored against the standard reference diagnosis.

Measurements were calculated for each of the readers and for the group. These measurements included diagnostic accuracy, sensitivity, specificity, positive predictive values, and negative predictive values. The McNemar test was used to analyze difference in the data sets by examining the dichotomous variables. The compressed and uncompressed data sets were analyzed by calculating the relevant probability values. A one-tailed P value less than .05 was considered statistically significant.

ROC analysis was performed with the use of the computer program CORROC (University of Chicago, IL), which analyzes correlated categorical rating-scale data sets. This program calculates the maximum-likelihood estimates of a "bivariate binormal" model for correlated ordinal category data from two diagnostic sets. ROC curves were calculated for each reader in compressed and uncompressed formats. Aggregate curves were

also calculated comparing the readings for compressed and uncompressed images.

The resulting data were analyzed with the computer program by generating 14 sets of ROC curves. The reported information included the area under the ROC curve (A_z), the standard deviation of A_z , and a one-tailed P value to test the significance of the difference between the compressed images (10:1 JPEG) and the uncompressed images.

RESULTS

A total of 900 cases were interpreted by six radiologists in both compressed and uncompressed forms. Table 1 shows the interpretation results for all six readers and Table 2 indicates the best performances by individual readers and their level of experience.

In five of the six readers, the test outcome measurements were higher for the uncompressed images than for the compressed images, but the difference was not significant ($P > .1111$). ROC curve measurement for the combined performance of all readers demonstrated a comparatively better performance for the uncompressed images (Fig 1), but the result was, again, not statistically significant ($P > .1430$). Statistical significance (one-tailed P value) was also not attained in any of the ROC curve calculations for each of the readers individually. The reader who scored better interpreting the compressed images performed well with both sets of images as measured by the area under the ROC curve (Fig 2), but the difference between the area measurements for the compressed and uncompressed images was negligible.

As demonstrated, there was little correlation between the total amount of experience or the extent of soft-copy interpretation experience and the best film interpretation performance (Tables 2 and 3). The best interpretation performances for individual readers as determined by the test outcome measurements were comparatively better for the uncompressed images than for the compressed images (Table 2), but the difference was not statistically significant ($P = .2013$). There was also no correlation, statistical or otherwise, between

Table 1. Combined Interpretation Measurements for Compressed and Uncompressed Images

Measurement	Compressed	Uncompressed
Sensitivity	81.5%	78.9%
Specificity	89.2%	94.5%
Accuracy	86.7%	89.3%
Positive predictive value	78.8%	87.7%
Negative predictive value	90.7%	90.0%

Table 2. Best Reader Performance Measurements Compared With Reader Experience Level

Reader	Experience: Months Reading Digital Images	Measurement	Reader With Best Performance	
			Compressed	Uncompressed
A	10	Sensitivity	C—90%	A—92%
B	6	Specificity	D—96%	C—99%
C	18	Accuracy	F—89.3%	F—93.3%
D	40	Positive predictive value	D—90%	C—97.5%
E	6	Negative predictive value	C—94.6%	F—96.9%
F	18			

fellowship training and the readers' image interpretation accuracy (Table 3).

DISCUSSION

Image Compression

In spite of a number of investigations and contributions to the literature on the subject of image compression, this topic continues to be vigorously debated. A 1990 study by Ishigaki et al¹ concluded that CR images with a compression ratio of 10:1 are acceptable for primary interpretation. A 1994 study by Mori and Nakata² examined chest radiographs with image compression ratios ranging from 10:1 to 25:1. Evaluation of these images revealed no significant differences in the radiolo-

gists' ability to detect pneumothoraces, pulmonary nodules or interstitial infiltrates. However, this study did note a deterioration of image quality with images compressed at 25:1. A 1996 study by Cox et al,³ examining the ability of radiologists to detect pulmonary nodules on compressed chest radiographs, showed a measurable decline in detection for each compression ratio.

More recently, a 1998 military study by Parisi et al⁴ compared uncompressed laser-printed chest radiographs from a neonatal intensive care unit to compressed soft-copy images. Despite 10:1 JPEG compression, no loss in anatomic detail was seen in five of six anatomic locations analyzed.

A 1997 study by Erickson et al⁵ examined the

Fig 1. Compression v no compression (combined reader data). TPF, true positive fraction; FPF, false positive fraction.

Fig 2. Compression v no compression (reader C).

effect of compression ratios from 10:1 to 80:1 on the diagnostic quality of chest radiographs. Eleven anatomic structures were evaluated and overall quality assessments rendered on both compressed and uncompressed images. This study found that lossy compression of at least 40:1 could be used without perceptible loss of detail in the anatomic structures assessed. Shortly thereafter Savcenko et al⁶ concluded that compression ratios of up to 40:1 can be used without decreased diagnostic accuracy in the detection of noncalcified pulmonary nodules and pulmonary fibrosis.

Many evaluations of various compression techniques have shown promising results for various image compression techniques,^{5,7,8} but, despite these findings, only the Joint Photographic Expert Group's algorithm has been approved by the Digital Imag-

ing and Communications in Medicine (DICOM) committee for compression of medical images.

Image Compression

A significant portion of the initial use of JPEG image compression was seen in the early PACS systems deployed by the US Military. These systems were designed to use JPEG compression partly due to the results of early evaluations of image compression. A 1984 assessment of image compression undertaken by the DOD and the Public Health Service (PHS) evaluated lossy compression and the teleradiology systems transmitting this information. Image compression was necessary to increase the speed of transmission over the standard telephone lines used for the study. The results reported discrepancies of less than 3% for

Table 3. Training Experience Versus Image Interpretation Accuracy

Reader	Experience		Interpretation Performance		
	Years Postresidency	Fellowship Training	Measurement	Compressed	Uncompressed
A	8.5	None	Accuracy	82%	91.3%
B	3.0	Neuroradiology	Accuracy	86.7%	85.3%
C	1.5	None	Accuracy	86.7%	92%
D	7.0	None	Accuracy	87.3%	86.7%
E	1.5	Interventional Radiology	Accuracy	85.3%	87.3%
F	1.5	None	Accuracy	89.3%	93.3%

findings of clinical significance. These results were reported at the 1985 meeting of the Medical Radiation Advisory Committee Meeting of the Center for Devices and Radiologic Health. Despite these findings, a subsequent PACS implementation contract established by the DOD mandated lossless compression. The contract, known as the Digital Imaging Network and Picture Archive and Communication System (DIN-PACS), is an agreement between the two vendors identified in the contract and the DOD for the proliferation of PACS throughout military health care facilities.

Before the establishment of the DIN-PACS contract, the original contract to implement this technology was the MDIS system. Under these guidelines, the long-term storage archive maintains images that are stored on optical discs with 10:1 JPEG lossy compression. Today, the DOD continues to operate approximately 15 MDIS sites that have successfully used long-term lossy compression since their inception.

Chest Radiograph Interpretation

Measurement of the image quality degradation caused by compression is essential to stratify the effects of image compression on the conspicuousness of anatomic structures and the overall diagnostic quality of the image. The ROC technique is a widely accepted and used technique that is readily applied to the differing degrees of certainty involved in the interpretation of radiographs.^{9,10} For this reason, the ROC analysis was integral to our comparison of reader evaluations.

The performance measurements for both the compressed and uncompressed images did not seem to vary with differences in experience (Table 2). Although the difference between the reader with the least experience and the reader with the most experience was only slightly less than 3 years, the best performances were consistently produced by readers with an intermediate amount of experience rather than by the reader with the greatest amount of experience in digital image interpretation. Additionally, the amount of experience postresidency did not seem to have an effect on image interpretation performance and those radiologists with fellowship training fared no better than those without a fellowship (Table 3).

Overall, the absence of significant differences between the compressed and uncompressed images in regard to the diagnostic accuracy (Figs 1 and 2)

agrees with much of the existing data from various investigations. Despite this agreement and the adequate statistical analysis of this investigation, one portion of this study that could be open to criticism is that the compressed and uncompressed images were interpreted from different media (laser-printed images for the uncompressed images and digital soft copy for the compressed images). Although this scenario is not unusual throughout the existing body of literature comparing compressed and uncompressed images, it is somewhat less than optimal for a more exact comparison. However, this study limitation was beyond the control of the investigators, as many of the compressed images could not be printed onto laser film. To attain the best possible data comparison, our future investigations will involve analogous image demonstration media to allow for the best possible comparison.

CONCLUSION

Our investigation demonstrated that although five of six radiologists had a higher diagnostic accuracy when interpreting uncompressed chest radiographs as compared with the same images modified by 10:1 lossy compression, the difference was not statistically significant. The lack of statistical significance agrees with the majority of the existing literature regarding the effect of image compression on the diagnostic accuracy of the interpretation.

As the amount of information about image compression and lesion detection increases, a paradigm shift may occur that results in the adoption of other image compression algorithms by the DICOM committee and the widespread use of image compression when storing and transmitting images. Until that time, the tradition of using lossless compression for primary interpretation will continue along with the use of JPEG as the only approved compression algorithm.

REFERENCES

1. Ishigaki T, Sakuma S, Ikeda M, et al: Clinical evaluation of irreversible image compression: Analysis of chest imaging with computed radiography. *Radiology* 75:739-794, 1990
2. Mori T, Nakata H: Irreversible data compression in chest imaging using computed radiography: An evaluation. *J Thorac Imaging* 9:23-30, 1994
3. Cox GG, Cook LT, Insana MF, et al: The effects of lossy compression on the detection of subtle pulmonary nodules. *Med Phys* 23:127-132, 1996
4. Parisi SB, Mogel GT, Dominguez R, et al: The effect of

10:1 compression and soft copy interpretation on the chest radiographs of premature neonates with reference to their possible application in teleradiology. *Eur Radiol* 8:141-143, 1998

5. Erickson BJ, Manduca A, Persons KR, et al: Evaluation of irreversible compression of digitized posterior-anterior chest radiographs. *J Digit Imaging* 10:97-102, 1997

6. Savchenko V, Erickson BJ, Palisson PM, et al: Detection of subtle abnormalities on chest radiographs after irreversible compression. *Radiology* 206:609-616, 1998

7. MacMahon H, Doi K, Sanada S, et al: Data compression: Effect on diagnostic accuracy in digital chest radiography. *Radiology* 178:175-179, 1991

8. Lyriboz TA, Zudoski MJ, Hopper KD, et al: A comparison of wavelet and Joint Photographic Experts Group lossy compression methods applied to medical images. *J Digit Imaging* 12:14-17, 1999 (suppl 1)

9. Swets JA: ROC analysis applied to the evaluation of medical imaging techniques. *Invest Radiol* 14:109-121, 1999

10. Metz CE: Some practical issues of experimental design and data analysis in radiological ROC studies. *Invest Radiol* 24:234-245, 1989

11. Kido S, Ikezoe J, Kondoh H, et al: Detection of subtle interstitial abnormalities of the lungs on digitized chest radiographs: Acceptable data compression ratios. *AJR* 167:111-115, 1996