The Effect of Compression on Clinical Diagnosis of Glaucoma Based on Non-analyzed Confocal Scanning Laser Ophthalmoscopy Images

Knowledge of the effect of compression of ophthalmic images on diagnostic reading is essential for effective tele-ophthalmology applications. It was therefore with great anticipation that I read the article “The Effect of Compression on Clinical Diagnosis of Glaucoma Based on Non-analyzed Confocal Scanning Laser Ophthalmoscopy Images,” by Bélair et al. in the July/August 2005 issue of *Ophthalmic Surgery, Lasers & Imaging.* The authors compared diagnostic accuracy of independent observers based on HRT images in different compressed and uncompressed formats, including lossy JPEG compression, lossy JPEG 2000 compression, lossy GIF compression, and lossless PNG and TIFF compression. Their choice to include PNG, a format not supported by most popular image viewers, allows an unexpected verification of their results.

An image, after being compressed to PNG format and then decompressed (using the lossless Lempel-Zif compression algorithm that is also used for TIFF compression and ZIP file compression), contains exactly the same information as the original image. The diagnosis obtained by reading PNG compressed–decompressed HRT images is therefore expected to be the same as that obtained by reading TIFF compressed–decompressed images, excluding random intraobserver and interobserver variability. In other words, kappa should be close to 1 for both TIFF and PNG in a study such as theirs.

The authors’ finding that the largest difference in kappa agreement was between TIFF compressed–decompressed and PNG compressed–decompressed images therefore means that intraobserver and interobserver variability was high, higher than any difference caused by a difference in compression technology. In other words, no difference between any technology is detectable in their study.

As a sideline, Receiver Operator Characteristics (ROC) curves and Free Response Receiver Operator Characteristics (FROC) curves are much better suited to compare diagnostic accuracy between multiple observers for multiple diagnostic classes than kappa.

Michael D. Abràmoff, MD, PhD
Department of Ophthalmology and Visual Sciences
University of Iowa Hospital and Clinics
Iowa City, Iowa

REFERENCE

Response

We would like to thank Dr. Abràmoff for his comments in relation to our recently published article. With the projected increase in glaucoma prevalence, screening for glaucomatous optic nerve damage is becoming an important public health priority. We evaluated whether using the images generated by the HRT II could be useful in identifying cases of glaucoma, and more specifically whether different compression formats affected the interpretation.

The Portable Network Graphic, or PNG, is a patent-free image format with many features, such as “48-bit” (16 bits per channel) true color or 16-bit grayscale, lossless compression, gamma correction, two-dimensional interlacing, and alpha transparency, which makes it an excellent alternative to TIFF and GIF.
Unfortunately, as noted by Dr. Abràmoff, PNG is not properly supported by most popular image viewers or browsers. In our study, the use of Adobe Photoshop resulted in different images when TIFF (Figure, right) and PNG (Figure, left) compression were used, with the PNG images being much paler and therefore not identical to the original.

Furthermore, PNG is not intended to replace JPEG. JPEG compression not only produces smaller files than PNG for photographic images, but is preferable for photographic images that are not to be altered, whereas PNG is better suited to computer-generated graphics.³

Finally, although Receiver Operator Characteristics (ROC) curves may be good to compare diagnostic accuracy between multiple observers, a quantitative value for the observers’ degree of certainty (a number from 1 to 5) is generally required. Additionally, a typical ROC study could require more than 300 images to obtain a reasonable statistical confidence level, five or more observers to view these images, and a full-time statistician to analyze the data.⁴

The authors therefore believe that, based on our study design, the conclusions are still valid.

Marie-Lyne Bélair, MD
Alvine Kamdeu Fansi, MD
Denise Descovich, MD
Paul Harasymowycz, MD
Centre de Recherche Guy-Bernier
Hôpital Maisonneuve-Rosemont
and
A.-R. Leblanc, PhD
Université de Montréal
Montréal, Québec, Canada

REFERENCES