Reproducibility Validation Study Comparing Analog and Digital Imaging Technologies for the Measurement of Intima-Media Thickness

Damiano Baldassarre, PhD; Elena Tremoli, PhD; Mauro Amato, PhD; Fabrizio Veglia, PhD; Alighiero Bondioli, MD; Cesare R. Sirtori, MD, PhD

Background and Purpose—New advances in B-mode imaging technologies have led to improved quality in the detection of minute changes in the surface of intima-media thickness (IMT) and plaques. The new digital systems, with increased numbers of imaging channels, multiple frequency probes, and increased microprocessing speeds, now generate images comparable to those of the analog predecessors. Can these digital systems have reproducibility comparable to that of a pure analog system? We compared the Biosound 2000II (analog) system with the Esaote AU4 (digital) system.

Methods—Twenty-two subjects were chosen who had varying degrees of IMT on the far wall of the common carotid artery. Common carotid IMT was determined twice: the first time with the analog system and the second time with the digital system. With each system, replicate scans were made within 2 weeks.

Results—The intramethod agreement was high with the analog system, with a bias between readings of \(-0.010\pm0.033\) mm, mean absolute difference of \(0.027\pm0.020\) mm, repeatability coefficient of 0.067, and correlation coefficient of 0.97. The digital system provided the highest reproducibility with a bias between readings of \(-0.002\pm0.016\) mm, mean absolute difference of \(0.012\pm0.011\) mm, repeatability coefficient of 0.033, and correlation coefficient of 0.99. When the analog and digital systems were compared, the bias between readings was \(-0.011\pm0.024\) mm with good agreement between the 2 systems; the repeatability coefficient was 0.047, with all points within \(\pm2\) SDs of the mean difference. The mean absolute difference between the 2 measurements was \(0.018\pm0.015\) mm with a correlation coefficient of 0.98.

Conclusions—The digital system for IMT evaluation compares well with the more widely used analog system and provides a reliable technology for common carotid IMT measurement that can be applied to clinical trials. (Stroke. 2000;31:1104-1110.)

Key Words: carotid arteries ■ intima-media thickness ■ ultrasonography

Intima-media thickness (IMT) of the carotid artery is increasing in importance and acceptance as an end point in large multicenter clinical trials. Clinical trials, including Pravastatin, Lipids, and Atherosclerosis in the Carotid arteries (PLAC-II), Kuopio Atherosclerosis Prevention Study (KAPS), Carotid Atherosclerosis Italian Ultrasound Study (CAIUS), Asymptomatic Carotid Artery Progression Study (ACAPS), Regression Growth Evaluation Statin Study (REGRESS), LIPID, and others, have shown that long-term studies can be successfully carried out while maintaining a high reproducibility value. Most of these studies were completed with the use of analog-based technology, which has long been the gold standard in IMT imaging. The new clinical trials, however, must also adopt new technologies as analog-based systems slip further and further into the past. The problem with the new digital processing systems is that this technology has no history or background to validate its ability in long-term studies and to maintain a very high reproducibility factor. To supplement this loss of background information, short-term system trials must be performed to ensure a high standard of credibility when this new technology is used. The present study was designed to perform a comparison between analog and digital B-mode ultrasound systems. To this aim, the within-method repeatability in common carotid IMT (CC-IMT) determination and the agreement between the 2 technologies were investigated.

Subjects and Methods
The study was performed in 22 healthy subjects who were recruited from the medical staff of E. Grossi Paoletti Center for the Study of Metabolic Disorders. Oral informed consent was obtained from all subjects.

Received July 9, 1999; final revision received January 25, 2000; accepted January 25, 2000.

From the E. Grossi Paoletti Center, Institute of Pharmacological Sciences, University of Milan, and the Unit of Biostatistics, Institute H. San Raffaele (F.V.), Milan, Italy.

Correspondence to Prof C.R. Sirtori, Institute of Pharmacological Sciences, University of Milan, Via Balzaretti, 9, 20133, Milan, Italy. E-mail cesare.sirtori@unimi.it

© 2000 American Heart Association, Inc.

Stroke is available at http://www.strokeaha.org
TABLE 1. Variability of CC-IMT Measurements According to Analog or Digital Imaging Technology

<table>
<thead>
<tr>
<th></th>
<th>Analog System: Scan 1 vs Scan 2</th>
<th>Digital System: Scan 1 vs Scan 2</th>
<th>Analog vs Digital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias between readings*</td>
<td>-0.010±0.033</td>
<td>0.002±0.016</td>
<td>-0.011±0.024</td>
</tr>
<tr>
<td>Repeatability coefficient*</td>
<td>0.067</td>
<td>0.033</td>
<td>0.047</td>
</tr>
<tr>
<td>Absolute differences</td>
<td>0.027±0.020</td>
<td>0.012±0.011</td>
<td>0.018±0.015</td>
</tr>
<tr>
<td>r</td>
<td>0.97</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>CV, %</td>
<td>4.1</td>
<td>2.1</td>
<td>2.7</td>
</tr>
<tr>
<td>Error, %</td>
<td>5.3</td>
<td>2.0</td>
<td>3.4</td>
</tr>
</tbody>
</table>

*Calculated as recommended by Bland and Altman (mean relative differences between readings).

The data extracted from the CC-IMT measurements summarize the ability of each system to reproduce and measure IMT. The intramethod repeatability and the agreement between the 2 technologies were evaluated through an estimate of the consistent bias between readings, as recommended by Bland and Altman. The intramethod comparison is made between the CC-IMT values, obtained with each system during the initial and the respective replicate scan. The coefficient of repeatability was calculated according to the method of the British Standards Institution and corresponds to 2 SDs of the relative differences between replicate measurements.

The second and equally important comparison was made between the 2 systems during measurement of the same IMT. Specifically, the agreement was evaluated through an estimation of the consistent bias between the mean value of the 2 replicate readings obtained with the first method versus the mean value of the 2 replicate readings obtained with the second method. To facilitate a comparison of our findings with those provided in other reproducibility studies, the intramethod and the intermethod absolute differences, the correlation coefficients, the coefficient of variation, and the percent error between replicate scans also are provided.

Results

Of the 22 subjects recruited for the study (mean±SD age 39±10.3 years, age range 26 to 55 years, mean total serum cholesterol 5.3±0.8 mmol/L), 14 (63.6%) were men and with the exception of 2 smokers, none had significant cardiovascular risk factors (eg, high blood pressure, lipids, diabetes).

Repeatability of Analog System (Biosound 2000II): Initial Versus Replicate CC-IMT Values

The comparison of CC-IMT values between the initial and replicate scans performed with the Biosound 2000II analog system is shown in Table 1 and Figure 1 (left). The relative differences in CC-IMT values remained constant as the IMT increased from 0.4 to 0.8 mm. Although there was some variability in the measurements of CC-IMT, the intramethod agreement was high, with a bias between readings of -0.010±0.033 mm and limits of agreement that ranged from -0.076 to 0.056 mm, with all points within these limits. The mean absolute difference was 0.027±0.020 mm (range 0.000 to 0.060 mm), the correlation coefficient was 0.97 (y=0.889x+0.072; P<0.0001; Figure 2, left), and the coefficient of repeatability was 0.067 mm.

Repeatability of the Digital System (Esaote AU4): Initial Versus Replicate CC-IMT Values

The comparison of CC-IMT values between the initial and replicate scans performed with the Esaote AU4 digital system are shown in Table 1 and Figure 1 (right). As in the case of the analog system, relative differences in CC-IMT values remained constant as the IMT increased from the lowest to the highest value. Although in this case there also was some
variability in CC-IMT measurements, the intramethod agreement was even better than that observed with the analog system, with a bias between readings of $-0.002 \pm 0.016$ mm, limits of agreement that ranged from $-0.030$ to $0.040$ mm, and all points but 1 within these limits. The mean absolute difference was $0.012 \pm 0.011$ mm (range 0.000 to 0.040 mm), and the correlation coefficient was $0.99$ ($y = -0.951x + 0.026$; $P < 0.0001$; Figure 2, right), and the coefficient of repeatability was $0.033$ mm.

**Agreement Between Analog and Digital Processing in the CC-IMT Determination**

CC-IMT ranged from 0.39 to 0.80 mm, with a mean value of $0.56 \pm 0.12$ mm, when measured with the analog system and from 0.41 to 0.78 mm, with a mean value of $0.57 \pm 0.12$ mm, when measured with the digital system. The bias between readings was $-0.011 \pm 0.024$ mm with an excellent agreement between the 2 systems (limits of agreement from $-0.056$ to $0.030$ mm; Figure 3, left). The mean absolute difference between the 2 measurements was $0.018 \pm 0.015$ mm (range 0.000 to 0.056 mm), and the correlation coefficient was $0.98$ ($y = 0.963x + 0.032$; $P < 0.0001$; Figure 3, right).

**Effects of Frequency Transducer**

A comparison of CC-IMT values between the initial and replicate scans performed with the analog system and with the digital system, equipped with 7.5- or 10-MHz probes, is shown in Table 2. Here, the CC-IMTs ranged from 0.46 to 1.12 mm ($0.69 \pm 0.21$ mm) when measured with the analog system and from 0.48 to 1.04 mm ($0.67 \pm 0.20$ mm) and from 0.47 to 1.02 mm ($0.67 \pm 0.19$ mm) when measured with the digital system equipped with the probe of 7.5 or 10 MHz, respectively. In this group of 10 subjects, the absolute differences between replicate scans observed with the analog system was $0.028 \pm 0.017$ mm, whereas those observed with the digital system were $0.019 \pm 0.008$ and $0.014 \pm 0.012$ with the 7.5- and 10-MHz probes, respectively. The limits of agreement ranged from $-0.075$ to $0.054$ mm with the analog system and from $-0.032$ to $0.048$ mm and from $-0.036$ to $0.041$ with the digital system equipped with the 7.5- and 10-MHz probes, respectively. Thus, independent of probes and frequency transducer, the intramethod agreement of CC-IMT measurements performed with the digital system was always better than that observed with the analog system. Indeed, in a comparison of the coefficients of repeatability obtained with the analog system, those obtained with the digital system were $1.56$ and $1.72$ times smaller with the 7.5- and 10-MHz probes, respectively.

**Discussion**

**Intramethod and Intermethod Accuracy and Reproducibility**

The present study shows that the within-method biases, evaluated with a performance of repeated measurements, were small for both analog and digital systems. In addition, because both methods showed a constant SD of 0.4 to 0.8 mm, no obvious systematic relationship between replicate scan differences and their mean values was observed.

Intramethod variability also clearly shows that compared with analog technology, digital technology can improve both the accuracy and reproducibility of carotid IMT measurements. In fact, although the analog system provides high-quality results (bias between readings $-0.01$ mm, repeatability coefficient $<0.07$ mm, and mean absolute differences $<0.03$ mm), reproducibility was further improved with the use of digital technology, with a bias between scans that was 5-fold less and a repeatability coefficient, a mean absolute difference, a percent CV, and a percent error that were 2-fold less than the corresponding values observed with the analog system. Because the

---

**Figure 1.** Intraobserver repeatability of IMT determination with an analog system (left) and with a digital system (right). Differences are shown between the first and second scans of the far wall CC-IMT measurements plotted against their mean values. The mean difference (dashed line) and the limits of agreement (continuous line) are also indicated.

**Figure 2.** Relationship between scan 1 and scan 2 for IMT measurements obtained with an analog system (left) or with a digital system (right).
comparison was carried out by the same sonographer (D.B.) and reader (M.A.) and with all other possible variables kept constant, including methods for image processing. The observed improvement may be due to a higher performance of digital instrumentation, probably determined through differences in imaging frequency (10 MHz digital versus 8 MHz analog), which may result in an improved axial resolution (Table 2). In addition, the digital electronics, with increased microprocessing speeds and the possibility of setting the most adequate imaging characteristics, such as probe frequency or scan depth, may be relevant for the reproducibility improvement that was observed.

The comparison of a new method against an established method has often been evaluated inappropriately through the use of the correlation coefficient between the results of the 2 methods as an indicator of agreement. Correlation coefficient measures the strength of a relation between 2 variables, not the agreement between them. Moreover, correlation depends on the range of the variables measured, with wider ranges leading to better correlation. More often, 2 methods measuring the same variable will be related, and thus the test of significance may be irrelevant regarding the question of agreement. Thus, a very high between-methods correlation coefficient, such as 0.98, as we determined (Table 1), may still harbor sufficient disagreement to render the new digital method an unsuitable substitute for the established analog standard. The analytical approach of Bland and Altman is the most appropriate approach to the evaluation of the consistency of a new method of measurements compared with an established method. Through the use of this approach, the analysis of between-method agreement has shown that the 2 methods can result in a discrepancy in IMT measurements of $\pm 0.011$ mm (3.4%). On this basis, the 2 methods can be, at least apparently, considered interchangeable. However, because the repeatability coefficient of the analog system (0.067 mm) turned out to be greater than that obtained with the between-method comparison (0.047 mm), which in turn was greater than that obtained with the digital system (0.033 mm), we might argue that the replacement of an analog system with a digital system is always possible without an effect on the study results, whereas the replacement of a digital system with an analog system is possible only after enlargement of the sample size, which must be recalculated on the basis of analog variability.

**Comparison With Other Reproducibility Studies**

Several studies that evaluated the reproducibility of IMT measurements have been performed, and because of different protocols for image acquisition, methods for IMT analysis, and methods for variability quantification, a comparison of the results among these studies may be difficult. As far as variability quantification is concerned, the analytic approach of Bland and Altman is the most appropriate method to evaluate between-method agreement as well as the within-method repeatability. However, because several authors carried out other analytical approaches to allow a comparison with our results, we also provide all parameters most frequently used in these studies, also to facilitate future users. Specifically, we provide the correlation coefficient, the CV, the percent error, and the absolute differences for the within-method analysis as well as for the between-method agreement (Table 1).

**TABLE 2. Variability of CC-IMT Measurements According to Analog or Digital Imaging Technology and Probe Frequency**

<table>
<thead>
<tr>
<th></th>
<th>Analog System 8 MHz: Scan 1 vs Scan 2</th>
<th>Digital System 7.5 MHz: Scan 1 vs Scan 2</th>
<th>Digital System 10 MHz: Scan 1 vs Scan 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias between readings*</td>
<td>0.010±0.032</td>
<td>0.008±0.020</td>
<td>0.002±0.019</td>
</tr>
<tr>
<td>Repeatability coefficient*</td>
<td>0.064</td>
<td>0.041</td>
<td>0.037</td>
</tr>
<tr>
<td>Absolute differences</td>
<td>0.028±0.017</td>
<td>0.019±0.008</td>
<td>0.014±0.012</td>
</tr>
<tr>
<td>$r$</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>CV, %</td>
<td>3.3</td>
<td>2.1</td>
<td>2.0</td>
</tr>
<tr>
<td>Error, %</td>
<td>4.5</td>
<td>3.1</td>
<td>2.2</td>
</tr>
</tbody>
</table>

*Calculated as recommended by Bland and Altman (mean relative differences between readings).
The intraobserver absolute differences of the far wall CC-IMT assessment between replicate scans of the present study was compared with those obtained in several other reproducibility studies (Figure 4). This figure indicate that the digital system provides even better results than the analog system. It should be emphasized, however, that an adequate comparison of our results can be performed only versus those studies in which, as in our case, the reproducibility of IMT measurements was determined in the far wall of the CCAs of healthy subjects. Touboul et al. performed a reproducibility study among 14 subjects with the use of an analog system, and a mean difference (SD) in IMT within observers of 0.06 (0.06) mm was found by using the same reading methods to assess carotid IMT as we used in the present study. Similar data were obtained, again with analog systems, in the 2 reproducibility studies performed by Persson et al. and Salonen et al., who found mean differences (SD) between replicate scans of 0.08 (0.07) and 0.09 (0.11) mm, respectively. In the present study, a significantly better reproducibility was observed with both analog and digital systems. This is probably due to the scan protocol we used, with ≥18 measurement sites considered for each subject. Slightly better results were found by Gariepy et al. and Wendelhag et al., who found a mean difference (SD) between replicate scans of 0.02 (0.02) and 0.007 (0.019) mm, respectively, which are very similar to those obtained in the present study with the digital technology.

In the present study, the participants were randomly selected from the medical staff of our lipid clinic, and IMT values ranged from only 0.4 to 0.8 mm. We fully recognize the importance in reproducibility studies of data obtained in patients with pathological IMTs. However, we are also convinced that these sources of variability are independent of image collection instrumentation and that they affect analog and digital systems with a similar variability. Thus, because the main objective of the present study was to investigate the variability induced by image collection instrumentation, all other possible sources of variability (eg, sonographer, reader, irregularity of vessel wall profile, instrumentation for image measurements) have been kept constant. Obviously, because in the present study only the IMT of CCA was considered, the reproducibility findings presented cannot be extended to either plaques or bulb or internal carotid arteries. However, we believe that an eventual reduction in measurement repeatability regarding plaques or bulb and internal carotid arteries should not be attributed to the instrumentation for image acquisition but rather to other sources of error, such as biological variability and sonographer/reader subjectiveness.

Practical Implication
We also analyzed the impact that an improved precision in IMT measurements might have on sample size and duration of follow-up required for clinical trials based on IMT progression evaluation. With the assumption of a comparison of the mean slopes in a treated and a control group, as, for example, in the CAIUS study, sample size depends on the magnitude of the expected treatment effect and on the within-group SD of the individual slopes. We estimate that

**Figure 4.** Intraobserver absolute differences in the present and in other reproducibility studies.

**Figure 5.** Sample size required in a clinical trial based on IMT progression evaluation, as a function of the expected treatment effect. The estimated treatment effect is reported as a percent of the effect of pravastatin observed in the CAIUS study: with the same measurement variability and the same follow-up length (3 years) as in the CAIUS study (○); with digital equipment, with 3 years' follow-up (●); and with digital equipment, with 2 years' follow-up (□).
Further Advantages of Digital Image Processing Systems

Apart from the higher accuracy and reproducibility, digital systems present a number of technical advantages with respect to the corresponding analog systems. Digital systems can be equipped with multifrequency probes, thus increasing the range of depth at which vessels can be investigated, and they can adapt the frequency and the scan depth used according to the vessel depth, thus the highest possible resolution is always obtained.

The characteristics of digital-based echograms can also be of clinical interest because when an image is frozen, the machine automatically stores not just 1 but a number of images, each of which can be zoomed, thus allowing the best image suitable for plaque or IMT measurements to be chosen. In addition, the presence and extent of atherosclerosis, with measurement and automatic calculation of percent of stenosis of transversal images, can be delineated, again in real time, thus highly improving the reliability of clinical reports. Finally, digital technology allows the storage of images not only on videotape but also on optical disc, thus providing the possibility of sending images via electronic mail, which is a very interesting feature for multicenter clinical trials with centralized readings.

Conclusions

Although the present study was performed on a relatively small sample size, the results strongly support the idea that digital systems may reduce the effect of variability determined with image collection instrumentation.

The variability of IMT measurements is determined by the sonographer and reader, with instrumentation for image collection and for image measurements, and with biological differences between subjects (in patients with increased IMT or plaques, measurements become more inaccurate because of tortuous arteries, eccentric plaques, and irregularities). Because biological differences between subjects cannot be influenced, it is important to reduce, whenever possible, the effect of other factors that determine variability. The present results show not only that the digital system is a reliable method for the IMT assessment in clinical trials but also that compared with the analog system, the gold standard technology for IMT imaging, it can further improve accuracy and precision so that an adequate statistical power can be achieved with a smaller sample size and, thus, with lower costs.

In conclusion, the digital system for IMT evaluation compares well with the more widely used analog system and provides a reliable technology for CC-IMT measurement that can be applied to clinical trials.

References

14. Touboul PJ, Prati P, Scarabino PY, Adrai V, Thibout E, Ducimetiere P. Use of monitoring software to improve the measurement of carotid


Reproducibility Validation Study Comparing Analog and Digital Imaging Technologies for the Measurement of Intima-Media Thickness
Damiano Baldassarre, Elena Tremoli, Mauro Amato, Fabrizio Veglia, Alighiero Bondioli and Cesare R. Sirtori

Stroke. 2000;31:1104-1110
doi: 10.1161/01.STR.31.5.1104

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/31/5/1104