

1 useful.

2 DR. WASNICH: That is correct. I don't think that
3 is a terribly difficult thing. It is already being done and
4 it's available.

5 DR. GARRA: I have one quick question myself. I
6 saw these curves, and the question is, is everyone in
7 agreement that these curves are reliable based on the data
8 that they were generated from is reliable, or does an
9 additional lengthy study need to be performed to verify the
10 shape of these curves and their slopes?

11 DR. WASNICH: I think the consistency between
12 these 21 different studies in all these locations is rather
13 remarkable and that we don't need more studies. We can
14 always perfect the paradigm, but, in fact, 10 years ago,
15 when the WHO came up with the guidelines, these data were
16 not available or were not considered.

17 It is 10 years later. I think we have a wealth of
18 prospective data with absolute fracture rates, and they are
19 even more consistent than I thought before I started to
20 prepare for this, and any error involved in this has to be
21 compared to the data we know, and it may not be perfect, and
22 undoubtedly there will always be some error involved in
23 absolute fracture, but it is a lot less than the somewhat
24 artificial system that we use right now.

25 DR. FAULKNER: Dr. Garra, just to comment on that,

1 since you asked, the thing I did notice is that we had a lot
2 of heel data, and the bone size effects at the heel are I
3 think not as maybe significant defects of geometry,
4 particularly at the hip, not hip axis length, but things
5 like just the size and shape of the femoral neck, those
6 things do not get picked up by DXA, and that may be very
7 important for differences between men and women in different
8 ethnic databases.

9 DR. WASNICH: But as long as whatever measurement
10 the machine spits out, as long as that curve has been
11 established, it doesn't matter. It has already taken that
12 into account, it has related it to certain absolute
13 fracture.

14 DR. FAULKNER: But it may not be the same, most of
15 the similarities with gender that we saw, a lot of that was
16 heel data. In fact, the hip data that you showed, didn't
17 show quite the same--these are a little bit more noise in
18 the data, and I think there may be some bone size effects
19 that could really be important.

20 DR. WASNICH: The reason for that is the intervals
21 reported by the authors were different, so I wouldn't over-
22 interpret that little noisy data.

23 DR. DESTOUET: I am confused, and I have a quick
24 question.

25 When one mentions absolute fracture risk, are the

1 other risk factors taken into account, high bone turnover,
2 that has already been taken into account to generate the
3 curves that you used as your reference data, or do you need
4 other bits of information to determine absolute fracture
5 risk?

6 DR. WASNICH: What I tried to show was that if you
7 only have bone density, you can convert that to absolute
8 risk. If you have additional data, by going to absolute
9 risk, you are able to use all the data you have. With bone
10 density alone, you can't really take into account the fact
11 that the patient has high bone turnover, because the
12 classification system doesn't make that feasible.

13 But that curve--maybe I should back up and show
14 that again--

15 [Slide.]

16 This, I really think is the crux of the issue.
17 What we are trying to do is interpret absolute fracture risk
18 based on what information was actually available, and if you
19 use bone density alone, you will get one answer. If you
20 throw in bone turnover, at least in our data shows an
21 independent relationship to fracture risk, then, you have
22 just perfected your prediction.

23 If you put in patient history, which I think is
24 vital, in which T-scores do not really do effectively, that
25 is probably the single largest risk factor we are missing.

1 DR. DESTOUET: May I have a follow-up question?
2 So, my question to you then, the clinician sitting at the
3 desk making a decision how to treat this patient, cannot
4 rely just on looking at the curve and determining absolute
5 fracture risk, you have to take into account, then, other
6 bits of data?

7 As a follow-up, then, how easily is that measured,
8 and what kind of additional costs are we talking about in
9 making that determination?

10 DR. WASNICH: This is really just software that
11 takes into account the fracture coefficient from each of
12 these, so you do need software to do this. Obviously, this
13 is for one age.

14 For age 80 it will look different, for age 85,
15 different, so it gets complicated, but if you feed in the
16 right data to the software, it will spit out the absolute
17 fracture rate. It is actually a complicated process, but
18 the software can make this very simple.

19 DR. GARRA: Judy, my interpretation of that is
20 that the bone density machine is going to spit out an
21 absolute risk based on bone density alone, and then you
22 would have to incorporate additional software to incorporate
23 the other risk factors unless the vendor has supplied that
24 software, but it would be complex.

25 DR. DESTOUET: What is the measurement for high

1 bone turnover, is that a urinalysis or a blood test?

2 DR. WASNICH: In this case, it was a blood test,
3 but there are also urinalyses that are potentially very
4 useful, but currently they are not being used very well
5 because it is not feasible to incorporate that information
6 and make it understandable.

7 DR. DESTOUET: So, then, you really do have an
8 additional test that the patient is subjected to, to get
9 that?

10 DR. WASNICH: That is not going to be every
11 patient, I think. Clearly, you always should take a patient
12 history, and who should have the additional measurements, I
13 think is not the issue I have tried to address, but it may
14 be that other factors will instigate those measurements
15 based on the initial level of risk.

16 DR. GARRA: Dr. Genant, did you have another
17 question or comment?

18 DR. GENANT: Yes. Dick, with regard to your
19 demonstration of the relatively comparable relationship
20 between fracture and BMD independent of race or gender, you
21 did indicate that you had corrected for other covariance.

22 I assume that you probably have corrected for body
23 size, weight, or height in some fashion there in order for
24 those relationships to be relatively constant.

25 DR. WASNICH: In those studies where those were

1 found to be significant independent variables, yes.

2 DR. GENANT: You have to keep in mind that
3 obviously, the measurements that we obtained typically are
4 not body size corrected, and perhaps they should be, but
5 they would have to be in order for that relationship that
6 you have shown to exist.

7 DR. WASNICH: Even if you have a measurement that
8 is not corrected, it is now reported as a bone mineral
9 density, and that is where these data came from, I would
10 suggest that as long as you know what that relationship is,
11 whether it be BMC or BMD, as long as you know the shape of
12 the curve, you can interpret it correctly.

13 DR. GARRA: Any other questions?

14 Thank you very much.

15 We are going to move on to the next speaker who is
16 Dr. Charles Turner, who is going to be speaking to us on the
17 physical bases for the noninvasive assessment of bone
18 strength.

19 Dr. Turner.

20 **The Physical Bases for the Noninvasive**
21 **Assessment of Bone Strength**

22 DR. TURNER: I was asked to speak on a topic that
23 is a little bit more toward the basic science on really what
24 the scientific basis is for the measurements that we are
25 discussing here.

1 [Slide.]

2 As an introduction, I would like to start with a
3 slide that shows really, at the bottom, what we are
4 interested in, and that is increased fracture risk, and that
5 is what we refer to, what osteoporosis really is, and there
6 is really two ways that you can get here, and this has been
7 emphasized previously, that there is more to this than just
8 simply bone mass.

9 You could possibly have an increase in trauma or a
10 change in the skeletal fragility. Now, I will give an
11 example of how lifestyle effects could affect fracture risk.
12 There is a study done in Japan, published recently, that
13 showed that Japanese women sleeping in western style beds
14 had a higher fracture incidence than Japanese women sleeping
15 in traditional Japanese households on futons.

16 You can interpret that how you like. There is
17 certainly differences in how Japanese people live, and it
18 does affect their fracture risk.

19 In the words of Dr. Wasnich, though, this is not
20 the issue for today. The issue today is really how do bone
21 measurements predict bone fragility, and then how does bone
22 fragility predict fracture risk. What I would like to do is
23 cover this issue, how can you take measurements of bone mass
24 or ultrasound and predict bone fragility or bone strength.

25 [Slide.]

1 I should tell you my background is more from an
2 engineering standpoint and actual bone strength
3 measurements. Now, this is a typical bone strength
4 measurement. If you were to take a specimen and apply
5 force, the specimen would displace and eventually it would
6 break, and there is a number of things that can be measured
7 from this type of measurement, like stiffness, the strength,
8 the displacement of brittleness of the material or the
9 energy absorbed. All of these are very important.

10 The problem is that you have to break a specimen
11 to learn all this information, and what we would like to do
12 is provide measurements that can be done noninvasively that
13 give this similar type information.

14 [Slide.]

15 The measurement that has probably become the gold
16 standard is bone mineral density.

17 [Slide.]

18 This is a measurement. It takes the bone mineral
19 content determined by x-ray, and normalizes it, and the
20 reason you want to normalize it, if you only looked at bone
21 mineral content, the smaller people would always have less
22 bone mineral, and they would always be diagnosed as
23 osteoporosis, so we need some kind of normalization, and
24 what is used here is kind of an imperfect solution. It is a
25 projected area of, say, the vertebral body or the hip,

1 whatever it is measuring.

2 Now, since this is an area, not a volume, this
3 isn't a true density measurements. It is grams per square
4 centimeter as opposed to grams per cubic centimeter. So, it
5 is somewhat imperfect, but it does, in fact, work.

6 [Slide.]

7 This is an example. In my work, I work with
8 specimens from a number of animals, and this is just to show
9 you a comparison across species. Bone mineral density can
10 predict vertebral strength in rats, in monkeys, in humans at
11 about the same degree.

12 You notice our values are 0.83 in two instances
13 and 0.85 in the other. These are very good correlations for
14 biomechanical data, and it shows that bone mineral density
15 does provide the basic information that we need, similar for
16 strength of the hip, femoral neck strength, again, rats,
17 monkeys, and humans.

18 One interesting thing you see here, though, is
19 that when we are looking at femoral neck strength, but
20 vertebral BMD in the case of the primate study I show here,
21 the correlation is less than if we were to look at vertebral
22 strength from vertebral BMD, and this is very common in all
23 biomechanical studies, than you can take a measurement from
24 any site in the skeleton, and it will correlate with
25 strength at other sites of the skeleton, but the measurement

1 made at the site of interest will always give you the best
2 correlation, and that may be germane to our discussion this
3 afternoon.

4 [Slide.]

5 We did a little exercise some years ago. We
6 attempted to provide a better assessment of fracture risk by
7 using engineering analysis with the idea that we were
8 smarter than these x-ray densitometers, and that a simple
9 measure of bone mass couldn't be nearly as useful as a
10 sophisticated analysis of the strength in hip.

11 This was done for hip fracture, and we created
12 this simulation where you calculate a force on a trochanter
13 from a fall to the side, and work out all the engineering
14 equations, and this requires that you add in a lot of
15 different factors. You have to add in the weight because
16 the more heavy you are, the harder you fall, the height
17 being because you would fall further, the bending moment on
18 the hip, what is called the cross-sectional moment of
19 inertia, which tells you how the mass is distributed in the
20 hip, the cross-sectional area of the hip, and even the
21 effect of aging.

22 In this exercise, we went in with great
23 enthusiasm, and after developing tools to do all of these
24 measurements, and we did a case-controlled study of a
25 population, fairly small population where there were a

1 number of individuals with hip fractures, and then we saw
2 how our analysis did. To our dismay, it didn't do very
3 well.

4 This is what I called a fracture index, which is
5 what came out of all those calculations that I showed you.
6 Here is an odds ratio, which shows the odds of fracture with
7 1 standard deviation drop in the value.

8 We see that our measurement had an odds ratio of
9 1.66. This was barely statistically significant, but this
10 BMD measure, which we were trying to improve upon, did
11 better than twice as well, and this was fairly consistent.
12 Each time we tried this type of analysis, we always came up
13 with the same result. BMD always did better.

14 It does better than BMC, this is well known, but
15 due to the normalizing effect, and this was eye-opening.

16 [Slide.]

17 Our conclusion from this is, yes, you can do all
18 these engineering analyses, but you add in error from each
19 measurement that you take, and there are a large number of
20 them as you can see here, and by adding in all this error,
21 you create a very lousy diagnostic tool, and it turns out
22 this BMD measurement has been around for 20 or 30 years, has
23 proven over and over to be the best way of predicting bone
24 strength clinically.

25 [Slide.]

1 I would also like to spend a few minutes talking
2 about the ultrasound, and there are a number of different
3 flavors of ultrasound. I am going to concentrate on the
4 ultrasound through the heel, which is, in my mind, the gold
5 standard of the sonometers that are available.

6 First, there are several different measurements
7 that these machines give, and one is the acoustic velocity
8 or speed of sound, SOS, as it is called by the
9 manufacturers. That is just how fast the wave goes through
10 the heel.

11 We have broad-band ultrasonic attenuation. This
12 is based on a finding that the amount of acoustic energy
13 absorbed increases as the frequency of the acoustic wave
14 increases, and you can measure a slope to this curve, and it
15 actually means something. That has been called broad-band
16 acoustic attenuation.

17 Finally, there are measurements that are dry
18 parameters, that have become the primary output of these
19 sonometers, either quantitative ultrasonic index or
20 stiffness index. What they are, are some weighted average
21 of the two measurements, and the values that go into this
22 averaging are manufacturer-specific, so this number isn't
23 necessarily the same for every machine or every
24 manufacturer.

25 [Slide.]

1 Now, the physical basis for ultrasound providing
2 us any information at all is this equation at the top of the
3 slide, that shows the acoustic velocity is related to
4 Young's modulus or stiffness of the material divided by the
5 density to the square root. This is a well-known equation
6 in physics, and it does give us a basis for why ultrasound
7 should work.

8 What we can see here, I have plotted a number of
9 different materials and how they fit and how their
10 ultrasonic velocity and density relate, and what we see is
11 there is not--in most materials if we look at iron and steel
12 and concrete--there is not really any relationship between
13 acoustic velocity and density, but when you look down here
14 at bone specimens, there is a nice relationship.

15 It just turns out that this material is
16 appropriate material for analyzing using ultrasonic methods,
17 and, in fact, you see very good correlations between bone
18 strength and acoustic velocity. The same holds true for
19 bone strength and the BUA measurement.

20 [Slide.]

21 What was hoped in the beginning when we started
22 looking at ultrasonic methods is that these measurements
23 would give us information that is different than the
24 information provided by bone density, and there is evidence
25 of this. In any joint, you see trabecular organization and

1 trabecular trajectories where you can see a very clear
2 organization alignment structure and architecture in the
3 bone.

4 [Slide.]

5 Now, this architecture can't necessarily be
6 quantitated using bone densitometry, but what has been shown
7 over and over is that this structure is quantitated using
8 ultrasound, and this a cartoon depiction of the struts of
9 trabeculae aligned. This structure is called anisotropic
10 because it has a directional alignment.

11 In this specimen with this degree of alignment, we
12 would expect the stiffness along this axis to be about five
13 times the stiffness along this axis. In fact, the
14 ultrasound sonic velocity will mirror this by showing an
15 over twofold difference between the a axis and the b axis
16 where density will give us the same in any direction that it
17 is measured.

18 So, this fact shows, one, that ultrasound measures
19 something different than bone density even though the two
20 are correlated, and, two, that it may measure something, in
21 the structure of the bone, give us information that is
22 independent of density.

23 [Slide.]

24 There is some clinical evidence that this may be
25 true. What I am showing here is again a case-controlled

1 study for both spine and hip fracture that was done in
2 Indiana. I think these studies are consistent with other
3 studies that have been done, as well, and I don't think
4 there is anything controversial here.

5 But see that the BMD measurement in the spine and
6 the ultrasonic velocity, or SOS--this UTV is the same as
7 SOS--and the broad-band acoustic attenuation all give fairly
8 similar receiver-operator characteristic curves for
9 specificity and sensitivity. In other words, they predict
10 the event in a similar fashion.

11 [Slide.]

12 If we look at hip fracture, there is some evidence
13 that--well, the ultrasound certainly does about the same as
14 BMD--there is some evidence in this particular measure of
15 SOS that it actually did a little bit better, and, in fact,
16 this study showed that there was an independent prediction
17 of hip fracture with ultrasound, independent of bone
18 density, that is, so that it may give us some new and
19 different information, but it certainly does as good a job
20 as bone density in predicting fracture.

21 [Slide.]

22 Just to summarize, fracture risk is certainly
23 determined by lifestyle issues, as well as skeletal
24 fragility, but when we talk about skeletal fragility, BMD
25 has certainly become the gold standard for radiological

1 assessment of fracture risk. It certainly is least
2 expensive, easiest to apply, and it does work very well.

3 Ultrasound, while it gives information that
4 correlates very well with bone strength, and probably, even
5 better than BMD in some cases, and it probably provides us
6 with some structural information about the bone that is more
7 than what is given by bone density alone.

8 So, these are different techniques. They are
9 techniques that work. Each work well, and I guess what is
10 germane to our discussion as we go on is that since they are
11 different, they may need to be treated differently in the
12 way that we use them clinically, and I guess we will follow
13 up on that with the speakers this afternoon.

14 That concludes what I have to say.

15 DR. GARRA: Thank you very much.

16 Are there any questions? Yes.

17 DR. FAULKNER: To what degree that we are
18 measuring with the x-ray absorbed geometric techniques in
19 area density hamper us or influence the question at hand due
20 to differences in bone size that we see among the different
21 genders and races?

22 DR. TURNER: If we can presume that our bones are
23 developed in size to fit our stature, then, the appropriate
24 measurement to tell us about the integrity of bone would be
25 more of a true density measurement. This true density or

1 grams per cubic centimeter would be size independent, and
2 would tell us only about the bone, not about the stature, so
3 there is some complexity put into the current situation in
4 that we have a measurement that is partially normalized for
5 size. It is normalized by area, therefore, it is normalized
6 for the most part for size, but there is still some size
7 effect carried with the measurement.

8 DR. FAULKNER: But is bone bone, I guess is the
9 question. If you take a sample of male bone and a sample of
10 female bone--

11 DR. TURNER: The simple answer is yes bone is
12 bone.

13 DR. FAULKNER: If we measured true density, then,
14 bone would be bone across genders--

15 DR. TURNER: Presuming that there is no other
16 disease state that would change--

17 DR. FAULKNER: Okay, because there may be some
18 porosity due to other disease states.

19 DR. TURNER: Right. All other things being equal,
20 yes.

21 DR. FAULKNER: So, it may be necessary, then, to
22 maintain some--as long as we are measuring surrogates of
23 true density, we can expect there will be differences among
24 genders and races or can we?

25 DR. TURNER: Surrogates of true density?

1 DR. FAULKNER: Yes, we are not measuring, but we
2 are getting grams per square centimeters.

3 DR. TURNER: If we had a true surrogate, then,
4 yes, the bone densities may be different between ethnic
5 groups and genders, but that difference would mean
6 something, in other words, it would reflect directly the
7 differences in risk.

8 DR. GARRA: Any other questions?

9 DR. DESTOUET: Yes. Are you saying, then, that we
10 really do need independent measures of different ethnic
11 groups, we can't just use Caucasian females as the standard
12 across the board?

13 DR. TURNER: No, no, I am not saying that. That
14 really wasn't my place to address that question. I guess
15 what I said is that if the bone mineral density works as a
16 true surrogate, then, we don't, in other words, assuming
17 that lifestyle issues are the same, assuming that it
18 reflects the integrity of the structure, it should then
19 directly be associated with the risk, and if somebody has a
20 higher bone density, they therefore should have a lower risk
21 regardless of ethnic group or gender.

22 DR. GARRA: Any other questions?

23 Thank you very much.

24 This is the time that we will break for lunch.
25 Before we leave for lunch, I would like to remind everyone

1 that open committee deliberations will begin again at 1:00
2 p.m., and that between 12:15 and 12:45, the panel here will
3 meet in closed session, and that panel session is not open
4 to the public.

5 I request the panel members to be back by 12:15,
6 so we can begin the closed session. Thank you. We will see
7 you all at 1:00.

8 [Whereupon, at 12:05 p.m., the panel proceedings
9 were recessed, to resume, in closed session, at 12:15 p.m.]

10

- - -

A F T E R N O O N P R O C E E D I N G S

[1 o'clock p.m.]

DR. GARRA: If everybody could be seated, I would like to call the meeting back to order and would like to remind public observers of the meeting that, while this portion of the meeting is open to public observation, public attendees may not participate unless specifically requested by the chairman.

We are going to continue on with the open committee discussion which we began in the morning. Our first speaker for this afternoon is Professor Harry Genant from University of California, San Francisco, Professor of Radiology, Medicine, Epidemiology, Orthopedic Surgery and many other specialties, I'm sure. It's quite impressive. I will have to see his calling card.

He is going to report on a comparison of common methods used for assessing bone strength, classification and misclassification.

DR. GENANT: I tried for pediatrics and OB-GYN but they wouldn't have me.

A Comparison of the Common Methods Used**for Assessing Bone Strength:****Classification and Misclassification**

DR. GENANT: Good afternoon.

[Slide.]

1 My charge this afternoon was to compare some of
2 the common BMD methods in the diagnosis and also estimation
3 of fracture risk as well as in relationship to issues of
4 classification and misclassification.

5 [Slide.]

6 Of the numerous methods of noninvasive assessment
7 of bone, the two that I will focus on in this context will
8 be DXA and SXA, with DXA representing the central
9 measurements and SXA the appendicular-peripheral
10 measurements.

11 [Slide.]

12 Some of the work that I am going to present is
13 based upon data from the study of osteoporotic fractures and
14 analyzed by our group in San Francisco by some of my
15 colleagues as listed here.

16 You will notice that I am indicating the we are
17 talking about, then, classification of individual patients.
18 This is going to be based upon the BMD measurement and we
19 are going to be considering the risk of hip and/or spine
20 fractures combined. There is, perhaps, some rationale for
21 focussing on hip and spine fractures as potentially the most
22 important of the osteoporotic fractures.

23 [Slide.]

24 The data will be drawn from the study of
25 osteoporotic fracture. Steve Cummings is the principle

1 investigator, as you all know, of this study. We evaluated,
2 in this case, around 5500 women, all age over 65. The
3 baseline BMD measurements of the hip and spine were by DXA
4 and of the radius and the calcaneus by SXA. And then
5 incident hip and/or spine fractures were assessed over a
6 five-year follow-up period with a total of about 450
7 fractures, half of which were hip fractures.

8 [Slide.]

9 The women were classified by two methods. This
10 gets to some of the aspects that we discussed this morning
11 with regard to the use of T-scores, WHO criteria. The first
12 criteria was the standard WHO criteria of less than -2.5 in
13 reference to the manufacturer normative data at age 30.
14 This is actually age 20 to 30. So that is kind of standard.

15 Then the other approach is somewhere between a T-
16 score and a Z-score because it is in reference to the SOF
17 population, itself, the 456 women aged 65. We looked at
18 this and it appeared as though there was sufficient
19 statistics with this number of patients for this to be a
20 reliable reference group.

21 We selected at T-score relative to this 65-year-
22 old group of -1, as you can see.

23 [Slide.]

24 In addition to standard statistics, we looked at
25 the percent agreement and the kappa statistics to evaluate

1 the consistency and classification based on these different
2 BMD measures and then we also used logistic regression
3 analyses to study the risk of incident hip and/or spine
4 fracture for the five-year follow up.

5 [Slide.]

6 Just to point out again, the differences in these
7 measurement approaches, the standard AP or PA spine,
8 measuring largely trabecular bone but also substantial
9 components of the posterior elements, as you can see here,
10 the pedicles and spinous processes, et cetera.

11 [Slide.]

12 In the hip, we actually had four regions that we
13 measured, the standard four regions, the neck, the Ward's
14 triangle--misnomered, I guess one would say--the
15 trochanteric area and total hip.

16 [Slide.]

17 The radius; two sites, the more proximal site
18 which is virtually 100 percent cortical bone and the more
19 distal site which is about 40 to 50 percent cortical bone,
20 the rest trabecular bone.

21 [Slide.]

22 And the calcaneus or weight-bearing site,
23 obviously, which is about 95 percent trabecular bone.

24 So we have those comparisons, central, peripheral,
25 cortical, trabecular.

1 [Slide.]

2 Here we are looking at the 5500 women from this
3 study and the mean age was 71. You can see the standard
4 deviation here. These are, then, the four regions in the
5 hip that were measured, their mean value, standard deviation
6 and the coefficients of variation for the entire population
7 and then distal, proximal, radius, calcaneus and spine.

8 [Slide.]

9 If we look at the results with regard to the
10 relative risk calculated from the logistic regressions based
11 upon simply a one standard deviation decrease relative to
12 the entire SOF population, the results are shown here. I
13 wanted to illustrate this slide because it gets back to a
14 point that I made this morning, at least, in reference to
15 Dick Wasnich's presentation, that in some prospective
16 studies with large data where you have a consistent
17 database, that the differential strength for predicting risk
18 of hip fracture, for example, here, by measuring at the hip
19 may be substantially higher than measuring at, for example,
20 in this case, spine and two sites at the radius with the
21 calcaneus in between.

22 Similarly, if we look at the incident spine
23 fractures, we see some differential although the
24 differential is not as great, but, in general, with the
25 three peripheral measurements here showing somewhat less

1 power than the spine and the hip.

2 [Slide.]

3 It is not surprising, then, that there are some
4 differences in the strength of fracture prediction, that the
5 information derived from these different sites is, in fact,
6 somewhat different. And that is born out by looking at the
7 correlations between these different anatomic sites. There
8 are no surprises here. Across this age range at 65 to about
9 80, we see that, within the hip, correlations run about 0.8
10 to 0.9 and then, as we go to disparate anatomic sites, the
11 correlations are on the order of about 0.5 to 0.6

12 This has some significant ramifications with
13 regard to classifying patients based upon specific
14 thresholds.

15 [Slide.]

16 So if we look at the percentage agreement, and
17 these are Christmas trees. I have shown similar slides of
18 this at a number of meetings so some of you are familiar
19 with these brightly colored Christmas-tree variants. But
20 here we are looking at the percentage agreement using the T-
21 score relative to the -2.5 manufacturers' young normal. So
22 this would sort of the standard WHO criteria.

23 When expressed at percent agreement, you will see
24 that the percentage agreement ranges from a low at some
25 disparate sites of about 30 to a high on the order of about

1 70 percent. That looks reasonable, but there is not a
2 highly strong percent agreement.

3 [Slide.]

4 In contrast to that, when you use the SOF
5 population, itself, in which you are then deriving the
6 reference population, so you are doing away with potential
7 differences in the reference population from different
8 manufacturers. You are also eliminating the impact of
9 variable age loss on the classification.

10 We see that, indeed, the percentage agreement now
11 rises to on the order of about 60 to 80 percent and
12 substantially better.

13 [Slide.]

14 But another way that one can look at this
15 agreement is by using the kappa statistics which takes into
16 account chance agreement. Here, if we look at the kappa
17 statistics for the standard WHO criteria, we see, in fact,
18 that we have kappa statistics ranging from about 0.2 to
19 about 0.4. Even amongst the hip sites, it is not very high.
20 These would be considered poor to modest.

21 If we look at the kappa scores for the other
22 method, then, of deriving our classification, SOF, 65, T-
23 score -1 relative to that, even here the kappa scores are
24 only in the modest range of about 0.3 up to about 0.6. So
25 this indicates that, despite the elimination of the

1 differences in age-related loss and the differences in
2 manufacturer normative data, while we see a substantial
3 improvement in agreement, there still is disparity.

4 So one does not eliminate that entirely by
5 deriving a large, young normal database across all
6 instruments. But it certainly does improve the situation.

7 [Slide.]

8 Let's look at the classification--again we are
9 back to the T-score of -2.5 to the manufacturers' young
10 normal, and we are looking at the prevalence of disease that
11 would be detected in this SOF population and then the
12 sensitivity and the specificity for incident hip and/or
13 spine fracture since we, in fact, have those patients that
14 had incident hip and/or spine fracture.

15 So the first thing that is impressive to look at,
16 if we look at prevalence in purple at the various anatomic
17 sites, we see that even amongst the hip sites, that the
18 prevalence ranges from about 30 on up to 80 percent. This
19 is, in part, due to the differential in rates of loss at,
20 for example, the Ward's triangle which is quite high during
21 this interval and from young normals to, say, 70-year-olds
22 and a slower loss at some of the other sites.

23 You can also see, within the hip site, itself,
24 that there are tradeoffs, then, between the sensitivity here
25 and the specificity but considerable non-homogeneity in the

1 number of patients identified and the sensitivity and
2 specificity.

3 Even within the radius, distal radius, proximal
4 radius, we see a marked difference in the prevalence of
5 disease, sensitivity and specificity, and the same for the
6 other sites. So this helps to illustrate some of the
7 magnitude of disparate results that one could get when we
8 are referencing young, normal data.

9 [Slide.]

10 There is a considerable improvement here if we
11 look at these results now for prevalence, sensitivity and
12 specificity, using the SOF population, itself, and -1 T-
13 score. Here we can see a relatively good balance across all
14 of the anatomic sites in the prevalence with a slightly
15 lower prevalence, for example, here at the spine and,
16 perhaps, at the trochanter. Specificity, slight
17 differences, and also a little bit of difference in
18 sensitivity but much more homogenous than the young normal
19 reference population.

20 [Slide.]

21 So that would indicate at least a greater degree
22 of comparability with regard to the populations identified.
23 It still doesn't eliminate the misclassification or altered
24 classification on an individual patient basis.

25 [Slide.]

1 Another way to look at this that, to some extent,
2 magnifies the differences in fracture prediction is with the
3 odds ratio, as shown here. This is a little bit of a
4 complicated slide. We are using, now, only the T-score -1
5 relative to SOF 65 recognizing that this gives, in general,
6 a more homogenous assessment.

7 If we look at the odds ratios, first off, of the
8 measurement at each of these individual sites, we again see
9 substantially higher odds ratios, on the order of about 3,
10 for the hip sites here in defining incident hip and/or spine
11 fractures, spine being relative high, distal-proximal radius
12 and calcaneus somewhat lower, calcaneus intermediate.

13 But this is another important point to recognize
14 here and that is let's look at the red box. This is a
15 further analysis where, if we identify those individuals who
16 have a low proximal radius, in this case 1400,
17 approximately, we are looking, then, at the added odds ratio
18 by, for example, performing a total hip measurement.

19 We see that even those with a low proximal radius,
20 one can enhance the discriminatory capability by this
21 additional measurement. Even if we began, for example, with
22 a low neck measurement, there is still some added
23 information here in green, although it is less than the
24 added information when one does it the other way--that is,
25 starting with a low proximal radius. There, the added

1 information from the central measurement is considerably
2 higher.

3 This simply points out that if you have
4 measurements at different sites that are giving somewhat
5 different information but that they all have some fracture-
6 prediction capability, that adding measurement may have some
7 utility, at least in selected cases.

8 [Slide.]

9 That same point is further pointed out in this
10 slide where we are looking at, on this axis, the number of
11 low BMD sites. We have got eight sites measured. This is
12 the percentage of patients with 0, 1, 2, 3 all the way up 8
13 low sites. The red represents the age-adjusted odds ratio
14 for the hip and/or spine fracture.

15 But one can see that, as you have increasing
16 numbers of low BMD sites that, in fact, the odds ratio does
17 go up. This is, perhaps, not surprising. I don't mean to
18 imply by this that one is going to make eight measurements
19 but simply that there may be algorithms that can be
20 developed whereby when a patient has a measurement that has
21 a marginal value that is near a threshold, one might benefit
22 from information from another measurement

23 [Slide.]

24 Or by simply looking, in this case, at the hip and
25 looking if one site is low, two sites are low, three sites,

1 four sites. The odds ratios go from, perhaps, about 2 up to
2 6 in this case.

3 So this is just to point out that even when we
4 have a measurement, for example, at the hip, to focus only
5 on the neck may not be as advisable as looking at all of
6 other parameters at the hip.

7 [Slide.]

8 What I would like to conclude from this is that
9 the correlation between these BMD measures is good but the
10 agreement in threshold-based classification is only modest
11 and particularly when referenced to young normals.

12 The threshold-based classifications depend heavily
13 on the reference data that are used. Standardization of
14 this reference data will, in fact, reduce but it will not
15 entirely eliminate the inconsistency between the
16 classifications.

17 [Slide.]

18 Finally, given a low BMD at one site, a low BMD at
19 another site may further increase the risk of hip and/or
20 spine fractures. Classification of an individual patient
21 based on one BMD measure--and I want to say here one BMD
22 measure that is near a threshold is not highly definitive.
23 So then, using a specific threshold on that basis may not be
24 reliable and potentially reporting multiple BMD, say a
25 second measure, may have some clinical relevance

1 particularly in those borderline cases.

2 Thank you.

3 DR. GARRA: Thank you very much.

4 Are there any questions for Dr. Genant? I have
5 one. Since I don't do this routinely in my clinical
6 practice, what are the constraints against doing multiple
7 sites? Obviously time, exposure, maybe. Can you just
8 enumerate those for us?

9 DR. GENANT: Sure. Time, money. I think those
10 are the major things. But keep in mind, for example, when
11 we do a hip measurement, we, in fact, have, like, three or
12 four measures. I am just suggesting that one can
13 potentially make use of all of that information rather than
14 just kind of discarding or disregarding the other measures.
15 Even at the radius, typically, we have at least two results
16 and some devices may give three results.

17 Perhaps integrating that information may help to
18 maximize our ability to predict fracture risk.

19 DR. MALCOLM: I was thinking the same issue. Do
20 we have any sense of what is going on "in the community"
21 that physicians are measuring more than one place, or they
22 take one measurement and that's it and they take that and
23 run with it rather than doing what you are suggesting? I am
24 a little concerned about that.

25 DR. GENANT: Again, I am not suggesting that all

1 patients need to have multiple measurements but what I am
2 suggesting is that if you do a measurement on a patient and
3 it is a somewhat borderline measure, and if the other
4 clinical parameters that you are using also are not
5 compelling one way or the other, then a second measure might
6 be useful.

7 I think that this would apply, if you believe that
8 fracture prediction for the most important hip fracture and,
9 perhaps, maybe second most important, vertebral fracture--
10 that is, perhaps, somewhat debatable, but if you believe
11 that those are particularly important, then at least some of
12 the data that I showed from SOF would suggest that the
13 central measurements do give stronger risk prediction.

14 So if one is working in the community with a
15 peripheral measuring device and end up with a result that is
16 quite high or quite low, I think that one can feel quite
17 comfortable in going forward with definitive treatment,
18 particularly along with clinical parameters.

19 But if one has a measure that is, perhaps,
20 somewhere in the middle, then it is a bit difficult, on that
21 single measurement, to have a high level of confidence that
22 this patient does not have a more significant risk or,
23 perhaps, a less significant risk if measured, for example,
24 at the hip.

25 DR. GARRA: Again, I am actually surprised that

1 people don't do more than one site because generally
2 multiple features do lead to more reliable estimates.

3 DR. GENANT: Let me say it is very common, I
4 think, for those centers that are doing central measures
5 with DXA to combine a hip and a spine measurement. The
6 patient is on the table. These are fast measurements now
7 and, typically, they will only be reimbursed for one. But,
8 in fact, many centers are measuring spine and hip.

9 DR. GARRA: Thank you.

10 DR. MCGOWAN: I might ask, Dr. Genant, it seems,
11 in the analysis of the risk factor cholesterol, for
12 cardiovascular disease, there is a two-tier system in place
13 with peripheral cholesterol measurements the first tier and
14 then, as you suggest, the follow up would be a complete
15 lipid analysis.

16 Is there anything thinking in people who study
17 bone densitometry that suggesting that an "iffy" peripheral
18 measurement should be followed by measurement of other sites
19 centrally.

20 DR. GENANT: Certainly, that is under discussion
21 and is, I think, a very reasonable approach to go. There is
22 another, perhaps, indication, for going, say, from a
23 peripheral measurement to a central measurement and that is
24 if, based upon the peripheral measurement, you have a fairly
25 definitive answer or quite a low value and you are going to

1 initiate treatment, I think that there is, in fact,
2 reimbursement for a second central measurement as a baseline
3 for following and monitoring the treatment.

4 DR. FAULKNER: Currently the Bone-Mass Measurement
5 Act would not permit you to scan at one site and then do a
6 confirming measurement at a second site and get reimbursed
7 for both of them. It has to be for the purposes of
8 monitoring.

9 I think that has confused people because there are
10 some scenarios. But, on the other side of the coin, one of
11 the educations that we need to do in the field is to help
12 people understand low bone density at any site predicts
13 fracture. You don't have to measure a hip to predict
14 fracture.

15 DR. GENANT: That's right.

16 DR. GARRA: Thank you very much.

17 We will move on to the next speaker. Dr. Anne
18 Looker is going to be discussing NHANES surgery results of
19 bone measurements on blacks and Hispanics.

20 **NHANES Surgery Results of Bone Measurements**
21 **on Blacks and Hispanics**

22 DR. LOOKER: Thank you very much to the panel for
23 inviting me today.

24 [Slide.]

25 In my brief time, what I am going to do is

1 describe for you the results of prevalences of osteoporosis
2 and osteopenia as defined by the WHO T-score approach that
3 are observed in the U.S. population when race and sex-
4 specific cutoff values are applied to femur bone-density
5 data that were collected in the third National Health and
6 Nutrition Examination Survey or NHANES III.

7 I am going to compare those results with the
8 results we have previously published for the U.S. population
9 that were generated using a single set of cutoffs derived
10 from young white women.

11 As we have already heard from Dr. Wasnich, you
12 will expect that these prevalences will differ. I am going
13 to show you just how much they differ. Although I am going
14 to focus on the results at the population level, I think
15 they can highlight for you some relevant issues to consider
16 in using race and sex-specific cutoffs for individual
17 patients as well.

18 [Slide.]

19 Before I start showing you results, though, I just
20 want to say a word or two about what NHANES is for those of
21 you who may never have heard of NHANES. NHANES III is the
22 most recently completed cross-sectional survey of a long
23 list of cross-sectional surveys that are done periodically
24 by the National Center for Health Statistics of the CDC to
25 assess the health and nutritional status of the non-

1 institutionalized civilian U.S. population.

2 You can see that actually NHANES III has been
3 over, now, for almost five years. But we just started in
4 the field with the next HNANES.

5 [Slide.]

6 One unique feature about NHANES has to do with the
7 sample design. I don't want to spend time here today
8 talking about how we draw this sample. What I want to point
9 out is that, in fact, it is a representative sample of the
10 U.S. population and this makes it unique relative to most of
11 the other studies that are done.

12 The other point to emphasize here today is that in
13 NHANES III, we oversampled individuals from the two largest
14 minority groups in the United States, those being African
15 Americans and Mexican Americans.

16 [Slide.]

17 In NHANES III, we also included measurements of
18 BMD or bone-mineral density in the proximal femur using DXA
19 or dual-energy X-ray absorptiometry. We performed these
20 measurements on all men and women ages 20 and over. So we
21 ended up with a very large dataset of about 14,600
22 individuals with usable femoral bone-mineral density at the
23 conclusion of NHANES III.

24 That is what I want to talk about today.

25 [Slide.]

1 Before I go on, though, I want to just briefly
2 show you some of the abbreviations I will be using so that
3 there is no confusion. I am going to be showing you
4 estimates for the three major race ethnic groups that we can
5 provide estimates for from NHANES III with abbreviations.
6 Those will be non-Hispanic, white, blacks and Mexican-
7 Americans. Of course, I think everybody recognizes those
8 particular abbreviations.

9 [Slide.]

10 Also, as I mentioned earlier, the definitions of
11 low bone density that I am going to be using are, in fact,
12 based on the WHO approach, the T-score approach, using, in
13 our case, a young reference group mean derived from the 20
14 to 29-year-old age group.

15 As Paul Miller mentioned to you this morning, the
16 WHO panel proposed these definitions for post-menopausal
17 white women. They made a few recommendations, or not
18 actually recommendations. They discussed a little bit about
19 how you might apply these to men but, in fact, they did not
20 make any firm recommendations on how to do that and they did
21 not make any recommendations on how to apply these
22 definitions to non-whites at all.

23 This left, for us, a dilemma at NCHS because we
24 needed to provide estimates for these conditions in
25 population groups in the United States other than just for

1 white women. We had to respond to questions from other
2 federal agencies and from the Congress and from other
3 researchers as well.

4 So we had to decide how and if we could go beyond
5 the WHO approach. What I want to show to you today are just
6 some of the results that we used in making a decision on how
7 to proceed.

8 [Slide.]

9 To start out, I will just simply start to show you
10 how does the bone density play out in the groups that we can
11 provide estimates for. This is showing you mean total femur
12 bone-mineral density levels by age and race in women. I
13 think it is no surprise to most of you here that, of the
14 groups we can look at in NHANES III, black women have the
15 highest mean bone-density values, white women have the
16 lowest at all ages. Mexican-American women fall somewhere
17 in between.

18 I want to point your attention in specific to this
19 20 to 29-year-old age group because, as I said, that is the
20 reference group that we use to derive these cutoff values,
21 the T-scores. You can see that the means differ, as I just
22 described. What is not shown here are the standard
23 deviations. Recall that that is part of the T-score
24 definition.

25 I will tell you that the standard deviations

1 differ in these groups as well. The standard deviation for
2 these young black women is about 10 percent higher than the
3 standard deviation for the young white women whereas the
4 standard deviation for the Mexican-American women in this
5 age group is about 8 percent lower.

6 [Slide.]

7 Looking at the data for men, you see the same
8 pattern by race and ethnicity, also at the 20 to 29-year-old
9 group.

10 [Slide.]

11 No surprise that bone density at the total femur
12 comparing between men and women, men have higher values.
13 This is showing you an overall mean adjusted for age, but
14 this would hold true in the 20 to 29-year-old group that,
15 within race, men are going to have somewhere in the range of
16 10 to 13 percent higher mean bone density levels than women.

17 Also not shown here, but, again, I will tell you
18 that the standard deviations in the 20 to 29-year-old men
19 did, indeed, differ from the standard deviation. They were
20 all larger than the standard deviations observed in the 20
21 to 29-year-old white women.

22 In white men, they were about 15 percent higher.
23 They were almost 30 percent higher in the black young men
24 and they were about 10 percent higher in the young Mexican-
25 American men.

1 So we are faced with a situation here that we have
2 different means and different standard deviations for these
3 particular young reference groups.

4 [Slide.]

5 I will show you now how that translates out into
6 the actual cutoff values, starting with that category
7 osteopenia. This would actually be the upper bound for
8 osteopenia, the mean minus 1 standard deviation. The bar
9 here, in orange, is the cutoff value, the absolute BMD value
10 that you will get from the 20 to 29-year-old white women if
11 you take their mean and subtract off 1 standard deviation.

12 You can see that, in fact, it is the lowest value.
13 All of the other groups have a somewhat higher, being
14 trivial here in the young Mexican-American women, all the
15 way to a 19 percent higher absolute value for bone density
16 defining the upper bound of osteopenia in young black men
17 relative to young white women.

18 [Slide.]

19 For osteoporosis, you see a similar pattern with
20 the young white women producing the lowest absolute BMD
21 value of the groups that we can look at, ranging, again,
22 from a trivial difference up to about 15 percent. You may
23 notice that these differences relative to the young white
24 cutoff for women, are slightly smaller than what I just
25 showed you for osteopenia.

1 I think this is where the fact that these groups
2 have, in the most part, larger standard deviations is
3 starting to be felt because, for this definition, you are
4 multiplying the standard deviation by a factor of 2.5. But
5 they are still all higher.

6 [Slide.]

7 What happens when you apply these cutoff values to
8 the bone-density distributions of older Americans? I am
9 going to start with total femur osteopenia and we will be
10 looking at the bone-density values of individuals ages 50
11 and older starting with the prevalences that you get if you
12 apply the cutoff values derived from the young white women
13 to all of the groups.

14 The pattern that you see is that the older white
15 women have the highest prevalence of osteopenia. The
16 Mexican-Americans women are next highest. The black women
17 and the white men are anywhere from 50 to 70 percent as high
18 as the white women and the black and Mexican-American men
19 have the lowest prevalences of femoral osteopenia.

20 [Slide.]

21 What happens if you use race and sex-specific
22 cutoffs instead? You get a very different pattern. In
23 fact, one of the most notable things here to me is the fact
24 that the prevalence is now, in these older black men and
25 older black women, almost as high as they are in older white

1 women.

2 There is now a higher prevalence of osteopenia in
3 older black men than there is in older white men. So this
4 pattern here seems somewhat at odds with the usual pattern of
5 risk that we would think to see in men versus women and in
6 blacks versus whites.

7 [Slide.]

8 What about osteoporosis? Perhaps, this is the
9 category that you are really more concerned about, the mean
10 minus 2.5 standard deviations. First of all, of course, the
11 overall prevalences are lower. This is, again, going back
12 to using white women's cutoffs. The overall prevalences are
13 lower. But, in women, you see roughly about a half as high
14 prevalence in the black women as you do in the white women.
15 This is somewhere in the 70 percent range for the Mexican-
16 American women. So this seems reasonable consistent with
17 what you would expect from fracture data.

18 This, however, these prevalences in men now seem
19 very low. These are now below 5 percent. In fact, they are
20 so low that the NHANES surveys are really not well designed
21 to estimate prevalences that low. That could account for
22 some of our discrepancies here, but I don't think it
23 accounts for all of it.

24 By discrepancy, I mean, for example, the
25 prevalence here in these older white women is now eight

1 times as high as it is in older white men. That seems a
2 little too much of a difference between men versus women.

3 [Slide.]

4 Does the situation improve if you use race and
5 sex-specific cutoffs? Well, yes, but no. Certainly the
6 discrepancy between men and women is now becoming less. It
7 is about four times different here. But the differences now
8 between, for example, white and black women are now becoming
9 less as well. And now we are almost estimating that black
10 men have a similar prevalence of osteoporosis or slightly
11 higher than white men.

12 [Slide.]

13 How does this play out in terms of actual numbers?
14 What really kind of impact are we talking about at the
15 population level. Here I am showing you what the estimated
16 numbers of individuals with total femur osteopenia are if
17 you use young white women's cutoffs or cutoffs derived from
18 their data, I should say, versus race and sex-specific
19 cutoffs and then what is the difference.

20 Really, the important number to focus on is right
21 here. If we apply race and sex-specific cutoffs to the U.S.
22 population, for the total femur, we are going to identify
23 close to 5 million more individuals with this condition.

24 [Slide.]

25 For osteoporosis, we are going to identify close

1 to a million additional individuals. So, if you sum this
2 number with that previous 4.8, we are talking about a
3 difference of close to 6 million individuals in the United
4 States being identified at some level of reduction of bone
5 density at their total femur.

6 I submit that that is probably a number that is
7 big enough to be felt within the healthcare system.

8 [Slide.]

9 Just one little aside that I will show you. You
10 may be noticing that the impact of the race and sex-specific
11 cutoff seems to be bigger on that osteopenic category than
12 it was on the osteoporosis. I think this is for two
13 reasons.

14 One is that the cutoff values--I am illustrating
15 this with older white men's data--are further apart than
16 they are for osteoporosis but, also, these cutoff values for
17 osteopenia hit this bone-density distribution in the area
18 where the number of people is changing very rapidly.

19 So you would add this shaded amount of people if
20 you moved the cutoff from the white women's cutoff to this,
21 cutoff which is actually for this group, would be the white
22 men's cutoff.

23 [Slide.]

24 This is in contrast to the osteoporosis cutoffs
25 where the cutoffs are now more similar. Again, this is

1 where the larger standard deviations in this white male
2 group made the two cutoffs more similar between white women
3 and white males but, also, these cutoffs hit the
4 distribution out in the tail. So when you start to move the
5 cutoff around, you are not getting as many people being
6 added or subtracted.

7 You will have to excuse me. That is a little
8 statistical aside here. I have to get my statistics in for
9 NCHS.

10 [Slide.]

11 What can we conclude from this kind of
12 information? Certainly, this kind of information is not
13 going to resolve the question of whether it is more
14 appropriate to use race and sex-specific cutoffs than some
15 single set of cutoffs. What this kind of analysis can show
16 you is some of the practical implications.

17 It is already self-evident that if we use race and
18 sex-specific cutoffs, you will be using cutoff values for
19 men and black or Mexican-American women that are at higher
20 absolute bone-density values than they are for white women.
21 This is despite the fact that these groups tend to have
22 lower fracture occurrence than white women.

23 I showed you what the implication was at the
24 population level, at 6 million additional people for that
25 particular femoral site. I think the implication at a

1 clinical level is that, to the extent that some kind of
2 action or medical decision is going to be made when a person
3 passes one of these thresholds, you will be making that
4 decision at a higher absolute bone-density value for men and
5 Mexican-American and black women than you are for white
6 women.

7 Secondly, we felt that the pattern of these
8 prevalences of low bone-mineral density was less consistent
9 with fracture patterns by race and sex when we used race and
10 sex-specific cutoffs. We recognize, as Dr. Wasnich has
11 already told us, that fracture is related to more than just
12 bone density. But it does provide one very ecological way
13 to try to assessment the impact of this.

14 We did feel that the race and sex-specific
15 cutoffs, particularly for osteopenia, produced ratios of
16 prevalences--for example, between men and women and between
17 blacks and whites--that didn't seem very consistent with
18 what you observe in fracture patterns.

19 This was also true for osteoporosis in women, not
20 so true in men. So our group, when we published estimates
21 for the U.S. population in 1995 and 1997, we chose to apply
22 cutoffs derived from the white reference groups to the non-
23 whites that we can estimate for in NHANES III.

24 But we left open the question of whether it was
25 more appropriate to use male or female cutoffs when

1 assessing these conditions in men.

2 Finally, I will just conclude by saying that, of
3 course, with these particular groups using the white data,
4 and particularly the white women's cutoffs, is, in fact, the
5 most conservative approach because that is the lowest cutoff
6 value of the groups that we looked at.

7 In light of the lack of data, we felt that this
8 was reasonable and, also, was clinically more simply than
9 having--we would have had twelve different sets of cutoffs.
10 I think Dr. Wasnich showed us it would have ultimately been
11 some 3 million or 8 million.

12 Thank you. Are there any questions?

13 DR. GARRA: Thank you very much.

14 Questions from that panel?

15 DR. FAULKNER: Just to try to help me understand.
16 The reason that the white reference population gets used to
17 estimate fracture risk--it was done in the NORA program; we
18 heard from Dr. Siris and you were saying that was done as
19 well for NHANES.

20 It is not because of a preferential treatment for
21 race. It is because that is the population we know that
22 fractures.

23 DR. LOOKER: That's right.

24 DR. FAULKNER: That is where bone fragility is.
25 So, in some sense, I think we could remove some of the

1 political stigma just by saying we were comparing this
2 against a from population.

3 DR. LOOKER: That's true. That is a good point.

4 DR. FAULKNER: That is maybe just a comment. But
5 the question is the differences that you saw in prevalence
6 when you used a consistent normal database, especially
7 between men and women, I think you mentioned it briefly,
8 could just be due to the different risk factors.

9 It may be that the fractures in men are not as
10 much a density-mediated event.

11 DR. LOOKER: That's true. That could be true. We
12 felt we had to proceed, even though we can't really address
13 all those unanswered question. We felt it was still more
14 reasonable to proceed with estimates than to simply say we
15 can only do estimates for U.S. white women.

16 DR. GENANT: I wonder if you have looked at these
17 relationships factoring in the body size impact on BMD in
18 these aerial projection measurements and whether that leaves
19 you with results that are more plausible or less plausible.
20 Alternatively, I guess one could use something like the so-
21 called BMAD approach which incorporates that correction.

22 DR. LOOKER: We haven't looked at the prevalences
23 specifically adjusted for something like body size. We have
24 looked at the impact on just the mean values and the
25 differences in the mean values between races and genders if

1 you adjust for something like body-mass index.

2 It does tend to reduce the differences but it
3 didn't completely take them all away. We did look at the
4 prevalences based on BMAD and got a very surprising, very
5 confusing, picture. Depending on which femur site you
6 looked at, you actually ended up with cutoff values
7 generated for men that were way below those for women.

8 So something about that approach wasn't working
9 well in our dataset at all. So I think you are right. I
10 think that certainly has an impact but we haven't been able
11 to identify the best way to deal with that.

12 DR. GENANT: When you corrected for body weight,
13 did it reduce the biological variation of coefficient of
14 variation in the young normals?

15 DR. LOOKER: I don't recall what it did on the
16 standard deviation. I would anticipate that it would,
17 though.

18 DR. GARRA: Any additional questions from the
19 panel? If not, thank you very much.

20 We are going to move on to the next talk which is
21 a description of the Canadian multicenter study of 7,500
22 males and females presented by Dr. Alan Tenenhouse.

23 Dr. Tenenhouse?

24 **Canadian Multicenter Study of 7,5000 Males and Females**

25 DR. TENENHOUSE: Thank you very much for having

1 invited me and for the interest by an American group in a
2 Canadian study.

3 [Slide.]

4 Unfortunately, the study is still in its early
5 stages and so I will not be able to present any prospective
6 data which I had hoped to be able to do. What I will show
7 you is really what the study is all about and some of the
8 cross-sectional and prevalence data that we have and,
9 hopefully, convince you that, as the study proceeds, we will
10 provide a lot of the data which, I think, will go some way
11 to answering some of the questions that have been raised
12 here.

13 [Slide.]

14 Approximately ten years ago, now, the Canadian
15 government decided that the cost of fracture to the Canadian
16 Medicare system and the resulting of loss of independence
17 among the elderly, in particular, was the second largest
18 drain on the Canadian healthcare system. The only thing
19 that cost them more money was dementia.

20 As I am sure you all know, the Canadian government
21 pays for all healthcare in Canada. We are a single payer.
22 Actually, we are ten insurance companies that all draw their
23 money from a single source, the federal government. So they
24 decided that, perhaps, they ought to do something to try and
25 prevent the fractures.

1 After a lot of workshops and questioning, they
2 came up with what is now the Canadian Multicenter Study.
3 The objectives are sort of very, very briefly outlined here,
4 and that is, first of all, to get a Canadian reference for
5 DEXA, one that was more appropriate and one that would more
6 accurately detect people at risk to fracture, then to
7 measure the prevalence of low bone density and fracture
8 risk, to actually measure prevalence and incidence of
9 fracture, detect the risk factors and to determine if there
10 is any ethnic variability across the country.

11 At about the time this was designed, there were
12 European studies which suggested that fracture rates and
13 bone density differed across Europe by a very large amount.
14 If you look at Canada, at the variations in climate, the
15 origins of the population, the ethnicity of the population
16 at various parts of the country, it is almost as different
17 as Europe.

18 It was thought that there may be some very
19 significant variations which would impact on the kind of
20 preventive strategies that were developed. There were also
21 political reasons for making sure all the provinces were
22 represented.

23 [Slide.]

24 What CaMos really became is a five-year
25 prospective study using a random sample--the total sample is

1 actually 9500--but only approximately 7500 are Caucasian who
2 have had bone density and, therefore, they are the ones that
3 are used in the analysis that I will be describing today.

4 Recruitment began in January, '96 and lasted
5 approximately 18 to 20 months. It includes both men and
6 women, 25 years and older. The data collected included
7 questionnaire data with something like 900 variables that
8 relate to potential risk factors and issues of bone health.

9 DEXA was done on the lumbar spine and three sites
10 at the hip. We did not do the total hip at that time
11 although most of the sites did acquire the necessary
12 software and equipment to allow them to do it. We do have
13 the data and we will analyze for that eventually.

14 We also did ultrasound. On everybody over the age
15 of 50, we did a lateral X-ray of spine. These X-rays were
16 digitized and analyzed for detection of vertebral deformity.

17 [Slide.]

18 This is the distribution between male and female,
19 approximately two-and-a-half times as many females as males.

20 [Slide.]

21 With an age distribution as is illustrated here.
22 As you can see, it is weighted toward the older people, as
23 you might expect. Our oldest is, actually, about 102 years
24 old, but there aren't very many at that age, I might point
25 out.

1 [Slide.]

2 The sites stretch from St. John's through to
3 Vancouver. The only two sites that I will point out because
4 they are distinctly different ethnically from the rest of
5 the country is St. John's which is almost pure Anglo-Irish.
6 It is a location where the population was established
7 approximately 300 or 400 years ago and nobody moves in. The
8 only movement is out.

9 For any of you who have been to Newfoundland, you
10 would probably understand why.

11 Quebec City was another one of our sites. Again,
12 it is exactly the same thing except that this location is
13 about 99 percent French-Canadian, and, again, a very, very
14 stable population where practically nobody moves in.

15 Otherwise, the populations are more or less
16 ethnically similar to what you would see across the United
17 States. Vancouver and Toronto both have large Chinese
18 populations and it was hoped, in our random sampling, we
19 would pick up a large enough cohort of Chinese to be able to
20 say something specifically about them. But, unfortunately,
21 it did not work out quite that way.

22 The only other thing I would say is that the way
23 the sampling was done, these centers represent approximately
24 40 percent of the entire population of Canada, the nine
25 centers we chose. That is a 50-kilometer radius around

1 these centers. So we think we have a very representative
2 random population.

3 [Slide.]

4 The first thing we did was to establish peak bone
5 mass which we could use as a reference standard. The
6 results are shown here for the spine and for the femoral
7 neck. The reason the n values are so small is that we
8 determined that, for the lumbar spine, bone density is very
9 constant from 25 to 39 years of age. No matter how you
10 analyze the data and look for break points, there is not
11 change. So we used that entire age spectrum to determine
12 the peak bone mass at lumbar spine in men and women.

13 The femoral neck, the bone density begins to
14 decline right after age 30. At the 25 to 29 age group,
15 beyond there, there is a very discernible and definite
16 decline. So we used only the 25 to 29-year subjects to
17 determine this parameter. Unfortunately, we didn't
18 anticipate that so we didn't include a large enough cohort
19 in that age group.

20 [Slide.]

21 But the 95 is pretty good and our test was to see
22 how we compared to NHANES. Although NHANES doesn't have the
23 lumbar-spine data so we can't compare that, you will see
24 that the femoral-neck data, the peak bone mass as well as
25 standard deviations--and this is from, I think, the '97

1 NHANES publication--and the trochanter data, both in men and
2 women, are very, very similar if not identical. The data
3 for women is virtually identical.

4 So we had great confidence, then, that we were, in
5 fact, doing things properly.

6 [Slide.]

7 We then asked what factors impact on peak bone
8 mass and we modeled for a whole series of things. The only
9 things that we found that were determinants of peak bone
10 mass were center--that is, geographic center--BMI and, even
11 after correcting for BMI, height was an important
12 determinant.

13 Several other things you saw, weight and calcium,
14 vitamin D intake, fractures, various obstetrical and
15 reproductive history as well as smoking had absolutely no
16 effect.

17 [Slide.]

18 What we then did was look at what this center
19 effect really was. I hope you can see this. This is my
20 attempt at powerpoint. I almost had it worked out except I
21 could not get the computer to print the names of the
22 centers. The computer insisted that they should be
23 numbered. But it is geographically precise, running from
24 Vancouver to St. John's. If you remember the map of Canada,
25 you know what each of the points are.

1 This tenth one is the national average of the
2 whole thing and the bars are 95 percent confidence interval.
3 I would like to point out that Quebec City is here and
4 Kingston, Ontario, which is only 150, 200 miles away along
5 the St. Lawrence River, is up here.

6 This difference is greater than one standard
7 deviation. The echo what Dr. Wasnich said earlier,
8 Caucasians are not all the same. These are all Caucasians.
9 This difference is real. I don't know how many of you are
10 up to date with recent Canadian politics, but to get up and
11 say that Quebec population of different from the rest of
12 Canada is very dangerous. So I have made very sure that
13 this, in fact, is correct.

14 What this means, I do not know but I can tell you
15 that, in terms of prevalence of vertebral deformity, there
16 is no significant difference as we go across the out center
17 to center. I can also tell you, with less confidence
18 because it hasn't been analyzed in detail, that we already
19 have the one-year fracture data from each of the centers.
20 As far as we could tell, there is no significant difference
21 center to center at the rate of all fractures.

22 This includes all fractures that have been
23 confirmed by X-ray. So precisely what this means, I am not
24 sure, but there clearly are significant differences between
25 various populations that are supposed to be ethnically the

1 same.

2 [Slide.]

3 We then used this data to estimate prevalence of
4 osteoporosis and osteopenia by the WHO criteria. We
5 compared it do what we got when we used the manufacturers'
6 standards. That is DEXA across the top. I am going to show
7 you only the femoral neck because it is the most striking.
8 Here you can see that with our standards the prevalence of
9 osteoporosis decreased by almost a fracture of four.

10 This is what the male prevalences looked like when
11 you compared the manufacturers' standards versus the women's
12 reference versus the male reference. So there is a dramatic
13 difference.

14 [Slide.]

15 When we compared our data to the published NHANES
16 data, and this, I think, came out of the '95 publication,
17 you can see that at the femoral neck, our 7.5 percent
18 prevalence compares to 20 percent in the United States.

19 I have no idea what that is so given that the peak
20 bone mass, the references, are virtually identical. One
21 would have to conclude that, although the two populations
22 start losing their age-dependent bone loss at approximately
23 the same level of bone mass, Americans lose their bone mass
24 a lot faster than Canadians.

25 There is one possible explanation for at least

1 part of it. It appears that our cohort of postmenopausal
2 women, a much larger proportion of them are on hormone-
3 replacement therapy than is usually believed. It may be as
4 high as 45 to 50 percent who have ever been on hormone
5 therapy with a number somewhere close to 30 percent who are
6 still on hormone therapy.

7 I believe that is a lot higher than what you
8 expect in the United States. So that may, in part, explain
9 it but I can't believe it explains all of it.

10 [Slide.]

11 We then looked at the prevalence of vertebral
12 deformity. As you can see here, in our entire cohort--and I
13 should say that better than 80 percent of the people over
14 the age of 50 who entered this study and who were eligible
15 for X-rays actually got the X-rays and the X-rays were
16 usable.

17 As you can see, approximately 26 to 27 percent of
18 the total population, men or women, have a vertebral
19 deformity. In our system, this is defined as a change in
20 vertebral height of greater than three standard deviations
21 which amounts to a change of approximately 25 percent. That
22 would be our grade I deformity. So 26 to 27 percent of the
23 entire population has deformities of greater than 25
24 percent.

25 [Slide.]

1 If you look at the distribution with age--and,
2 remember, again, this is prevalence. The red is women and
3 the yellow is men. You can see that women start with a
4 somewhat lower vertebral fracture prevalence than men, at
5 about 12 to 14 percent, and from the age of approximately 60
6 on, the prevalence increases in an almost linear fashion.

7 With men, they start at about 18 percent and from
8 about age 50 to age 70 or 75, the proportion of men with
9 detectable vertebral deformity remains relatively constant.
10 It goes up from about 18 to 25 percent.

11 After age 75, the number of men with deformity
12 begins increasing rather dramatically.

13 [Slide.]

14 The other thing that is easy to demonstrate with
15 this group is that the mean bone-mineral density of those
16 with deformity, in red, is consistently less than the bone-
17 mineral density--this is at the lumbar spine--of those with
18 no deformity.

19 [Slide.]

20 The same holds true if you use the bone-mineral
21 density of the femoral neck, that there is a very strong
22 association between bone-mineral density and fracture.

23 As I have said, we have collected the first year
24 follow-up data but have analyzed it only in a very
25 superficial fashion so I am not prepared to discuss it. In

1 fact, I don't know much of what the results are.

2 We are now in year 3 of the study. In year 3, we
3 plan to bring back all people who were between the age of 40
4 and 60 at the time of entry into the study for a complete
5 questionnaire and repeat bone-density measurements and
6 ultrasound measurements. This was built into the protocol
7 because it was felt that the perimenopausal stage was the
8 time in life when age-dependent bone mass began and was
9 accelerating.

10 We wanted to determine, as much as possible, how
11 early and fast that was. We hope that, very soon, we will
12 have some incident data and, with that prospective data not
13 only density changes and fracture but also on risk factors,
14 we will be able to answer a lot of questions that you people
15 put forth.

16 Thank you.

17 DR. GARRA: Thank you very much.

18 Are there any questions?

19 DR. FAULKNER: You seem to have been forward
20 thinking up in the northern regions there to do this
21 national kind of a study linked up with insurance. That is
22 a very interesting model. But I kind of wonder what else is
23 happening. Can you tell us, are there similar discussions
24 like this going on in Canada? Is that a touchy question?

25 DR. TENENHOUSE: No; it is not a touchy question.

1 Let me respond by saying two things. First of all, I must
2 give credit where credit is due and I forgot. With Connie
3 Johnston sitting here, I can't ignore it. We had a panel of
4 advisors of which Dr. Johnston was one who helped us out.
5 So there is American input based on U.S. experience. Also
6 John Kanis was in the group so we had the benefit of his
7 experience.

8 A discussion like this can't happen in Canada,
9 really, because of the provincial-federal division. Health,
10 in Canada, is a provincial issue and policy such as this, as
11 to how to deal with a technology, is a provincial affair.
12 Densitometry is approved. I think only this year is there
13 an approval process for new technologies.

14 X-ray has been approved so you could do anything
15 you want with X-ray. Ultrasound, medical ultrasound, has
16 been approved so you can do anything you want with
17 ultrasound providing you use the similar frequencies of
18 detection and all the rest.

19 And then it goes to the provinces. They are the
20 ones that decide what to do. We really have ten different
21 policies with respect to densitometry.

22 DR. FAULKNER: Are they T-score based?

23 DR. TENENHOUSE: Most of them are T-score based;
24 yes. One of the problems is that we have so well educated
25 our family practitioners that they see a T-score, they begin

1 treating. The government who has a perfect record of
2 prescriptions written and tests done is panicking.

3 So, for instance, BC has put a moratorium of
4 DEXAs. Quebec had one on DEXAs for a long time. The fees
5 that are reimbursed to radiologists are shrinking at an
6 enormous rate. So they are not taking away the technology
7 in many places. They are just reducing the reimbursement to
8 the physician so there is less incentive to do it.

9 DR. FAULKNER: That is out of concern of
10 overtreating, based on a simple application of a criteria?

11 DR. TENENHOUSE: Of an economic model.

12 DR. GENANT: I thought that it was interesting
13 when you showed the data on vertebral deformities in men and
14 women, you had about comparable percentages, 26, 27 percent
15 or so. But then when you showed the relationship to BMD and
16 that there was a differential between those who were
17 fractured and were not fractured for both women, that
18 differential became much greater as you got into the older
19 age group.

20 The likelihood is that the overall comparability
21 relates to residual deformities, trauma, Shoreman's disease
22 and all those phenomena that occur more frequently,
23 probably, in men during earlier life.

24 DR. TENENHOUSE: Yes. That's possible. I must
25 say we attempted to eliminate as much of the artifact that

1 is introduced by these changes and that is very, very
2 serious in men. In fact, if you look at raw BMD in our
3 population with age at the lumbar spine, for men, the women
4 give you the very nice curve we have all been taught to
5 expect. With men, it is a straight line. In fact, there
6 may be an upward slope as you get older.

7 So the artifact must be extremely common.

8 DR. GARRA: Any other questions?

9 Thank you very much, then.

10 Next, Dr. Dennis Black is going to be speaking to
11 us on a comparison of T-scores, Z-scores and other various
12 measurements for the assessment of fracture risk. We have
13 all been waiting for this one.

14 **A Comparison of T-scores, Z-scores and other Measurements**
15 **for Assessment of Fracture Risk**

16 DR. BLACK: Thank you.

17 [Slide.]

18 I would also like to thank the panel for the
19 opportunity to speak to you.

20 [Slide.]

21 I am going to actually talk about three distinct
22 but related topics. The first is to talk a little bit about
23 T-scores versus Z-scores and then ask the question of
24 whether T-scores are comparable across devices. That is
25 going to be my focus is what do we do with different

1 devices, different sites, in terms of T-scores and then talk
2 to you about a preliminary proposal which Dr. Johnston
3 mentioned to alter the way we use BMD to diagnose
4 osteoporosis.

5 [Slide.]

6 I am sure we have heard about T-scores and Z-
7 scores all day. Basically, we have three options in terms
8 of how we present bone-density data, raw values, T-scores or
9 Z-scores. I will say just a couple of words, really, about
10 Z-scores and then focus more on T-scores versus raw values.

11 [Slide.]

12 Here you see the prevalence of osteoporosis by age
13 if we use a Z-score or a T-score of -2.5. The Z-scores are
14 in orange. What you see is the prevalence based on Z-score,
15 because a Z-score is an age-specific comparison, the
16 prevalence doesn't change with age, by design. And so the
17 Z-scores have this really not very nice epidemiologic
18 property in that the prevalence of the disease does not
19 increase if you use Z-scores. That is probably one of the
20 main reasons why people have gone to T-scores and other
21 things like T-scores because they have the nice property
22 that the prevalence of the disease increases with age, just
23 as we would expect.

24 [Slide.]

25 Are Z-scores useful at all? I think they are in

1 some clinical situations. We have talked about this already
2 today--in situations, particularly in younger people where
3 we have less of an idea of kind of what is happening with
4 age. For example, a 25-year old man or, say, a younger
5 African-American woman, an amenorrheic woman. In those
6 kinds of situations where you really don't know what to
7 expect I think it is very useful to have a comparison to the
8 age-specific normal data. How does this patient compare to
9 average for their age?

10 So, while it is not maybe the ultimate diagnostic
11 tool, I think it is still a useful measure in clinical
12 situations.

13 [Slide.]

14 Let me go on, then, and talk more about T-scores
15 versus raw values. This is part of the underlying issue
16 that we talked about all day. Of course, this formula, if
17 you are on your toes, is backwards, the numerator. But, in
18 any case, what it shows is that a T-score is simply a linear
19 transformation of the raw value.

20 So if you are asking the question of should you
21 use T-scores versus raw values, the answer is it really
22 doesn't matter because if you plot raw BMD by corresponding
23 T-scores, you will find a correlation of 1.0, just a linear,
24 straightforward transformation from raw to T-scores.

25 Another way of looking at that is if we look here,

1 on this axis, we have fracture risk. Here we have bone-
2 mineral density in terms of grams per centimeter squared,
3 raw units. Here are T-scores. It doesn't matter which of
4 these scales you use. You pretty much will find the same
5 thing.

6 So there is really no difference between the two.
7 So, why have we foisted T-scores upon the world? What is
8 the advantage?

9 [Slide.]

10 I think there are a couple of really important
11 rationales for why T-scores were adopted. Firstly, of
12 course, there is some kind of sense that the T-score
13 measures the loss from peak which is some kind of measure of
14 absolute osteopenia, the amount of bone that is lost.

15 But I think the more important rationale is that
16 the T-score potentially provides a means of standardizing
17 these measurements across multiple sites and across multiple
18 devices so it provides a way of saying, "A BMD value at the
19 spine, what does that mean in terms of equivalence to the
20 hip or to an ultrasound at the calcaneus?"

21 [Slide.]

22 In terms of how we use these T-scores, I think it
23 is worthwhile going back to the history a little bit. Dr.
24 Miller talked about this in a lot more detail, but I just
25 wanted to remind the group of what the original development

1 was of T-scores. Dr. Johnston can correct me later if I am
2 wrong here, but the purpose really was to compare how many
3 people have osteoporosis in different countries.

4 It was a gross epidemiologic comparison so we
5 could say, "Is there more osteoporosis in Italy or in
6 Sweden?" It was really meant as a very sort of a tool just
7 to assess what is going on. There wasn't a lot of thought,
8 I think, in the WHO report about different sites and
9 different devices because there weren't very many different
10 sites and devices at that point.

11 I have heard a number of individuals who were on
12 the WHO panel get up and object at various osteoporosis
13 sessions and be very emphatic about the fact that the
14 original use of T-scores was not really meant for individual
15 diagnosis.

16 [Slide.]

17 But, in fact, we have done that. If we look at
18 what is going on in 1999, we see that T-scores are very
19 commonly used. In fact, it is our primary tool for
20 individual diagnosis. Even more, the NOF, recently, in
21 their treatment guidelines suggested the use of T-score
22 cutpoints of -2.0 and -1.5 depending on whether there are
23 risk factors or not to make individual treatment decisions.

24 So that is even kind of a step up. The NOF
25 report, if you read it carefully, is based on hip BMD. They

1 wave their hands a little bit and say, "We think the same
2 kinds of considerations apply to other sites." But it is
3 specifically focussed on hip BMD.

4 Again, I think the general assumption in the
5 clinical community now is that a T-score at one site or
6 device is comparable to a T-score from another site or
7 another device. The key word here is "comparable."

8 [Slide.]

9 So what do we mean when we say "comparable?" What
10 is the assumption here. I think most people would say that
11 there are really two different assumptions that are
12 implicitly made--I am going to look at the example
13 specifically looking at -2.5, but these same considerations
14 apply to any T-score cutpoint--that is that if we measure
15 two sites in the same population, we would see two things.

16 First of all the same proportion would have BMD
17 less than 2.5. People have been talking about prevalence of
18 osteoporosis. So the assumption here, I think, in general,
19 is that if you use two sites or two devices, you would see
20 the same proportion with less than 2.5.

21 But secondly, and I think less often stated, but I
22 think it is also true that there is some assumption that the
23 fracture risk would be similar in those with BMD less than
24 2.5. These are really two separate assumptions that have
25 been looked at. Specifically, the first has been looked at

1 and you have heard some data saying that that is not the
2 case. But the second one, I think, is less closely looked
3 at.

4 [Slide.]

5 I am going to look at those two assumptions for
6 three example devices; hip BMD by DXA, spine BMD by DXA and
7 BMD or calcaneal measurement by ultrasound. In doing these
8 prevalence comparisons, I am going to use age-specific means
9 and standard deviations for Caucasian women. For hip DXA, I
10 am going to use the NHANES data. And for spine DXA and
11 calcaneal ultrasound, I am going to use the specific
12 manufacturers' data.

13 In order to calculate these prevalences from these
14 data, you have assume a normal distribution which is
15 probably a pretty good assumption from the data that I have
16 seen.

17 [Slide.]

18 So, for these three example devices which, again,
19 are only examples--we have done this exercise on about
20 fifteen different sites and devices--here you see the BMD
21 values that correspond to a T-score of -2.5. I think it is
22 worth mentioning in passing that when you use T-score
23 cutpoints, you generate one specific BMD value that is used,
24 in general, for diagnosis.

25 So a T-score of -2.5 translates to, say, a hip DXA

1 value, the hologic value, of 0.58, a spine DXA of 0.77,
2 ultrasound 39.7 for these particular devices. If you apply
3 that cutpoint to age 70, you see these are the prevalences
4 of osteoporosis; 25 percent at the hip, 35 at the spine,
5 6 percent at calcaneal ultrasound.

6 So you see that the prevalences here are not
7 comparable across devices at the same T-score cutpoints. In
8 fact, there is a very large difference here between these
9 two, about a six-fold difference. So, as a general rule,
10 there is no way in our mechanism to assure that the
11 prevalences will be equal at the same T-score cutpoints.

12 [Slide.]

13 So what about the second assumption? What about
14 fracture risk among those with low BMD? Also, we can go
15 ahead and calculate this for the same three example sites.
16 What I used to calculate this, we needed to look at the
17 relative risk per SD. As Dr. Wasnich said, you need some
18 estimate here to calculate risk.

19 I used 2.6 for hip BMD and that is based on the
20 Marshall review in BMJ, spine DXA of 1.6, and calcaneal
21 ultrasound, 2.0. That is based on two large epidemiologic
22 studies which recently reported relative risk for SD of 2.0.
23 These are relative risks per SD for hip fractures.

24 Then the other assumption necessary in this
25 calculation is that I assumed that BMD and fracture risk

1 follow a logistic function, a very common epidemiologic
2 function, that people assume about the relationship, say,
3 between lipids and heart disease, blood pressure and heart
4 disease. So it is a fairly general functional form.

5 [Slide.]

6 So, making those assumptions, we can ask the
7 question of whether hip-fracture risk in those with low BMD
8 is equal at the three sites. What you see here--this first
9 set of columns is what I already showed you for prevalence,
10 but here is what we see for risk; hip BMD, 6 percent, spine
11 BMD, 3.8 percent, calcaneal BMD about 8 percent. So there
12 is almost a twofold difference in hip fracture risk between
13 these two sites.

14 Again, these are not chosen as the most extreme
15 sites. These are sort of representative of all the sites
16 that we might look at. So the hip fracture risk is also not
17 comparable between the T-score cutpoint. So we really don't
18 have comparability for either of those two things. The T-
19 score cutpoints do not yield comparable prevalence of low
20 BMD across sites or comparable fracture risk, at least for
21 hip fracture, across these two sites.

22 [Slide.]

23 There has been a lot of discussion already today
24 about these discrepancies, how can they be fixed.
25 Certainly, part of the discrepancies are due to the fact

1 that the young-normal data is collected on different
2 populations and the samples use different criteria, et
3 cetera.

4 So probably we could decrease these discrepancies
5 a bit if we did a study in which we used the same large
6 population. And there is a study, of course, that Dr.
7 Miller and others are working on to try and do this for
8 existing devices.

9 But I think that much of the discrepancy, as has
10 also been mentioned today, is probably due to differential
11 bone loss, that bones seem to lose bone mineral at different
12 rates. If this is the case, then there is really nothing
13 you can do to fix the young-normal data.

14 I will show you some data in just a minute about
15 this.

16 Lastly, the risk discrepancies are also due to the
17 fact that some sites of BMD have a stronger relationship to
18 fracture than other sites. So there are a number of
19 different causes. The first can be addressed directly and
20 probably will be addressed, but I think much of the
21 discrepancy will still remain due to problems of the second
22 two.

23 [Slide.]

24 Let me just show you some longitudinal data. This
25 is from the study of osteoporotic fracture. You have seen

1 data in Dr. Genant's presentation here. What this
2 particular data is looking at is bone loss after age 65.
3 There are longitudinal measurements where you have a
4 baseline and then a follow-up measure for four different
5 sites; femoral neck, BMD, lumbar spine, calcaneal--BMD here
6 is not ultrasound--and radial BMD.

7 And then we looked at the average bone loss by age
8 from longitudinal data. What you see is very different
9 rates of bone loss. These are projected out over twenty
10 years assuming people start at the same place at age 65. On
11 average, we see about a 5 percent gain over twenty years at
12 the spine and there is a range here. The biggest loss is at
13 the calcaneus where there is -33 percent estimated over
14 twenty years.

15 The point I would like to make with this is even
16 if you had bone loss being absolutely parallel from age 30
17 to age 65, even if that were the case, which is not the
18 case, then, after age 65, you would see very large
19 discrepancies developing in T-score cutpoints.

20 So I think the problem with these discrepancies is
21 inherent to the use of the T-scores.

22 [Slide.]

23 These T-score discrepancies of differential bone
24 loss will guarantee that these discrepancies will continue
25 to exist. Basically, you can think about this way. Our

1 bones just don't know that they are supposed to lose mineral
2 at the same rate, especially in cross-sectional standard-
3 deviation units.

4 There is really no reason to believe that they
5 should. So I think that these discrepancies are sort of
6 inherent to the use of T-scores.

7 [Slide.]

8 Other problems. We have heard about some of the
9 problems of precisely and accurately estimating young-normal
10 estimates. I know, for example, if you look at the NHANES
11 data, part 1 of the NHANES and part 2 from the NHANES III
12 survey, there actually are small differences in the standard
13 deviations and in the means just between those two parts of
14 a very, very good survey.

15 When we project those small standard-deviation
16 differences over 50 years, between, say, age 30 and age 80,
17 you come out with very large differences. So that is one of
18 the problems.

19 Of course, the whole concept of whether this is a
20 generalizable concept that can be applied to other
21 ethnicities and to males is not known. So I think those are
22 two important problems with T-scores. Based on those
23 problems, I think there is a growing recognition in the bone
24 community in the U.S. that there is a need to possibly
25 revise the diagnostic cutpoints and a growing recognition

1 that T-score cutpoints can't really be applied uniformly to
2 all sites, devices and technologies and they may not be the
3 best way to generalize across gender-race groups.

4 [Slide.]

5 That is the problem. What is the solution? I'm
6 sorry; I'm out of time here. No, no; just kidding. Based
7 on this recognition, the National Osteoporosis Foundation
8 and the International Society of Clinical Densitometry have
9 set up a joint committee to evaluate the problem.

10 I think this group is likely to be joined very
11 soon by the ASBMR. So we want to put all the groups
12 together, look at the problem. Specifically, what this
13 committee is doing is evaluating BMD reporting and
14 diagnostic problems and it is going to recommend a revision
15 to address some of these problems, a number of revisions,
16 probably.

17 The group is sort of chaired jointly by Dr.
18 Johnston, Dr. Miller and myself and includes about a dozen
19 academic investigator-researcher types. We have also been
20 working very closely with the manufacturers directly and
21 through the densitometry standardizations committee.

22 The hope is that this committee will come up with
23 some recommendations, final recommendations, which will be
24 submitted for approval to all these organizations by the
25 fall of this year.

1 [Slide.]

2 What I would like to do, then, in the next few
3 minutes is just present the outlines and the overview of our
4 current proposal for revision of diagnostic cutpoints.
5 General principles; the first general principle is that most
6 people believe, my clinical friends believe, that we need to
7 have a single diagnostic value for each device that is
8 applied at least across ages.

9 That is, again, what you get with a T-score. You
10 get a particular value that is applied across all the ages.
11 Now, Dr. Wasnich, this morning, talked about maybe a risk-
12 specific value. I would think that, in general, that is
13 probably a good idea but, in terms of clinical simplicity,
14 we probably need to have a specific value, to say the value
15 for DXA at the hip measured on a hologic is 0.60, for
16 example.

17 So that is one of the principles underlying this
18 proposed revision. Secondly, there is a recognition that
19 there is a huge impetus toward T-scores, that we have been
20 very successful in a lot of ways in educating the clinical
21 community about T-scores. They know about them from the WHO
22 guidelines and also from the NOF. The manufacturers as well
23 as the pharmaceutical companies have been very successful in
24 promulgating T-scores.

25 So I think there is a feeling that we can't

1 abandon those T-scores totally. We have to somehow adapt
2 them. Therefore, the solution is to create device-specific
3 diagnostic cutpoints which are anchored to T-scores. I will
4 explain in a few minutes what I mean by that and possibly
5 call these T-score-equivalent cutpoints.

6 So we are going to develop specific device or site
7 diagnostic cutpoints which are anchored to existing T-
8 scores.

9 [Slide.]

10 What is the starting point here? The starting
11 point is we have to anchor to a very specific T-score. What
12 we are proposing to do anchor the T-score at the hip,
13 specifically the femoral neck, and use age 70 as the
14 reference. So we are going to start with BMD at the femoral
15 neck at age 70 as the reference.

16 Why BMD at the hip? For a number of reasons. I
17 think it is probably the most well-studied site. The NHANES
18 study, for example, looked at hip BMD. The NOF clinical
19 guidelines were also developed based on hip BMD. And then
20 the large epidemiologic studies, particularly the study of
21 osteoporotic fractures and the EPIDOS study in France--there
22 is a large study in Rotterdam, in the Netherlands--have all
23 used hip BMD.

24 There is a lot of information relating hip BMD to
25 fracture risk. And so we have a lot of data to start with.

1 [Slide.]

2 The other point, and it has been looked at already
3 today, is that if we look at these different BMD sites, say,
4 the radius, spine and the hip, and how well they predict
5 fractures, specifically hip fractures, what you see is that
6 hip BMD is the strongest predictor of hip fracture.

7 It is pretty much the same as the other sites were
8 predicting risk fractures and spine fractures, it doesn't
9 much matter, but if you want to predict hip fractures which
10 are, by far, the most important types of fractures, hip BMD
11 seems like the best measurement to use.

12 [Slide.]

13 I would disagree a little bit with Dr. Wasnich.
14 Here is some data looking at the quartile of bone density
15 and fracture risk for hip fracture. On this side, you see
16 hip BMD. On this side, you see spine BMD. I think the
17 gradient of risk for hip BMD is quite steep. It is almost a
18 ten-fold increase in risk from the highest to the lowest
19 quartile of hip BMD, whereas there is still a gradient for
20 spine BMD but it is only about double across those
21 quartiles.

22 So I think a relative risk of 2.7 is really a much
23 more important gradient of risk than 1.5. Anyway, that is
24 sort of a digression. The point here is that we think this
25 is one of the reasons why hip BMD is probably the best thing

1 to anchor to in these revisions.

2 [Slide.]

3 The other question is why age 70. Again, this is
4 a provisional decision. All these decisions are still being
5 discussed. But there were really two reasons why we have
6 chosen age 70. The first is that hip fracture incidence
7 starts to increase at about age 70. We have an exponential
8 increase with age. It starts to go up at about age 70.

9 Secondly, and I think more importantly, age 70 is
10 about the midpoint of the postmenopausal period. If you
11 think about a 50-year-old woman, she has about a 30 to 35-
12 year life expectancy, so age 70 is about in the middle. The
13 idea is if we can synchronize the measurements about at that
14 point, they will probably be pretty well synchronized across
15 the postmenopausal period.

16 [Slide.]

17 Here you see our anchor values; femoral-neck BMD,
18 age 70. These are, of course, data from Caucasian women.
19 These are femoral-neck BMD T-scores, -2.5, the WHO cutpoint,
20 -2.0, -1.5. But I really want to focus on the -2.5 because
21 that is the data I am going to show you today.

22 There are two aspects of the -2.5. One is the
23 prevalence of osteoporosis, percent below T-score -2.5, at
24 age 70, the femoral neck for Caucasian women, that is about
25 25 percent, the risk of hip fracture, the five-year risk of

1 hip fracture in those below that cutpoint is about 6
2 percent.

3 So I am going to use these numbers over and over
4 again. 25 percent prevalence at -2.5 and 6 percent
5 incidence of hip fracture over five years at -2.5, both of
6 these at age 70.

7 [Slide.]

8 So what we are going to do is to create diagnostic
9 cutpoints for other sites and devices which are equivalent
10 to a T-score of -2.5 at the hip. Again, there are two
11 possible criteria for equivalence; there is the prevalence
12 of low BMD which, again, was 25 percent for age 70, -2.5, or
13 fracture risk which was 6 percent.

14 I am going to show you what happens when you use
15 prevalence and when you use fracture risk, very briefly.

16 [Slide.]

17 Again, these are just the assumptions. I have
18 already mentioned all these in passing, what you need to
19 calculate these T-score equivalent cutpoints. Here is what
20 you get if you calculate an equivalence cutpoint based on
21 prevalence. As I said, the prevalence of low BMD at the
22 hip, femoral neck, was 25 percent. So we want to find the
23 same BMD value, or a BMD value which corresponds to a
24 prevalence of 25 percent at the spine or at the calcaneus
25 with ultrasound.

1 These are the corresponding values you get. These
2 are the corresponding T-scores. So, to generate a
3 prevalence of 25 percent at age 70, you need a T-score of -2
4 at the spine and -1.6 for calcaneal ultrasound. So you can
5 see these are not the same as 2.5. It just points out that
6 you need a different cutpoint.

7 [Slide.]

8 We have created a equivalence cutpoint here based
9 on prevalence at age 70. Here you see 25 percent, 25
10 percent, 25 percent. So what happens if we look at age 50,
11 what is the prevalence. It is not too bad. It is a little
12 bit off for the spine, but it is pretty similar at age 50.
13 Also, at age 85, there is a little bit of discrepancy, but
14 it is still pretty close.

15 So if we set age 70 as the age we want to
16 synchronize the measurements based on prevalence, the
17 prevalence stays pretty constant from age 50 to age 85. So
18 that seems to work pretty well.

19 [Slide.]

20 If we go through the same exercise based on
21 fracture risk at age 70--in other words, find the cutpoint
22 that yields a hip-fracture risk below that cutpoint of
23 6 percent for spine BMD or calcaneal ultrasound, these are
24 the cutpoints and these are the T-scores which correspond to
25 those cutpoints.

1 So, again, you can see that -2.5 doesn't work
2 across the board. You have to take different cutpoints to
3 get the same hip-fracture risk. But it can be done.

4 [Slide.]

5 Again, if you look at the same thing, here we made
6 the cutpoints in sync at age 70. If you look at age 50, or
7 age 85, the five-year fracture risk is pretty much the same
8 at age 50 and at age 85. It is very low, of course, at age
9 50 but at age 85, they are pretty much the same.

10 So, again, the point is if we synchronize in the
11 midpoint of the postmenopausal years, we can pretty much
12 create comparability across those years.

13 [Slide.]

14 Risk versus prevalence-based equivalence; what is
15 the best? Of course, you can't have both. There is some
16 advantage to prevalence-based equivalence; less information
17 required and you also maximize diagnostic concordance. If
18 you are a clinician, if you have the same prevalence of
19 disease by different devices, the measurements are going to
20 want to average, tend to agree. They are not always going
21 to agree because we have different bones. Some are going to
22 be low, some are high.

23 But if you measure a thousand people on two
24 devices with this prevalence-based equivalence, they will
25 agree. On the other hand, I think there are a lot of

1 arguments for risk-based equivalence.

2 Since we really want to treat people that are at
3 highest risk, using risk-based equivalence will maximize the
4 cost-effectiveness of treatments and, also, I think, very
5 importantly, hopefully since risk would be incorporated in
6 the cutpoint, it would encourage the development of devices
7 that are more predictive and, conversely, discourage less
8 predictive devices.

9 We have some pretty good devices already, the hip
10 BMD for hip fracture, even calcaneal ultrasound and BMD are
11 all pretty good predictors. I think it would be better for
12 the field to continue to have devices that are at least as
13 predictive as those types of devices.

14 [Slide.]

15 But regardless of whether we use risk or
16 prevalence-based, there are a lot of advantages to these T-
17 score equivalents in terms of creating a basis and a
18 framework for established comparability across sites and
19 devices for existing devices and, very importantly, for new
20 devices.

21 So we kind of assure that anything that comes down
22 the road can somehow be used in conjunction with our current
23 devices.

24 [Slide.]

25 Let me just say a couple of words about what do we

1 need if we adopted something like this. We really need to
2 focus on samples of women, say, or men 65 to 75-years-old if
3 we are going to make them equivalent based on age 70.

4 If we are going to use risk-based equivalence, we
5 have to have some information about BMD and hip fracture.
6 So we might need to incorporate some kind of case-control
7 studies that would be done so that we could assess how well
8 the BMD and the new device predicts fracture, hip fracture,
9 specifically.

10 It has been proposed by Dr. Huey that we
11 incorporate hip BMD as well in these case-control studies so
12 that we can calibrate the case-control studies. We know hip
13 BMD and hip fracture has a certain relative risk so we can
14 then compare our new device to that to kind of make sure
15 that the study was done right.

16 Of course, since we are probably going to want to
17 keep some measure of age-specific equivalence, we probably
18 want to do also samples at other ages. But those samples
19 could be quite a bit smaller. The largest samples would be
20 in the older ages, around age 65 to 75.

21 [Slide.]

22 Getting back to the original problem, will these
23 T-score equivalents solve gender, race, BMD diagnosis
24 problems. I think that really hasn't been the focus of our
25 group so far. I think we are trying to come up with some

1 kind of framework and then I think we figure out how well it
2 would solve this problem.

3 But I think the answer is, if risk is equal at the
4 same BMD, and Dr. Wasnich showed data and there is other
5 data that I am aware of as well that suggest, at least for
6 men and women at the same BMD value, once you adjust for
7 age, the risk is about the same. If this is the case, then
8 a risk-based equivalent might provide a very nice framework
9 for generating a specific BMD value that could at least be
10 used for men and for women.

11 I think the jury is still out with respect to
12 different race groups but, at least with men and women, it
13 looks like, at the same BMD value, they tend to have the
14 same fracture risks. So that might, to some extent, go to
15 the first step towards solving the problem.

16 I want to conclude, but I also want to remind you
17 that what I am showing you is preliminary. For example, we
18 assumed in this analysis that we were going to do a risk-
19 based equivalence if we do it based on hip fracture at age
20 70 using femoral neck hip BMD. All these assumptions are
21 going to be discussed and they may change but the concept
22 remains the same.

23 In addition, there are other issues that have been
24 brought up. For example, will these techniques work for
25 women between 50 and 55 years old, 50 to 60 years old.

1 Maybe we need to look at something other than hip fracture.
2 So these are all issues that will be discussed prior to the
3 main report that we are planning to get out in the fall.

4 But, in any case, if I can conclude, fixed T-score
5 cutpoints such as T at -2.5 for all sites and all devices
6 creates a lack of comparability between diagnoses across
7 devices. We know that we can develop site and device-
8 specific cutpoints that can be anchored to either prevalence
9 of low BMD or to fracture risk and that these types of
10 cutpoints can provide a uniform framework for rational
11 incorporation of new devices.

12 Thank you very much.

13 DR. GARRA: Thank you.

14 Are there questions?

15 DR. GENANT: Dennis, as our group started to work
16 with some of these problems, the assumption had been that if
17 one balanced for prevalence that, in fact, the risks would
18 be quite disparate. I thought that you had generated some
19 data in the last two weeks that suggested that when you
20 balanced for prevalence at age 70 that, in fact, the risks
21 were surprisingly comparable.

22 DR. BLACK: Actually, what happens is if you
23 balance for prevalence, then the risks are actually fairly
24 comparable. If you balance for risk, the prevalences are
25 not too far off, either. They are much better than if you

1 just use the T -2.5 base cut scores. So, either way, you
2 are doing much better.

3 I think part of that is conceptually because you
4 are developing equivalence in the years that you really care
5 about as you showed with your data at age 65. When we try
6 to balance at age 30, and then project this out 50 years,
7 things have a long time to get out of balance.

8 So I think that, no matter what we do in that
9 period, we are going to do better.

10 DR. FAULKNER: So this is what I am struggling
11 with a little bit. I am struggling with it a lot. This is
12 great data showing that it is possible to produce similar
13 risk populations across BMD measurement sites. We heard
14 from Dr. Wasnich this morning that that is something that he
15 thinks is possible, as well.

16 But we have got this concept of T-score. I guess
17 this is maybe not a question so much for you, but from the
18 NOF and maybe from Dr. Johnston and Dr. Miller if they want
19 to respond about the importance of keeping the T-score
20 around.

21 The NOF guidelines, the ink is still wet and they
22 are based on T-scores. So I guess I am looking for some
23 guidance here about can we throw out a T-score. I guess
24 that is what Dr. Wasnich is suggesting.

25 DR. BLACK: I don't know. I think that is a

1 question to be dealt with by the clinical community. I will
2 let Paul and Connie respond. But I think the point here,
3 with this particular thing, it is kind of a compromise
4 because we have started with the concept of femoral-neck T-
5 scores and used the risk and the prevalence from them. And
6 then you could at least call things a T-score equivalent.
7 So then you would have a transition period.

8 But, in fact, these new things have nothing to do
9 with T-scores.

10 DR. FAULKNER: Really, what you just presented was
11 a report, a format, of providing absolute risk.

12 DR. BLACK: Yes.

13 DR. FAULKNER: Because everything is linearly.

14 DR. BLACK: Absolute risk with a cutoff of a
15 certain value because I guess, again, the clinical people
16 say in the world where doctors have six minutes to see a
17 patient, they need to have a particular value that they
18 would use. That would be the reason why some people would
19 say that a continuous risk measurement or a continuous any
20 measurement is not as useful.

21 DR. FAULKNER: I don't know if we can use some of
22 your time to let Dr. Johnston and Dr. Miller address that
23 question. Or maybe later.

24 DR. GARRA: We can do that in the open discussion
25 in a few minutes.

1 DR. GENANT: Another comment, Dennis. You
2 indicated right at the end that this approach, while it
3 might be useful for women in kind of the older group, the
4 say 70 and on, but, perhaps, in the first five to ten years
5 after the menopause where hip fracture is not a major
6 concern, vertebral fracture is more of a concern and,
7 likely, trabecular bone, that perhaps one could focus on,
8 say, vertebral fracture and spine measurement to determine
9 the threshold for that particular group.

10 DR. BLACK: Certainly that is a possibility, to do
11 something different in the 50 to 60-year-olds. I am a
12 little pessimistic because we have so many definitions of
13 spine fractures and so little data. But you could do
14 something like that; yes.

15 DR. GARRA: I have a question, myself, regarding
16 new instrumentation that might be developed and coming on
17 line. It looked to me like the scheme that was being
18 bandied about or at least proposed was that you would
19 produce these T-score equivalents. They would really be
20 pegged to BMD values that were based, perhaps, on the
21 manufacturer doing a study of fracture risk and pegging
22 those exactly to BMD scores; is that correct?

23 DR. BLACK: Yes. Obviously, ideally, you do a
24 prospective study.

25 DR. GARRA: Is it really necessary for each new

1 manufacturer to do their own prospective study, BMD versus
2 fracture risk, or can they calibrate their instrument
3 against an existing instrument for which it has already been
4 done?

5 DR. BLACK: I think there are a couple of
6 possibilities. Firstly, it is not practical to do a
7 prospective study. That is not a practical thing for every
8 new device. So what I had mentioned here is doing a case-
9 control study which could be done fairly small and fairly
10 cheaply. I think the manufacturers could do something like
11 that. The Marshall analysis has shown that the case-control
12 relative risks seem to be pretty close to the prospective
13 one. So that is one possibility.

14 But a second possibility is I think someone could
15 establish that if a device has a certain minimum correlation
16 with an existing device, say a 0.95, another measurement of
17 radial BMD--we already know the relationship. You could
18 just develop the cutpoints based on that. So I think that
19 that is a way to do it.

20 I think that it is very important that if you do
21 that, to specify the correlation be within a narrow age
22 range because everything is correlated between age 10 and
23 99.

24 DR. GARRA: Thank you.

25 Any other questions from the panel?

1 Thank you very much. I will get back to the issue
2 of the T-scores again in the open discussion.

3 Now we move to a phase of some industry
4 presentations. The first presentation will be given by Dr.
5 Christian Langton from McCue PLC discussing fracture risk
6 better predicted by physician properties than by T-score.

7 **Industry Presentations**

8 **Fracture Risk Better Predicted by Physical Properties**
9 **Than T-Score**

10 DR. LANGTON: Thank you for the opportunity to
11 address the panel.

12 [Slide.]

13 My post is senior lecturer in the Center for
14 Metabolic Bone Disease at the University of Hull and Royal
15 Hull Hospitals and also consultant to McCue PLC. I am going
16 to restrict my presentation to quantitative ultrasound and
17 the belief I have that we need a consensus in terms of its
18 application.

19 Some of the ideas I am going to describe to you
20 are sort of a general hypothesis that I believe, from what
21 we have already heard today, could be applied to other
22 technologies.

23 [Slide.]

24 I propose that there is now a strong body of
25 evidence for the scientific basis of quantitative

1 ultrasound. As Dr. Turner explained earlier today, if we
2 look at ultrasound velocity measurements which is volumetric
3 then it is strongly related to the elasticity and the
4 density. We have shown r-square values approaching sort of
5 95, 96 percent on this basis.

6 Looking at other in vitro fundamental studies for
7 the attenuation and particularly the broad-band ultrasound
8 attention that we first described in 1984, this is typically
9 an aerial rather than volumetric parameter because we don't
10 normalize the data for the bone width.

11 We have shown that, for the human calcaneus, there
12 is a very high correlation between apparent density and the
13 BUA measurement. However, if you look at other tissue types
14 such as equine or bovine that has a more structural
15 component to it, then you actually get a dependence. What I
16 have always said is that if there is structural variability,
17 BUA will pick that up. Otherwise, you will be simply left
18 with the density function.

19 What we have recently shown is a strong dependence
20 between attenuation and the fractile dimension of bone
21 which, again, you are sort of beginning to fit the whole
22 thing together in terms of quantity and quality.

23 If you look at in vivo studies that have been done
24 in the clinical area, then I propose to you that there is
25 strong evidence that has demonstrated that quantitative

1 ultrasound is sensitive to age-related changes, that
2 normative data have been defined for several devices, that
3 ultrasound can discriminate osteoporosis subjects and that
4 there is fracture data that has already been described today
5 that is commensurate with that obtained with axial DEXA
6 measurement.

7 [Slide.]

8 I believe that there are several clinical roles
9 for quantitative ultrasound. The two major ones that I
10 would like to propose to you--the first one has already been
11 discussed today--is to predict fracture risk independent of
12 other established bone densitometry. So, in this case, the
13 performance criterion is the ability of ultrasound to give a
14 prospective indication of fracture risk.

15 The question that has already been raised today is
16 the option of using it as a case-finding referral tool for
17 subsequent conventional densitometry to specially center.
18 The performance criterion here will be sensitivity and
19 specificity to discriminate subjects as defined by bone
20 densitometry derived T-score data. That is with current
21 practice.

22 We have published data on this in a cohort of
23 women aged 60 to 69 years and looked at cost-effectiveness.
24 We have a second paper now in press looking at age 50 to 54
25 which shows similar analysis.

1 [Slide.]

2 In terms of the quantitative ultrasound devices, I
3 think we all now are aware that the quantitative ultrasound
4 is more diverse than conventional ultrasound densitometry.
5 Devices can measure both cortical and/or cancellous bone,
6 noting that these have dissimilar pathophysiological
7 behavior. There are now a plethora of devices available to
8 the clinician, some with FDA PMA approval.

9 We have two fundamental parameters, the
10 attenuation and velocity. Some of these have different
11 device-specific implementations. There is a consensus in
12 terminology definition and I chaired a European community
13 biomed workshop program that actually looked at this and
14 published the findings.

15 Also, as has already been said, some manufacturers
16 now offer a combined proprietary parameter so the message
17 here is that there is a great deal of variability in this
18 simple term, ultrasound.

19 [Slide.]

20 A question that is often asked is can you apply
21 WHO criteria to quantitative ultrasound noting that T-score
22 is the number of standard deviations below your normal, then
23 the WHO criteria, as we know, is BMC or BMC, osteopenia
24 below -1 but higher than -2.5 and osteoporosis 2.5 or below
25 with established osteoporosis having a fracture.

1 The question that I am still working on is what do
2 we mean by young, what do we mean by normal. We have BMD,
3 BMC. Is that of the hip? Is it of the spine? So, again,
4 as has already been said, there are a number of options of
5 what do we really mean by T-score using WHO criteria.

6 [Slide.]

7 The comparison hypothesis I am going to offer to
8 you I believe can be offered at different sites, between
9 different techniques and between different devices of a same
10 technique, same site. The example I will choose will have a
11 gold standard and, for example, we could use BMD of the
12 spine. We are going to compare the performance of a
13 surrogate measurement that, in this case, is BUA of the
14 calcaneus.

15 [Slide.]

16 The first step is that we assume that there is
17 proportional BMD or BMC. So, what we are saying here is
18 that a subject's BMC or BMD is proportional throughout the
19 skeleton. This means that if a person has a low reading on
20 one measurement, they will have a low reading on the other.
21 If it is high on one, it will be high on the other.

22 In this case, I have proposed that the T-scores
23 for each anatomical site should be equal even if the rates
24 of loss between those different sites are different. So you
25 will have a correlation coefficient of 1 but you will have

1 this proportional change even allowing for bone loss as well
2 as absolute values.

3 [Slide.]

4 If, however, we have what I believe occurs in
5 practice which is we have different skeletal sites that have
6 a non-proportional subject BMD or BMC, we are saying, for
7 example, that with our two BMD devices or sites, et cetera,
8 that we could be low on one and higher with respect to the
9 population than the other measurement, so we have this
10 variability.

11 Therefore, the correlation between the two
12 measurements is not one. Then we have to describe new T-
13 score thresholds, lower here in magnitude than in absolute
14 value, based on sensitivity and specificity analysis. This
15 variability for a particular individual from one site to
16 another or one technique to another will actually be
17 compounded if there is a variable loss.

18 So, if we take a lady who is postmenopausal and
19 one person loses a different percent at hip as opposed to
20 calcaneus compared to another subject, then you are not
21 going to be able to get the same T-score between the two
22 different techniques at the two different measurement sites.

23 [Slide.]

24 So, what we have to do is do surrogate
25 discrimination. If we took here our gold-standard

1 measurement in terms of absolute values and we apply the WHO
2 criteria, so we say anybody who has BMD lower than this
3 level here, a T-score of less than -2.5, is considered
4 osteoporotic, then, using your performance criteria, whether
5 it be accuracy of PPV, then you are actually going to define
6 a different threshold that optimizes either your true
7 positives and true negatives to get accuracy, or the number
8 of people referred for your positive predictive value.

9 I believe that what will happen is that you will
10 actually get a different threshold value and the value of
11 1.5, -1.6, has already been shown today that will actually
12 optimize this performance. So I believe that one of the
13 factors why we are seeing these values used here, different
14 to the -2.5, is for the reason that I have described in the
15 previous two slides.

16 [Slide.]

17 The next question is should we use device-specific
18 parameters or should we use T-scores. I believe that the
19 danger for ultrasound with the device-specific parameter
20 approach is that, at the moment, there is a lack of
21 uniformity between parameters and even between equivalent
22 devices, for example, devices that measure BUA and velocity
23 in the calcaneus.

24 The advantage of the T-score is that it normalizes
25 the data to population and device-parameter variability but

1 it doesn't normalize for anatomical variability and
2 anatomical variability of bone loss as well.

3 [Slide.]

4 When we were looking at the information provided
5 for today's meeting, the terms "ethnicity" and sort of
6 male/female subjects was discussed. I would propose that,
7 for ethnicity, that if fracture risk and disease diagnosis,
8 sort of example osteoporosis, are absolute rather than
9 population-prevalence driven, then there should be a single
10 normative database for each device and parameter.

11 In terms of male subjects as opposed to female
12 subjects, I would propose that we probably are back to the
13 pathophysiological question looking at the difference in the
14 behavior of the two types of osteoporosis.

15 [Slide.]

16 The strategy that I would like to propose for the
17 future is we have to start by defining our purpose. Are we
18 primarily assessing fracture risk or are we primarily
19 diagnosing disease, for example, osteoporosis, or example
20 the WHO criteria.

21 I think we need to define an unambiguous as
22 possible gold standard. This means we have a narrow
23 benchmark. So, instead of having BMC or BMD at hip or
24 spine, et cetera, that we try and narrow this down. I
25 believe that we are going to have to introduce health

1 economics and cost-effectiveness into our future
2 deliberations which may lead us, for example, to consider
3 more hip fracture than other types of fracture.

4 This would lead me, then, to propose that one
5 option would be, perhaps, that a decision to treat would be
6 based upon an age-weighted risk of hip fracture rather than
7 an absolute measurement. This already has been discussed
8 today.

9 I say "age weighted" because I believe that there
10 will be some ages, for example 50 to 60 years of age, that
11 you will want to treat as opposed to 70 or 80 years of age.

12 Thank you.

13 DR. GARRA: Thank you very much.

14 Are there any questions?

15 I have one for you. As far as moving to something
16 like No. 4, age-weighted risk of hip fracture, would it
17 still not be possible--it looks like we have enough data
18 from all the data I have seen presented today, that the
19 physician could still decide the treatment threshold based
20 on whether he was more interested in spine or hip and that
21 those values could be available for--and you could have a
22 complete set of data and calibrations for instruments that
23 have several different types of fracture risk included in
24 them.

25 DR. LANGTON: Yes. I mean I certainly, as a