Improved analysis of brachial artery ultrasound using a novel edge-detection software system

R. J. WOODMAN,1 D. A. PLAYFORD,1 G. F. WATTS,1 C. CHEETHAM,2 C. REED,3 R. R. TAYLOR,1,4 I. B. PUDDEY,1 L. J. BEILIN,1 V. BURKE,1 T. A. MORI,1 AND D. GREEN2,4

Departments of 1Medicine and 2Human Movement and Exercise Science, The University of Western Australia, Perth 6001; and Departments of 4Cardiology and 3Medical Physics, The West Australian Heart Research Institute Royal Perth Hospital, Perth 6000, Western Australia, Australia

Received 29 December 2000; accepted in final form 20 March 2001

Woodman, R. J., D. A. Playford, G. F. Watts, C. Cheetham, C. Reed, R. R. Taylor, I. B. Puddey, L. J. Beilin, V. Burke, T. A. Mori, and D. Green. Improved analysis of brachial artery ultrasound using a novel edge-detection software system. J Appl Physiol 91: 929–937, 2001.—Brachial artery ultrasound is commonly employed for noninvasive assessment of endothelial function. However, analysis is observer dependent and susceptible to errors. We describe studies on a computerized edge-detection and wall-tracking software program to allow more accurate and reproducible measurement. In study 1, three purpose-built Perspex phantom arteries, 3.00, 4.00, and 6.00 mm in diameter, were measured with the software. There was a mean bias of 11 μm (P < 0.001 at each level) between known and measured values; the mean resolving power of the software was estimated as 8.3 μm. In study 2, the mean intraobserver coefficient of variation of repeated measures of flow-mediated dilation (FMD) using the software (6.7%) was significantly lower than that for traditional manual measurements using the intima-lumen interfaces (24.8%, P < 0.05) and intima-media interfaces (32.5%, P < 0.05). In study 3, 24 healthy volunteers underwent repeat testing twice within 1 wk; the coefficients of variation for betweenvisit reproducibility of FMD and response to glyceryl trinitrate using the software were 14.7 and 17.6%, respectively. Assuming 80% power and an α of 0.05, eight subjects with matched controls would be required, in a parallel designed study, to detect an absolute 2.5% change in FMD. In summary, we have developed a semiautomated computerized vascular ultrasound analysis system that will improve the power of clinical intervention studies to detect small changes in arterial diameter.

flow-mediated dilation; glyceryl trinitrate; high-resolution ultrasound; wall tracking

THE IMPORTANCE OF THE ENDOTHELIUM in maintaining a healthy vasculature has been increasingly recognized, particularly with respect to endothelial release of nitric oxide and its physiological functions. Endothelium-dependent vasodilation, largely dependent on nitric oxide, is not only impaired with overt vascular disease (7, 12, 18, 36) but is also associated with conventional vascular risk factors (3) and may improve with appropriate interventions (17, 20–22). Endothelial dysfunction is an early manifestation of atherosclerotic disease (3, 5, 26, 28, 39), and recent follow-up studies have demonstrated its usefulness in predicting cardiac events (1, 19, 27).

Brachial arterial flow-mediated dilation (FMD), assessed by high-resolution ultrasound, reflects endothelium-dependent vasodilator function (4), which may be used as a surrogate for function of the coronary circulation (2). Although the techniques for measurement of brachial FMD are well established (33), analysis has generally relied on manual assessment of vessel diameter using visual inspection of single frames and placement of ultrasonic calipers (33). Manual assessment is subject to significant observer error (4, 10, 32) and, hence, may be inappropriate for measurement of the brachial artery diameter response to increased flow, which is generally on the order of 0.2–0.4 mm and is even less in the presence of impaired endothelial function. Computer-assisted analysis utilizing edge detection permits multiple measurements along the vessel wall and would be expected to increase the precision of the measurements.

Although computerized analysis of arterial morphology has been previously described, it has most commonly been used in the ultrasonographic estimation of carotid intima-media thickness (29, 37, 38), in association with intravascular ultrasound (23, 30, 31), and with quantitative coronary angiography (13). We have developed a computerized edge-detection and wall-
tracking software program to allow accurate and reproducible measurement of arterial diameter for clinical studies of endothelial function in the peripheral circulation.

**METHODS**

We describe the analysis system and its laboratory and clinical validation. Ultrasonographic images acquired on phantom models and during clinical assessment of postischemic and glyceryl trinitrate (GTN)-induced dilation of the brachial artery were analyzed using the software system (SS).

**Hardware**

The analysis system consisted of a Dual Pentium II 450-MHz personal computer with 256-MB random access memory, G200 8-MB accelerated graphics port Matrox Millenium graphics card, and a 22-GB hard disk running Windows NT4 SP5, used in conjunction with a standard National Instruments IMAQ-PCI-1407, single-channel, 8-bit monochrome image-acquisition board. The board was accessed through a National Instruments NI-IMAQ 2.1 application program interface, which handled all of the necessary timing, gain control, and memory transfers. The personal computer’s internal serial port was used as the RS-232 interface for controlling the S-VHS JVC SR-3888E video recorder. High-resolution ultrasound images were acquired (Acuson Aspen, Mountain View, CA) by using a 12-MHz linear array probe.

**SS**

LabVIEW 5.1 is an icon-based graphical programming language in which developers build software programs called virtual instruments (VIs). The IMAQ vision tool kit has its own specific set of VIs to handle all the image-acquisition and analysis aspects of the system. Ultrasound studies were recorded on an S-VHS tape inside the ultrasound machine and then played back on a separate S-VHS video recorder for analysis. The IMAQ-PCI-1407 card digitized the analog data back into a digital gray-scale image in real time, and the software analyzed the real-time data at ~16 frames/s.

**Edge-detection analysis.** To perform the analysis, the operator selects three regions of interest (ROI) (Fig. 1A). The “electrocardiogram” (ECG) ROI is used for detecting each R wave, which allows the recording of the vessel diameter to be synchronized with end diastole. An algorithm based on the highest and lowest values within the ECG ROI is employed, with visual feedback provided to the operator on detection of each R wave.

The “calibration” ROI allows the observer to convert the image size on the computer, measured in pixels, to the actual diameter of the artery (mm). This is performed by drawing a box on the screen between two marks on the ultrasound image that are known to be 1 cm apart.

Finally, the “vessel” ROI is used by the operator to select the most stable portion of the artery automatic diameter analysis. An IMAQ automatic binary thresholding VI is used to convert the 256-gray-scale image into black and white. Each column of pixels within the ROI is then analyzed to determine the longest sequence of binary zeros (i.e., black). Once all columns have been processed, they are sorted into numerical order, and the median column length of binary zeros is then taken to be the diameter of the artery. Depending on the size of the ROI chosen by the observer, ~200–400 individual pixel columns are used to create the overall diameter for each frame. With the use of this diameter, a rectangle is then drawn that corresponds to the overall internal vessel diameter to provide visual feedback to the observer, who can verify that the system is accurately tracking the arterial walls during frame processing. Data corresponding to end diastole (the peak of an R wave) and/or continuous data (i.e., from every frame) can be stored for further analysis.

**Display of results.** Once the study has been analyzed, the software plots a graph of the arterial diameter against time (Fig. 1B). Cursors corresponding to each phase of the study (Fig. 1B, vertical lines), e.g., baseline and reactive hyperemia, are adjusted by the observer to the specific times of the study. Clearly erroneous data points may be removed by the

---

**Fig. 1.** A: Still frame of image-acquisition software. Electrocardiogram (ECG), vessel, and calibration regions of interest (ROI) are shown by arrows. B: output screen. The vertical cursors are placed at times corresponding to baseline, flow-mediated dilation (FMD), and glyceryl trinitrate (GTN) administration. FMD and GTN responses are calculated by fitting a third-order polynomial function to the data.
Validation Studies

We carried out three studies to validate the analysis system described above.

Study 1: phantom studies employing SS. Phantom studies were performed to determine the accuracy and resolving power of the SS. Three Perspex phantoms were constructed, employing a reaming tool, to diameters of 3.0, 4.0, and 6.0 mm, as determined by micrometry. Fluid flow through the tubes was obtained with saline solution and the aid of a hemodialysis pump; plastic sleeves attached the pump to the ends of each phantom. The diameters of the phantoms were imaged directly by using the imaging protocol described in study 3, with images being optimized to obtain the greatest contrast between the Perspex and saline solution interfaces. A simultaneous ECG signal from a volunteer was recorded to enable R-wave gating for analysis. Analysis was performed over a period of 1 min by one observer using the software. Because artifactual borders are apparent when vessels are imaged (15), to obtain an estimate of the true luminal diameter, the software was modified for this study to measure from the top of the near-wall artifact to the top of the far-wall artifact (15).

Study 2: comparison of software and manual systems of measurement. Twenty scans, stored on super-VHS videotape, were selected from our database of clinical studies and carried out as described in study 3. The scans were selected to cover a wide range of brachial artery diameters and responses to FMD and GTN. Three observers then analyzed each of the scans on two separate occasions separated by at least 1 wk. Three methods of analysis were performed by each observer: two manual and one with the software. All measurements were performed in a blinded fashion.

SOFTWARE (SS) ANALYSIS. Analysis of FMD and GTN responses of the brachial artery were performed using the edge-detection software as described above.

MANUAL ANALYSIS. Measurements were taken at end diastole (corresponding to an R wave). In contrast to the software analysis where FMD diameter measurements were taken as the peak response after cuff deflation, measurements for FMD diameter commenced at the traditional manual analysis time point of 60 s after cuff deflation, and GTN diameter responses took place 300 s after GTN administration (6). Ultrasonographic images with good intimal definition were chosen for both basal and stimulated responses. Two diameter measurements, within close proximity (~2–8 mm apart), of the near-to-far intima-media interface (MM line) (4) and of the near-to-far intima-lumen interface (EE line) were then obtained on two consecutive R waves. The four measurements were averaged, and the percent increases from baseline diameter were calculated for both FMD and GTN responses.

Study 3: clinical study of biological variability. FMD and the GTN responses were assessed in 24 healthy volunteers on two occasions, 7 days apart. The 24 subjects were aged 55 ± 6 yr, with a body mass index of 26.9 ± 2.8 kg/m², systolic blood pressure of 125 ± 14 mmHg, diastolic blood pressure of 75 ± 9 mmHg, total cholesterol of 5.5 ± 0.9 mmol/l, low-density lipoprotein cholesterol of 3.6 ± 0.7 mmol/l, high-density lipoprotein cholesterol of 1.3 ± 0.2 mmol/l, and triglycerides of 1.3 ± 0.9 mmol/l.

CLINICAL PROTOCOL FOR ULTRASONOGRAPHIC MEASUREMENT OF ENDOTHELIAL FUNCTION. The detailed clinical protocol has been described elsewhere (4, 24). Briefly, subjects lay supine for 20 min before the first ultrasound scan was recorded with three-lead ECG monitoring. Ultrasound images of the brachial artery were obtained in the distal one-third of the upper arm. Ultrasonic parameters were set to optimize longitudinal, two-dimensional B-mode images of the lumen-arterial wall interface with the focus zone positioned on the near wall. Once set, the operating parameters remained constant throughout each session while the probe was held by a stereotactic clamp and its precise location was recorded for the repeat session. A rapid inflation-deflation pneumatic cuff, placed around the forearm immediately distal to the humeral epicondyles, was used to provide a stimulus for FMD. After the initial 20-min resting period, baseline scans assessing vessel diameter were recorded over 2 min. The proximal forearm cuff was then inflated to 200 mmHg for 5 min. Recordings commenced 30 s before cuff deflation and continued for 3 min after cuff deflation. After 15 min, a second baseline scan was recorded. This was followed by sublingual administration of GTN (400 μg).

Analysis of these studies was performed by using the SS, and the results were used to determine biological reproducibility and sample sizes required for use in clinical trial design.

Statistical Methods

Study 1. The accuracy or systematic error of the SS in measuring the three phantom arteries was assessed by performing a t-test to determine whether the mean of the measured diameters differed from the known diameters of each artery. The mean bias of the software was assessed by pooling the data of the three phantoms. The resolving power was calculated for each diameter as the SD of replicate measurements over 1 min. The mean resolving power of the software was calculated as the mean of the resolving powers obtained from the three phantoms.

Study 2. BIAS. Bland-Altman plots of SS vs. EE-line and SS vs. MM-line data were constructed for baseline diameter, %FMD, and %GTN response, by using the pooled data of the three observers’ measurements. Systematic error was assessed as the mean difference between methods for each variable using a paired t-test. Proportional bias was examined using simple linear regression. Mean values with the use of manual and software methods were also compared by two-way (observer and method) repeated-measures ANOVA for individual observer data and also by one-way (method) repeated-measures ANOVA for pooled observer data.

VARIABILITY. Intraobserver variability was assessed by using coefficients of variation (CV) for each of the three observers. CVs were calculated as the pooled SD for the replicate mean (11). On each of their three sets of measurements, a CV for each method was also calculated from the pooled results of the three observers. CVs were compared by using an F test of variance ratios. General linear models (SAS 6.12) were used to estimate the between-observer, between-subject, and random-error contributions to the total variance in the measures of baseline diameters, %FMD, and %GTN response for each of the three methods. In these analyses, we employed the mean of two measurements from each observer.

Study 3. Within-subject variability was calculated as the CV for two visits of FMD and response to GTN. Sample size tables for FMD for clinical trials were calculated by using MS DOS-based software (8). The SD for determining sample size was calculated by using the SD of the mean of visits 1 and 2.
was estimated as 8.3
0.56%, and the mean resolving power of the software
and measured values was, therefore, 0.011 mm or
respectively. The overall mean bias between known
the %FMD were significantly lower (P,

, 0.01), respectively, for SS analysis vs.
EE-line analysis (P < 0.001 for both). Proportional bias between methods (by linear regression) was not statistically significant. Table 1
displays the mean and SD data for baseline diameter, FMD, and GTN measured by the three methods for
each observer and for all observers combined. Repeated-
measures ANOVA confirmed that the mean baseline
diameters were significantly lower and the mean re-
response to GTN significantly higher when measured with
SS compared with MM line, both for individual
observers (P < 0.01) and for all observers combined
(P < 0.01). Repeated-measures ANOVA for the indi-
vidual observers also showed that there was a border-
line significantly greater FMD (P = 0.055) when mea-
sured with SS compared with MM-line analysis.
In these analyses, there were no significant differences
between SS and EE line.

Variability. The intraobserver CVs for baseline bra-
chial artery diameter, %FMD, and %GTN-mediated
dilation measured by the three methods, by each ob-
server and by all observers combined, are shown in Fig.
3, A, B, and C, respectively. The intraobserver CV was
significantly lower for brachial artery baseline diam-
ter and FMD with the software compared with both
manual methods (P < 0.05 for both). For GTN-mediated
dilation, however, the pooled intraobserver CV was
lower for SS compared with MM line (P < 0.05),
but not for SS compared with EE line. The average of
the intraobserver CVs measured by the software for
brachial artery diameter, %FMD, and %GTN-mediated
dilation were 0.36, 6.7, and 3.9%, respectively, and the
pooled data intraobserver CVs were 0.16, 3.6, and
2.2%, respectively. For the EE line, the corresponding
intraobserver CVs were 1.15, 24.8, and 6.9%, and for
the MM line these values were 2.03, 32.5, and 10.3%,
respectively. Table 2 shows results of the variance
components using general linear models. It can be seen
that, after accounting for the large between-subject
variance, the between-observer and random-error vari-
ance components for both brachial artery diameter and
%FMD were least for SS compared with the other
methods. For %GTN, the combined between-observer
and random-error variance component was least for SS.

Study 3: Clinical Studies

The mean (±SD) %FMD values for visits 1 and 2
were 6.6 ± 2.7 and 6.4 ± 3.2%, respectively. The
combined mean was 6.5 ± 2.8%, and the mean absolute
difference between visits was 1.6 ± 1.0%. The between-
visit CV for %FMD was 14.7%, and the Pearson corre-
lation coefficient was 0.821. The mean %GTN for visits
1 and 2 were 19.2 ± 8.5 and 18.5 ± 9.0%, respectively.
The combined mean was 18.8 ± 8.4%, and the mean
absolute difference between visits was 3.8 ± 3.3%. The
between-visit CV for %GTN was 17.6%, and the Pear-
son correlation coefficient was 0.836.

Power calculations for sample sizes required to show
significant changes in %FMD, for both crossover and
parallel designed intervention studies, are shown in
Table 3. For example, assuming a power of 80% (α = 0.05), a parallel designed study would require 16 sub-
jects in total to detect a 2.5% change in FMD over the
study duration and 24 subjects to detect a 2% change,
whereas five and six subjects, respectively, would be
required in a crossover design.

DISCUSSION

We have described and validated an accurate and
reproducible edge-detection and wall-tracking SS for
use in ultrasound studies of the brachial artery. The
dge-detection software reduces observer error signifi-
cantly and will allow clinical studies to be capable of
detecting significant changes in endothelial function
with substantially fewer subjects.

Reproducibility of measurements was significantly
better than with manual methods; intraobserver CVs
were significantly better than traditional EE-line and
MM-line methods in the analysis of baseline brachial
artery diameter and %FMD and also compared with
MM-line manual analysis for response to GTN. These
findings are important because previous reproducibil-
ity studies suggest that observer error for analysis of
FMD may account for as much as 60% of the within-
subject variation (4, 33). In addition, because the use
of the software allows continuous monitoring of arterial
diameter, true peak response is more reliably assessed.
Fig. 2. Bland-Altman plots of software system vs. intima-lumen interface (EE line; A, C, E) and intima-media interface (MM line; B, D, F) for baseline diameter (A and B), %FMD (C and D), and %GTN (E and F). Dashed lines represent means ± 2 SD for estimated difference between methods.
Previous reports of observer reproducibility have not always been expressed appropriately (33). In a study of repeat scans on 21 subjects, with repeat observations on 127 scans, Celermajer et al. (4) reported an interobserver CV using nested ANOVA of 1.4% for MM-line manual analysis. However, FMD was expressed as a percentage of the baseline diameter, rather than as a percent increase above baseline diameter. Expressed this way, our intraobserver CVs using the software ranged from 0.2 to 0.45% for the three individual observers and from 0.95 to 1.15% using MM-line manual analysis. This improvement in CV associated with use of the software occurred with a mean FMD in our study that was similar to that of Celermajer et al. (5.15 vs. 4.85%), an important consideration because the CV is inversely proportional to the mean FMD. Whereas interobserver variability is likely to be higher than intraobserver variability, especially when nonautomated methods of analysis are used, general linear model analysis demonstrated that between-observer variance contributed only a small degree to the total variance using the software. The between-observer variance component of baseline diameter using the software was only 3.8 and 4.5% of that for the EE-line and MM-line manual methods, respectively. The between-observer variance component of %FMD was 27.0 and 42.5% of that for the EE-line and MM-line manual methods, respectively. Interobserver error was thus also substantially reduced, which would assist in studies involving more than one observer.

There have also been several studies of within-subject biological variability of brachial artery FMD using manual analysis techniques (16, 33). However, this study is the first to assess comprehensively the reproducibility and power of the test using a semiautomated edge-detection algorithm. For example, Stadler et al. (34) assessed resting diameter repeatability using their own edge-detection algorithm but not FMD repeatability. This is an important point as CVs for FMD repeatability (33) are 10- to 20-fold higher than resting diameter repeatability (6, 35). Indeed, our own resting diameter CV of 0.36% using the software vs. the CV of 6.7% for FMD demonstrates the importance of assessing the reproducibility of changes in diameter rather than just baseline diameter measurements alone. The between-visits CV of 14.7% compares favorably with most other studies that have used manual methods of measurement. Although CVs as low as 2.3 and

### Table 1. Comparison of means

<table>
<thead>
<tr>
<th>Method of Analysis</th>
<th>Software</th>
<th>EE line</th>
<th>MM line</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline diameter, mm</td>
<td>3.80 ± 0.59</td>
<td>3.67 ± 0.59</td>
<td>4.33 ± 0.67</td>
</tr>
<tr>
<td>%FMD</td>
<td>5.00 ± 3.81†</td>
<td>4.39 ± 5.07</td>
<td>3.84 ± 5.04</td>
</tr>
<tr>
<td>%GTN</td>
<td>19.32 ± 4.63*</td>
<td>18.54 ± 4.80*</td>
<td>15.78 ± 4.52</td>
</tr>
<tr>
<td>Observer 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline diameter, mm</td>
<td>3.78 ± 0.58*</td>
<td>3.83 ± 0.58*</td>
<td>4.43 ± 0.79</td>
</tr>
<tr>
<td>%FMD</td>
<td>5.12 ± 3.83†</td>
<td>4.63 ± 4.03</td>
<td>3.26 ± 3.16</td>
</tr>
<tr>
<td>%GTN</td>
<td>19.27 ± 4.63*</td>
<td>18.16 ± 4.40*</td>
<td>15.09 ± 3.48</td>
</tr>
<tr>
<td>Observer 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline diameter, mm</td>
<td>3.77 ± 0.59*</td>
<td>3.83 ± 0.52*</td>
<td>4.31 ± 0.72</td>
</tr>
<tr>
<td>%FMD</td>
<td>5.32 ± 3.67†</td>
<td>4.30 ± 3.41</td>
<td>4.11 ± 2.90</td>
</tr>
<tr>
<td>%GTN</td>
<td>19.24 ± 4.31*</td>
<td>17.5 ± 4.98*</td>
<td>15.2 ± 3.49</td>
</tr>
<tr>
<td>Pooled data (all observers)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline diameter, mm</td>
<td>3.78 ± 0.59*</td>
<td>3.78 ± 0.55*</td>
<td>4.36 ± 0.68</td>
</tr>
<tr>
<td>%FMD</td>
<td>5.15 ± 3.85</td>
<td>4.44 ± 3.92</td>
<td>3.73 ± 3.20</td>
</tr>
<tr>
<td>%GTN</td>
<td>19.28 ± 4.64*</td>
<td>18.07 ± 4.42†</td>
<td>15.36 ± 3.76</td>
</tr>
</tbody>
</table>

Values are means ± SD. Shown are comparison of means for baseline diameter, %flow-mediated dilatation (FMD)- and %glyceryl trinitrate (GTN)-mediated dilatation. EE line, intima-lumen interface; MM line, intima-media interface. Significantly different from MM-line manual analysis, *P < 0.01, †P = 0.05.

![Fig. 3. Intraobserver coefficient of variation (CV) for baseline brachial artery diameter (A), %FMD (B), and %GTN-mediated dilation (C) measured by the software and the 2 manual methods of analysis (EE line and MM line) by each observer and by all observers combined (pooled data). *Significantly different from corresponding analysis with software, P < 0.05.](http://jap.physiology.org/Downloaded from http://jap.physiology.org/)
1.8% have been reported (4, 33), these were again obtained by expressing FMD as a percentage of the baseline diameter. As discussed above, more appropriately, the CVs in the studies quoted would have been 33 and 28%, respectively (with mean FMDs of 7.4 and 7.0%), when FMD is expressed as a percent change. Other recent studies have also reported apparently low but inappropriate CVs (9, 35). Liang et al. (16) obtained a true CV of 10.8% in a group of 30 healthy subjects, with a larger mean FMD of 10.8%. Had the mean response in our study also been 10.8% and our within-subject SD remained the same, our between-visit CV would have been 8.9%. We estimated that, in an intervention study, we would have a power of 80% to detect a 2.5% change in FMD in a crossover design with only five subjects enrolled, and with only eight subjects enrolled in each arm in a parallel design, demonstrating that even small changes during intervention should be detectable. These sample sizes are much smaller than those of 10 and 45, respectively, estimated previously (33) using manual analysis methods and FMDs of a similar magnitude to those in this study and are also smaller than the sample sizes of 11 and 19, respectively, obtained with manual analysis and much larger FMDs (16). No studies using edge-detection algorithms have reported on the power of the test.

Computer-assisted measurement of carotid intima-media thickness has been employed for some time (29, 37, 38). More recently, several groups have used semi-automated methods for determining brachial artery diameter (14, 25, 34, 35). However, most of the latter systems rely on initial manual identification of the near and far walls (14, 25, 34) and/or are not designed for continuous frame analysis, which is necessary to determine the time course of the response and the maximum dilation. In addition, appropriate validation of these other methods against traditional manual analysis either has not been reported (14, 25, 35) and/or has only been assessed under resting conditions (34, 35). Perhaps more importantly, the mean intraobserver CVs for resting artery diameter of 0.36% measured using our software compare favorably with the intraobserver CVs reported for other semiautomated baseline diameter analyses of 2.9% (35) and 1.5% (34). Furthermore, it is noteworthy that both of these were worse than the EE-line manual CV of 1.1% achieved in this study. Sonka et al. (32) recently published a report of a semiautomated system similar to our own, but regression analysis between the resting arterial diameters measured by semiautomated and manual methods only suggested that they were not different. In addition, comparisons were only made during rest rather than hyperemia, the importance of which has been pointed out above.

Our SS had a mean bias of 11 μm for the three phantoms employed and had a mean resolving power of 8.3 μm. The accuracy compares favorably with the edge-detection algorithm reported by Stadler et al. (34) of 52 μm, and the resolving power is much lower than the resolution of ultrasonic calipers of most ultrasound machines (~100 μm). In practical terms, however, the resolving power is limited by the confines of the measurement system, i.e., one screen pixel or 50 μm. Study 2 indicated a measurement bias between the software and MM-line manual analysis for baseline diameter, %FMD, and response to GTN. This was expected, because the software measures the smaller EE-line diameter. Despite there being no differences in baseline diameter between the software and EE-line analyses, measures of both %FMD and response to GTN tended to be higher with the software. This indicates that the software may be measuring a truer maximum than that obtained at the arbitrary, predetermined time points for manual analysis of 60 and 300 s, respectively. Analyzing with the software at a fixed time point of 60 s may have lowered the intraobserver CVs still further, but we felt it was preferable to ascertain as closely as possible the true peak for FMD.

Table 3. Estimated sample sizes

<table>
<thead>
<tr>
<th>Power, %</th>
<th>α-Error</th>
<th>Improvement in FMD, %</th>
<th>Sample Size of Study Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>0.05</td>
<td>0.5</td>
<td>182</td>
</tr>
<tr>
<td>80</td>
<td>0.05</td>
<td>1.0</td>
<td>46</td>
</tr>
<tr>
<td>80</td>
<td>0.05</td>
<td>1.5</td>
<td>21</td>
</tr>
<tr>
<td>80</td>
<td>0.05</td>
<td>2.0</td>
<td>12</td>
</tr>
<tr>
<td>80</td>
<td>0.05</td>
<td>2.5</td>
<td>8</td>
</tr>
<tr>
<td>80</td>
<td>0.05</td>
<td>3.0</td>
<td>6</td>
</tr>
<tr>
<td>80</td>
<td>0.05</td>
<td>3.5</td>
<td>5</td>
</tr>
<tr>
<td>80</td>
<td>0.05</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>90</td>
<td>0.05</td>
<td>0.5</td>
<td>243</td>
</tr>
<tr>
<td>90</td>
<td>0.05</td>
<td>1.0</td>
<td>61</td>
</tr>
<tr>
<td>90</td>
<td>0.05</td>
<td>1.5</td>
<td>28</td>
</tr>
<tr>
<td>90</td>
<td>0.05</td>
<td>2.0</td>
<td>16</td>
</tr>
<tr>
<td>90</td>
<td>0.05</td>
<td>2.5</td>
<td>11</td>
</tr>
<tr>
<td>90</td>
<td>0.05</td>
<td>3.0</td>
<td>8</td>
</tr>
<tr>
<td>90</td>
<td>0.05</td>
<td>3.5</td>
<td>6</td>
</tr>
<tr>
<td>90</td>
<td>0.05</td>
<td>4.0</td>
<td>5</td>
</tr>
</tbody>
</table>

Shown are estimated sample sizes required for crossover and parallel study designs to detect significant changes in flow-mediated dilatation using the results of study 3.
Whereas reducing measurement error is important, it is also desirable to eliminate other potential sources of variation. Optimal imaging techniques, use of imaging landmarks such as branch points or veins, use of a stereotactic clamp for the ultrasound probe, and ECG gating should all be employed. Sources of variation in biological CV should be kept to a minimum by ensuring that ambient temperature, exercise status, fasting state, and psychological status all remain unchanged. It has not yet been determined whether correcting for flow significantly improves reproducibility. Sonka et al. (32) assessed FMD with or without correction for blood flow. Linear regression coefficients for repeat scans were lower for flow-corrected FMD when baseline diameter was taken as the arterial diameter during cuff inflation but not when normal resting arterial diameter was used. In addition, no statistical analysis was used to compare the coefficients.

The computerized approach described here obtains an accurate, continuous assessment of diameters, thus allowing evaluation of the time course of diameter and blood flow changes, as well as facilitating the detection of peak response. It is limited only by image quality, the processing power of the computer (16–17 frames/s with the hardware described), and/or the frame rate of the recording medium (25 frames/s). An automated approach should also lower costs by obviating the need for repeat measurements by two observers, as demonstrated by the low intraobserver CVs and the small contribution of observer variance to total variance. Digitization of ultrasound images would improve picture quality and resolution presently lost in the transfer of images from digital to analog and back to digital. This would also obviate the need for a video and frame grabber and would accelerate analysis. Development of a real-time analysis package to track brachial artery diameter as scanning is taking place may facilitate the technique of brachial artery ultrasound to be used in intervention studies with realistic sample sizes.

We are indebted to Mark Trotman at Icon Technologies, Perth, Western Australia, for expert programming in the LabVIEW environment and helping us overcome numerous technical problems during the development; Lotteries Commission of Western Australia for providing the ultrasound machine used for data collection in this study; and Renate Zilkens, Lisa Rich, and Dr. Jonathan Hodgson for permission to use some of their data for study 3.

National Health and Medical Research Council of Australia provided the program grant that enabled development of the software. Extensive literature searches by ourselves and patent attorneys in Perth appointed by the University of Western Australia established that there are no reports of similar techniques for luminal diameter measurement. Accordingly, we have obtained a provisional patent (Patent PQ 7882 issued on 30/5/2000) entitled, Method and Apparatus for Establishing a Vessel Diameter.

REFERENCES


