

## Precision and Discriminatory Ability of Calcaneal Bone Assessment Technologies\*

SUSAN L. GREENSPAN,<sup>1</sup> MARY L. BOUXSEIN,<sup>2</sup> MARY E. MELTON,<sup>3</sup> ANDREA H. KOLODNY,<sup>3</sup>  
JAMES H. CLAIR,<sup>3</sup> PAUL T. DELUCCA,<sup>3</sup> MICHAEL STEK, JR.,<sup>3</sup> KENNETH G. FAULKNER,<sup>4</sup>  
and ERIC S. ORWOLL<sup>5</sup>

### ABSTRACT

To determine if measuring skeletal status at the calcaneus is a potentially valuable technique for diagnosing osteoporosis, we examined five calcaneal assessment techniques in 53 young normal women and 108 postmenopausal women with osteoporosis and compared these measurements to dual-energy X-ray absorptiometry (DEXA) at the calcaneus, hip, and spine. The five instruments, including single-energy X-ray absorptiometry (SEXA) and four quantitative ultrasound (QUS) instruments, were evaluated for precision, ability to discriminate osteoporotic from young normal subjects, and correlation to the other instruments. The coefficient of variation (%CV) for instrument, positioning, interobserver, and short-term precision of the five calcaneal instruments ranged from 1.34–7.76%, 1.63–7.00%, 1.84–9.44%, and 1.99–7.04%, respectively. The %CVs for positioning, interobserver, and short-term precision were similar for calcaneal DEXA, calcaneal SEXA, and stiffness (as measured by Achilles). The %CVs for instrument precision were similar between calcaneal DEXA and SEXA. The ability of the five calcaneal instruments to discriminate osteoporotic from young normal subjects was similar based on the analysis of area under the receiver operating characteristic curves (range 0.88–0.93) and equivalent to DEXA of the calcaneus and hip (0.88–0.93). The correlations between the measurements of five calcaneal instruments were strong ( $0.80 \leq r \leq 0.91$ ,  $p < 0.001$ ). These data suggest that although the precision is variable, the calcaneal QUS and SEXA instruments can discriminate between osteoporotic patients and young normal controls and appear to be a useful technique for assessment of osteoporosis. (J Bone Miner Res 1997;12:1303–1313)

### INTRODUCTION

OSTEOPOROSIS IS A growing concern due to the increased aging of our population. In the United States, there are more than 250,000 hip fractures annually at a cost of over \$10 billion.<sup>(1,2)</sup> Because women over the age of 65 are the fastest growing segment of the U.S. population,<sup>(3)</sup> it is

expected that the number of hip fractures alone will double or triple within the next quarter century.<sup>(4)</sup> Along with established agents, such as estrogen replacement therapy, additional therapeutic alternatives are available to treat osteoporosis.<sup>(5)</sup> It has therefore become imperative to evaluate and diagnose patients with osteoporosis so that preventive or therapeutic measures can be instituted as soon as possible.

Although there are several techniques to assess bone mineral density (BMD) or content (BMC), including quantitative computerized tomography, dual photon absorptiometry, single photon absorptiometry, and radiographic ab-

\*This work was presented in part at the Second Annual Scientific Meeting of the Society for Clinical Densitometry, January 18–21, 1996, Colorado Springs, Colorado, U.S.A.

<sup>1</sup>Osteoporosis Prevention and Treatment Center, Beth Israel Deaconess Medical Center, Boston, Massachusetts, U.S.A.

<sup>2</sup>Orthopedic Biomechanics Laboratory, Beth Israel Deaconess Medical Center, Boston, Massachusetts, U.S.A.

<sup>3</sup>Merck & Co., Inc., West Point, Pennsylvania, U.S.A.

<sup>4</sup>Oregon Osteoporosis Center, Portland, Oregon, U.S.A.

<sup>5</sup>Bone and Mineral Research Unit, Portland VA Medical Center and Oregon Health Sciences University, Portland, Oregon, U.S.A.

sorptiometry, dual-energy X-ray absorptiometry (DEXA) of the hip and spine has become the most widely accepted technique for evaluation of skeletal status.<sup>(6)</sup> DEXA measurements have good precision and accuracy, low radiation exposure, and are associated with hip and vertebral fracture risk.<sup>(6)</sup> However, because these machines require dedicated office space and can be expensive, they are not always accessible and tend to be located mainly in urban areas.

Quantitative ultrasound (QUS) and single-energy X-ray absorptiometry (SEXA) are potentially useful techniques to evaluate skeletal integrity at peripheral sites such as the calcaneus, patella, tibia, finger, and forearm. While SEXA measures BMC (g/cm), QUS uses sound waves to assess skeletal status. Two parameters are typically measured by QUS: ultrasound velocity or speed of sound (SOS), which reflects the speed of the ultrasound wave; and broadband ultrasound attenuation (BUA), which reflects the frequency dependence of ultrasound attenuation.<sup>(7)</sup> An additional parameter, stiffness, is calculated as a linear combination of SOS and BUA. Both QUS and SEXA are associated with fracture risk.<sup>(8-12)</sup> Although calcaneal devices have been developed that are small, portable, and relatively inexpensive compared with larger, stationary DEXA machines, these techniques have not received widespread clinical acceptance in the U.S.

This study was undertaken to evaluate five calcaneal bone assessment instruments for precision, correlation, and ability to discriminate osteoporotic from normal subjects. The investigation allowed for comparisons between the various calcaneal instruments and comparisons with standard DEXA measurements.

## MATERIALS AND METHODS

### Subjects

Two U.S. centers participated in the study. Both centers received Institutional Review Board approval, and each subject gave written informed consent prior to participation. The study population consisted of two groups, young normal and osteoporotic. The subjects were recruited from either the study sites' current patient population or through community-based advertising. The young normal group ( $n = 53$ ) was comprised of non-Hispanic, Caucasian females aged 25-35 who were not pregnant or lactating and without a history of osteoporosis or medical disorders associated with low bone mass. The osteoporotic group ( $n = 108$ ) was comprised of non-Hispanic, Caucasian, postmenopausal females at least 55 years of age with a BMD more than 2.5 standard deviations (SD) below the mean young normal bone mass at the femoral neck or trochanter ( $<0.645$  or  $<0.497$  g/cm<sup>2</sup>, respectively) as determined by the Hologic normative database. The osteoporotic group was comprised of two subgroups: subjects with no history of osteoporotic fracture (osteo non-Fx) ( $n = 52$ ) and subjects with osteoporotic fracture (osteo Fx) ( $n = 56$ ). An osteoporotic fracture was defined as a fracture that (1) occurred after menopause, (2) was due to no more than a moderate trauma (i.e., energy less than or equal to a fall from stand-

ing height), and (3) involved 1 of 15 fracture sites associated with low bone mass or mild/moderate trauma.<sup>(13)</sup>

Excluded from the osteoporotic group were subjects with a history or evidence of metabolic bone disease (other than postmenopausal bone loss), those ever treated with fluoride, or those who took high doses of calcium ( $>1500$  mg/day) or vitamin D ( $>800$  IU/day) within the past year. Subjects who initiated use of the following medications within the prior year were excluded: estrogen, glucocorticoids, anticonvulsants, anticoagulants, bisphosphonates, calcitonin, androgen, or vitamin D. Subjects on these medications at a stable dose during the entire year prior to the study and who planned to continue use throughout the study were eligible.

### Bone assessment technologies

Four QUS instruments were examined and included: Achilles, Lunar Corporation (Madison, WI, U.S.A.); CUBA, McCue Ultrasonics, LTD. (Hampshire, U.K.); QUS-IX, Osteo Sciences Corporation (Beaverton, OR, U.S.A.); and Ultrasonic Bone Analyzer (UBA575+), Hologic, Inc. (Waltham, MA, U.S.A.) (Table 1). A single calcaneal SEXA instrument was used: OsteoAnalyzer, Dove Medical Systems, Inc. (Newbury Park, CA, U.S.A.). DEXA was used to assess calcaneal, hip, and spine BMD (QDR-1500 and QDR-2000; Hologic, Inc.). All of the QUS instruments provided multiple measurement parameters; however, the single measurement parameter considered by the corresponding instrument manufacturer as the most clinically relevant was used in the primary analyses. Secondary analyses with the other QUS measurement parameters are presented in Appendix 1.

Measurements were conducted by two technicians at each study site. One technician at each site was identified as the primary technician for the Achilles, OsteoAnalyzer, and UBA575+, and the other technician was the primary technician for the CUBA, calcaneal DEXA, and QUS-IX. The technicians received training on each machine and were instructed to conduct all measurements according to the manufacturers' operating manuals.

### Study design

This cross-sectional study was conducted during a 7-month period. Qualification of each subject was determined by obtaining a medical history including prior medications and a disease-directed physical examination on the subjects who reported any ailments. All subjects of child-bearing potential were required to have a negative urine pregnancy test. Vital signs, height, and weight were collected. Subjects who met all inclusion and exclusion criteria underwent a DEXA scan of the hip and posteroanterior spine.

Within each group (young normal, osteo non-Fx, osteo Fx), qualified subjects were randomly assigned to one of four groups. All measurements were conducted by the primary technician for the instrument with the exception of the positioning/interobserver precision group. The baseline group underwent one measurement on each of the five

TABLE 1. CALCANEAL INSTRUMENTS

<i>Instrument (manufacturer) Measurement parameters</i>	<i>Units</i>	<i>Instrument type</i>	<i>Weight (lb)</i>	<i>Water bath</i>
Achilles (Lunar Corporation) *stiffness (0.67 [BUA] + 0.28 [SOS]) - 420 broadband ultrasound attenuation (BUA) speed of sound (SOS)	% dB/MHz m/s	QUS	46	yes
CUBA (McCue Ultrasonics, Ltd.) *BUA velocity of sound (VOS)	dB/MHz m/s	QUS	28	no
QUS-1X (Osteo Sciences Corporation) *ultrasonic bone index-4N (UBI-4N) <sup>†</sup> BUA ultrasonic bone index-4 (UBI-4) <sup>†</sup>	dB/MHz/cm dB/MHz dB/MHz	QUS	3	no
UBA575+ (Hologic, Inc.) *BUA bone velocity (Vb) SOS	dB/MHz m/s m/s	QUS	25	yes
OsteoAnalyzer (Dove Medical Systems, Inc.) *bone mineral content (BMC)	mg/cm <sup>2</sup>	SEXA	72	yes
DEXA calcaneus (Hologic, Inc.) *bone mineral density (BMD)	g/cm <sup>2</sup>	DEXA	1250	no

QUS, quantitative ultrasound; SEXA, single-energy X-ray absorptiometry; DEXA, dual-energy X-ray absorptiometry.

\* Measurement parameter selected by the instrument manufacturer as being the most clinically relevant, and therefore used in the primary analyses. Analyses with the other measurement parameters are presented in Appendix 1.

<sup>†</sup> UBI-4 is a variant of BUA. UBI-4N is normalized for heel size (UBI-4/est. heel width).

calcaneal instruments plus calcaneal DEXA within a single day. The instrument precision group underwent a total of five measurements on each instrument without repositioning; measurements for all instruments were completed within a 1-week period. The positioning/interobserver precision group underwent a total of four measurements on each instrument within a 1-week period. The first three measurements were conducted with repositioning by the primary technician; the fourth measurement was conducted by the secondary technician. The short-term precision group underwent one measurement by each of the instruments on 5 separate days within a 2-week period.

Throughput time was defined as the time inclusive of preparing the subject for the measurement (e.g., swabbing the heel with alcohol, positioning the foot), entering the subject demographics into the computer, performing the measurement, and analyzing the measurement.

An in vivo cross calibration of the two units of each of the four QUS instruments was conducted at the beginning of the study (at one study site) and at the end of the study (at the other study site). A total of 10 subjects were selected at each study site with T scores at the femoral neck or trochanter ranging from -3.0 to +3.0 on a recent DEXA scan (within the past 6 months). Each subject underwent five measurements (without repositioning) on each instrument by the primary technician. In vitro cross calibration was performed at the beginning and end of the study using instrument-specific phantoms supplied by the manufacturers.

### Statistical analysis

*Precision:* Precision was assessed using two parameters, coefficient of variation (%CV) and standardized coefficient of variation (%SCV), for the five calcaneal instruments and compared with the precision of calcaneal DEXA. For testing and estimation purposes of the %CVs for instrument, positioning, interobserver, and short-term precision, the logarithmic transformation was used. The SD of the logarithmic transformed data was computed for each subject on each instrument. The estimated %CVs were pooled across subjects to obtain estimates for each instrument. The estimated %CVs of the instruments were compared by taking the square root of the 95% confidence intervals on the ratios of the squared %CVs for all pairs of instruments using the F-distribution. The %SCV was calculated using the ratio of the pooled within individual variation to the 95% range of the data. Ninety-five percent confidence intervals were generated for all precision estimates. The lowest precision indicates the best precision.

*Measurements, T scores, and throughput time:* Descriptive statistics were calculated on the first measurement of each instrument and subject. For each instrument, the mean measurements from the young normal, osteo non-Fx, and osteo Fx group were compared by F-test from two-way analysis of variance (ANOVA) tests with effects for center and cohort. T scores were calculated for each subject and calcaneal instrument using the young normal group from the study as the reference database:

TABLE 2. DEMOGRAPHICS OF STUDY POPULATION

	Total	Young normal	Osteoporotic all	Osteo non-Fx	Osteo Fx
Subjects enrolled	161	53	108	52	56
Age (years)					
mean $\pm$ SEM	57.4 $\pm$ 1.60	30.2 $\pm$ 0.43	70.8 $\pm$ 0.78	68.8 $\pm$ 1.02	72.5 $\pm$ 1.14
range (min-max)	24-92	24-35	55-92	55-83	55-92
Postmenopausal (years)					
mean $\pm$ SEM	NA	NA	24.1 $\pm$ 0.99	22.8 $\pm$ 1.46	25.3 $\pm$ 1.33
range (min-max)	NA	NA	3-56	6-45	3-56
Height (cm)					
mean $\pm$ SEM	161.1 $\pm$ 0.55	165.7 $\pm$ 0.72	158.8 $\pm$ 0.63	159.1 $\pm$ 0.96	158.5 $\pm$ 0.83
range (min-max)	141-176	152-176	141-173	144-173	141-172
Weight (kg)					
mean $\pm$ SEM	61.3 $\pm$ 0.77	63.2 $\pm$ 1.41	60.4 $\pm$ 0.90	60.9 $\pm$ 1.36	60.0 $\pm$ 1.21
range (min-max)	40-100	49-100	40-88	40-88	41-83

NA, not applicable.

$$T \text{ score} = \frac{\text{Measurement} - \text{Young normal group mean}}{\text{Young normal group SD}}$$

In addition, with the young normal group as the reference database, T scores were calculated for DEXA femoral neck and trochanter. Statistical similarity of T scores was assessed by Duncan's multiple range test.

**Discriminatory ability:** Receiver operating characteristic (ROC) curves and area under the curve (AUC) were used to evaluate the sensitivity and specificity of each calcaneal instrument, DEXA femoral neck and trochanter to (1) identify osteoporotic subjects from all study subjects and (2) discriminate between osteoporotic subjects with and without fracture, using the first measurement by each instrument and subject.

**Correlation:** Correlations of the measurements from all calcaneal instruments and hip and spine DEXA were assessed using Pearson's product moment correlation coefficients. The first measurement by each instrument and subject was used to calculate the correlation coefficients.

Analyses were conducted using both an "all scans" and a "per protocol" approach. The all scans approach included the measurements from all subjects who qualified for the study and underwent the protocol specified procedures. The per protocol approach was the result of a post hoc quality assurance audit of all calcaneal measurements conducted for the study (3005 measurements). This quality assurance audit identified 72 measurements (2.4%) that did not meet the manufacturers' specifications (OsteoAnalyzer,  $n = 29/502$  measurements [5.8%]; calcaneal DEXA,  $n = 6/501$  [1.2%]; QUS-1X,  $n = 37/500$  [7.4%]). These measurements were excluded from the per protocol analyses. Reasons for exclusion included incorrect positioning, movement, artifacts, and wrong foot scanned (left or right). The 72 individual measurements that were excluded from the per protocol analyses were distributed among all of the four subject groups. In addition, the measurements of the subjects initially recruited into the osteoporotic group who did not qualify as osteoporotic ( $n = 3$ ) and the measurements of the subjects who did not have all of their measurements

completed in 31 days as specified by the protocol ( $n = 2$ ) were excluded from the all scans and per protocol analyses of measurements, T scores, and discriminatory ability. The measurements from the latter two subjects were also excluded from the correlation calculations. Unless otherwise noted, data presented are from the per protocol analyses.

#### Cross calibration

For each study site and QUS instrument, the data were modeled using a repeated measures ANOVA and 95% confidence intervals on the mean differences between the units were constructed. In addition, descriptive statistics were used to summarize measurements using the phantoms for each unit, and in cases where the same phantom was used at both sites, the difference between sites was also summarized.

## RESULTS

### Demographics

A total of 161 subjects were enrolled in the study (young normal,  $n = 53$ ; osteo non-Fx,  $n = 52$ ; osteo Fx,  $n = 56$ ). In the osteo Fx group, 32 women (57%) had fractures of the hip, vertebrae, and/or wrist and 24 women (43%) had other osteoporosis-associated fractures. Demographics for both of the osteoporotic subgroups were similar (Table 2) with the exception that, on average, the subjects in the osteo Fx group were 4 years older and postmenopausal 2.5 years longer than those in the osteo non-Fx group. A comparison of the subject demographics between the two study sites revealed that the osteo non-Fx group subjects at one site were on average 6.5 years older and 6.1 cm shorter than those at the other study site. These differences were not considered clinically meaningful.

### Precision

Table 3 summarizes instrument, position, interobserver, and short-term precision. The %CV and %SCV for instru-

TABLE 3. PRECISION

Instrument (parameter)	Instrument precision		Positioning precision		Interobserver precision		Short-term precision	
	%CV* (95% CI)	%SCV (95% CI)	%CV* (95% CI)	%SCV (95% CI)	%CV* (95% CI)	%SCV (95% CI)	%CV* (95% CI)	%SCV (95% CI)
Achilles (stiffness)	2.74 <sup>c</sup> (2.44, 3.11)	2.69 (1.74, 3.64)	1.78 <sup>a</sup> (1.51, 2.17)	2.03 (1.53, 2.52)	2.48 <sup>a</sup> (1.99, 3.32)	2.47 (0.59, 4.35)	2.35 <sup>a</sup> (2.08, 2.69)	2.51 (1.94, 3.08)
CUBA (BUA)	4.37 <sup>d</sup> (3.90, 4.97)	3.21 (0.79, 5.63)	4.44 <sup>b</sup> (3.77, 5.40)	3.87 (2.24, 5.50)	7.58 <sup>c</sup> (6.06, 10.13)	6.04 (2.57, 9.51)	5.21 <sup>b</sup> (4.62, 5.96)	4.31 (3.52, 5.11)
QUS-1X (UBI-4N)	7.76 <sup>e</sup> (6.90, 8.86)	8.13 (5.62, 10.64)	7.00 <sup>c</sup> (5.90, 8.63)	6.84 (3.87, 9.81)	9.44 <sup>c</sup> (7.47, 12.86)	9.16 (0.91, 17.42)	7.04 <sup>c</sup> (6.22, 8.10)	6.81 (5.56, 8.06)
UBA575+ (BUA)	2.01 <sup>b</sup> (1.80, 2.29)	1.61 (0.67, 2.54)	3.99 <sup>b</sup> (3.39, 4.86)	3.12 (1.54, 4.70)	4.97 <sup>b</sup> (3.97, 6.65)	3.98 (1.71, 6.24)	6.64 <sup>c</sup> (5.89, 7.60)	5.33 (4.40, 6.26)
OsteoAnalyzer (BMC)	1.34 <sup>a</sup> (1.19, 1.53)	1.16 (0.96, 1.36)	1.63 <sup>a</sup> (1.37, 2.02)	1.68 (1.16, 2.20)	1.84 <sup>a</sup> (1.45, 2.55)	1.72 (0.85, 2.58)	1.99 <sup>a</sup> (1.77, 2.29)	1.76 (1.33, 2.19)
DEXA calcaneus (BMD)	1.23 <sup>a</sup> (1.09, 1.40)	1.07 (0.76, 1.37)	1.63 <sup>a</sup> (1.39, 1.99)	1.68 (0.25, 3.12)	1.91 <sup>a</sup> (1.53, 2.55)	1.96 (0.84, 3.07)	2.02 <sup>a</sup> (1.79, 2.31)	2.35 (1.84, 2.87)

BUA, broadband ultrasound attenuation; BMC, bone mineral content; BMD, bone mineral density; %CV, coefficient of variation; %SCV, standardized coefficient of variation (ratio of the pooled within individual variation to 95% range of the data).

\* Identical letters denote precision that is not statistically different as determined by the confidence intervals of the pairwise ratios within each precision parameter.

ment precision (i.e., repeat measurements with no repositioning) of the five calcaneal instruments ranged from 1.34 to 7.76% and 1.16 to 8.13%, respectively, as compared with precision of the calcaneal DEXA (%CV, 1.23%; %SCV, 1.07%). The OsteoAnalyzer had the lowest %CV, which was statistically similar to calcaneal DEXA.

The positioning precision (i.e., repeat measurements on same day with repositioning; %CV and %SCV) of the five calcaneal instruments ranged from 1.63 to 7.00% and 1.68 to 6.84%, respectively, as compared with precision of the calcaneal DEXA (1.63 and 1.68%). The Achilles and OsteoAnalyzer had the lowest %CVs, which were statistically similar to calcaneal DEXA.

The %CV and %SCV for interobserver precision (i.e., repeat measurements on same day by different technicians) ranged from 1.84 to 9.44% and 1.72 to 9.16%, respectively, as compared with precision of the calcaneal DEXA (1.91 and 1.96%). The Achilles and OsteoAnalyzer had the lowest %CVs, which were statistically similar to calcaneal DEXA.

The short-term precision (i.e., repeat measurements within 2 weeks) ranged from 1.99 to 7.04% and 1.76 to 6.81% (%CV and %SCV, respectively) as compared with precision of the calcaneal DEXA (2.02 and 2.35%). The Achilles and OsteoAnalyzer had the lowest %CVs, which were statistically similar to calcaneal DEXA.

When the analyses were performed including all scans (i.e., adding the 72 measurements that had been excluded during the quality assurance audit), the %CV and %SCV

were similar to the per protocol results for all instruments except the OsteoAnalyzer. For the OsteoAnalyzer, the instrument precision (%CV and %SCV) increased from 1.34 to 6.71% and 1.16 to 5.75%, respectively, positioning precision from 1.63 to 8.02% and 1.68 to 7.54%, interobserver precision from 1.84 to 5.79% and 1.72 to 5.38%, and short-term precision from 1.99 to 5.86% and 1.76 to 5.16%.

#### Measurements, T scores and throughput time

Higher mean measurements ( $p < 0.001$ ) were obtained in the young normal group than in either of the osteoporotic groups for all instruments and body sites (Table 4). Similarly, higher mean measurements were obtained in the osteo non-Fx group than in the osteo Fx group for all technologies ( $p \leq 0.024$ ) except DEXA of the spine ( $p = 0.450$ ).

Mean T scores for the osteoporotic subjects ranged from -1.79 to -2.43 (Table 5). T scores from DEXA femoral neck (-2.43), OsteoAnalyzer (-2.36), QUS-1X (-2.31), and Achilles (-2.30) were statistically similar.

Throughput time (mean minutes  $\pm$  SD) was shortest for the CUBA (4.09  $\pm$  1.38) followed by QUS-1X (5.69  $\pm$  1.28), OsteoAnalyzer (6.04  $\pm$  1.47), Achilles (8.55  $\pm$  3.14), UBA575+ (10.89  $\pm$  1.48), and calcaneal DEXA (11.40  $\pm$  1.81). Throughput time was assessed using the all scans analysis.

TABLE 4. MEASUREMENTS BY INSTRUMENT

Instrument (parameter) Cohort	n	Mean $\pm$ SD*
Achilles (stiffness)		
young normal	50	106.8 $\pm$ 15.39
osteo non-Fx	49	75.6 $\pm$ 11.85
osteo Fx	54	67.7 $\pm$ 13.04
CUBA (BUA)		
young normal	52	86.3 $\pm$ 16.58
osteo non-Fx	49	56.8 $\pm$ 13.89
osteo Fx	55	49.1 $\pm$ 14.18
QUS-1X (UBI-4N)		
young normal	46	21.68 $\pm$ 2.802
osteo non-Fx	47	16.10 $\pm$ 3.516
osteo Fx	50	14.40 $\pm$ 3.838
UBA575+ (BUA)		
young normal	52	93.5 $\pm$ 18.01
osteo non-Fx	49	65.8 $\pm$ 19.28
osteo Fx	55	57.1 $\pm$ 16.84
OsteoAnalyzer (BMC)		
young normal	52	428.6 $\pm$ 64.04
osteo non-Fx	47	294.0 $\pm$ 59.07
osteo Fx	53	263.1 $\pm$ 76.33
DEXA calcaneus (BMD)		
young normal	52	0.570 $\pm$ 0.0799
osteo non-Fx	49	0.430 $\pm$ 0.0911
osteo Fx	54	0.387 $\pm$ 0.0979
DEXA trochanter (BMD)		
young normal	52	0.699 $\pm$ 0.0987
osteo non-Fx	49	0.529 $\pm$ 0.0686
osteo Fx	55	0.489 $\pm$ 0.0799
DEXA femoral neck (BMD)		
young normal	52	0.854 $\pm$ 0.1223
osteo non-Fx	49	0.573 $\pm$ 0.0510
osteo Fx	55	0.543 $\pm$ 0.0727
DEXA spine (BMD) <sup>†</sup>		
young normal	52	1.044 $\pm$ 0.1240
osteo non-Fx	49	0.783 $\pm$ 0.1228
osteo Fx	54	0.762 $\pm$ 0.1407

BUA, broadband ultrasound attenuation; BMC, bone mineral content; BMD, bone mineral density.

\* Statistical significance demonstrated between young normal and osteo non-Fx groups and between young normal and osteo Fx groups within each instrument ( $p < 0.001$ ). Statistical significance demonstrated between osteo non-Fx and osteo Fx groups within each instrument ( $p \leq 0.024$ ) except DEXA of the spine ( $p = 0.450$ ).

<sup>†</sup> Subjects with at least two evaluable vertebrae were included.

#### Discriminatory ability

The data from all five calcaneal instruments and DEXA calcaneus, femoral neck, and trochanter were examined with ROC curves to assess the discriminatory ability between the osteoporotic and young normal subjects. The eight curves were similar and exhibited discriminatory ability (AUC ranged from 0.88 to 0.93) (Fig. 1). ROC curves for all techniques could not discriminate between osteo Fx subjects and osteo non-Fx subjects (Fig. 2).

TABLE 5. T SCORES\* FROM ALL OSTEOPOROTIC SUBJECTS

Instrument (parameter)	n	Mean	SD	Duncan's multiple range test <sup>†</sup>	
Achilles (Stiffness)	103	-2.30	0.848	A	B
CUBA (BUA)	104	-2.02	0.877		C
QUS-1X (UBI-4N)	97	-2.31	1.351	A	B
UBA575+ (BUA)	104	-1.79	1.026		C
OsteoAnalyzer (BMC)	100	-2.36	1.097	A	
DEXA calcaneus (BMD)	103	-2.04	1.212		B C
DEXA trochanter (BMD)	104	-1.93	0.783		C
DEXA femoral neck (BMD)	104	-2.43	0.527	A	

BUA, broadband ultrasound attenuation; BMC, bone mineral content; BMD, bone mineral density.

\* T scores calculated with the reference range established by the young normal group for the study.

<sup>†</sup> Duncan's multiple range test; identical letters denote instruments exhibiting T scores that are not statistically different.

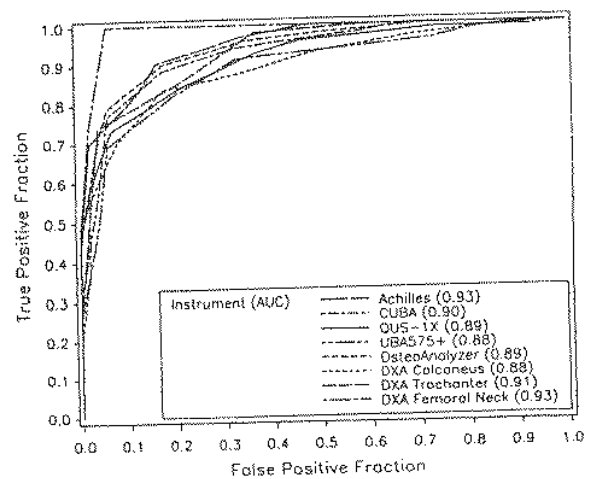


FIG. 1. Receiver operating characteristic curves: discrimination between osteoporotic and young normal. True positive fraction, proportion of osteoporotic subjects correctly classified as osteoporotic; false positive fraction, proportion of young normal subjects incorrectly classified as osteoporotic; AUC, area under the curve.

#### Correlations

The correlations between the five calcaneal measurements (QUS and SEXA) were strong ( $0.80 \leq r \leq 0.91$ ,  $p < 0.001$ ) and correlations between the six calcaneal measurements (including calcaneal DEXA) and DEXA of the femoral neck, trochanter, and spine were generally weaker ( $0.58 \leq r \leq 0.82$ ,  $p < 0.001$ ) (Table 6).

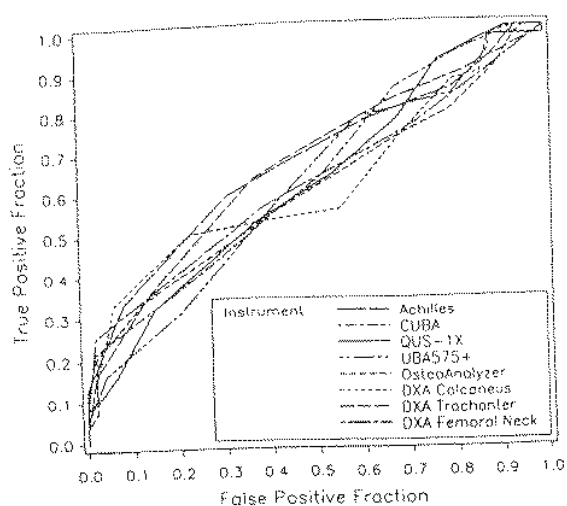


FIG. 2. Receiver operating characteristic curves: discrimination between osteo Fx and osteo non-Fx. True positive fraction, proportion of osteo Fx subjects correctly classified as osteoporotic with a fracture history; false positive fraction, proportion of osteo non-Fx subjects incorrectly classified as osteoporotic with a fracture history.

#### Cross calibration

Review of the initial and final cross calibration results from the QUS instruments did not indicate any clinically significant differences between the instruments (data not shown).

## DISCUSSION

In this study, designed to examine the precision and discriminatory ability of five calcaneal bone assessment technologies, we found that all calcaneal instruments were able to discriminate osteoporotic patients from young normal controls in a manner similar to DEXA of the hip. In general, all of the calcaneal instruments appeared to have adequate precision for identifying patients with osteoporosis, since the precision errors were much smaller than the difference between young normals and osteoporotics (approximately 24–34%) (Table 4).

For all types of precision studied, SEXA performed in a similar manner to calcaneal DEXA with %CV approximately 2% or less. The most clinically relevant precision categories are likely the positioning and short-term precision. For these categories, the %CV for Achilles was statistically similar to that of calcaneal DEXA and SEXA. For positioning precision, CUBA and UBA575+ were intermediate and QUS-1X was the highest (Table 3). Although we were unable to perform hypothesis testing for %SCV, the performance trends were similar to the %CV. Comparisons with the literature are problematic because many studies have not clearly defined the type of precision assessed. Nonetheless, the %CVs that we obtained were similar to those previously reported.<sup>(14–24)</sup> The precision results re-

ported here are specific for the instruments studied. These results cannot be applied to other peripheral instruments measuring other sites (e.g., patella, tibia, or phalanx).

For instrument and interobserver precision, the best precision among the ultrasound machines was seen for those using a water bath. This was not true for positioning precision, where the CUBA and UBA575+ were similar, and short-term precision, where the CUBA performed better than the UBA575+. In choosing an appropriate ultrasound device for a particular application, many factors need to be taken into consideration, such as precision, size, cost, ease of use, patient comfort, and portability. For diagnostic use, all instruments demonstrate adequate precision. However, if the goal is to monitor changes over time, then an instrument with good short-term precision would be preferable.

It should be noted that the precision error calculated as %CV is influenced by the QUS variable used. Techniques that report SOS or stiffness often have lower %CVs than those that report BUA. Therefore, using stiffness (defined as a combination of BUA and SOS) as the most clinically relevant measurement for the Achilles instrument may explain why its %CV was lower than that of the other QUS techniques which used BUA as the most clinically relevant measurement (Table 3 and Appendix 1). Because of difficulty in comparing %CVs, some investigators<sup>(23,25)</sup> have suggested using a %CV adjusted for the range of clinical values known as the standardized %CV (%SCV). This eliminates the favorable bias on instruments that offer a small range of clinical values when compared with instruments that offer a large range of clinical values. We found that, whereas the %SCV for calcaneal SEXA and calcaneal DEXA were nearly equivalent to %CV for each precision studied, the %SCV for QUS measurements ranged from nearly equivalent to the calcaneal DEXA precision values to nearly eight times greater than the calcaneal DEXA precision values (Table 3).

There were no significant differences for the precision determinations using an "all scans" analysis versus a "per protocol" analysis, with the exception of the calcaneal SEXA, which improved significantly when only properly performed scans were analyzed. This was primarily due to misplacement of the heel and heel movement during the scan. Per protocol analyses provide reliable information since this reflects the performance quality when operated as specified by the manufacturer. In contrast, the all scans analysis provides "real world" use of the machines when quality assurance of each scan cannot be performed routinely. Since both analyses provide information that could be pertinent to several different medical settings and the results were relatively similar, with the exception of the SEXA instrument (OsteoAnalyzer), the per protocol analysis is provided in the body of this paper.

For all calcaneal measurements, the ability to discriminate patients with osteoporosis from young normal controls was similar to that of DEXA of the calcaneus and hip. As expected, the discriminatory ability at the femoral neck was good since DEXA of the femoral neck was a criteria used to classify subjects as osteoporotic. Other studies have demonstrated that QUS techniques can discriminate between controls and patients with hip fracture,<sup>(14,16,26–28)</sup> vertebral

TABLE 6. CORRELATION OF MEASUREMENTS BETWEEN INSTRUMENTS FOR ALL SUBJECTS\*

Instrument <sup>†</sup>	Achilles	CUBA	QUS-1X	UBA575+
Achilles	1.000			
CUBA	0.906 (0.873, 0.931)	1.000		
QUS-1X	0.861 (0.812, 0.898)	0.815 (0.752, 0.863)	1.000	
UBA575+	0.882 (0.841, 0.913)	0.840 (0.787, 0.881)	0.862 (0.813, 0.899)	1.000
OsteoAnalyzer	0.906 (0.873, 0.931)	0.832 (0.776, 0.875)	0.834 (0.776, 0.878)	0.798 (0.733, 0.849)
DEXA calcaneus	0.859 (0.811, 0.895)	0.822 (0.764, 0.867)	0.804 (0.738, 0.855)	0.787 (0.719, 0.840)
DEXA trochanter	0.770 (0.697, 0.827)	0.696 (0.606, 0.768)	0.697 (0.603, 0.772)	0.659 (0.561, 0.739)
DEXA femoral neck	0.800 (0.735, 0.850)	0.715 (0.629, 0.783)	0.711 (0.620, 0.783)	0.697 (0.607, 0.769)
DEXA spine <sup>‡</sup>	0.678 (0.583, 0.755)	0.640 (0.538, 0.724)	0.655 (0.551, 0.739)	0.580 (0.466, 0.675)

\* All correlations  $p < 0.001$  testing whether correlation equals 0.

<sup>†</sup> Quantitative ultrasound parameters assessed: UBA575+ and CUBA, broadband ultrasound attenuation; Achilles, stiffness; QUS-1X, UBI-4N.

<sup>‡</sup> Subjects with at least two evaluable vertebrae were included.

fracture,<sup>(26,28-30)</sup> wrist fracture,<sup>(31)</sup> and osteopenia.<sup>(32,33)</sup> In practice, this would be a clinically relevant use for the calcaneal instruments, in that a diagnosis of osteoporosis can be established and appropriate therapy initiated. Although there were significant differences in mean values between osteoporotic subjects with and without fractures by ANOVA, using ROC analysis we could not distinguish between the two groups. This may be because all of our patients were enrolled with the primary inclusion criterion of a low hip BMD. In addition, it is well established that factors independent of BMD (e.g., fall type, body mass index, and use of medications associated with falls, etc.) are associated with hip fractures<sup>(8,34-36)</sup>; these factors were not assessed.

We found a strong correlation between all calcaneal measures ( $0.79 \leq r \leq 0.93$ ). We observed a moderately strong correlation between calcaneal measurements and DEXA measurements of the hip or spine. These correlations are similar to what has been reported when comparing measurements at peripheral sites to central site measurements.<sup>(37)</sup>

There were several strengths of this study. It is one of the few to compare the precision and discriminatory abilities of five different calcaneal instruments. In addition, we compare their performance to DEXA of the calcaneus, hip, and spine. Furthermore, we used our own reference group of young normal controls, basing our T score and ROC curves on this group rather than the manufacturer's normative database.

However, our study also has several limitations. Because we defined our study population (osteoporotic versus young normal) based on hip BMD measurements provided by the manufacturer's database, it would be anticipated that we

would be able to discriminate between osteoporotic and nonosteoporotic patients based on hip BMD measurements. Although we used the manufacturer's database for classification of osteoporosis based on WHO criteria, use of other young normal databases such as NHANES III<sup>(38,39)</sup> may have resulted in a slightly different study population. In fact, the mean and standard deviation values for the hip femoral neck and trochanter derived from the young normal population used in this study are different from the values in use by the manufacturer. This may be explained by differences in the selection of the normative population, geographic differences, and the age ranges used. We only included non-Hispanic, Caucasian, osteoporotic women in the postmenopausal group. Precision values among the technologies may be different for men, patients with metabolic bone disease, and non-Caucasians. In addition, we only examined precision over a short period of time. We did not examine whether these techniques would be appropriate for long-term follow-up with or without therapy. Because therapy today is primarily directed at maintaining or improving BMD at the hip and spine, follow-up of calcaneal measurements could potentially be misleading if instruments have differential effects at various skeletal sites.<sup>(40,41)</sup>

In summary, we found that the five calcaneal instruments can discriminate osteoporotic subjects from young normal controls. The techniques have variable precision but appear to be adequate for detection of osteoporosis. Because these calcaneal instruments provide rapid measurements, do not require dedicated office space, and are more affordable than DEXA of the hip and spine, they offer a reasonable alternative for the initial evaluation of patients with osteoporosis. However, future studies are needed to determine



TABLE 6. CONTINUED

OsteoAnalyzer	DXA calcaneus	DXA trochanter	DXA femoral neck	DXA spine <sup>†</sup>
1.000				
0.930 (0.905, 0.949)	1.000			
0.816 (0.756, 0.863)	0.749 (0.671, 0.810)	1.000		
0.802 (0.738, 0.852)	0.729 (0.646, 0.795)	0.878 (0.837, 0.909)	1.000	
0.740 (0.659, 0.804)	0.685 (0.592, 0.760)	0.796 (0.731, 0.847)	0.788 (0.721, 0.841)	1.000

whether these techniques can be used to monitor disease progression or response to therapeutic intervention.

#### ACKNOWLEDGMENTS

We thank Lauren Ferguson and Michelle MacCallum (Osteoporosis Prevention and Treatment Center, Beth Israel Deaconess Medical Center, Boston, MA, U.S.A.), and Elizabeth Allen and Sandra Veith (Bone and Mineral Research Unit, Oregon Health Sciences University, Portland, OR, U.S.A.) for performing and analyzing the subject measurements. This study was funded and conducted by Merck & Co., Inc.

#### REFERENCES

- Melton III LJ, Chrischilles EA, Cooper C, Lane AW, Riggs BL 1992 How many women have osteoporosis? *J Bone Miner Res* 7:1005-1010.
- Praemer A, Furner S, Rice DP 1992 Costs of musculoskeletal conditions. In: *Musculoskeletal Conditions in the United States*. American Academy of Orthopaedic Surgeons, Park Ridge, IL, U.S.A., pp. 142-170.
- Dans PE, Kerr MR 1979 Gerontology and geriatrics in medical education. *N Engl J Med* 300:228-232.
- Kelsey JL, Hoffman S 1987 Risk factors for hip fracture. *N Engl J Med* 316:404-406.
- Lieberman UA, Weiss SR, Broll J, Minne HW, Quan H, Bell NH, Rodriguez-Portales J, Downs Jr RW, Dequeker J, Favus M, Seeman E, Recker RR, Capizzi T, Santora II AC, Lombardi A, Shah R, Hirsch LJ, Karpf DB 1995 Effect of treatment with oral alendronate on bone mineral density and fracture incidence in postmenopausal osteoporosis. *N Engl J Med* 333:1437-1443.
- Kellie SE 1992 Diagnostic and therapeutic technology assessment (DATTA). *JAMA* 267:286-288, 290-294.
- Langton CM, Porter SB, Porter RW 1984 The measurement of broadband ultrasound attenuation in cancellous bone. *Eng Med* 13:89-91.
- Cummings SR, Nevitt MC, Browner WS, Stone K, Fox KM, Ensrud KE, Cauley J, Black D, Vogt TM 1995 Risk factors for hip fracture in white women. *N Engl J Med* 332:767-773.
- Hans D, Dargent-Molina P, Schott AM, Sebert JL, Cormier C, Kotzki PO, Delmas PD, Pouilles JM, Breart G, Meunier PJ 1996 Ultrasonographic heel measurements to predict hip fracture in elderly women: The EPIDOS prospective study. *Lancet* 348:511-514.
- Ross P, Huang C, Davis J, Imose K, Yates J, Vogel J, Wasnich R 1995 Predicting vertebral deformity using bone densitometry at various skeletal sites and calcaneus ultrasound. *Bone* 16:325-332.
- Porter RW, Miller CG, Grainger D, Palmer SB 1990 Prediction of hip fracture in elderly women: A prospective study. *Br Med J* 301:638-641.
- Black DM, Cummings SR, Genant HK, Nevitt MC, Palermo L, Browner W 1992 Axial and appendicular bone density predict fractures in older women. *J Bone Miner Res* 7:633-638.
- Secley DG, Browner WS, Nevitt MC, Genant HK, Scott JC, Cummings SR 1991 Which fractures are associated with low appendicular bone mass in elderly women? *Ann Intern Med* 115:837-842.
- Schott AM, Weill-Engerer S, Hans D, Duboeuf F, Delmas PD, Meunier PJ 1995 Ultrasound discriminates patients with hip fracture equally well as dual-energy X-ray absorptiometry and independently of bone mineral density. *J Bone Miner Res* 10:243-249.
- Faulkner KG, McClung MR, Coleman LJ, Kingston-Sandahl E 1994 Quantitative ultrasound of the heel: Correlation with densitometric measurements at different skeletal sites. *Osteoporos Int* 4:42-47.
- Mautalen C, Vega E, Gonzalez D, Carrilero P, Otano A, Silberman F 1995 Ultrasound and dual x-ray absorptiometry densitometry in women with hip fracture. *Calcif Tissue Int* 57:165-168.
- Massic A, Reid DM, Porter RW 1993 Screening for osteopo-

- rosis: Comparison between dual energy x-ray absorptiometry and broadband ultrasound attenuation in 1000 perimenopausal women. *Osteoporos Int* 3:107-110.
18. Salamone LM, Krall EA, Harris S, Dawson-Hughes B 1994 Comparison of broadband ultrasound attenuation to single x-ray absorptiometry measurements at the calcaneus in postmenopausal women. *Calcif Tissue Int* 54:87-90.
  19. Baran DT, Kelly AM, Karellas A, Gionet M, Price M, Leahey D, Steuterman S, McSherry B, Roche J 1988 Ultrasound attenuation of the os calcis in women with osteoporosis and hip fractures. *Calcif Tissue Int* 43:138-142.
  20. Bauer DC, Gluer CC, Genant HK, Stone K 1995 Quantitative ultrasound and vertebral fracture in postmenopausal women. *J Bone Miner Res* 10:353-358.
  21. Herd RJM, Blake GM, Ramalingam T, Miller CG, Ryan PJ, Fogelman I 1993 Measurements of postmenopausal bone loss with a new contact ultrasound system. *Calcif Tissue Int* 53:153-157.
  22. Brooke-Wavell K, Jones PRM, Pye DW 1995 Ultrasound and dual x-ray absorptiometry measurement of the calcaneus: Influence of region of interest location. *Calcif Tissue Int* 57:20-24.
  23. Miller CG, Herd RJM, Ramalingam T, Fogelman I, Blake GM 1993 Ultrasonic velocity measurements through the calcaneus: Which velocity should be measured? *Osteoporos Int* 3:31-35.
  24. Gluer CC, Vahlensieck M, Faulkner KG, Engelke K, Black D, Genant HK 1992 Site-matched calcaneal measurements of broadband ultrasound attenuation and single X-ray absorptiometry: Do they measure different skeletal properties? *J Bone Miner Res* 7:1071-1079.
  25. Moris M, Peretz A, Tjeka R, Negaban N, Wouters M, Bergmann P 1995 Quantitative ultrasound bone measurements: Normal values and comparison with bone mineral density by dual x-ray absorptiometry. *Calcif Tissue Int* 57:6-10.
  26. Turner CH, Peacock M, Timmerman L, Neal JM, Johnston Jr CC 1995 Calcaneal ultrasonic measurements discriminate hip fracture independently of bone mass. *Osteoporos Int* 5:130-135.
  27. Stewart A, Reid DM, Porter RW 1994 Broadband ultrasound attenuation and dual energy X-ray absorptiometry in patients with hip fractures: Which technique discriminates fracture risk. *Calcif Tissue Int* 54:466-469.
  28. Gluer CC, Cummings SR, Bauer DC, Stone K, Pressman A, Mathur A, Genant HK 1996 Osteoporosis: Association of recent fractures with quantitative U.S. findings. *Radiology* 199:725-732.
  29. Heaney RP, Avioli LV, Chesnut III CH, Lappe J, Recker RR, Brandenburger GH 1989 Osteoporotic bone fragility detection by ultrasound transmission velocity. *JAMA* 261:2986-2990.
  30. Ott SM, Kilcoyne RF, Chesnut III CH 1987 Ability of four different techniques of measuring bone mass to diagnose vertebral fractures in postmenopausal women. *J Bone Miner Res* 2:201-210.
  31. Kroger H, Jurvelin J, Arnala I, Penttila K, Rask A, Vainio P, Alhava E 1995 Ultrasound attenuation of the calcaneus in normal subjects and in patients with wrist fracture. *Acta Orthop Scand* 66:47-52.
  32. Agren M, Karellas A, Leahey D, Marks S, Baran D 1991 Ultrasound attenuation of the calcaneus: A sensitive and specific discriminator of osteopenia in postmenopausal women. *Calcif Tissue Int* 48:240-244.
  33. Rossman P, Zagzebski J, Mesina C, Sorenson J, Mazess R 1989 Comparison of speed of sound and ultrasound attenuation in the os calcis to bone density of the radius, femur and lumbar spine. *Clin Phys Physiol Meas* 10:353-360.
  34. Grisso JA, Kolsoy JL, Strom BL, Chiu GY, Maislin G, O'Brien LA, Hoffman S, Kaplan F 1991 Risk factors for falls as a cause of hip fracture in women. *N Engl J Med* 324:1326-1331.
  35. Greenspan SL, Myers ER, Maitland LA, Resnick NM, Hayes WC 1994 Fall severity and bone mineral density as risk factors for hip fracture in ambulatory elderly. *JAMA* 271:128-133.
  36. Nevitt MC, Cummings SR 1993 Study of Osteoporotic Fractures Research Group—Type of fall and risk of hip and wrist fractures: The study of osteoporotic fractures. *J Am Geriatr Soc* 41:1226-1234.
  37. Genant HK, Grampp S, Mathur A, Gluer C, Jergas M, Takada M, Lang P, Lu Y, Chavez M 1995 Comparisons of non-invasive bone mineral measurements in assessing age-related loss, fracture discrimination, and diagnostic classification. Proceedings of the 11th International Bone Densitometry Workshop, Glenden Beach, OR, U.S.A.
  38. Looker AC, Johnston Jr CC, Wahner HW, Dunn WL, Calvo MS, Harris TB, Heyse SP, Lindsay RL 1995 Prevalence of low femoral bone density in older U.S. women from NHANES III. *J Bone Miner Res* 10:796-802.
  39. Looker AC, Wahner HW, Dunn WL, Calvo MS, Harris TB, Heyse SP, Johnston Jr CC, Lindsay RL 1995 Proximal femur bone mineral levels of US adults. *Osteoporos Int* 5:389-409.
  40. Naessen T, Mallmin H, Ljunghall S 1995 Heel ultrasound in women after long-term ERT compared with bone densities in the forearm, spine and hip. *Osteoporos Int* 5:205-210.
  41. Gonnelli S, Cepollaro C, Pondrelli C, Martini S, Rossi S, Gennari C 1996 Ultrasound parameters in osteoporotic patients treated with salmon calcitonin: A longitudinal study. *Osteoporos Int* 6:303-307.

Address reprint requests to:

Mary E. Melton, M.D.

Merck & Co., Inc.

Sumneytown Pike and Broad Street

P.O. Box 4, HM-216

West Point, PA 19486 U.S.A.

Received in original form November 27, 1996; in revised form March 21, 1997; accepted March 31, 1997.