

NEWS & Commentary

Note: The level of evidence indicated for each study is based on a grading system that evaluates the scientific rigor of the study design as developed by the U.S. Preventive Services Task Force. A synopsis of the levels is presented at the end of these items.

Stopping Hormone Therapy Reverses Most Beneficial Effects on Bone

Gallagher JC, Rapuri PB, Haynatzki G, Detter JR. Effect of discontinuation of estrogen, calcitriol, and the combination of both on bone density and bone markers. *J Clin Endocrinol Metab* 2002;87:4914-4923.

Elderly women who discontinue postmenopausal hormone therapy and/or calcitriol therapy lose much of the bone mineral density (BMD) gains from treatment, according to this 5-year, randomized, prospective, placebo-controlled trial. A total of 489 women (aged 65-77) were randomized to one of four treatment groups: hormone therapy (HT) alone, calcitriol alone, HT plus calcitriol, or placebo. Hormone recipients received estrogen (0.625 mg/day conjugated equine estrogens) either alone or with 2.5 mg/day of medroxyprogesterone acetate. Calcitriol, a synthetic vitamin D analog, was dosed at 0.25 µg twice daily. The treatment phase lasted 3 years and was followed by a 2-year follow-up off drug. Compared with placebo, all treatment groups at year 3 had statistically significant BMD increases in the spine, total body and total hip; femoral neck and trochanter BMD increased in all three groups but was significant only for the HT recipi-

ents. During years 4 and 5 (years 1 and 2 posttherapy follow-up), spinal, femoral neck, trochanter, total body and total hip BMD decreased in all three treatment groups, with the exception of spinal BMD which increased during year 5 in the calcitriol group. Statistical significance persisted at year 5 for all three groups for total body BMD, but only the HT plus calcitriol group had statistically significant differences for spinal and total hip BMD. Compared with baseline, only spinal BMD in the combination group was significantly higher at 5 years. Treatment withdrawal reversed the increase in calcium absorption and the decrease in serum parathyroid hormone levels in the calcitriol groups, and it reversed the decrease in bone markers in the hormone-treatment groups.

Level I evidence

Comment. This well-designed study is extremely timely given the number of women who are discontinuing hormone therapy following the termination of the estrogen-progestin arm of the Women's Health Initiative. It has previously been well-established by numerous studies that the cessation of hormone therapy in early postmenopausal women results in either an accelerated or normal rate of bone loss. The effect of discontinuation of calcitriol therapy on BMD has not been as widely studied. This study specifically evaluated elderly (late) postmenopausal women. Discontinuation of all treatments resulted in rapid bone loss at all sites, with the majority of the loss during the first year. Despite this rapid bone loss, there was still some benefit even 2 years after discontinuation of treatment, especially in the combination therapy group. Nevertheless, this report reinforces the need for clinicians to monitor women after discontinuing hormone therapy and to provide treatment alternatives for the prevention or treatment of bone loss in women at risk.

Risa Kagan, MD
Co-Medical Director
Foundation for Osteoporosis Research
and Education
Oakland, CA
Associate Clinical Professor
Department of Obstetrics and Gynecology
and Reproductive Sciences
University of California at San Francisco

Estrogen Alone Does Not Increase Secondary Cardiovascular Risk

The ESPRIT team. Oestrogen therapy for prevention of reinfarction in postmenopausal women: a randomised placebo controlled trial. *Lancet* 2002;260:2001-2008.

Estradiol valerate does not reduce the overall risk of additional cardiac events in postmenopausal women who have suffered a myocardial infarction, according to data from the estrogen in the prevention of reinfarction trial (ESPRIT). In this randomized, placebo-controlled trial, 1,017 postmenopausal women (aged 50-69 years) who had survived a first myocardial infarction received either estradiol valerate (2 mg) or placebo daily for 2 years. Women with a uterus did not receive concomitant progestogen therapy. Primary outcomes were reinfarction or cardiac death as well as all-cause mortality. At 24 months, the frequency of reinfarction or cardiac death did not differ between the two groups (risk ratio for estrogen recipients, 0.99; 95% CI, 0.70-1.41). All-cause mortality was also statistically similar between the two groups (risk ratio, 0.79%; 95% CI, 0.50-1.27). Among the secondary endpoints (stroke or other embolic event, endometrial cancer, breast cancer and fracture), no outcome reached statistical significance.

Level I evidence

Comment. This study adds to the growing body of randomized clinical

(continued on page 25)

Does Bone Densitometry Predict Mechanical Competence Throughout the Skeleton?

Eva-Maria Lochmüller, MD

Ralph Müller, PhD

Felix Eckstein, MD

Several methodologies are available today for diagnosing osteoporosis and evaluating the risk of bone fracture. These methods display substantial differences in physical background, anatomical measurement location, X-ray exposure, reproducibility and costs involved. However, it has remained ill-defined which technique clinicians should use to diagnose and select patients who require antiosteoporotic treatment.

Introduction

Approximately 30% of all postmenopausal and elderly women in North America and Europe are affected by osteoporosis,¹ defined as a loss of bone mass and a deterioration of bone microstructure (Fig. 1) associated with an increase in susceptibility to fracture. The risk for osteoporotic fracture is as high as the combined risk of breast, endometrial and ovarian cancer, and 60% of all women have encountered at least one osteoporotic fracture by age 80. These fractures affect quality of life as well as mortality;² and the associated costs have been estimated at more than \$14 billion in the United States alone.³ At least one prospective study has demonstrated hormone therapy's (HT) effectiveness in reducing bone fracture risk,⁴ although some controversy remains about the treatment's value and whether the benefits outweigh the risks.⁵⁻⁷ At the same time, there are other treatments available that are proven to significantly reduce the incidence of fractures.⁸ Despite the availability of effective treatments, a major challenge remains—accurately identifying those patients who will benefit most from

treatment. Of equal importance is identifying individuals who do not require treatment.

Bone densitometry has been commercially available for over 20 years to conduct non-invasive evaluation of bone mineral status. Dual energy X-ray absorptiometry (DXA), in particular, has shaped the field of diagnostics in osteoporosis. In fact, the current operational definition of osteoporosis from the World Health Organization (WHO) is based on this measurement

technique.⁹ Of course, several other methods of measuring bone density also are available. These include quantitative computed tomography (QCT), peripheral quantitative computed tomography (pQCT), quantitative ultrasound (QUS) and magnetic resonance imaging (MRI). These have been advocated to be highly predictive (e.g., QCT and pQCT), involve no ionizing radiation (e.g., QUS and MRI) or be less costly (e.g., pQCT and QUS) than other techniques. However, it is impor-

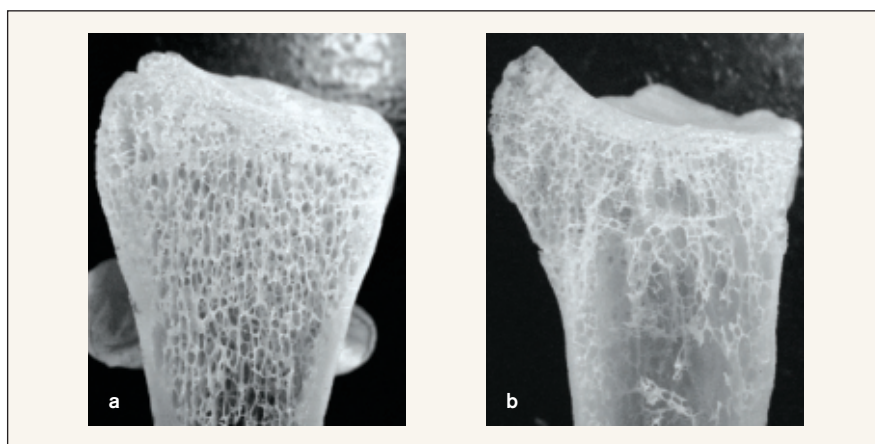


Figure 1. Bone mineral content and trabecular microstructure in the distal metaphysis of the radius
a) healthy case, b) osteoporotic case
Images courtesy of Dipl. Ing. Volker Kuhn, Department of Gynecology and Institute of Anatomy, LMU München)

tant to remember that QCT and QUS cannot be used to classify patients according to the WHO criteria of osteoporosis (> -2.5 SD from healthy, young subjects of the same gender).¹⁰ Spinal QCT yields lower T-scores than central DXA and would overdiagnose osteoporosis with regard to WHO criteria. In particular, this applies to patients aged 50 to 65, for whom no prospective fracture data exist as a function of T-scores.

To clarify which of these techniques can predict mechanical competence of the skeleton (as a surrogate of fracture risk) most accurately, we recently undertook a large experimental study in which we related these techniques to mechanical failure loads of the hip, spine and distal radius of elderly subjects, as obtained from biomechanical testing experiments in cadavers.

Clinical Studies on Bone Densitometry and Fracture Risk

A large number of case-control studies have tested densitometric methods. However, this review will focus primarily on prospective clinical and experimental investigations.

Cummings and Black,¹¹ as well as Marshall and colleagues,¹² found that DXA of the hip was best suited to predict hip fracture, although other studies have indicated that other sites also may be of predictive value.^{13,14} In contrast, these studies have indicated no clear preference for site-specific measurements for spinal and radial fractures.

Studies found the predictive value of bone densitometry to be equal to or better than that of blood pressure measurements for predicting stroke. Huang and colleagues¹⁵ showed that measurements obtained 10 years prior to fracture were almost as predictive as those obtained within 3 years of fracture, suggesting that bone densitometry has long-term predictive ability.

Only one prospective study has been conducted with QCT.¹⁶ This one suggested that QCT allows for better prediction of spinal fractures than DXA.

However, case-control studies involving QCT versus DXA have been inconsistent. There are no prospective studies published on pQCT. Case-control studies that have tested the ability of pQCT to predict radial and spinal fractures (in relation to radial and spinal DXA) have been contradictory. However, one study did find that pQCT of the distal radius was clearly inferior to femoral DXA in differentiating subjects with and without hip fracture.¹⁷

Several large prospective trials have evaluated the ability of calcaneal QUS to predict fracture in relation to DXA. One of these studies concluded that QUS was as predictive as DXA and added significant independent information to site-specific bone mass.¹⁸ However, another study found calcaneal QUS to be slightly inferior to femoral DXA and determined that it did not contribute additional information when corrected for calcaneal DXA.¹³ A more recent prospective report¹⁹ suggested that QUS was effective for predicting distal radius fractures but not hip fractures. Small prospective studies with phalangeal QUS²⁰ and radial QUS²¹ have indicated that these techniques may be useful in predicting non-spinal fractures. Tibial QUS, in contrast, has been found to be ineffective in discriminating between patients with and without fractures.¹⁷

No prospective clinical trials have been performed with MRI. However, case-control studies have indicated that measurement of microstructural properties with high-resolution MRI may add significant information on fracture risk in combination with DXA.^{22,23}

Experimental Studies on Bone Densitometry and Bone Strength

Experimental (biomechanical) studies are ideally suited for determining which fraction of a skeletal structure's mechanical competence (failure strength) can be predicted with a certain method at a given skeletal site. Results can be obtained at much lower costs and at much shorter intervals than

with prospective clinical trials. Therefore, experimental studies play an important role in the pre-clinical evaluation of new techniques. Specific advantages and disadvantages of different methods can be determined directly and efficiently, but both the densitometric measurements and the mechanical tests should simulate realistic clinical conditions.

Recent studies have shown that the densitometric, geometric and mechanical properties of bones are highly variable throughout the skeleton in elderly individuals.^{24,25} Site-specific measurements allow one to predict between 50% and 65% of the variability in mechanical failure loads at all relevant sites, including the hip, spine and radius. Site-specific densitometry was shown to perform significantly better than non-site-specific analysis (Fig. 2), which only predicted between 20% and 35% of failure strength variability at remote sites. Lateral scanning at the lumbar spine did not involve an advantage over anterior/posterior (AP) scans in predicting mechanical failure loads of the thoracic spine, the predominant site of clinical fracture.²⁶

When comparing lumbar QCT to DXA, a combination of cortical and trabecular density from QCT improved the ability to predict failure loads of lumbar vertebral bodies, but not thoracic vertebral bodies.²⁷ Structural information from high-resolution CT improved the prediction of bone strength only in combination with density.²⁸ At the same time, pQCT of the distal radius and of the lower limb, as well as QUS of the calcaneus, displayed significantly lower correlations with mechanical failure loads than site-specific DXA and QCT²⁷ (Fig. 2).

When examining mechanical competence of the distal radius, pQCT and DXA (at the radius) were found to be of similar value, whereas QCT, femoral and spinal DXA and calcaneal QUS displayed significantly lower correlations²⁹ (Fig. 2). No improvement in predictive value versus DXA was seen when using

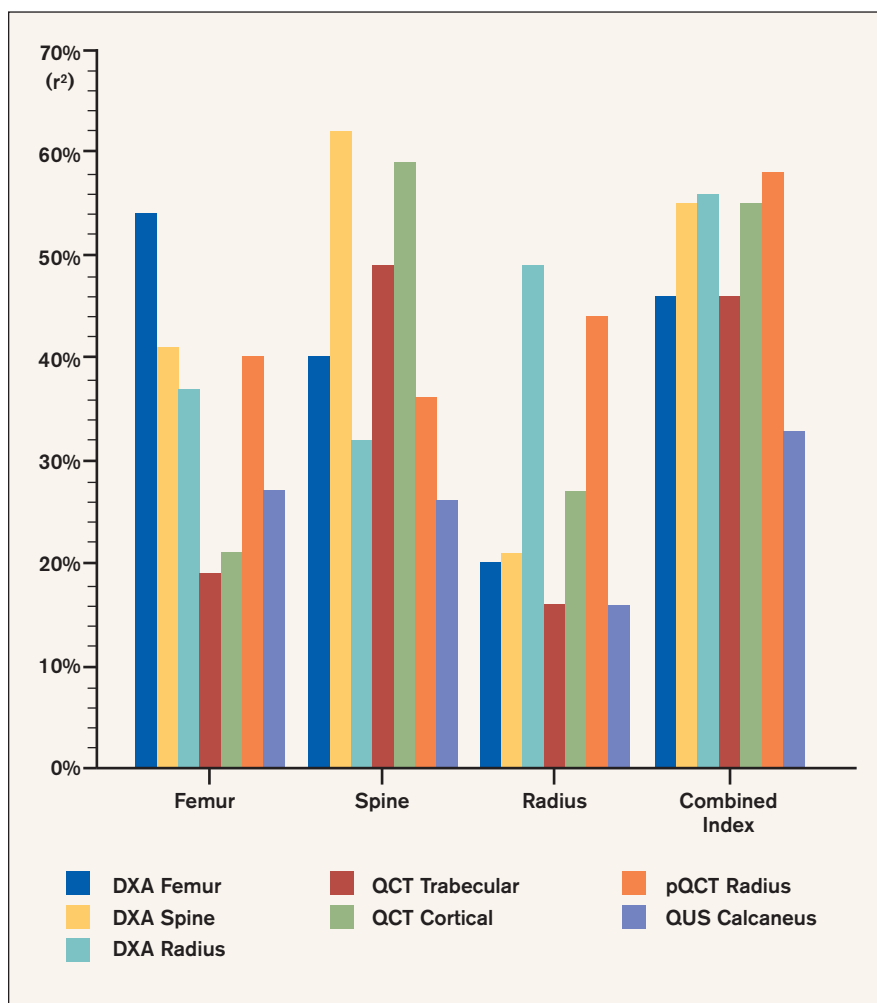


Figure 2. Graph showing coefficients of determination (r^2) of different densitometric methods for the prediction of bone failure loads of the hip (side impact configuration), spine (combined failure index of thoracic vertebra 6 and 10, and lumbar vertebra 3 in axial compression), distal radius (fall simulation) and a combined failure index at the hip, spine, at radius.³¹

high-resolution pQCT alone to determine microstructural parameters of trabecular bone. However, the predictive value was substantially improved ($r^2 = 75\%$ vs. 50%) when constructing microstructural finite element models from these scans and when simulating failure in a sophisticated computer analysis³⁰ (Fig. 3).

The best prediction of femoral failure loads was achieved by femoral DXA (Fig. 2), whereas spinal QCT, pQCT of the distal radius and calcaneal QUS displayed substantially lower correlation coefficients.³¹ A pQCT of the lower extremities (tibia and distal femur) achieved no advantage over pQCT at the distal radius.²⁴ Therefore, no improvement in femoral fracture prediction is to be expected from measure-

ments taken at the lower limb.

When relating densitometric techniques to a combined index of femoral, spinal and radial failure, most techniques performed equally well, with the exception of calcaneal QUS which showed a lower correlation than the other methods (Fig. 2). Calcaneal QUS did not add significant information to site-specific DXA at any site.³¹ In other experimental studies, tibial QUS was found to be inferior to calcaneal QUS in predicting femoral strength.³² Phalangeal QUS was somewhat more predictive than calcaneal QUS when using specialized parameters from the ultrasound transmission signal.³³ Although phalangeal QUS displayed significantly lower correlations with failure loads than site-specific DXA, its predictive

ability of vertebral and femoral failure strength was similar to that of peripheral (non-site-specific) DXA (e.g., the distal radius).

Conclusions Based on Available Clinical and Experimental Data

It is evident from experimental studies that mechanical competence of bones varies dramatically throughout the human skeleton, with some sites being much weaker and others much stronger. Generally, this will be in accordance with the given mechanical function (i.e., conditions) of the bone at a specific site. However, it is important to remember the differences between the sites. Also notable is that the way mechanical competence is acquired and/or lost over time at those sites can vary substantially in individual patients. For this reason, bone-densitometry at the site of interest is vastly superior to densitometry scans at remote skeletal sites (Fig. 2).

Clinical data suggest that femoral fractures are best predicted from femoral DXA, and results of experimental studies support the notion that site-specific measurements also are significantly superior at the spine and radius. So far, there has been no consistent evidence that QCT is superior to spinal DXA in predicting vertebral fracture or that pQCT is superior to radial DXA in the prediction of radial fracture. Although some clinical studies have indicated that calcaneal QUS is equivalent to femoral DXA for predicting hip fracture, the results of experimental studies are in obvious disagreement with these observations. Studies have shown tibial QUS to be inferior to calcaneal QUS, but there are some indications that phalangeal or radial QUS may perform similar to or better than calcaneal QUS. Larger prospective studies will be necessary to show whether these techniques are ultimately useful in the diagnosis of osteoporosis and whether evaluation from trabecular bone structure with high-resolution QCT, pQCT or MRI adds

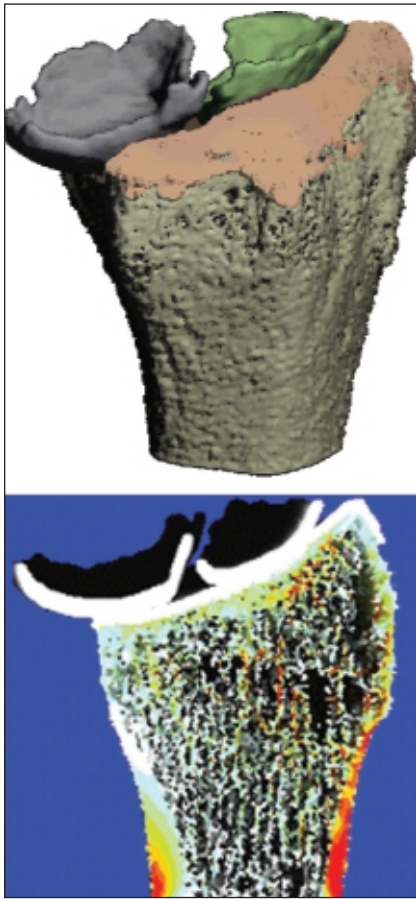


Figure 3. Computer simulation (microstructural finite element model) of failure in the distal radius.²⁶ (used with permission)

predictive value to measurement of bone mass or areal bone density (BMD) by DXA alone.

One general problem with peripheral techniques, however, is that the heterogeneity of bone properties restricts the predictive ability for more central skeletal sites, such as the femur and spine, independent of the level of technical sophistication. For the time being, therefore, site-specific DXA remains a “gold standard” for predicting mechanical strength and fracture risk in osteoporosis. Because the site of future fracture is unknown, all relevant clinical sites must be screened to optimally predict individual patients’ fracture risk. If only one measurement is performed, it is important to emphasize that this may limit prediction for other skeletal sites. Because hip fractures involve deleterious consequences for both patients and society, the proximal femur may be the

site of preference if only one site can be measured.

Some of the discrepancies between techniques could be eliminated with greater uniformity in WHO classification and fracture prediction via creation of a standardized and consistent reference population database.³⁴⁻³⁶ In particular, because a large overlap still exists in bone densitometry results for patients with and without fractures,¹² further efforts are needed to improve the non-invasive prediction of bone strength and fracture risk. These efforts should allow for more accurate identification of patients likely to benefit from antiosteoporotic treatments than currently is possible. ■

Eva-Maria Lochmüller, MD, is with the Department of Gynecology, Ludwig-Maximilians-University in Munich, Germany. Ralph Müller, PhD, is with the Institute for Biomedical Engineering, at the University of Zürich, Switzerland, and Felix Eckstein, MD, is with the Musculoskeletal Research Group, Institute of Anatomy, Ludwig-Maximilians-University, Munich.

References

- Melton LJ. How many women have osteoporosis now? *J Bone Miner Res* 1995;10: 175-77.
- Schürch MA, Rizzoli R, Mermillod B, et al. A prospective study on socioeconomic aspects of fracture of the proximal femur. *J Bone Miner Res* 1996;11:1935-42.
- Ray NF, Chan JK, Thamer M, Melton LJ. Medical expenditures for the treatment of osteoporotic fractures in the United States in 1995: report from the National Osteoporosis Foundation. *J. Bone Miner Res* 1997;12:24-35.
- The Writing Group for the Women’s Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women. *JAMA* 2002;288:321-33.
- Randell KM, Honkanen RJ, Kroger H, Saarikoski S. Does hormone-replacement therapy prevent fractures in early postmenopausal women? *J Bone Miner Res* 2002;17:528-33.
- Cauley JA, Black DM, Barrett Connor E, et al. Effects of hormone replacement therapy on clinical fractures and height loss: The Heart and Estrogen/Progestin Replacement Study (HERS). *Am J Med* 2001;110:442-50.
- Hulley S, Furberg C, Barrett Connor E, et al. Noncardiovascular disease outcomes during 6.8 years of hormone therapy: Heart and Estrogen/progestin Replacement Study follow-up (HERS II). *JAMA* 2002;288: 58-66.
- Delmas PD. Treatment of postmenopausal osteoporosis. *Lancet* 2002;359:2018-26.
- World Health Organization. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. WHO technical reports series 843;1994. Geneva, Switzerland.
- Faulkner KG, von Stetten E, Miller P. Discordance in patient classification using T-scores. *J Clin Densitom* 1999;2:343-50.
- Cummings SR, Black DM. Bone mass measurements and risk of fracture in Caucasian women: a review of findings from prospective studies. *Am J Med* 1995;98:24s-28s.
- Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *BMJ* 1996;312:1254-59.
- Bauer DC, Gluer CC, Cauley JA, et al. Broadband ultrasound attenuation predicts fractures strongly and independently of densitometry in older women. A prospective study. Study of Osteoporotic Fractures Research Group. *Arch Intern Med* 1997;157:629-34.
- Miller PD, Siris ES, Barrett-Connor E, et al. Prediction of fracture risk in postmenopausal women with peripheral bone densitometry: evidence from the National Osteoporosis Risk Assessment. *J Bone Miner Res* 2002;17:2222.
- Huang C, Ross PD, Wasnich RD. Short-term and long-term fracture prediction by bone mass measurements: a prospective study. *J Bone Miner Res* 1998;13:107-13.
- Ross PD, Genant HK, Davis JW, et al. Predicting vertebral fracture incidence from prevalent fractures and bone density among non-black, osteoporotic women. *Osteoporos Int* 1993;3:120-26.
- Augat P, Fan B, Lane NE, et al. Assessment of bone mineral at appendicular sites in females with fractures of the proximal femur. *Bone* 1998;22:395-402.
- Hans D, Darquent Molina P, Schott AM, et al. Ultrasonographic heel measurements to predict hip fracture in elderly women: the EPIDOS prospective study. *Lancet* 1996;348:511-14.
- Bauer L, Palermo L, Black DM, et al. A prospective study of dry calcaneal quantitative ultrasound and fracture risk in older women: the study of osteoporotic fracture. *J Bone Miner Res* 2002;16:166.
- Mele R, Masci G, Ventura V, et al. Three-year longitudinal study with quantitative ultrasound at the hand phalanx in a female population. *Osteoporos Int* 1997;7:550-57.
- Gnudi S, Ripamonti C, Malavolta N. Quantitative ultrasound and bone densitometry to evaluate the risk of non-spine fractures: a prospective study. *Osteoporos Int* 2000;11:518-23.
- Wehrli FW, Gomberg BR, Saha PK, et al. Digital topological analysis of in vivo magnetic resonance microimages of trabecular bone reveals structural implications of osteoporosis. *J Bone Miner Res* 2001;16:1520-31.
- Majumdar S, Link TM, Augat P, et al. Trabecular bone architecture in the distal radius using magnetic resonance imaging in subjects with fractures of the proximal femur. *Magnetic Resonance Science Center and Osteoporosis and Arthritis Research Group. Osteoporos Int* 1999;10:231-39.
- Lochmuller EM, Groll O, Kuhn V, Eckstein F. Mechanical strength of the proximal femur as predicted from geometric and densitometric bone properties at the lower limb versus the distal radius. *Bone* 2002;30:207-16.
- Eckstein F, Lochmuller EM, Lill CA, et al. Bone strength at clinically relevant sites displays substantial heterogeneity and is best predicted from site-specific bone densitometry. *J Bone Miner Res* 2002;17:162-71.
- Burklein D, Lochmuller E, Kuhn V, et al. Correlation of thoracic and lumbar vertebral failure loads with in situ vs. ex situ dual energy x-ray absorptiometry. *J Biomech* 2001;34:579-87.
- Lochmuller EM, Burklein D, Kuhn V, et al. Mechanical strength of the thoracolumbar spine in the elderly: predic-

tion from in situ dual-energy X-ray absorptiometry, quantitative computed tomography (QCT), upper and lower limb peripheral QCT, and quantitative ultrasound. *Bone* 2002;31:77-84.

28. Link TM, Majumdar S, Lin JC, et al. Assessment of trabecular structure using high resolution CT images and texture analysis. *J Comput Assist Tomogr* 1998;22:15-24.

29. Lochmuller EM, Lill C, Kuhn V, et al. Radius bone strength in bending, compression, and falling and its correlation with clinical densitometry at multiple sites. *J Bone Miner Res* 2002;17:1629-38.

30. Pistoia W, van Rietbergen B, Lochmuller EM, et al. Estimation of distal radius failure load with micro-finite element analysis models based on three-dimensional peripheral quantitative computed tomography images. *Bone* 2002;30:842-48.

31. Lochmuller EM, Muller R, Kuhn V, et al. Can novel clinical densitometric techniques replace or improve DXA in predicting bone strength in osteoporosis at the hip and other skeletal sites? *J Bone Miner Res* (in press) 2002.

32. Bouxsein ML, Coan BS, Lee SC. Prediction of the strength of the elderly proximal femur by bone mineral density and quantitative ultrasound measurements of the heel and tibia. *Bone* 1999;25:49-54.

33. Barkmann R, Lochmuller EM, Eckstein F, et al. Associations between Quantitative Ultrasound of finger phalanges and failure loads of spine and femur in vitro. *Calcif Tissue Int* 2002;70:281.

34. Greenspan SL, Bouxsein ML, Melton ME, et al. Precision and discriminatory ability of calcaneal bone assessment technologies. *J Bone Miner Res* 1997;12:1303-13.

35. Faulkner KG, Roberts LA, McClung MR. Discrepancies in normative data between Lunar and Hologic DXA systems. *Osteoporosis Int* 1996;6:432-36.

36. Lu Y, Genant HK, Shepherd J, et al. Classification of osteoporosis based on bone mineral densities. *J Bone Miner Res* 2001;16:901-10.

News & Commentary (continued)

trial data documenting that estrogen replacement with either estradiol or conjugated equine estrogens with or without concomitant progestin treatment offers no cardiovascular benefits to those women who already have coronary artery disease. The non-significant trends toward lower cardiovascular and total mortality should be viewed with great caution due to the likely play of chance when considering so few events. Women with established heart disease should be encouraged to use proven forms of therapy for secondary prevention, including aspirin, beta blockers and lipid-lowering therapies when not contraindicated.

David M. Herrington, MD, MHS
Professor of Internal Medicine/Cardiology
Wake Forest University School of
Medicine
Winston-Salem, NC

Gabapentin Shows Promise as Treatment for Hot Flashes

Loprinzi CL, Barton DL, Sloan JA, et al. Pilot evaluation of gabapentin for treating hot flashes. *Mayo Clin Proc* 2002;77:1159-1163.

Gabapentin, an anticonvulsant drug marketed as Neurontin, significantly reduced hot flashes in a small, prospective, single-arm pilot trial. Investigators enrolled 24 women with bothersome hot flashes, defined as at least 14 occurrences per week severe enough to require treatment. All women received 4 weeks of therapy with gabapentin, dosed from 300 to 900 mg/day. Of the 16 women who completed the trial, gabapentin treatment reduced the hot flash frequency by a mean of 66% and reduced the hot flash score (frequency times average severity) by a mean of 70%. Therapy appeared to be well tol-

erated; and at the end of the study period, 88% of the women chose to continue therapy.

Level II-2 evidence

Comment. This uncontrolled trial of women evaluated the effect of gabapentin, a commonly prescribed anticonvulsant and treatment for neuropathic pain, on hot flashes. It is somewhat difficult to evaluate the effectiveness of therapies for hot flashes when the study does not have a placebo arm and is of relatively short duration (4 weeks). Placebo treatments are known to improve hot flashes, but their effectiveness decreases over time. By 12 weeks, the placebo effect begins to reverse. Thus, only limited conclusions can be drawn from this report, and the authors have appropriately designated it a pilot study. That said, encouraging aspects of the study are the large magnitude of reduction in both frequency and severity of hot flashes, and the relative tolerability of the drug. Clinicians might want to consider this low-toxicity, alternative treatment for hot flashes while we await randomized clinical trial data.

Description of the levels of evidence

Level I: Properly randomized, controlled trial.

Level II-1: Well-designed controlled trial but without randomization.

Level II-2: Well-designed cohort or case-control analytic study, preferably from more than one center or research group.