

Recalculation of the NHANES Database SD Improves T-Score Agreement and Reduces Osteoporosis Prevalence

Neil Binkley,¹ Gary M Kiebzak,² E Michael Lewiecki,³ Diane Krueger,¹ Ronald E Gangnon,⁴ Paul D Miller,⁵ John A Shepherd,⁶ and Marc K Drezner¹

ABSTRACT: In attempt to improve diagnostic agreement between manufacturers, a recent software update incorporated NHANES III data in GE Lunar densitometers. As a result, the femur neck and trochanter T-scores were lowered, and osteoporosis prevalence was increased. Use of a recalculated young-normal SD for the GE Lunar-adjusted NHANES III database improved diagnostic agreement and is recommended.

Introduction: Use of manufacturer-specific normative databases for T-score derivation leads to discordance in T-score values and differences in diagnostic classification. To address this issue, the International Committee for Standards in Bone Measurement (ICSBM) recommended the NHANES III database for femur T-score derivation. Acquired on Hologic (Hol) instruments, this database requires conversion equations for application to other DXA systems. NHANES III total femur (TF) conversions for GE Lunar (GE) have previously been available, and femoral neck (FN) and trochanter (TR) equations were reported recently. Per the ICSBM recommendation, GE Lunar incorporated these values into their female database. This should produce T-score and diagnostic agreement between Hol and GE instruments; however, this has not been evaluated.

Materials and Methods: We compared GE femur scans in 115 postmenopausal women using software before and after the NHANES III software update. Subsequently, T-scores derived from femur scans obtained on GE and Hol densitometers were compared in a different group of 89 postmenopausal women.

Results: The NHANES III software update had no effect on measured BMD (g/cm^2) at any femur region. However, because of changes in values used for T-score calculation (increase in the mean young-normal BMD at the FN and TR and a reduction in SD at the TR), the T-scores were lower (mean, 0.48 and 0.68, respectively) at the FN and TR using post-NHANES III software. Consequently, this update increased femur osteoporosis prevalence in these 115 women from 7.8% to 18.3%. Comparison of GE with Hol total proximal femur T-scores revealed a minimal difference (<0.1) and equal diagnoses of osteoporosis. FN and TR differences were larger, with mean GE T-scores lower than Hol ($p < 0.001$) by 0.17 and 0.50, respectively, thereby introducing osteoporosis diagnostic disagreement (13 [GE] versus 9 [Hol]). Our evaluation suggested that this disparity resulted from direct application of published NHANES III SDs at the FN and TR. As such, we applied the conversion formulae to the NHANES III published Hologic data and found the FN and TR SDs were greater than assumed by GE. Using our recalculated SD to derive T-scores reduced the mean GE/Hol T-score difference to 0.03 at the FN and 0.32 at the TR and resolved osteoporosis diagnostic disagreement.

Conclusion: The GE NHANES III software update leads to lower FN and TR T-scores than obtained with Hol or prior GE software. Recalculation of the young-normal SD reduces this difference and is recommended. Clinicians are advised to avoid using the TR for diagnosis or, at a minimum, use caution when making treatment decisions based solely on T-score at this site.

J Bone Miner Res 2005;20:195–201. Published online on November 16, 2004; doi: 10.1359/JBMR.041115

Key words: osteoporosis, diagnosis, T-score, database

INTRODUCTION

THE WORLD HEALTH Organization (WHO) criteria specify that osteoporosis is present in postmenopausal women when the BMD is 2.5 SD or more below that of the mean in a young adult population.⁽¹⁾ This approach is applied clinically as a T-score of -2.5 or below with the T-score

Dr Drezner served as a consultant for Aventis, GE Lunar, Novartis, NPS Pharmaceuticals, and Roche. Dr Lewiecki owns stock in GE, received funding from Amgen, Aventis, Eli Lilly and Company, GE Lunar, Kyphon, Merck, Novartis, Pfizer, Procter & Gamble, Roche, and Wyeth, and served as a consultant for Aventis, Eli Lilly and Company, GE Lunar, Kyphon, Merck, Novartis, Procter & Gamble, Roche, and Wyeth. All other authors have no conflict of interest.

¹University of Wisconsin Osteoporosis Clinical Center and Research Program, Madison, Wisconsin, USA; ²Center for Orthopaedic Research and Education, St Luke's Episcopal Hospital, Houston, Texas, USA; ³New Mexico Clinical Research and Osteoporosis Center, Albuquerque, New Mexico, USA; ⁴Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison, Wisconsin, USA; ⁵Colorado Center for Bone Research, Lakewood, Colorado, USA; ⁶Department of Radiology, University of California, San Francisco, California, USA.

TABLE 1. EFFECT OF DATABASE ON YOUNG-ADULT MEAN BMD AND SD IN WOMEN

Femur region	GE Lunar software before version 7			GE Lunar software version 7.0 and later			NHANES III recalculated		
	Database used	Mean BMD (g/cm ²)	SD	Database used	Mean BMD (g/cm ²)	SD	Database used	Mean BMD (g/cm ²)	SD
Total	NH	1.008	0.126	NH	1.0077	0.122	RC	1.008	0.126
Neck	GE	0.980	0.120	NH	1.038	0.120	RC	1.039	0.139
Trochanter	GE	0.790	0.110	NH	0.851	0.099	RC	0.850	0.115

NH, NHANES III; GE, GE Lunar; RC, Recalculated.

For software versions 7.0 and later, the mean femur neck and trochanter BMD was obtained by converting published Hologic values; however, the SD was not similarly converted. The SD was also converted for the "recalculated" database.

defined as follows: (individual's BMD – young-adult mean BMD)/SD of the young-adult normal population. As such, the population used to define the young-adult mean BMD and the variability within this group substantially impacts osteoporosis prevalence using this T-score-based approach.^(2–6) Because manufacturer-developed proprietary databases use different populations, the young-adult mean BMD and/or SD may differ.⁽²⁾ Thus, the diagnostic classification of an individual as normal, osteopenic, or osteoporotic may vary depending on the densitometer used.⁽³⁾

Ideally, all bone mass measurement devices would use the same population to define the young-normal mean BMD and SD, a process that does lead to derivation of similar T-scores with instruments of different manufacturers.⁽⁴⁾ Although use of a single large sample population to develop a uniform normative database for all densitometers has been suggested,^(5,6) this process has not been implemented. To improve harmonization of diagnostic classification, the International Committee for Standards in Bone Measurement (ICSBM) agreed on a universal reference database for the femur based on NHANES III,⁽⁷⁾ the only large standardized reference database ever published.^(1,8) Because the NHANES III data were acquired using Hologic densitometers, the ICSBM published formulae to convert measured BMD into standardized BMD at the total femur, thereby allowing use of the NHANES III database by other densitometer manufacturers.^(7,9–11) This methodology allowed for calculation of total proximal femur T-scores from a common young-adult database, thereby removing the T-score discrepancy between Hologic and GE Lunar that had previously existed at this skeletal site.⁽³⁾ Recently, Lu et al.⁽¹²⁾ published standardization equations allowing the NHANES III database to be used at the femur neck and trochanteric subregions. Subsequently, GE Lunar used these conversion equations to incorporate NHANES III data at the femur neck and trochanter into their software, beginning with version 7.0 in the fall of 2002. Because the earlier application of NHANES data did not materially change the total femur T-score from that derived using the GE Lunar normative database, it might be anticipated that this software update would not alter femur neck or trochanter T-scores. However, the impact of the software upgrade on T-scores as determined using GE Lunar densitometers has not been reported.

We initially noted that this software upgrade seemed to

lower femur neck and trochanter T-scores. If lower T-scores at the same BMD resulted from this software update, the number of women diagnosed as "low" using the WHO criteria would be increased. To investigate this possibility, we evaluated the effect of this update on osteoporosis prevalence in a group of postmenopausal women measured using a GE Lunar densitometer. Furthermore, it would be expected that use of a common normative database, in this case NHANES III, would produce T-score and diagnostic agreement between densitometer manufacturers. To investigate this, we subsequently evaluated whether the software upgrade produced diagnostic agreement between GE Lunar and Hologic densitometers in a second group of postmenopausal women.

MATERIALS AND METHODS

Terminology and databases used

A summary of software versions and normative databases used in this study with the female young-normal mean BMD and SD are presented in Table 1. Before the GE Lunar software update (instituted at software version 7.0 in November of 2002), NHANES III data were used only for total femur T-score calculation. The total femur values were previously derived by converting the published Hologic value to sBMD and converting this to a Lunar-equivalent BMD by rearranging the sBMD equation. Subsequently, GE Lunar incorporated NHANES III values by applying conversion equations to the published young-normal mean BMD as measured by Hologic instruments while employing the published SD of BMD, as measured by Hologic instruments, to derive T-scores at all hip regions.

We recalculated the young-normal mean BMD and SD for GE Lunar using published young-normal mean BMD and SD values at the hip from the 409 non-Hispanic white females 20–29 years of age in the NHANES III cohort.⁽¹⁾ Equations for converting Hologic BMD values to Lunar-equivalent BMD values were obtained from an author of the publication from Lu et al. (T Fuerst, personal communication, 2004), which are as follows:

Total Proximal Femur: GE Lunar = 0.038 + 1.030 × Hologic

Femur Neck: GE Lunar = 0.045 + 1.158 × Hologic

Trochanter: GE Lunar = 0.026 + 1.164 × Hologic

The young-adult GE Lunar-equivalent mean BMD val-

ues and SDs were determined using standard formulas for linear transformations. In particular, if we apply the linear transformation $y = a + b \times x$ to data with mean $\mu(x)$ and SD $SD(x)$, the transformed data have mean $\mu(y) = a + b \times \mu(x)$ and SD $SD(y) = b \times SD(x)$. The T-scores derived using this approach are referred to as “recalculated” results.

Subjects

Effect of GE Lunar NHANES III software update: A total of 115 postmenopausal women was randomly selected from a population that responded to newspaper advertising to obtain BMD measurement at the University of Wisconsin Osteoporosis Clinical Center and Research Program (Madison, WI). Each woman received a single lumbar spine and left hip measurement using standard scan protocols on a GE Lunar Prodigy. All of these women were white, with a mean age and weight of 65.6 years (range, 42.9–90.3 years) and 156 lbs (range, 100–275 lbs), respectively. On review by the University of Wisconsin IRB, acquisition of informed consent was waived because all subject identifiers were removed from these scans.

GE Lunar/Hologic comparison: As part of a different study to compare GE Lunar Prodigy and Hologic Delphi densitometers, 90 women (30 each from three osteoporosis centers) were recruited. Their mean age was 61.6 ± 1.0 (SE) years. The three clinical sites were the Department of Radiology, University of California at San Francisco, CA; the New Mexico Clinical Research and Osteoporosis Center, Albuquerque, NM; and the Colorado Center for Bone Research, Lakewood, CO. Each woman received a total of four hip scans, two each using a GE Lunar Prodigy and Hologic Delphi. After their initial scan, all patients stood up and were subsequently repositioned, and these scans were repeated in their entirety. IRB approval was obtained for this study, and all participants provided written informed consent.

Scan analysis

All scans were performed by International Society for Clinical Densitometry–certified technologists who used manufacturer-recommended acquisition and analysis techniques.

GE Lunar NHANES III software update: Each scan was initially auto-analyzed using either software version 6.60 or 6.70 to obtain the pre-NHANES III software update T-scores. No operator adjustment to region of interest placement or point typing was performed. Subsequently, all hip scan files were auto-reanalyzed using software version 7.53 to obtain post-NHANES III software update T-scores.

GE Lunar/Hologic comparison: Hologic Delphi software version 11.2 and GE Lunar Prodigy software versions 7.50 and 7.51 using the NHANES III update were used for all initial analyses. Subsequently, all GE Lunar scans were re-analyzed with software before the NHANES III update. For this repeat analysis, auto-reanalysis was performed on all femur scans, and no operator adjustment to region of interest placement or point typing was performed. The average left femur T-score of the two scans performed on each densitometer was used for comparison. One patient

was excluded because her anatomy and positioning precluded appropriate femur region of interest placement.

Statistical analyses

Bland Altman analyses (Analyze-it, Leeds, UK) and paired *t*-tests (Statview Abacus, Cary, NC, USA) were used to compare BMD and T-scores obtained with the GE Lunar “pre” and “post” NHANES III software update. These analyses were also performed on the GE Lunar/Hologic comparison data. Regression analysis was used to compare GE and Hologic T-scores. The effects of reference database on WHO diagnostic categorization were evaluated using McNemar’s test, a test for paired proportions that analyzes the number of disagreements (Analyze-it). Paired *t*-tests were used to compare BMD and T-scores for patients measured on both GE Lunar and Hologic DXA systems, and $p < 0.05$ was considered statistically significant.

RESULTS

Impact of database on average young-normal BMD and SD

As noted in Table 1, the NHANES III software update resulted in an increase in young-normal mean BMD of 5.9% at the femur neck and 7.7% at the trochanter compared with the GE Lunar normative database. Furthermore, the SD at the trochanter was decreased by 10%. Either of these changes (i.e., an increase in young-normal mean BMD or a reduction in SD) will increase the number of women diagnosed as low when using the WHO diagnostic criteria. Our recalculated NHANES III values also led to an increase in young-normal mean BMD at the femur neck and trochanter; however, the SD also increased at these sites compared with that used with the post-NHANES software.

GE Lunar NHANES III software update

Proximal femur BMD: The GE Lunar NHANES III software update did not alter measured BMD at the total proximal femur, femur neck, or trochanter (data not shown), with the exception of three individuals. In 3 of the 115 women, total proximal femur BMD differed by 0.001 g/cm^2 ; in 1 of these 3 women, the trochanter BMD also differed by 0.001 g/cm^2 . Measured BMD at all three femoral subregions was identical for the remaining subjects.

Proximal femur T-score and osteoporosis prevalence-comparisons: T-scores of the total proximal femur, although statistically different ($p < 0.0001$), were clinically unchanged, differing by only 0.02, when derived using software before (“pre”) and after (“post”) the NHANES III software update (Fig. 1A). However, at the femur neck (Fig. 1B) and trochanter (Fig. 1C), all subjects had a lower T-score using the post-NHANES III update software. The mean T-score decrease resulting from use of the updated software was 0.48 at the femoral neck and 0.68 at the trochanter. The T-scores were lower ($p < 0.0001$) because of an increase in young-adult mean BMD at the neck and trochanter and a concomitant reduction in trochanteric SD (Table 1) compared with the GE Lunar database. As a result of this T-score decrease, the number of women with

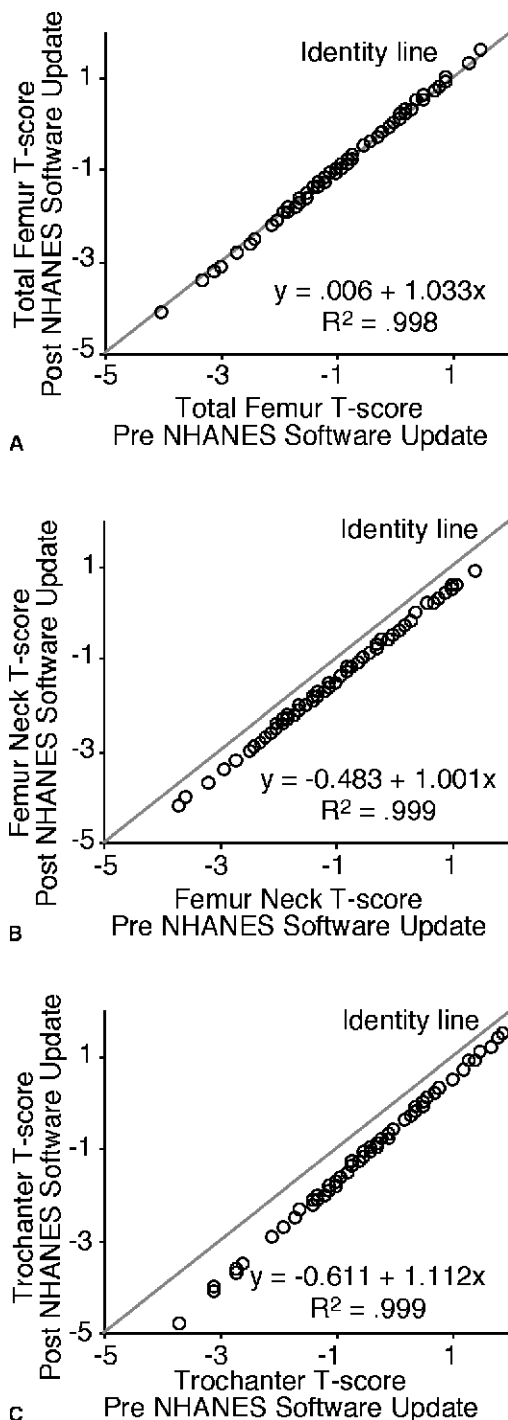


FIG. 1. (A–C) Effect of NHANES III software update on femur T-scores. Total femur T-scores in 115 postmenopausal women were very similar (bias of 0.02 T-score). However, substantial bias was present after the NHANES III update at the femur neck; the T-score for all subjects was 0.4 or 0.5 lower. This reflects an increase in young-normal mean BMD from 0.980 to 1.038 g/cm² (pre- to post-update) without change in SD. Similarly, the trochanter T-score was 0.4–1.1 lower because of an increase in young-normal mean BMD from 0.790 to 0.851 g/cm² (pre- to post-update) and an SD decrease from 0.110 to 0.099/cm².

osteoporosis at the femur according to WHO criteria increased ($p = 0.0005$) from 9 (7.8%) to 21 (18.3%) after the NHANES III software update. Additionally ~20% fewer women ($p < 0.001$) were classified as having normal bone mass after the NHANES III software update (Table 2; Fig. 3).

GE Lunar/Hologic comparison

As is widely recognized,^(9,12) the measured BMD (g/cm²) at all femur sites was lower ($p < 0.0001$) using Hologic densitometers than with GE Lunar instruments (Table 3). These differences between densitometers of various manufacturers reflect different methods of dual-energy X-ray generation, edge detection paradigms, and region of interest placement.⁽¹³⁾ Specifically, in this study, the mean Hologic-measured BMD was 7.8%, 17.8%, and 13.6% lower at the total proximal femur, femur neck, and trochanter, respectively, than that measured using GE Lunar densitometers.

T-scores at the total proximal femur were very similar, with a mean difference ($p < 0.01$) of 0.08. Specifically, the mean total proximal femur T-score in this population was –0.79 using GE Lunar and –0.87 with Hologic densitometers. This small difference likely does not have clinical significance. Consistent with absence of a clinically meaningful effect, in this population, an equal number of women ($n = 6$) were diagnosed as osteoporotic at the total proximal femur. Femur neck and trochanter differences ($p \leq 0.002$) were larger, and the mean GE Lunar T-scores were lower by 0.17 and 0.50 at these sites, respectively; mean femur T-scores are presented in Fig. 2. Given the lower T-scores, one would expect that a larger number of women would be diagnosed with osteoporosis using GE Lunar densitometers. Overall, using the lowest T-score of the total proximal femur, femur neck, and trochanter, 13 (14.6%) of these women were classified as osteoporotic with GE Lunar densitometers compared with 9 (10.1%) with Hologic instruments (Fig. 3).

Effect of SD recalculation on osteoporosis prevalence

Because our observations suggested that the demonstrated increase in osteoporosis diagnosis with the GE Lunar NHANES III software update might be caused by differences in SD, we evaluated T-score and diagnostic agreement using our recalculated values. Using our recalculated SD to derive T-scores reduced osteoporosis prevalence in the initial group of 115 women from 18.3% to 7.8%. In the GE/Hologic comparison group, the mean T-score difference was reduced to 0.03 at the FN and 0.32 at the trochanter (Fig. 2). As a result, an equal number of women (9) were diagnosed as osteoporotic at the femur (Fig. 3).

DISCUSSION

Previous observations demonstrate the importance of using a standard normative database for T-score derivation^(3,4) to enhance diagnostic agreement. However, acquisition of NHANES III data exclusively using Hologic densitometers⁽⁸⁾ initially precluded database standardiza-

TABLE 2. FEMUR DIAGNOSTIC CATEGORIZATION IN 115 WOMEN BEFORE AND AFTER THE NHANES III SOFTWARE UPDATE AND WITH RECALCULATION

Diagnostic category	Pre-NHANES III update: [no. (%)]	Post-NHANES III update: [no. (%)]	Recalculated NHANES III: [no. (%)]
Normal	45 (39.1%)	22 (19.1%)*	29 (25.2%) [†]
Osteopenia	61 (53.0%)	72 (62.6%)	77 (67.0%)
Osteoporosis	9 (7.8%)	21 (18.3%) [†]	9 (7.8%)

Diagnosis based on lowest T-score of the total proximal femur, femur neck, or trochanter.
 * $p < 0.0005$, [†] $p < 0.001$ compared to pre-NHANES III software update.

TABLE 3. LEFT FEMUR MEAN BMD, T-SCORES AND NUMBER OF WOMEN DIAGNOSED AS OSTEOPOROTIC USING HOLOGIC AND GE LUNAR SOFTWARE INCORPORATING THE NHANES III UPDATE

Region	BMD (g/cm ²)		T-score		No. with T-score ≤ -2.5	
	GE Lunar	Hologic	GE Lunar	Hologic	GE Lunar	Hologic
Total	.911 ± .128	*.840 ± .126	-0.79 ± 1.05	-0.87 ± 0.97	6	6
Neck	.869 ± .118	*.714 ± .105	-1.42 ± 0.98	[†] -1.25 ± 0.92	10	6
Troch	.731 ± .120	*.632 ± .110	-1.21 ± 1.21	*-0.71 ± 1.07	9	5

Total n = 89; data presented as mean ± SD
 * $p < 0.0001$, [†] $p < 0.001$

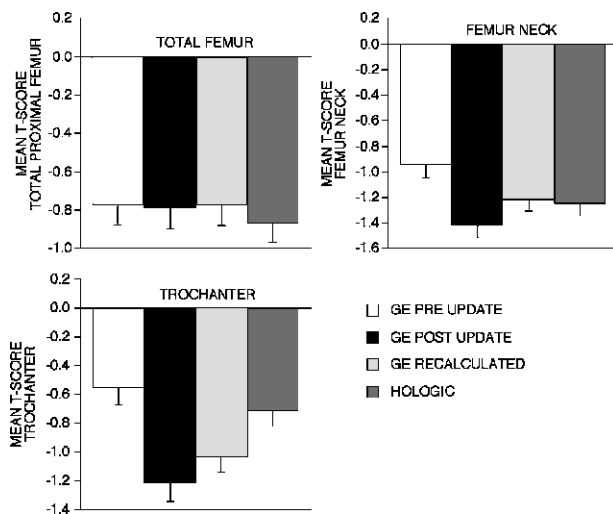


FIG. 2. Effect of database and densitometer on total proximal femur T-score. Although statistically different ($p < 0.01$), the mean total proximal femur T-score of these 89 women was quite similar (range, -0.87 to -0.75) regardless of densitometer or software used. Larger differences in mean T-score ($p < 0.0005$) between GE Lunar and Hologic were present at the femur neck and trochanter using either the pre- or post-NHANES III software update. Recalculation of the young-normal mean BMD and SD resolved the between-manufacturer difference at the femur neck but not at the trochanter.

tion among manufacturers. To rectify this problem, conversion equations were previously developed allowing incorporation of total proximal femur NHANES III data into other manufacturers normative databases. Therefore, before the recent update, when “USA/NHANES” was selected in GE Lunar software, T-scores were derived using converted NHANES III data at the total proximal femur but not at the femur neck or trochanter, where the manu-

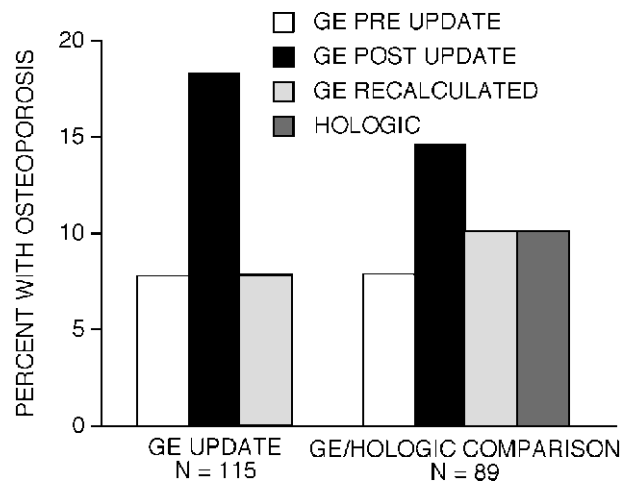


FIG. 3. Osteoporosis prevalence at the hip. Using the lowest T-score of the total proximal femur, femur neck, and trochanter, the T-score decrease that resulted from the GE Lunar NHANES III software update increased ($p = 0.0005$) osteoporosis prevalence from 7.8% to 18.3% in this population of 115 women postmenopausal women. In a second population of 89 women, use of our recalculated young-normal mean BMD and SD led to classifying an identical number (9) as osteoporotic with GE Lunar and Hologic instruments.

facturer’s normative database was used. Our study shows that the recent availability of equations allowing application of the NHANES III database to the femur neck and trochanter⁽¹²⁾ results in lower T-scores at these sites than was previously the case with the GE Lunar normative database. This reduction in T-score is caused by increase in mean BMD combined with reduction in SD of the young-normal population. As such, osteoporosis prevalence (using the WHO criteria) is increased compared with prior GE software and also compared with Hologic. Recalculation of the

SD from published NHANES III data reduces the T-score difference between GE Lunar and Hologic and yields an equal number of women diagnosed as osteoporotic. Although not tested in this study, it is reasonable to assume that the SD correction would improve T-score agreement not only for GE Lunar Prodigy but also pencil-beam systems. Thus, use of existing GE Lunar software that has incorporated converted NHANES III data at all femur subregions (versions 7.0 or higher) is not recommended.

In an attempt to minimize diagnostic discordance among manufacturers, the ICSBM previously endorsed use of the NHANES III database for T-score derivation at the hip.⁽⁷⁾ On recalculation of the young-adult SD, this study shows good agreement between GE Lunar and Hologic at the total proximal femur and femur neck. Thus, the expectation that use of a common database would lead to similar T-scores and diagnosis is corroborated, and use of NHANES III as the default hip database is affirmed. Recalculation and use of updated SD values at the femur neck and trochanter for GE densitometers is recommended.

In the interim, clinicians using GE Lunar densitometers must be aware of the clinical impact of these reference database changes that were incorporated into enCORE software versions 7.0 and later. Because the femur neck and trochanter T-scores are lowered by -0.5 and 0.7 , respectively, diagnostic classification may worsen, despite stable BMD, in women whose bone mass is being monitored over time. Such individuals should be advised that a change in the data used for T-score determination, not a worsening of their condition, has occurred. Additionally, some individuals receiving osteoporosis preventive therapy could, on follow-up BMD measurement, have a statistically significant increase in bone mass, but worsening of their diagnostic categorization from osteopenia to osteoporosis. Clinicians and their patients must be aware that this "worsening" of diagnostic status should not promote patient anxiety or lead to changes in therapeutic regimen, because it is simply a phenomenon of the software change. Additionally, it must be emphasized that therapeutic response should always be evaluated using change in BMD (g/cm^2) and not T-scores.⁽¹⁴⁾ Furthermore, if/when GE Lunar software upgrades incorporate recalculated NHANES III values, the above noted issues should be improved.

The observation that both the GE Lunar NHANES III software update and our recalculated values result in lower trochanter T-scores on GE Lunar instruments compared with the prior database, coupled with the observation that the T-score is now lower at this site than that observed with Hologic densitometers, is cause for concern. It is important to note that conversion equations for the hip regions were derived from data acquired with an earlier generation GE Lunar densitometer (DPX-L).⁽¹²⁾ Thus, it is possible that improvements in GE Lunar trochanter edge detection programs that have occurred since that time are cause for this discrepancy. Regardless, discordance between GE Lunar and Hologic exists at the trochanter, such that more women will be diagnosed as "low" at this site when using GE Lunar densitometers than with Hologic. It is suggested that formulas allowing NHANES III data application to non-Hologic densitometers be developed using current genera-

tion instrumentation. Additionally, until this discordance is resolved, it seems prudent for clinicians to avoid using the trochanter for diagnosis when using GE Lunar post-NHANES III update software, or at a minimum, use caution when making treatment decisions based solely on trochanter T-score.

It is important to recognize that these T-score changes associated with updating the GE Lunar NHANES III software apply only to women. Equations allowing conversion of NHANES III data in men at the femur neck and trochanter have not been published, and Lu et al.⁽¹²⁾ specifically caution that "...more research is needed to determine whether these results can be extended to other races and to men." As such, in all GE Lunar software versions, the manufacturer's proprietary normative database, not the NHANES III database, is used for men to derive T-scores at the femur neck and trochanter despite the statement on the printout that the "NHANES/USA" femur reference population is used.

T-score discrepancies, and therefore differences in diagnostic categorization, can be caused by multiple factors including differences in technology, sites at which BMD is measured, and normative database used.⁽¹⁵⁾ Additionally, as this study emphasizes, software updates may substantially alter T-score without affecting the measured BMD. In the case of this particular software update, it is clear that the "normal" populations are not identical as the young-normal mean BMD, and SDs differ between the NHANES III and GE Lunar normative databases. It is not surprising that the young-normal populations differ, because NHANES used age 20–29 and GE Lunar used age 20–40. Nonetheless, these differences are smaller at the total femur, possibly because of the larger amount of bone being measured, than at the femur subregions. Although not evaluated in this study, it is possible that similar diagnostic disagreement exists at other skeletal sites. In fact, available data suggest that T-score agreement between manufacturers is quite good at the lumbar spine in women,⁽¹⁶⁾ but not in men,⁽¹⁷⁾ and that diagnostic discordance exists at the radius.⁽¹⁷⁾ Attainment of a common database, by measuring BMD in a large number of young adults on all instruments, would rectify this situation. In the absence of such, it is not possible to determine which database or manufacturer is "right." Therefore, use of NHANES III after recalculation of the SD is the best available option to achieve diagnostic agreement at the femur at this time.

Finally, this study emphasizes that the T-score based system of diagnosis is not ideal. The prevalence of osteopenia and osteoporosis in a population, and these diagnoses in individual patients, may vary according to the mean BMD and SD of the young-normal population; values that may vary substantially if different populations are used to construct the normative database. As such, T-score values may vary substantially despite similar BMD. Efforts to standardize reference databases are encouraged in that establishment of a single young-normal reference database among all densitometry manufacturers would minimize such diagnostic discrepancies. Additionally, the use of fracture risk reporting may become a useful complement to diagnostic

classification by T-score, thereby allowing improved selection of patients for therapeutic intervention.

In conclusion, the GE Lunar software update that allows application of NHANES III data at the femur neck and trochanter leads to lower T-scores at these sites. Recalculation of the NHANES III young-normal SD reduces this difference, but discrepancy remains at the trochanter, raising questions regarding conversion equation accuracy at this site. Use of recalculated NHANES III values at the femur subregions for GE Lunar densitometers is recommended. Clinicians must be aware of this database impact on diagnosis and watch for a software upgrade correcting the SD to assure optimal patient care. Difficulties with the T-score-based diagnostic system, such as those noted here, should further prompt development of diagnostic criteria that include estimates of absolute fracture risk. However, as an absolute fracture risk-based system is likely to retain use of T-scores, standardization of their derivation will remain essential to assure uniform diagnosis of osteoporosis.

ACKNOWLEDGMENTS

The authors thank Nellie Vallarta-Ast, Cathy Brothers, Abby Erickson, Julie Montano, and Vesta March for excellent technical assistance in performing the DXA scans used in this study and Nicole Moore for assisting with data management.

REFERENCES

1. Looker AC, Wahner HW, Dunn WL, Calvo MS, Harris TB, Heyse SP 1998 Updated data on proximal femur bone mineral levels of U.S. adults. *Osteoporos Int* **8**:468–489.
2. McMahon K, Kalnins S, Freund J, Pocock N 2003 Discordance in lumbar spine T-scores and nonstandardization of standard deviations. *J Clin Densitom* **6**:1–6.
3. Faulkner KG, Roberts LA, McClung MR 1996 Discrepancies in normative data between Lunar and Hologic DXA systems. *Osteoporos Int* **6**:432–436.
4. Greenspan SL, Bouxsein ML, Melton ME, Kolodny AH, Clair JH, DeLuca PT, Stek M, Faulkner K, Orwoll ES 1997 Precision and discriminatory ability of calcaneal bone assessment technologies. *J Bone Miner Res* **12**:1303–1313.
5. Miller PD 2000 Controversies in bone mineral density diagnostic classifications. *Calcif Tissue Int* **66**:317–319.
6. Miller PD 2002 Controversial issues in bone densitometry. In: Bilezikian JP, Raisz LG, Rodan GA (eds.) *Principles of Bone Biology*, 2nd ed. Academic Press, San Diego, CA, USA, pp. 1587–1597.
7. Hanson J 1997 Standardization of femur BMD. *J Bone Miner Res* **8**:1316–1317.
8. Looker AC, Orwoll ES, Johnston CC, Lindsay RL, Wahner HW, Dunn WL, Calvo MS, Harris TB, Heyse SP 1997 Prevalence of low femoral bone density in older U.S. adults from NHANES III. *J Bone Miner Res* **12**:1761–1768.
9. Genant HK, Grampp S, Gluer CC, Faulkner KG, Jergas M, Engelke K, Hagiwara S, Van Kuijk C 1994 Universal standardization for dual x-ray absorptiometry: Patient and phantom cross-calibration results. *J Bone Miner Res* **9**:1503–1514.
10. Genant HK 1995 Letter to the Editor Universal standardization for dual x-ray absorptiometry: Patient and phantom cross-calibration results. *J Bone Miner Res* **10**:997–998.
11. Steiger P 1995 Standardization of spine BMD measurements. *J Bone Miner Res* **10**:1602–1603.
12. Lu Y, Fuerst T, Hui S, Genant HK 2001 Standardization of bone mineral density at femoral neck, trochanter and Ward's triangle. *Osteoporos Int* **12**:438–444.
13. Blake GM, Wahner HW, Fogelman I 1999 *The Evaluation of Osteoporosis: Dual Energy X-Ray Absorptiometry and Ultrasound in Clinical Practice*, 2nd ed. Martin Dunitz, London, UK.
14. Lenchik L, Kiebzak GM, Blunt BA 2002 What is the role of serial bone mineral density measurements in patient management? *J Clin Densitom* **5**:S29–S38.
15. Faulkner KG, von Stetten E, Miller P 1999 Discordance in patient classification using T-scores. *J Clin Densitom* **2**:343–350.
16. Kiebzak G, Lewiecki EM, Miller PD 2004 Good diagnostic agreement using T-scores between Delphi and Prodigy. *J Clin Densitom* **7**:229.
17. McMahon K, Nightingale J, Pocock N 2004 Discordance in DXA male reference ranges. *J Clin Densitom* **7**:121–126.

Address reprint requests to:

Neil Binkley, MD

2870 University Avenue

Suite 100

Madison, WI 53705, USA

E-mail: nbinkley@facstaff.wisc.edu

Received in original form August 2, 2004; revised form August 16, 2004; accepted September 14, 2004.