Only reasonable theory and well-defined concepts can underpin the knowledge on which new scientific applications and procedures (e.g., evidence-based treatments) can be built. Otherwise, any new concept or method, irrespective of its potential, may remain loose and not transferred into the form of truly working knowledge and useful applications.

The common notion that DXA-derived BMD would represent a valid predictor of bone strength is based on high correlation (r values up to 0.9 or even more) between these two variables in controlled laboratory experiments. However, when large multicenter clinical trials on the effects of antiresorptive bisphosphonate therapy (alendronate and risendronate) or selective estrogen receptor modulator therapy (raloxifene) reported consistently greater reductions in fracture incidence than could be anticipated from changes in BMD, speculations on possible additional effects of antiresorptive treatment on something yet unknown, called "Bone Quality", were inspired. "Bone Quality" was defined as "The sum total of characteristics of the bone that influence the bone's resistance to fracture", suggesting that BMD and "Bone Quality" together would independently account for bone fragility in totality (Figures 1A and 1B).

The "Bone Quality" concept has since been embraced among the clinical osteoporosis community, as it seems to offer a panacea for the clinical paradox that while clinical BMD measurements can predict the relative fracture risk at the population level, the predictive value of BMD in individual patients remains quite marginal; one will fracture but another will not, despite similar BMD values. Accordingly, it has been proposed that increased bone turnover (i.e., rapid bone remodeling resulting in negative balance in the bone multicellular unit) would underlie a loss in bone strength and lead to higher susceptibility to fractures through deteriorated bone microarchitecture, a trait that cannot be captured by the BMD measurement. However, as fascinating as this approach might be – filling the information gap in the fracture prediction not revealed by BMD – there are fundamental concerns that need to be elaborated. To begin with, we need to comprehend the complex interrelationship between BMD, bone strength, and fractures (bone fragility).

While there is indeed a high correlation (up to 0.9 or even more) between BMD and bone strength in laboratory testing, one should recall that:
1) Correlation does not mean straight agreement but only group-level association between these two measures; and, the wider the range in these variables, the stronger the group correlation despite unacceptably large individual differences. This means that in two individuals with the same BMD, the actual bone strength can differ considerably, even tens of percents.
2) The high correlation between BMD and bone strength has been obtained in vitro measurements, whereas in a real clinical setting, BMD measurements on individual patients are subject to sizable inaccuracies and substantial uncertainty.
3) Susceptibility to fractures ("bone fragility") is attributable not only to declined bone strength but especially to extraskeletal etiological factors of fractures (falling).

What is BMD?

Given the strong association between DXA-derived areal BMD and bone strength in cadaver biomechanical experiments, it is tempting to conclude that BMD is a valid measure of bone strength. However, considerable inconsistency between DXA-derived "density" and actual whole bone
strength arises from the fact that the DXA measurement is "planar" by nature; that is, DXA scans the three-dimensional structure of bone only from one direction (as if it were a two-dimensional sheet). Accordingly, DXA-derived BMD does not directly represent a volumetric density of any kind, but rather is an ambiguous, lumped parameter depending strongly on volumetric bone mineral apparent density (BMAD) and bone size (i.e., the larger the bone, the higher the BMD at given BMAD)\(^\text{12-15}\) and also on the scan projection (i.e., the thicker the bone in the scan direction, the higher the BMD at given BMAD and bone size)\(^\text{16-19}\). In fact, BMD reflects nothing but the mean thickness of bone mineral within the given bone region without knowledge of the true spatial bone mineral distribution in the depth direction – at its best (Figure 2).

The other crucial limitation of DXA pertains to its inherent inaccuracy which arises from the "two-component assumption" of the method. In short, taking bone material as the first component, DXA assumes that the composition and distribution of all extra- and intraosseus soft tissues and other body constituents within the scanned region constitute an absorptiometrically homogeneous second "component". Not surprisingly, this two-component model does not mirror true anatomy, and thus, DXA-derived BMD is inherently

**Figure 1.** Schematic representation of associations between different factors underlying bone strength and fragility. Traditionally it was simplicistically believed that the DXA-derived areal BMD, given the strong correlation between BMD and bone strength in biomechanical experiments of cadaver bones (r values up to 0.9 or even more), would represent a valid measure of whole bone strength (Figure 1A). Later it was observed in several large multicenter clinical drug trials that the reductions in fracture incidence and BMD are inconsistent indicating that BMD alone is not enough, but BMD together with "Bone Quality" would independently account for bone fragility in totality (Figure 1B). However, given the physical ambiguity of both of these factors\(^\text{15,48}\), it is obvious that BMD and most of the alleged quality characteristics, measurable in vivo, are largely inseparable (Figure 1C). Finally, it must be underscored that an elderly person’s bone fracture depends not only on whole bone strength but especially fall-induced external loading, the latter, in fact, accounting much more for the fragility fractures than bone strength (Figure 1D). Fall biomechanics and the concomitant bone loading is the major determinant of these fractures. For example, over 90% of hip fractures are a direct consequence of falling\(^\text{37,58}\). Of note, this schematic presentation is not an exhaustive description of all possible factors underlying bone strength, bone fragility, and their interactions but is intended to illustrate the complexity that needs to be taken into account when assessing skeletal fragility in general.
inaccurate, either under- or overestimating true BMD to some indeterminate extent in any given patient, the actual magnitude of this error ranging as high as 20 to 50%\textsuperscript{19-23}. The adverse effect of these inaccuracies on the validity of BMD as a measure of whole bone strength of an individual is shown in many biomechanical studies and on the BMD measurement per se in careful cadaver studies (Table 1).

### Table 1. Correlation (\(r\)) between the DXA-derived BMD as obtained in in situ (all soft tissues intact) and ex situ (water bath) measurements of cadaver bones and the failure strength of lumbar vertebrae and femoral neck.

<table>
<thead>
<tr>
<th>DXA site</th>
<th>In situ</th>
<th>Ex situ</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lateral spine</td>
<td>0.45</td>
<td>0.71</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>0.91</td>
<td></td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>0.83</td>
<td></td>
<td>51</td>
</tr>
<tr>
<td>AP spine</td>
<td>0.48</td>
<td>0.51</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>0.53</td>
<td></td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>0.72</td>
<td>0.77</td>
<td>51</td>
</tr>
<tr>
<td>Femoral neck</td>
<td>0.64</td>
<td>0.84</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>0.96</td>
<td></td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>0.89</td>
<td></td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>0.82-0.84</td>
<td>0.94</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.95</td>
<td>57</td>
</tr>
</tbody>
</table>

### BMD – a valid surrogate of whole bone strength?

Since BMD depends on bone cross-sectional size (~load-bearing area) and apparent density (~indicator of bone tissue strength), it is, in principle, mechanically plausible why BMD is a relatively good surrogate of bone strength in a laboratory setting. On the other hand, this dual nature of BMD also complicates its precise interpretation\textsuperscript{15}. Obviously, BMD reflects the amount of bone mass, and bone mass in itself cannot be indicative of actual bone structure, whereas many structural particulars (i.e., overall bone size, cortical thickness and porosity, trabecular thickness and number, mineralization) contribute directly to bone mass. This simply means that BMD and most of the structural (alleged quality) characteristics, measurable in vivo, are largely inseparable (Figure 1C). In other words, BMD is virtually an average measure of almost everything within the measured bone site, but nothing specifically. Accordingly, there is not much left for subtle structural particulars to account for in the statistical sense. So, why don’t we ask what would the bone mass add if we knew the bone structure? At the moment, the paradigm is upside down.

### Bone Quality – the newly introduced concept

As of today, we do not have a universally accepted measurement or indicator of "Bone Quality" (neither a single measure nor a combination of different bone parameters), or a unit for bone quality, or even a criterion for "good or bad..."
bone quality". According to the Webster’s dictionary, the word "quality" is defined as: 1) any of the features that make something what it is; characteristics element; attribute, 2) basic nature; character; kind, 3) the degree of excellence which a thing possesses, or 4) excellence; superiority. In business and industry, the concept "quality" has been classically defined as "fitness for use"24 or "conformance to requirements"25. Coupling the arguments behind the "Bone Quality" concept with different definitions of quality, one is inherently left with the impression that a good quality of bone equates to good bone strength, or even as broadly as high resistance to fractures.

**Bone Strength = BMD + Bone Quality?**

At present, "Bone Quality" denotes an obscure term incorporating a pool of various non-BMD indices of bone strength. Needless to say, it is impossible to measure directly the bone strength in patients in vivo, whereas in pre- clinical animal experiments the bone strength can be measured. Then, if "Bone Quality" had accounted independently for the bone fragility (as implied by several clinical fracture59 prevention studies), we should have observed the similar discordance between the changes in bone mineral density or mass and bone strength in numerous well-controlled preclinical studies. However, the discrepancy is not there! In fact, the effects of bisphosphonates on bone strength have been perfectly commensurate with changes in bone mass in virtually life-long animal experiments of different species26-32.

**Bone fragility (Resistance to fractures) = BMD + Bone Quality?**

More recently, the term "Bone Quality" has been defined as broadly as: "Bone quality recognizes the unpredicted portion of fracture risk with respect to the predicting variables"83. In essence, this "extended definition of Bone Quality" not only suggests that together BMD and "Bone Quality" would independently account for bone fragility in totality, but also implies that the bones are primarily designed and adapted to resist fractures. However, there are major conceptual flaws in this approach, too.

First, while the high correlation between BMD and bone strength in the laboratory setting, as well as the increased relative risk of osteoporotic fractures among patients with low BMD, are indeed established, one should keep in mind that at various skeletal sites the overall proportion of elderly people’s fractures attributable to low BMD remains modest (ranging from <10% to 44%)8. In other words, more than 50% of fragility fractures occur in the population which is not classified as osteoporotic in the sense of the current WHO operational definition of osteoporosis (i.e., BMD 2.5 standard deviations or more below the young sex-matched adult reference level). In this respect, we have to recall that BMD in itself is only a modest risk factor of fractures - some 85% of the contribution to the rise in fracture risk with age is unrelated to BMD84.

According to several well-designed studies on risk factors of fractures among elderly people36,43, falling with its determinants - not the BMD-based osteoporosis - has been shown to be the strongest single risk factor for a fracture, and when a person falls, the type and severity of falling (fall height and energy; fall direction; fall mechanics; anatomical site of the impact; and energy absorption capacity and impact force attenuation of the body-landing surface complex) are crucial in determining whether or not a fracture occurs85,40. Compared with the modest association between BMD and risk of fracture, the relative risk of hip fracture for a sideways fall is about 5, and if the fall impacts directly the greater trochanter of the proximal femur, the relative risk rises up to 3047,41,42. Similar results have been obtained for upper extremity fractures43,45. With this in mind, it is quite utopian to envision that a purely bone-derived measure (such as BMD or bone turnover markers) alone could explain the occurrence of fractures, as implied by the broadest definitions of "Bone Quality" at present.

Now, regarding the notion that bones would primarily be designed and adapted to resist fractures one should recall the following facts. The human skeleton is basically and continuously adapted to habitual locomotive loadings (Figure 3A), and is particularly fit for endurance running44, but not to loadings caused by falls onto the ground (or by other similar trauma-related events) that cause fractures45. Clearly, there is a tradeoff between the bone strength and other factors, including the metabolic pressure to keep the weight of these locomotive organs light. In support of this, there is overwhelming evidence that the variation in the apparent strength of human bones is attributable to variations in the loading environment the bones are subjected to during daily habitual activities85,46. As regards the capacity of the skeleton to resist fracture during accidents one must distinguish between two situations: whether the loading experienced during a traumatic incident is just a magnification of the loading experienced during habitual activities (but in this case just exceeding the bone’s capacity to withstand the loading) (Figure 3B), or the loading and its direction are completely different from that the bones are customarily adapted to. In many cases of older adults’ fractures (even those of vertebrae that are commonly considered “spontaneous”), they are indeed different (Figure 3C)45.

**Are we hostages to the ambiguous BMD?**

Since the whole bone strength provides the ultimate measure of bone mechanical competence, the clinical osteoporosis paradox seems to simply stem from our inherent inability to determine the actual bone strength or fragility of an individual in vivo. We should not remain hostages to the ambiguous BMD and cursorily reduce the bone fragility into two seemingly independent factors: 1) The familiar/conventional BMD, and 2) the new and fascinating "Bone Quality" denoting the vague provisions of (all) non-BMD factors,
which together would perfectly explain the bone fragility (Figure 1B).

As a properly defined concept, the "Bone Quality" should equate nothing but the capacity of bones to withstand a wide range of loading without breaking. Because bone structure is the ultimate determinant of whole bone mechanical competence and pertains to the endpoint of all interim biochemical processes within the bone tissue\(^4\), the adequately comprehensive in vivo assessment of the bone structure is what we should pursue.

"We measure things because we CAN. Accordingly, it is crucial to fully understand the context of one's measurements."

Anonymous

Bone Quality – An empty term

We can assess the whole bone strength quite well in a laboratory setting. In real life, our abilities to do so are much more limited. Nevertheless, the mere inability to do so with actual patients in vivo cannot be taken as a justification to introduce new, obscure and ill-defined terms that are anticipated to fill the bill. Launching a new ambiguous concept with loose attachment to the actual clinical problem will not facilitate its solution – on the contrary, it can cause confusion and exacerbate the situation at its worst. Prediction of a fracture of an individual (or any other medical event, such as heart attack or stroke) will be a formidable task given the
overwhelming dominance of chance over all measurable "risk factors". No matter how sophisticated methods we use and how hard we try, the life of an individual is often just too complicated to be predicted and it will always run its own, largely unforeseen paths.

Taken together, we see that the term "Bone Quality" is an empty term\(^\text{8}\) – identical to Emperor’s new clothes in the famous fairy tale – and thus, should be abandoned. If it really must be used, the term "Bone Quality" should, as said, refer only to the capacity of bones to withstand a wide range of loading without breaking. And, for such capacity we already have a proper term or the whole bone strength\(^\text{7}\). Our real challenge will be to reliably estimate the whole bone strength \textit{in vivo}.

Acknowledgements

The study was supported by the Competitive Research Funding of the Pirkanmaa Hospital District, the Research Council for Physical Education and Sports, Ministry of Education, Academy of Finland, and the AO Research Fund, Switzerland.

References


