Errors in Longitudinal Measurements of Bone Mineral: Effect of Source Strength in Single and Dual Photon Absorptiometry

William L. Dunn, Stephen H. Kan, and Heinz W. Wahner

Mayo Clinic and Mayo Foundation, Rochester, Minnesota

The effect of changing strength during the useful life of a radiation source was evaluated in studies performed on four dual photon (DPA) and two single photon (SPA) bone absorptiometry instruments. Two DPA units and one SPA unit did not show any systematic dependence of measured bone mineral content or bone mineral areal density (BMD) on source activity when evaluated over an entire source life. One DPA and one SPA instrument, however, showed significant time trends associated with source activity. The fourth DPA instrument had a significant linear decrease in BMD over a source life in the automatic mode but performed better in the manual mode.

J Nucl Med 28:1751–1757, 1987

Dual photon absorptiometry (DPA) of the lumbar spine has evolved as a major technique for quantitation of mineral in the axial skeleton (*1,2*). When properly performed, this technique has acceptable precision, provides good discrimination between normal and osteoporotic individuals, and is noninvasive with a low radiation dose (<200 μGy) (*3–5*). There is great interest in using the technique to measure the rate of bone loss in disease or during treatment in longitudinal studies performed over years. Physiologic bone loss generally is ~1% to 3% per year, depending on sex and age, and very stable instruments are needed. Quality control data given by manufacturers generally demonstrate accuracy and short-term precision. The latter usually is based on repeated measurements on phantoms or on subjects within a period of days, weeks, or a month. This paper reports a study of measurement precision over the lifetime of a source and beyond. Four instruments for DPA and two instruments for single-photon absorptiometry (SPA) are studied.

Repeated measurements over more than 1 yr were made on ashed bones with all instruments, and measurements were made of extrapolated air counts, an aluminum tube, a cadaver spine, and a volunteer with some. The results stress the need for an organized quality control program to determine long-term stability. The actual precision achieved in a given longitudinal study can only be determined from a retrospective analysis of quality control data obtained during the study period.

Received Aug. 25, 1986; revision accepted May 28, 1987.

For reprints contact: Heinz W. Wahner, MD, Mayo Clinic, 200 First St. SW, Rochester, MN 55905.

This study was presented in part at the 33rd Annual Meeting of the Society of Nuclear Medicine, Washington, DC, June 22 to 25, 1986.

METHODS

DPA Instruments

Four instruments, D1, D2, D3, and D4,* were evaluated during the useful life of one source of gadolinium-153 (^{153}Gd).† The source strength was 1.0 and 1.5 Ci (37 and 55 GBq) for the commercial and locally built instruments, respectively, at the time of installation.

On all instruments periodic measurements of the same ashed bone under 20 cm of water were made to assess the effectiveness of the manufacturer's recommended quality control program for the detection of lack of stability of the instruments. The calibration and standardization procedures used in the different instruments are described in Table 1.

The locally built instruments (D1 and D2) for DPA have been described (*1,3,4*). These instruments operate with a speed of 1.2 mm/sec. Counts are acquired and dumped into memory once per second. The scan increment is 0.5 cm. The detector collimator opening is 6.0 mm. The extrapolated in-air counts for the 44-keV channel were 139 kcps with the new source and 45 kcps at the time of source change. Calibration of these instruments to read bone mineral mass (g) or areal density (g/cm^2) was achieved by using a standard curve obtained by scanning five ashed bones of different mass in 20 cm of water

TABLE 1
Calibration and Quality Control Procedures

	Dual photon absorptiometry instruments			Single photon absorptiometry instruments	
	D1, D2	D3	D4	S1	S2
Calibration:					
Time	Every 3 wk	At installation	At installation	At source change	At installation
Phantom	5 ashed bones under 20 cm H_2O	Company reference standard (ashed bone)	5-step liquid bone-simulating phantom (K_2HPO_4)	4 Al tubes of known bone equivalent	Metal bar
Periodic quality control:					
Time	Daily	Daily	Weekly	Daily	Weekly
Phantom	Ashed bone under 17 cm of Lucite (block)	Lunar block phantom with 3 bone-equivalent regions Height 8.3 cm	Al tube under 16 cm H_2O	Al tube in 7 cm H_2O	Metal bar

every 3 wk. A block phantom consisting of an ashed bone of 14 g mass embedded in a 17-cm thick Lucite block was scanned daily to ensure instrument stability. Details have been described (*3,4*).

With one commercial increment (D3) the 8-mm detector collimator was used exclusively; scan speed was 2.5 mm/sec, and scan increment was 4.5 mm; counts were dumped into memory every 0.5 second. For calibration, the scan results of a three-bone-equivalent phantom were entered into the computer daily and added to the data obtained on the previous 4 days for averaging. The extrapolated in-air counts for the 44-keV channel ($^{44}I_0$) were 220 kcps initially and 62 kcps at the time of source change.

Data from another commercial instrument (D4) obtained in an observation period of 43 wk during which the ^{153}Gd source decayed from 18.5 to 7.4 GBq also were evaluated. Counts in the 44-keV channel were always greater than 2,500 cps in 16 cm of water, which is given as the lower limit of clinical operation by the manufacturer. This instrument had a speed of 4 mm/sec, a scan increment of 4.0 mm, and a 13-mm detector collimator. Acquisition interval was 1 sec.

SPA Instruments Using Iodine-125 Sources

A locally built instrument (S1)[‡] capable of measuring five paths across the radius 1 mm apart was tested. This instrument was calibrated with a set of aluminum tubes, of different wall thicknesses, in 7 cm of water; the aluminum tubes had previously been calibrated against ashed bone samples from a human radius (*3*). Scan speed was variable with count rate, with a low of 1.0 cm/min and a maximum of 2.0 cm/min. The detector collimator opening was 3 mm, the source was iodine-125 (^{125}I).[§] Source activity was 7.4 GBq initially and about 3.7 GBq at the time of removal.

A commercial instrument (S2)[‡] also was evaluated. This instrument was capable of scanning ultradistal and more proximal sites of both forearm bones. A water bath was used with this instrument to surround the arm with a tissue-equivalent medium of uniform thickness. The manufacturer's recommended quality control procedures consisted of scanning a metal bar phantom repeatedly in 8 cm of water over the life of a source. The instrument was calibrated at the factory. Scanning speed was 2 mm/sec; the detector collimator opening was 2 × 6 mm.

Data Analysis

Least-square regression analysis over the entire period of a source life was performed on data from the ashed bone on all instruments, on data from the extrapolated counts in air from D1 and D2, and on data from one commercial calibration phantom from D3. Polynomial models of various degrees (i.e., linear, quadratic, or cubic) were examined and the ones with the best fit to the data were chosen. In addition, data obtained prior to replacement of an old source and after installation of a new source (high count difference) from these instruments were compared. Lumbar spine measurements of a subject obtained with one commercial instrument (D3) before and after source change were compared. For these comparisons, two-sample t-tests for significance of differences were performed.

RESULTS

DPA Instruments

With the two locally built instruments, results from measurements (57 for D1; 51 for D2) of the ashed bone—expressed in arbitrary units (equivalent to total mass of the bone) and the ratio of arbitrary units to projected bone area (equivalent to areal bone density)—prior to calibration showed systematic dependence on time in the form of cubic and quadratic relationships (Fig. 1). This reflects drift of the instruments which is corrected by repeated calibration. It shows the magnitude of changes which are corrected by calibration. The bone mass and areal density measurements on a repeatedly scanned ashed bone over one source life (420 days) after calibration are given in Figure 2. There is no dependence of measured bone mineral ($r = 0.07$; $P = 0.58$) or areal density ($r = 0.09$; $P = 0.52$) on source

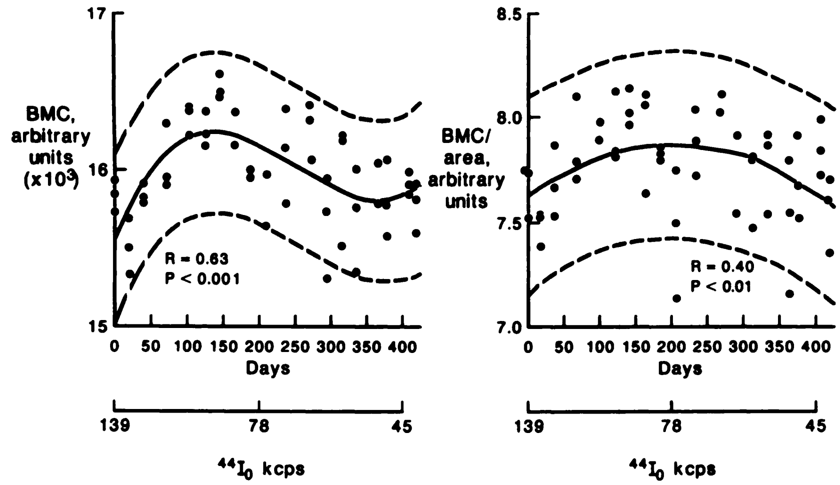

FIGURE 1
Left: Uncalibrated bone mineral content equivalent (BMC) (in computer units) of the standard bone observed over a period of 420 days with instrument D1. These data were taken prior to a new calibration and reflect the instrument drift in the 3 wk between calibration periods. The source strength declined from 139 to 45 kcps during this time. Regression equation (d, days): $y = 1.561 \cdot 10^4 + 10.585\ d - 5.305 \cdot 10^{-2}\ d^2 + 6.94 \cdot 10^{-5}\ d^3$ (n = 57; R^2 = 0.39, P = 0.0001; $S_{y \cdot x}$ = 246.3, 1.5% of $\bar{y}$). Upper and lower lines mark the 95% prediction interval. Right: Uncalibrated areal bone mineral density (computer units/cm²) of area scanned). Regression equation: $y = 7.636 + 2.30 \cdot 10^{-3}\ d - 5.8 \cdot 10^{-6}\ d^2$ ($n = 57$; $R^2 = 0.16$; $P < 0.01$; $s_{y \cdot x}$ = 0.22, 2.8% of $\bar{y}$).

activity or time. The coefficient of variations (CV) for bone mineral and areal density were 1.2% (mean bone mineral, 39.0 g) and 2.4% (mean density, 2.0 g/cm²), respectively. Consistent results were found with the second instrument, with CV 1.3% (mean, 39.0 g) and 1.8% (mean, 2.0 g/cm²) for bone mineral and areal density, respectively. The half-life of source decay was estimated from weekly counts in air (extrapolated) with both the 44- and 100-keV energies: 242.5 and 240.5 days, respectively. The true half-life of ^{153}Gd is reported to be 242 days (*6*). This confirms the system was functioning properly.

The 247 daily measurements of the three-bone-region phantom with instrument D3 showed strong time dependence ($P = 0.0001$) in linear and quadratic patterns between calibration results and time of use (source strength) for all three bone regions (Fig. 3). Bone mass and areal density from the ashed bone (Fig. 4) also showed a strong linear increase with decreasing source strength ($P = 0.0001$). Detailed analysis of these curves is given in Table 2. When the extrapolated air counts were plotted against time, the half-lives were short, 201.4 and 199 days for the 44- and 100-keV windows.

As shown in Figure 4, in each case the measured BMC and BMD increased with a decrease in source activity. To study the dependence of bone mineral value of D3 on source activity further, BMD measurements were performed immediately before and after replacement of a weak ^{153}Gd source with a new source; the objects measured were an ashed bone, an aluminum tube (both scanned in 20 cm of water), and a volunteer subject. Results showed increases in mean values—of 9.8%, 9.4%, and 3.8%, respectively—for the lumbar spine of the volunteer subject, aluminum tube, and ashed bone (Table 3). Despite the small number of measurements, these differences are statistically significant ($P < 0.05$).

The manufacturer of D3 developed a new software

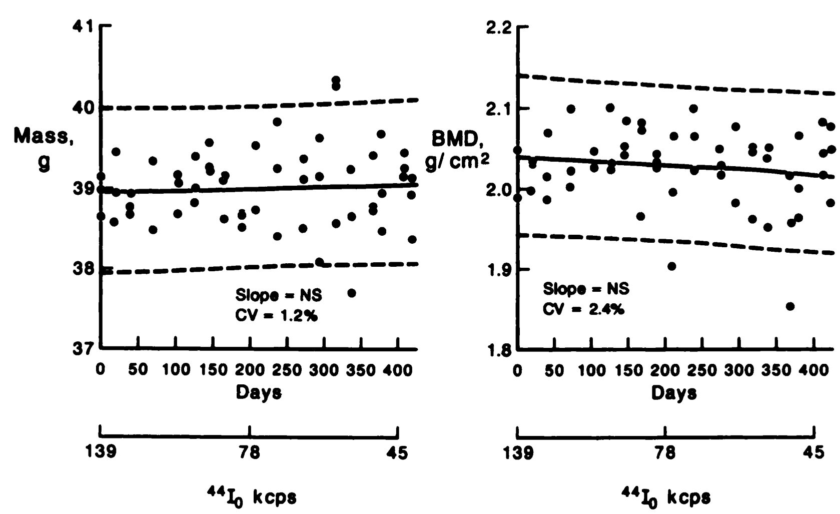

FIGURE 2
Calibrated bone mineral content (left) and areal density (right) obtained with instrument D1 from repeated measurements of an ashed bone sample over the lifetime of a source. The slopes are not significantly different from zero.

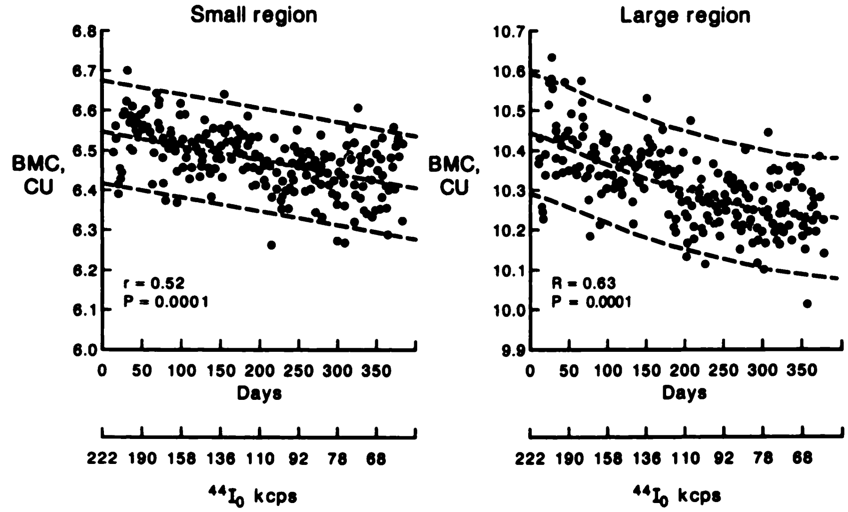

FIGURE 3
Daily measurements of calibration phantom with instrument D3 over the lifetime of one source. The source strength declined from 220 kcps to 62 kcps for the small (left) and large (right) bone equivalent regions in the phantom. Regression equations are shown in Table 2.

version (Version 1.4), which addressed this problem and included a correction designed to eliminate the measurement dependence on source activity. Table 4 presents measurements made with the new version on a cadaver spine (L1–4), an aluminum tube, and ashed bone with both a 7.4-GBq (at 44 keV, 46 kcps) and a 37-GBq (at 44 keV, 210 kcps) ^{153}Gd source. Differences in measured BMD with different source activities were negligible with the newer software.

Evaluation of the second commercial instrument (D4) showed that, when the data (n = 43) were processed in a mode that required operator interaction for setting of baselines and edge delimiters, the slope of the regression line was not significantly different from zero for both mass and BMD (r = 0.007; P = 0.66). In processing with no operator intervention (automatic mode), the BMD slope was significant (r = −0.58; $P <$ 0.0001), showing a substantial decrease (~10%) from the beginning to the end of the 43-week measurement period (Fig. 5).

SPA Instruments

Daily measurements (n = 47), over the life of a source, of an aluminum tube under 7 cm of water showed no dependence on source activity in the locally built instrument (S1). This instrument yielded a CV of 0.4%.

With the commercial SPA instrument (S2), measurements (n = 42) over a 138-day period on the metal bar standard increased quadratically and gradually with time (R = 0.84; $P <$ 0.0001) and leveled off at about day 98 (Fig. 6). The maximum increase in BMC was ~2%. A drastic change in data upon replacement of the source was apparent; the difference between the groups of five data points immediately before and after source change is highly significant ($P <$ 0.0001).

DISCUSSION

Loss of bone mineral from the lumbar spine with age in normal women is ~1% per year; it is higher at the

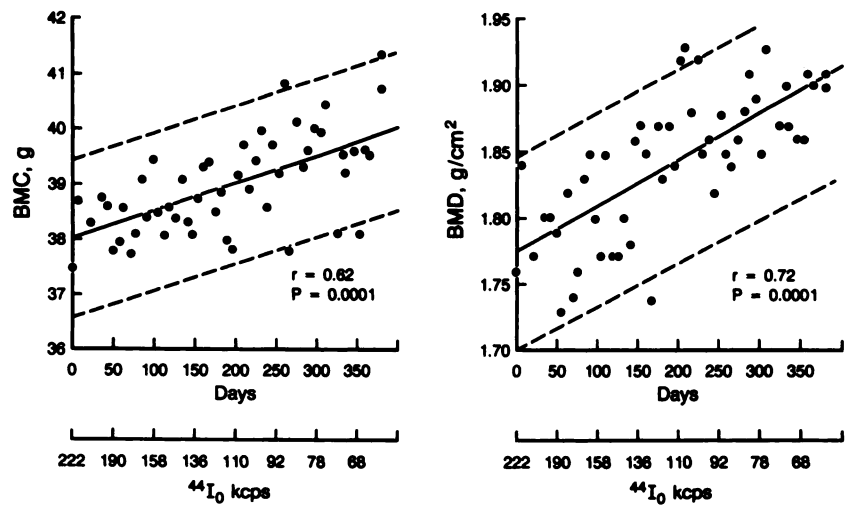

FIGURE 4
Bone mineral content (left) and areal density (right) of ashed bone sample in water measured repeatedly with instrument D3 over the lifetime of a source. Regression equations are shown in Table 2.

TABLE 2
Results of Regression Analysis of Data From Instrument D3 With a Commercial Quality Control Phantom and an Ashed Bone Standard in 20 cm of H_2O

Object*	Regression equation†	n	$s_{y \cdot x}$ (% of $\bar{y}$)	R^2	P
Small bone phantom, CU	$BM = 6.541 - 3.71 \cdot 10^{-4} d$	247	0.064 (1.0)	0.27	0.0001
Medium bone phantom, CU	$BM = 8.575 - 1.62 \cdot 10^{-3} d + 2.80 \cdot 10^{-6} d^2$	247	0.077 (0.9)	0.44	0.0001
Large bone phantom, CU	$BM = 10.438 - 9.26 \cdot 10^{-4} d + 1.00 \cdot 10^{-6} d^2$	247	0.074 (0.7)	0.39	0.0001
Ashed bone mass, g	$BMC = 38.036 + 4.85 \cdot 10^{-3} d$	52	0.687 (1.8)	0.38	0.0001
Ashed bone BMD, g/cm²	$BMD = 1.774 + 3.524 \cdot 10^{-4} d$	52	0.037 (2.0)	0.52	0.0001

* CU, units of measurement are computer units of the software. Conversion from day of use into count rate is possible such as the data for Figure 4.
† d, days.

time of the menopause and is lower for men at all ages (5). Higher-than-normal rates of loss (3% to 5% per year) may be seen in patients with metabolic bone disease including osteoporosis (7). To be considered effective, a treatment regimen should either decrease or abolish the rate of loss or even increase the BMC at the site of fracture risk.

To measure these rather small changes over a period of several years requires a rigorous quality control program to ensure long-term stability of the instruments. Even small systematic errors not recognizable by comparing the day-to-day data or monthly review of quality control results may significantly affect the outcome of longitudinal studies.

When repeated measurements made on the same patients are interpreted, the confidence interval to be applied to the measured difference is improved with the increasing long-term precision of the instrument.

Generally, precision measurements of instruments are based on short-term experiments in which phantoms, cadaver bones, or subjects are measured repeatedly over a period of days to months. These measurements are not relevant for long-term stability assessments when studies are conducted over several years and when sources are changed repeatedly during the period of the study. This is evident from the fact that all the tested instruments had a short-term precision of <3% when tested in our laboratory with phantoms (8). Daily or weekly measurements of a suitable standard

TABLE 4
BMD With Weak and Strong ^{153}Gd Sources Calculated With New Software (Version 1.4) for Instrument D3

^{153}Gd activity, GBq	Cadaver spine (L1–4) BMD, g/cm²	n	Al tube BMD, g/cm²	n	Ashed bone BMD, g/cm²	n
7.40	1.14	2	1.11	2	1.91	2
37.0	1.17	2	1.09	2	1.89	3

TABLE 3
BMD Immediately Prior to and After ^{153}Gd Source Replacement for Instrument D3

	BMD, g/cm²		
^{153}Gd activity, GBq	Volunteer subject L1–4	Al tube	Ashed bone
11.1	1.10	1.16	1.91
	1.13	1.16	1.90
	1.05		1.91
			1.90
Mean	1.093	1.160	1.905
37.0	0.99	1.06	1.82
	1.00	1.06	1.84
			1.84
			1.85
			1.86
			1.80
Mean	0.995	1.060	1.835
P*	<0.05	†	0.0003

* P values are for two-sample t-test between the two sources.
† No variation within each group.

TABLE 5
Quality Control Procedures for Longitudinal Studies

Measurement	Information
Extrapolated air values for 44 keV channel	Tests electronic hardware performance, particularly of counters and stability of energy windows; data obtained over several months should give half-life of 230–250 days
Ashed bone in 20 cm H_2O or equivalent standard	Tests entire system, including scanning mechanism, data acquisition, and processing; no slope should be detectable throughout one source life for total BM and for area
Cadaver spine or subject	Tests specifically the data processing part and edge detection efficiency; this is the definitive test

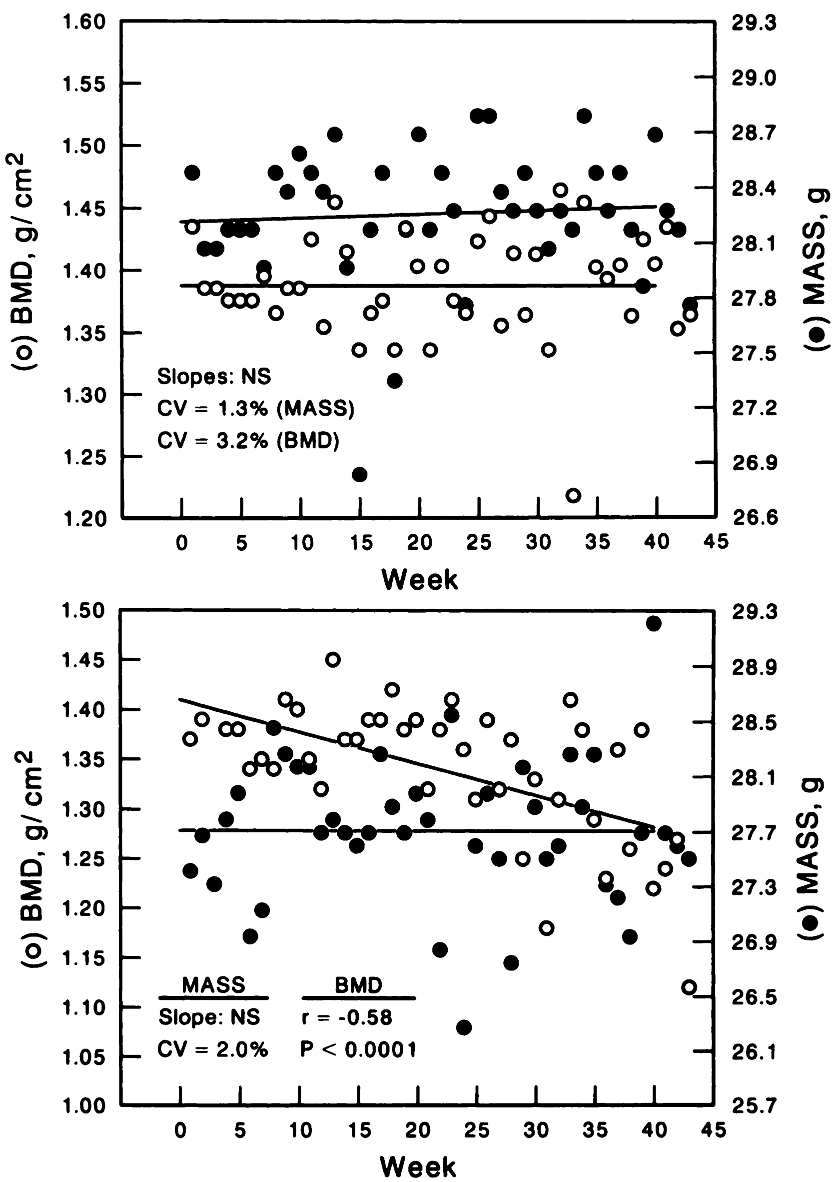

FIGURE 5
Bone mineral content (solid circles) and areal density (open circles) from an ashed bone scanned repeatedly over 45 weeks with instrument D4. Upper: Operator-interactive data. Lower: Processing with no operator intervention ($BMD = 1.41 - 3.15 \cdot 10^{-3}$ wk).

are necessary for detection of instrument failure, but a slight drift of the data may be overlooked until a retrospective review of data obtained for 1 yr or longer is conducted. This retrospective evaluation allows detection of possible systematic time trends.

None of the commercial instruments tested had such a phantom to test the effectiveness of the primary standardization. A summary of the tests used in this study is given in Table 5. Of all the instruments tested, one DPA instrument (D3) showed an error that would render it useless for longitudinal measurements. The magnitude of the error depended on the count rate. Because each of the two energy windows has a different count rate for any given point, the error would affect counts in both windows differently. In addition, a thin absorber (high count rates for I) would have a larger error than a thick absorber (low count rates for I). Therefore, one would not expect to find a constant error when measuring different objects. This explains why in the study of weak and strong sources the subject ($BMD = 0.995$ g/cm^2) had the greatest error (9.8%) and the bone sample ($BMD = 1.835$ g/cm^2) the lowest error (3.8%). In absolute values, however, differences in error are far less dramatic. The described error has been corrected in a new software version. The poor performance of the nonoperator-interactive mode in another commercial DPA instrument (D4) probably was due to sensitivity of the analysis algorithm to count rate. The error described for the SPA forearm instrument (S2) may be responsible for the need to use the correction procedure suggested by Nilas et al (*9*) for SPA of the forearm to achieve a precision of under 3% (CV).

Based on our experience we propose two approaches for longitudinal measurements of bone mineral by either DPA or SPA. In addition to the manufacturer's recommended quality control for calibration of the

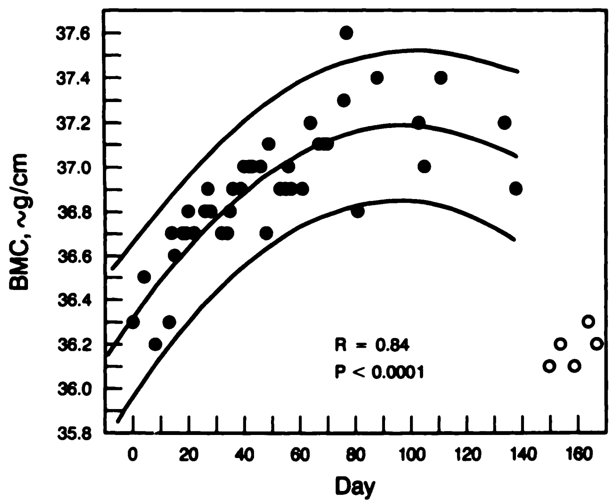

FIGURE 6
Bone mineral equivalent of a metal standard over the source life of 140 days and the first few days of a new source (open circles), measured with SPA instrument S2. Regression equation (d, days): $y = 36.325 + 1.798 \cdot 10^{-3} d - 9.170 \cdot 10^{-5} d^2$ ($n = 42$; $R^2 = 0.71$; $P < 0.0001$; $s_{y \cdot x} = 0.16$, 0.4% of $\bar{y}$).

instrument, first measurements of a standard (perhaps a bone) in 15–20 cm of Lucite or under 20 cm of water should be performed repeatedly, perhaps weekly. These results should be plotted and analyzed for possible time trends for the entire lifetime of a given source. If significant trends are discovered, the data should be adjusted based on the derived equations. This can be done, for example, by forming ratios of predicted mean from the equation at specific times to the predicted mean at the beginning of the source and using these ratios as adjustment factors. Second, data from the bone phantom obtained during the last week of scanning with an old source should be compared with data from the first week of the new source. This tests the performance of the instrument under extremes of count rates and confirms the data obtained from the first approach suggested above. This test, however, does not replace the longitudinal evaluation because, in the comparison of two sources (strong and weak), differences specific to the two sources also have to be considered. As an additional test, the half-life of ^{153}Gd can be determined from the counts in air for each energy.

We also propose that the results of these two approaches should be part of any report of a longitudinal study. At the end of the study these data should be evaluated for the entire study period. When measurements made on two different dates are compared the actual instrument precision should be calculated from measurements of the phantom made on corresponding dates.

Long-term precision of commercial or locally built instruments cannot be derived from short-term measurements made prior to the start of a longitudinal study.

NOTES

* D1 and D2, locally built dual-photon absorptiometry instruments, previously described in reference *4*; D3, Lunar DP3, Lunar Corporation, Madison, WI; D4, Novo BMC Lab 22a, Novo Diagnostics, Bagsveard, Denmark.

† From Oak Ridge National Laboratory; encapsulated by Gulf Nuclear Inc., Webster, TX.

‡ S1, locally built single-photon absorptiometry instrument, described in reference *2*; S2, System 1100, Nuclear Data, Inc., Medical Products Division, Schaumburg, IL.

§ ^{125}I sealed sources, from Atomic Energy of Canada.

ACKNOWLEDGMENT

The authors thank the manufacturers of the instruments used for their assistance and suggestions in this project.

REFERENCES

1. Wahner HW, Dunn WL, Riggs BL. Assessment of bone mineral. Part 2. *J Nucl Med* 1984; 25:1241–1253.
2. Wahner HW, Dunn WL, Riggs BL. Noninvasive bone mineral measurements. *Semin Nucl Med* 1983; 13:282–289.
3. Wahner HW, Dunn WL, Thorsen HC, et al. Bone mineral measurements: review of various techniques. In: Wahner HW, ed. Nuclear medicine. Quantitative procedures. Boston: Little, Brown and Company, 1983:127.
4. Dunn WL, Wahner HW, Riggs BL. Measurement of bone mineral content in human vertebrae and hip by dual photon absorptiometry. *Radiology* 1980; 136:485–487.
5. Riggs BL, Wahner HW, Melton LJ, III, et al. Rates of bone loss in the appendicular and axial skeletons of women: evidence of substantial vertebral bone loss before menopause. *J Clin Invest* 1986; 77:1487–1491.
6. Lederer CM, Shirley VS. Table of isotopes. Seventh edition. New York: John Wiley and Sons, 1978:866.
7. Seeman E, Wahner HW, Offord KP, et al. Differential effects of endocrine dysfunction on the axial and appendicular skeleton. *J Clin Invest* 1982; 69:1302–1309.
8. Dunn WL, Wahner HW. Operational comparison of commercial dual photon absorptiometry (DPA) systems, Read at the Fifth International Workshop on Bone and Soft Tissue Densitometry Using Computed Tomography, Bretton Woods, New Hampshire, October 14 to 18, 1985.
9. Nilas L, Borg J, Gotfredsen A, et al. Comparison of single- and dual-photon absorptiometry in postmenopausal bone mineral loss. *J Nucl Med* 1985; 26:1257–1262.