

---

**Imaging the femoral cortex: thickness,  
density and mass from clinical CT**

G.M. Treece, K.E.S. Poole and A.H. Gee

**CUED/F-INFENG/TR 665**

29 April 2011

Cambridge University Engineering Department  
Trumpington Street  
Cambridge CB2 1PZ  
England

Corresponding e-mail: [gmt11@eng.cam.ac.uk](mailto:gmt11@eng.cam.ac.uk)

---

## Abstract

There is growing evidence that focal thinning of cortical bone in the proximal femur may predispose a hip to fracture. Detecting such defects in clinical CT is challenging, since cortices may be significantly thinner than the imaging system's point spread function. We recently proposed a model-fitting technique to measure sub-millimetre cortices, an ill-posed problem which was regularized by assuming a specific, fixed value for the cortical density. In this paper, we develop the work further by proposing and evaluating a more rigorous method for estimating the constant cortical density, and extend the paradigm to encompass the mapping of cortical mass (mineral  $\text{mg}/\text{cm}^2$ ) in addition to thickness. Density, thickness and mass estimates are evaluated on sixteen cadaveric femurs, with high resolution measurements from a micro-CT scanner providing the gold standard. The results demonstrate robust, accurate measurement of peak cortical density and cortical mass. Cortical thickness errors are confined to regions of thin cortex and are bounded by the extent to which the local density deviates from the peak, averaging 20% for 0.5 mm cortex.

## 1 Introduction

Hip fractures are the most common cause for acute orthopaedic hospital admission in older people (Parker and Johansen, 2006), with their annual incidence projected to rise worldwide from 1.7 million in 1990 to 6.3 million in 2050 (Sambrook and Cooper, 2006). Bone mineral density (BMD) is currently the imaging biomarker of choice for assessing an individual's fracture risk, but although it is specific (Johnell et al., 2005; Kanis et al., 2008) it lacks sensitivity (Kanis et al., 2008; Kaptoge et al., 2008; Sanders et al., 2006), missing the majority who go on to fracture. There is now growing evidence that focal, structural weaknesses may predispose a hip to fracture (Mayhew et al., 2005; Poole et al., 2010; de Bakker et al., 2009) and there is a consequent need to develop novel imaging methods capable of detecting such weaknesses, with multi-detector computed tomography (MDCT) the favoured modality (Bouxsein and Delmas, 2008).

While trabecular bone undoubtedly plays some role, it is the distribution of cortical bone that is believed to be critical in determining a femur's resistance to fracture (Holzer et al., 2009; Verhulp et al., 2008). Compressive cracking of the cortex in the femoral neck or trochanter is often the first point of failure (de Bakker et al., 2009; Carpenter et al., 2005; Mayhew et al., 2005). Unfortunately, thin laminar structures such as the femoral cortex are not accurately depicted in clinical CT because of the images' limited spatial resolution. Consequently, straightforward thickness estimation techniques, such as those based on thresholding (Buie et al., 2007; Hangartner, 2007) or some measure of full-width half-maximum (Prevrhal et al., 1999, 2003), are unreliable when the cortex is thin in relation to the imaging resolution. With normal bore, clinical CT scanners, such methods are increasingly inaccurate below around 2.5 mm (Dougherty and Newman, 1999; Hangartner and Gilsanz, 1996), with errors exceeding 100% for sub-millimetre cortices (Prevrhal et al., 2003).

We have recently proposed a more sophisticated technique that produces good estimates of cortical thickness down to 0.3 mm (Treece et al., 2010). In common with other attempts at deblurring medical images of laminar structures (Streekstra et al., 2007), we adopt restrictive models of both the object being scanned and the imaging system, and then attempt to fit these models to the observed data. This process is inevitably ill-posed, since a dense, thin, blurred cortex might appear identical to a less dense, less thin, blurred cortex. It is therefore necessary to incorporate some prior knowledge about either the density or the blur, and we found in Treece et al. (2010) that assuming a specific, fixed value for the density was more successful than assuming a constant blur. We subsequently applied this technique to show precisely where in the proximal femur new bone is laid down following two years' treatment with bone anabolic drugs (Poole et al., 2011).

In this paper, we present a significant extension of this work to encompass not only thickness, but also density and mass estimation. In Section 2, we review the theory behind constant density thickness estimation and develop a more rigorous approach to establish the critical density on which the method relies. In so doing, we discover an alternative technique for thickness estimation, and show how both variants may be trivially

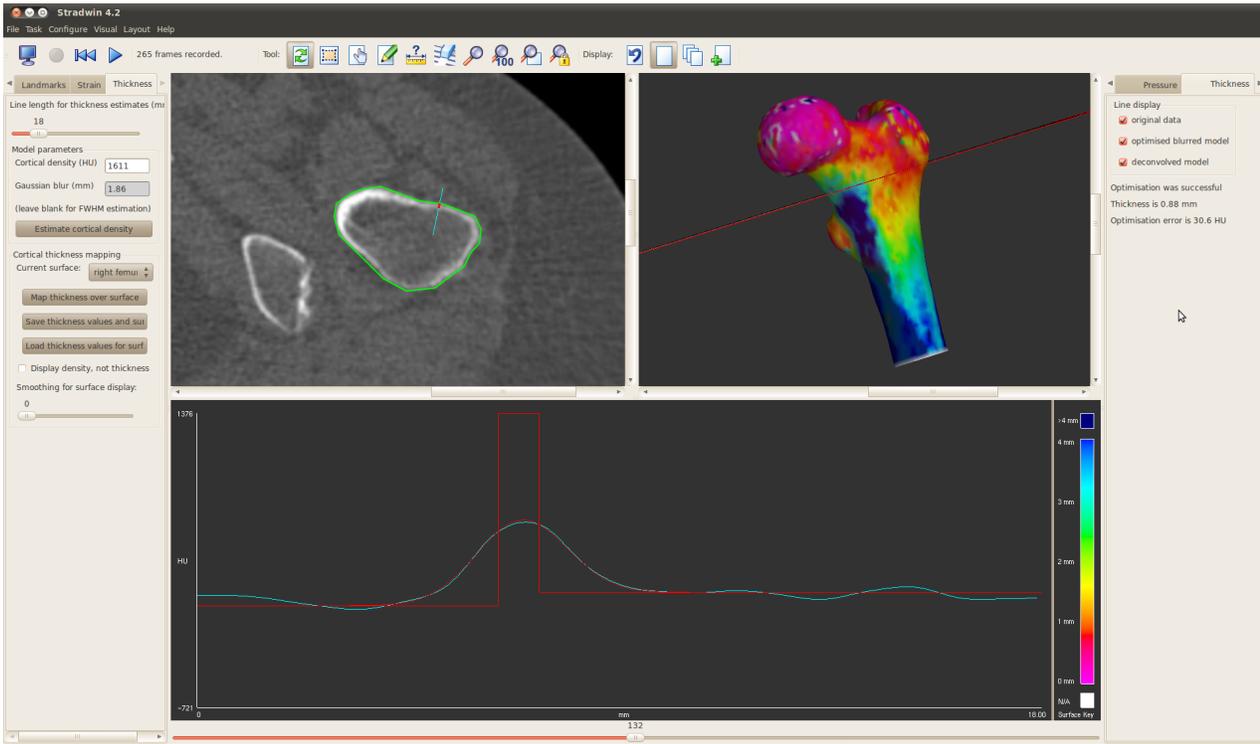


Figure 1: Cortical thickness estimation from clinical CT data. Given a prior segmentation of the proximal femur (green contour, top left), CT values are examined along short lines (cyan, top left) that straddle the cortex and are perpendicular to it. The sampled CT values are shown in cyan in the bottom panel. Cortical thickness is estimated by assuming, in this example, a constant cortical density  $\hat{y}_1$  of 1611 HU. The Levenberg-Marquardt algorithm (More, 1977) is then used to find the imaging blur  $\sigma$ , tissue density  $y_0^{\text{cd}}$ , trabecular density  $y_2^{\text{cd}}$  and thickness  $t^{\text{cd}}$  that best explain the data. The solution, and the idealised data  $y_{\text{blur}}$  that it implies, are shown in red in the lower panel. By repeating this process at a large number of points, the cortical thickness can be mapped across the entire surface (top right).

extended to measure cortical mass in addition to thickness. The methods are evaluated in Section 3 with the aid of sixteen cadaveric femurs which were imaged in a high resolution, peripheral quantitative CT (pQCT) system to establish ground truth thickness and mass. Section 3 also includes some simulations to confirm the causes of the subtle effects observed in the cadaveric studies. In Section 4, we discuss the relative merits of mass and thickness estimation, and how best to estimate density, in the illuminating context of our earlier work on anabolic drug responses. Finally, we draw some conclusions in Section 5.

## 2 Method

### 2.1 Estimating cortical thickness and mass in blurred data

Figure 1 illustrates the process of cortical thickness estimation, as implemented in our free-to-download Stradwin software<sup>1</sup>. Given a prior segmentation of the proximal femur, the CT data is sampled along short lines perpendicular to the femoral surface, at a large number of points on the surface. The lower panel in Figure 1

<sup>1</sup><http://mi.eng.cam.ac.uk/~rwp/stradwin>

shows the data from one such point (cyan), along with the model that best explains the data (red). In this model, the underlying density distribution  $y(x, t)$  along the line is assumed to be

$$y(x, t) = y_0 + (y_1 - y_0)H\left(x + \frac{t}{2}\right) + (y_2 - y_1)H\left(x - \frac{t}{2}\right) \quad (1)$$

where  $y_0$ ,  $y_1$  and  $y_2$  are the CT values in the surrounding tissue, cortex and trabecular bone respectively,  $H$  is the unit step function,  $t$  is the cortical thickness and  $x$  is measured from the centre of the cortex. The impulse response  $g(x)$  of the imaging system (image blur) is taken to be Gaussian with standard deviation  $\sigma$ ,

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}} \quad (2)$$

and hence the step response  $h(x)$  of the imaging system is

$$h(x) = \frac{1}{2} \left[ 1 + \operatorname{erf}\left(\frac{x}{\sigma\sqrt{2}}\right) \right] \quad (3)$$

Combining equations (3) and (1) gives the blurred cortex  $y_{\text{blur}}(x, t)$  as

$$y_{\text{blur}}(x, t) = y_0 + \frac{y_1 - y_0}{2} \left[ 1 + \operatorname{erf}\left(\frac{x + \frac{t}{2}}{\sigma\sqrt{2}}\right) \right] + \frac{y_2 - y_1}{2} \left[ 1 + \operatorname{erf}\left(\frac{x - \frac{t}{2}}{\sigma\sqrt{2}}\right) \right] \quad (4)$$

All that remains is to find the model parameters that minimize the sum of squared errors between  $y_{\text{blur}}$  and the actual CT data. As mentioned previously, this is an ill-posed problem that requires regularization, and we have achieved good results by adopting a fixed value  $\hat{y}_1$  for the cortical density (Trecede et al., 2010). The underlying assumption is that  $y_1$  does not vary significantly at different points on the proximal femur, at least in comparison with the much larger variation in cortical thickness. In all our work to date, we have estimated  $\hat{y}_1$  from regions of thick cortex, typically below the lesser trochanter on the femoral shaft, where the apparent peak density is not affected by the imaging blur.

We are now in a position to elaborate in more detail the aims and objectives of this paper. Firstly, we wish to further test the validity of the constant density assumption by measuring true density in high resolution (i.e. very low blur) CT scans of cadaveric femurs. While we have previously shown that cortical thickness estimates are relatively insensitive to the value of  $\hat{y}_1$  (Trecede et al., 2010), it would nevertheless be informative to examine how  $y_1$  varies with location on the proximal femur, and by how much. Secondly, we will develop and assess a more rigorous method for estimating  $\hat{y}_1$  using all the available data, not just a single measurement at a region of thick cortex. Finally, given that underestimating  $\hat{y}_1$  results in overestimating thickness  $t$ , we investigate whether we can measure cortical mass per unit area more accurately than thickness alone. A mass estimate of this nature is essentially the product of  $\hat{y}_1$  and  $t$ , though we additionally need to convert image density (HU) to mineral density ( $\text{mg}/\text{cm}^3$ ) using the calibration phantom included in every quantitative CT examination (Cann, 1988). Although the resulting measure has units  $\text{mg}/\text{cm}^2$ , we shall refer to it loosely as ‘‘mass’’ in the remainder of this paper. Dougherty and Newman (1999) speak of cortical mass as a potentially useful indicator of regional bone strength.

In the mathematical notation adopted throughout this paper, superscripts denote the estimation technique. Thus,  $t^{\text{cd}}$  is the cortical thickness estimate obtained using the constant density technique, and  $m^{\text{cd}} = t^{\text{cd}}\hat{y}_1$  is the corresponding mass estimate. As Figure 2 (right) shows, the fitted model also produces estimates of the tissue density  $y_0^{\text{cd}}$ , the trabecular density  $y_2^{\text{cd}}$ , and the average background density  $y_b^{\text{cd}} = (y_0^{\text{cd}} + y_2^{\text{cd}})/2$ . For comparison, we also consider the full-width half-maximum measure of cortical thickness (Prevrhal et al.,

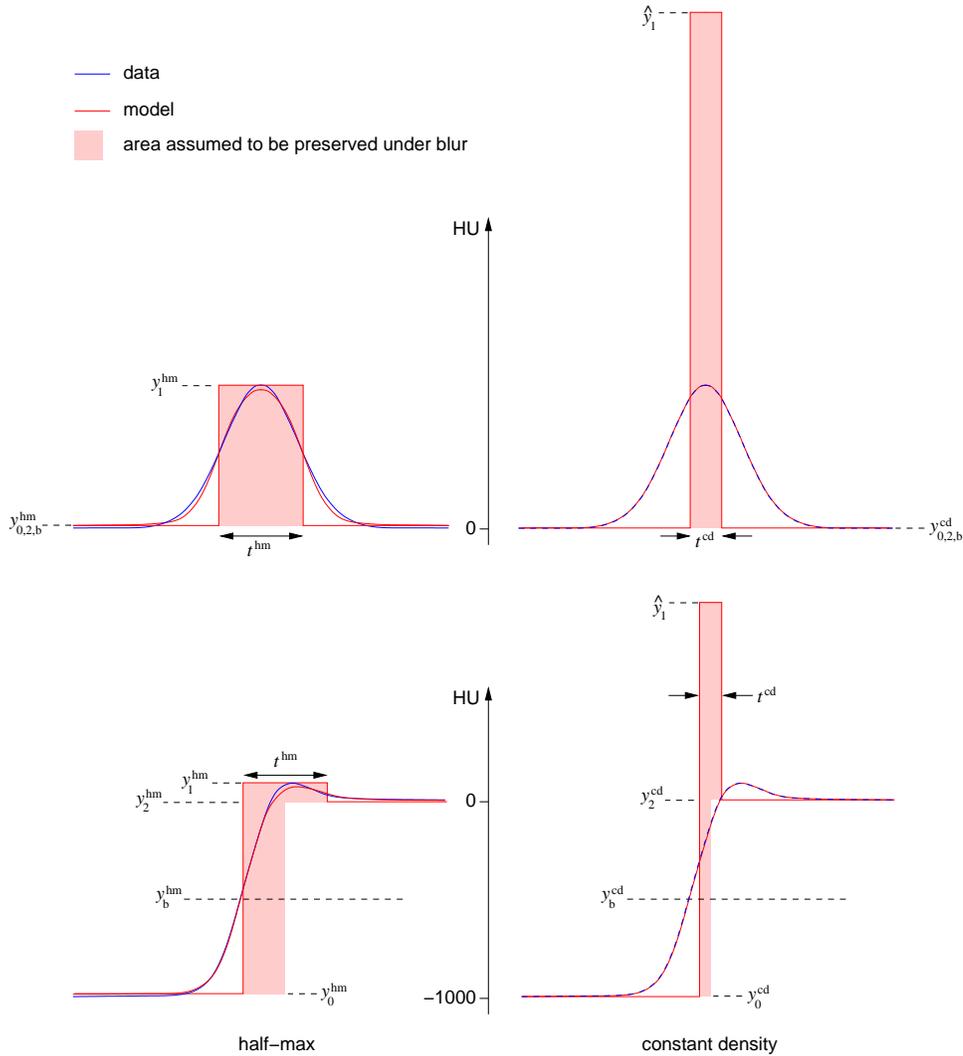


Figure 2: The constant density and half-max estimation techniques. The upper two sub-figures illustrate the simple case where  $y_0 = y_2$ , with a thin, dense cortex (red rectangle, top right) that appears less thin and less dense under image blur (blue). The constant density technique starts with an estimate  $\hat{y}_1$  of the true cortical density, and then finds the other model parameters that minimize the sum of squared errors between the data (blue) and  $y_{\text{blur}}$  (red curve). Details of the optimization process can be found in [Trece et al. \(2010\)](#). In this case the fit is perfect, since the simulated data is a pure Gaussian, but see Figure 1 for typical clinical data. The half-max technique uses a similar optimization process, except the true cortical density is assumed to be  $y_1^{\text{hm}}$ , the apparent peak density. The edges of the fitted cortex then coincide with the half-way points between  $y_b^{\text{hm}}$  and  $y_1^{\text{hm}}$ . The shaded areas are equivalent to the areas under the red Gaussians and are therefore approximately equal: the half-max area is actually around 6% smaller because the incorrect model does not precisely explain the data. This mass preservation property lies behind the hybrid constant density/half-max approach. When  $y_0 < y_2$  (lower two sub-figures), the apparent density peak lies some distance to the right of the actual cortex. While the half-max model still fits well to produce the  $t^{\text{hm}}$  estimate, the red areas are no longer equivalent since the half-max area is predicated on a  $y_0$ - $y_2$  transition at the wrong location. This leads to significant overestimation of mass and thickness in the hybrid approach, by approximately 60% in this case.

1999, 2003), which we shall henceforth abbreviate to the “half-max” method<sup>2</sup>. This estimates the thickness by looking at the half-way points between the apparent peak density  $y_1^{\text{hm}}$  and each of the background densities  $y_0^{\text{hm}}$  and  $y_2^{\text{hm}}$ . In practice, we estimate these parameters by optimizing a model fit, in exactly the same way as for constant density, except we fix  $y_1$  at the apparent peak density  $y_1^{\text{hm}}$  instead of  $\hat{y}_1$ : see Figure 2 (left).

The half-max method is known to produce highly inaccurate thickness estimates when the cortex is thin, since even the narrowest density spike is blurred to the extent of  $g(x)$ . For the straightforward case where  $y_0 = y_2$ , it is trivial to deduce that  $t^{\text{hm}} \rightarrow 2.35\sigma$  in the limit of small  $t$ . However, as previous studies have observed (Dougherty and Newman, 1999), the apparent extra mass of the cortex over and above the background is preserved under blur, and we might seek to exploit this property to improve the half-max estimate. For the simple case where the background is constant, so  $y_0 = y_2 = y_b$  (Figure 2, top), mass preservation implies

$$t^{\text{cd}}(\hat{y}_1 - y_b^{\text{cd}}) = t^{\text{hm}}(y_1^{\text{hm}} - y_b^{\text{hm}}) \quad (5)$$

Given  $\hat{y}_1$ , we can therefore adjust the half-max thickness estimate and also obtain a mass estimate, resulting in a hybrid constant density/half max technique:

$$t^{\text{cd-hm}} = t^{\text{hm}} \frac{(y_1^{\text{hm}} - y_b^{\text{hm}})}{(\hat{y}_1 - y_b^{\text{hm}})} \quad (6)$$

$$m^{\text{cd-hm}} = t^{\text{cd-hm}} \hat{y}_1 \quad (7)$$

The situation is not so straightforward when  $y_0 < y_2$ . Applying equation (5) in this case is equivalent to preserving the areas shaded red in Figure 2 (bottom), but the half-max area is predicated on the cortex being centred at the apparent density peak, whereas the cortex is actually located a little to the left of this, as in Figure 2 (bottom right). By shifting the  $y_0$ - $y_2$  step in the half-max area to the left, we would obtain an area that is indeed preserved under blur, but we have no way of knowing the required correction *a priori*, since it depends on the true cortical thickness  $t$ . Equations (6) and (7) therefore represent a best guess, though they clearly overestimate mass and thickness in situations like Figure 2 (bottom).

## 2.2 Estimating cortical density in blurred data

The constant density and hybrid techniques require an estimate  $\hat{y}_1$  of the true cortical density  $y_1$ . Our early approach was to sample the peak density at a region of thick cortex, but considering all the blurred data, and not just a single datum, should increase the accuracy and precision of the estimate. According to the model in equation (4), and noting that  $\text{erf}(x) = -\text{erf}(-x)$ , the apparent density at the centre of the cortex is given by

$$y_{\text{peak}} = y_{\text{blur}}(0, t) = y_b + (y_1 - y_b) \text{erf}\left(\frac{t}{2\sqrt{2}\sigma}\right) \quad (8)$$

Equation (8) describes the *expected* distribution of apparent peak density  $y_{\text{peak}}$  with thickness  $t$ , given  $y_b$ ,  $y_1$  and  $\sigma$ . If we could plot the *observed* distribution, by measuring apparent peak density and thickness at each point on the surface, we could then estimate  $y_b$ ,  $y_1$  and  $\sigma$  by minimizing the differences between the two distributions. In contrast with our earlier approach, we would be making use of all the data and might expect a reasonable estimate even when there are no regions of cortex sufficiently thick to reveal the true density  $y_1$  in the blurred data.

The details of the algorithm are as follows. The first step is to use the half-max technique to measure apparent thickness  $t^{\text{hm}}$ , apparent peak density  $y_1^{\text{hm}}$  and background density  $y_b^{\text{hm}}$  at all points on the femoral

<sup>2</sup>The reader is referred to Treece et al. (2010) for further comparisons with the alternative thresholding approach (Buie et al., 2007; Hangartner, 2007), which is less precise than the half-max method when the cortex is thick, and increasingly prone to missing the cortex altogether when it is thin.

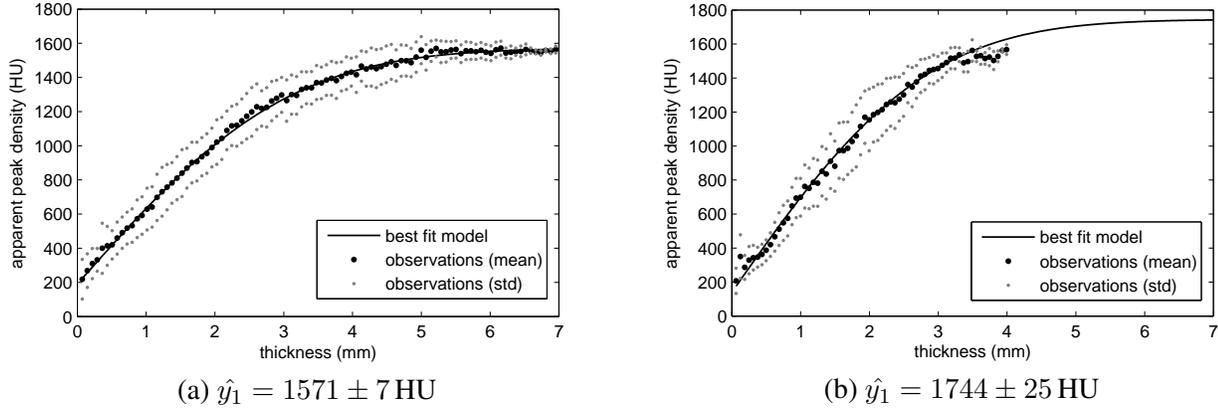


Figure 3: Estimating the cortical density. For representative low resolution CT data sets, the graphs show the observed distribution of apparent peak density with thickness (dots), and the best fit model according to equation (8) (black line). For each bin, the mean (black dots) and  $\pm$  one standard deviation (grey dots) of the observations are displayed. The three model parameters are found by minimizing the sum of the squared differences between the mean observations and the model. (a) In this case, there is sufficient data to constrain the model parameters within a narrow range. (b) This scan did not extend far enough below the lesser trochanter to encounter thick cortical bone, so there is more uncertainty in the model parameters.

surface. These observations are grouped into  $N$  bins indexed by the product  $p(i) = t^{\text{hm}}(y_1^{\text{hm}} - y_b^{\text{hm}})$ ,  $i \in \{1 \dots N\}$ , and the average value  $y_1^{\text{hm}}(i)$  is recorded in each bin. Estimates  $\hat{y}_b$ ,  $\hat{y}_1$  and  $\hat{\sigma}$  are then refined by minimizing the sum of the squared errors  $\epsilon(i)$  between the observed and expected distributions at each bin  $i$ :

$$t(i) = \frac{p(i)}{(\hat{y}_1 - \hat{y}_b)} \quad (9)$$

$$y_{\text{peak}}(i) = \hat{y}_b + (\hat{y}_1 - \hat{y}_b) \operatorname{erf}\left(\frac{t(i)}{2\sqrt{2}\hat{\sigma}}\right) \quad (10)$$

$$\epsilon(i) = y_1^{\text{hm}}(i) - y_{\text{peak}}(i) \quad (11)$$

Equation (9) converts the bin index  $p(i)$  into an estimate of true thickness  $t(i)$  using the hybrid technique of equation (6). Equation (10) calculates the expected apparent peak density  $y_{\text{peak}}(i)$  for this thickness. Finally, equation (11) calculates the difference between the observed and expected peak densities. These three equations define the errors which we minimize by Levenberg-Marquardt optimization (More, 1977) to obtain optimal estimates  $\hat{y}_b$ ,  $\hat{y}_1$  and  $\hat{\sigma}$ . A typical outcome is shown in Figure 3(a), indicating good agreement between the observations and the model. However, when the acquired data extends less far down the femur, or when the bone is osteoporotic, there may be insufficient thick cortex to fully constrain the model: such a case is shown in Figure 3(b). Scenarios (a) and (b) may be distinguished objectively by calculating the standard error of fit for the model parameter  $\hat{y}_1$ . In (a)  $\hat{y}_1 = 1571 \pm 7$  HU, in (b)  $\hat{y}_1 = 1774 \pm 25$  HU. The ranges are 95% confidence intervals, which are equal to twice the standard error of fit assuming the data is normally distributed and representative of the population (Press et al., 2002). It is, however, necessary to decide subjectively how much uncertainty to tolerate in any particular study. When the confidence intervals exceed this limit, the safest course of action is simply to exclude the scan from cortical thickness analysis.

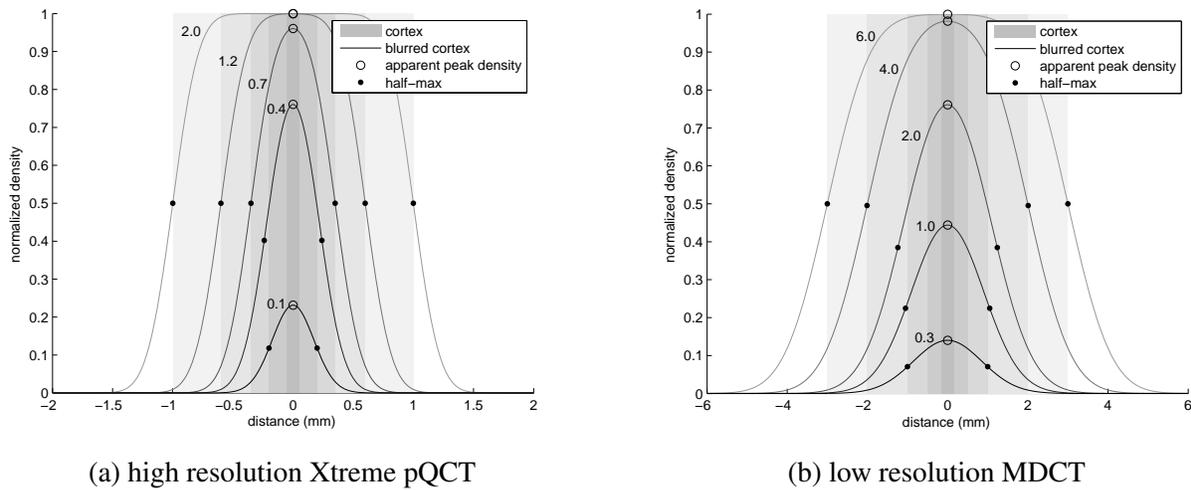


Figure 4: High and low resolution data. For the straightforward case of uniform background ( $y_0 = y_2 = y_b$ ), these figures illustrate the expected effects of Gaussian blur on half-max thickness estimates and apparent peak density. (a) High resolution Xtreme pQCT. (b) Low resolution MDCT. Each curve is labelled with the corresponding true cortical thickness (mm). Note the different scales on the two distance axes.

## 2.3 Evaluation

Sixteen cadaveric femurs were scanned in air at both high and low resolutions. The high resolution data was acquired on an Xtreme pQCT machine (Scanco Medical AG, Brüttisellen, Switzerland) at  $82 \mu\text{m}/\text{pixel}$  in-plane resolution and  $82 \mu\text{m}$  slice thickness. In these scans, the appearance of all but the thinnest cortices is largely unaffected by blur. We can therefore treat simple half-max thickness and density estimates, obtained from this high resolution data, as a gold standard. The half-max method is unbiased as long as the cortex is sufficiently wide compared with the imaging blur (Prevrhal et al., 1999), a condition that is satisfied down to around 0.4 mm thickness: see Figure 4(a). The low resolution data was acquired on a Siemens Somatom Sensation 64 MDCT machine (Siemens AG, Erlangen, Germany) at  $589 \mu\text{m}/\text{pixel}$  in-plane resolution and 1 mm slice thickness. This is typical of clinical scanning conditions and serves as a test bed for the constant density and hybrid estimation methods. Image blur is far more severe, as shown in Figure 4(b).

Several precautions are necessary when comparing cortical measurements in corresponding high and low resolution data. First, the two femoral surfaces must be aligned using some sort of registration technique: we use the iterative closest point algorithm described in Treece et al. (2010). Thickness and mass estimates are made at every vertex of the low resolution surface mesh, using CT values sampled on 18 mm lines (the cyan lines in Figure 1) oriented along the corresponding surface normals. These measurement locations and directions are then mapped onto the aligned high resolution surface, allowing us to compare the same number of thickness and mass estimates, taken at the same locations and using the same surface normals.

The next precaution concerns the spatial localization of the high and low resolution estimates. Figures 5(a) and (b) show corresponding patches of cortex from high and low resolution scans respectively. High resolution thickness estimates would vary rapidly from the top to the bottom of the image, picking up every small detail of the irregular cortex. This detail is blurred out in the low resolution data: thickness estimates at any particular location, however well adjusted for trans-cortical blur, are always going to reflect the *average* cortical thickness in the vicinity of the estimation location, not the *particular* thickness at that location. We therefore blur the high resolution data *in one dimension only, parallel to the cortex*, as shown in Figure 5(c), before estimating thickness and mass. Note that this blur does not compromise estimation accuracy, only

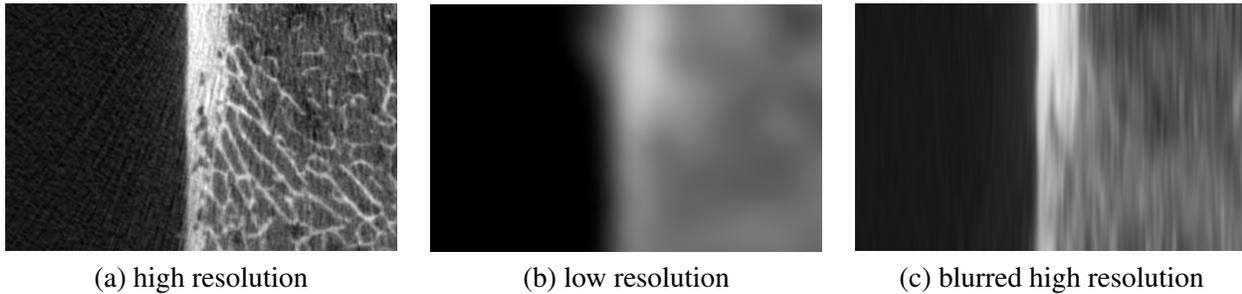


Figure 5: Blurring of high resolution data to match low resolution thickness localization. (a) High resolution data, (b) corresponding low resolution data and (c) the blurred high resolution data used in the comparisons. The extent of the smoothing kernel is chosen to match the spatial resolution of the low resolution data.

estimation localization, which is exactly what is required for a meaningful comparison between the high and low resolution data.

The final precaution concerns the sometimes ambiguous nature of the endocortical surface. Figure 6(a) shows some high resolution data where the location of the  $y_1$ - $y_2$  transition is far from clear. Is this a wide cortex encompassing a low density pore, or a narrow cortex with a nearby trabecular peak? For a fair comparison between the high and low resolution estimates, we must ensure that the two measurements opt for the same interpretation. We encourage this behaviour when fitting the model to the high resolution data, by blurring the *error* between the observed and modelled data within the Levenberg-Marquardt algorithm. Again, the extent of the blur is chosen to match the spatial resolution of the low resolution data. In the example in Figure 6(a), blurring the optimization error causes the algorithm to converge to the alternative solution, as it does with the low resolution data. There is little effect when the cortex is well defined, as in Figure 6(b).

What we have described is a refinement of the evaluation protocol described in our earlier study (Treece *et al.*, 2010), where we (correctly) attributed extreme errors to mismatched normals or different interpretations of the endocortical surface, and excluded outliers from the subsequent error analysis. With our new evaluation protocol, we no longer see extreme errors and outlier rejection is not required. In the next section, we present a series of experiments and results that utilize this protocol to: validate the derived distribution of apparent peak density with thickness across all sixteen femurs; compare subsequent thickness and mass estimates from the various estimators with the gold standard high resolution measurements; quantify typical intra-subject cortical density variation; and correlate the observed density variations with thickness estimation errors. Some of the results are best presented as colour maps on a representative (“canonical”) femur surface. To obtain these visualizations, each of the sixteen individual surfaces was first warped onto the canonical surface using a nonrigid, free form deformation. Having established a common morphology, the sixteen sets of results could then be averaged to show typical distributions across the femoral surface.

### 3 Experiments and results

#### 3.1 Density variation and estimation in low and high resolution data

Figure 7(a) shows how cortical density varies with thickness in the sixteen cadaveric femurs. Since these measurements were made using the half-max method on the high resolution data, they can be trusted down to around 0.7 mm, beyond which Figure 4(a) shows blur affecting the apparent peak density. The sharp density reduction at the far left of the distribution can therefore be attributed to measurement error. Above 0.7 mm, we observe fairly constant density where the cortex is relatively thick, but some decline below 3 mm. This is not a measurement artefact but a feature of the bones.

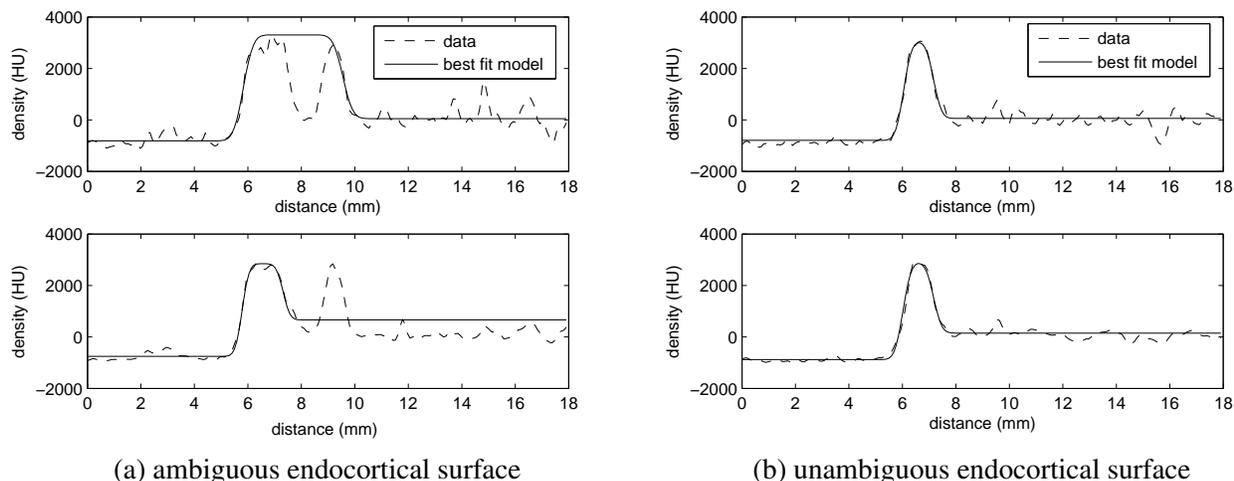


Figure 6: Consistent interpretation of the endocortical boundary in high and low resolution data. By blurring the *error* when fitting a model to the high resolution data, we encourage the same interpretation of the endocortical boundary as with the low resolution data. (a) The top graph shows sample data through a cortex from a high resolution scan, along with the best fit half-max model without error blurring. The bottom graph shows the same data after blurring as in Figure 5, and the best fit half-max model after blurring the optimization error. While the cortical extent is different, reflecting the alternative interpretation of the ambiguous data, the edge localization accuracy is unaffected. (b) When the cortex is well defined, neither of these blurring operations has a significant impact on the measured thickness.

Figure 7(b) shows a similar plot, but this time with the corresponding cortical densities observed in the low resolution data. As anticipated in Figure 4(b), image blur starts affecting the apparent peak density below around 3.5 mm, but the distribution appears to follow the form of equation (8) (c.f. Figure 3) and we can therefore expect successful estimation of  $\hat{y}_1$  using the model-fitting approach of Section 2.2. This hypothesis is confirmed in Figure 8, which shows the stability of the  $\hat{\sigma}$  and  $\hat{y}_1$  estimates as progressively more measurements, starting from the top of the femoral head, are fed into the model-fitting procedure. The estimates are highly variable when too few observations are used, but both  $\hat{\sigma}$  and  $\hat{y}_1$  have converged to stable values by the time measurements are incorporated from just below the lesser trochanter. Adding yet more data from lower down the femoral shaft appears to make very little difference for these sixteen femurs. With marginally sufficient data, the model-fitting procedure tends to overestimate  $\hat{y}_1$  by a few percent — see the cyan band above the lesser trochanter in Figure 8(c). While Figure 8 reflects the healthy norm, osteoporotic specimens may need scanning a little lower before sufficiently thick cortex is encountered to fully constrain the model parameters. Unreliable solutions are not an undetectable hazard: on the contrary, they are readily identified by examining the confidence intervals, as demonstrated in Figures 8(b) and (d).

### 3.2 Mass and thickness estimation using optimized density

Figure 9 summarises the performance of constant density thickness estimation. It is apparent, at first sight, that the low resolution thickness estimates in (b) are good surrogates for the gold standard, high resolution measurements in (a). As expected, estimation errors (d) are correlated with true cortical density (c). We have already seen in Figure 7(a) that the cortical density is not perfectly constant: at least in these sixteen femurs, it tends to drop off where the cortex is thinner. We can now see in Figure 9(c) where these less dense regions of cortex are located. Furthermore, since Figure 9(c) is normalized by the high resolution density that

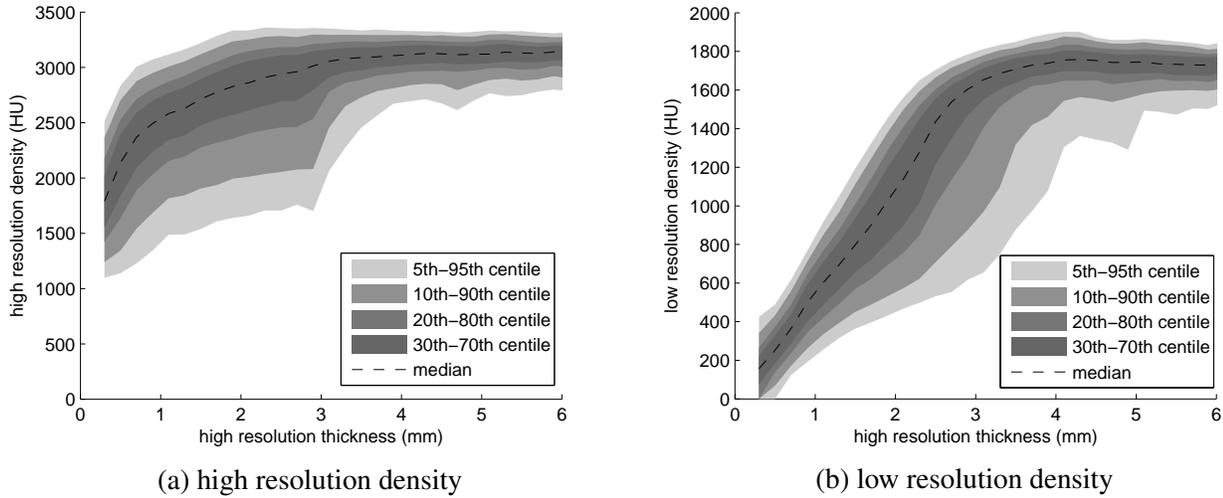


Figure 7: Distribution of apparent peak density with cortical thickness. All sixteen femurs contributed to the data in these graphs. (a) The high resolution distribution, with density and thickness estimated using the half-max method. (b) The  $y$ -axis now shows the corresponding peak CT values from the low resolution data.

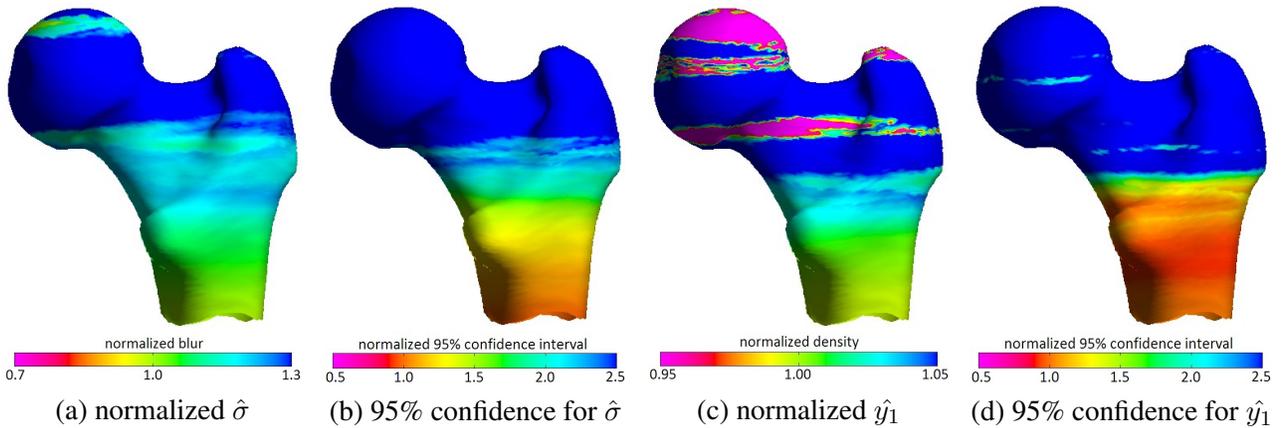


Figure 8: How much data is required to estimate image blur and cortical density? The algorithm described in Section 2.2 was run repeatedly on all sixteen femurs, starting with just a little data from the top of the femur, then adding in progressively more data down to below the lesser trochanter. After each run, the estimates  $\hat{\sigma}$  and  $\hat{y}_1$  were normalized by their “correct” values obtained using all the available data. The colour maps in (a) and (c) show the mean of the normalized estimates (a “correct” value being one), mapped onto the canonical femur to show how much data contributed to the estimate. (b) and (d) show the corresponding mean 95% confidence intervals, again normalized by the final values obtained using all the available data. For these sixteen femurs, the maps in (a) and (c) have settled to one below the lesser trochanter, indicating that the femur must be scanned at least this far for reliable estimation of density and blur. (b) and (d) confirm that the 95% confidence intervals are good indicators of estimation precision for  $\hat{\sigma}$  and  $\hat{y}_1$ .

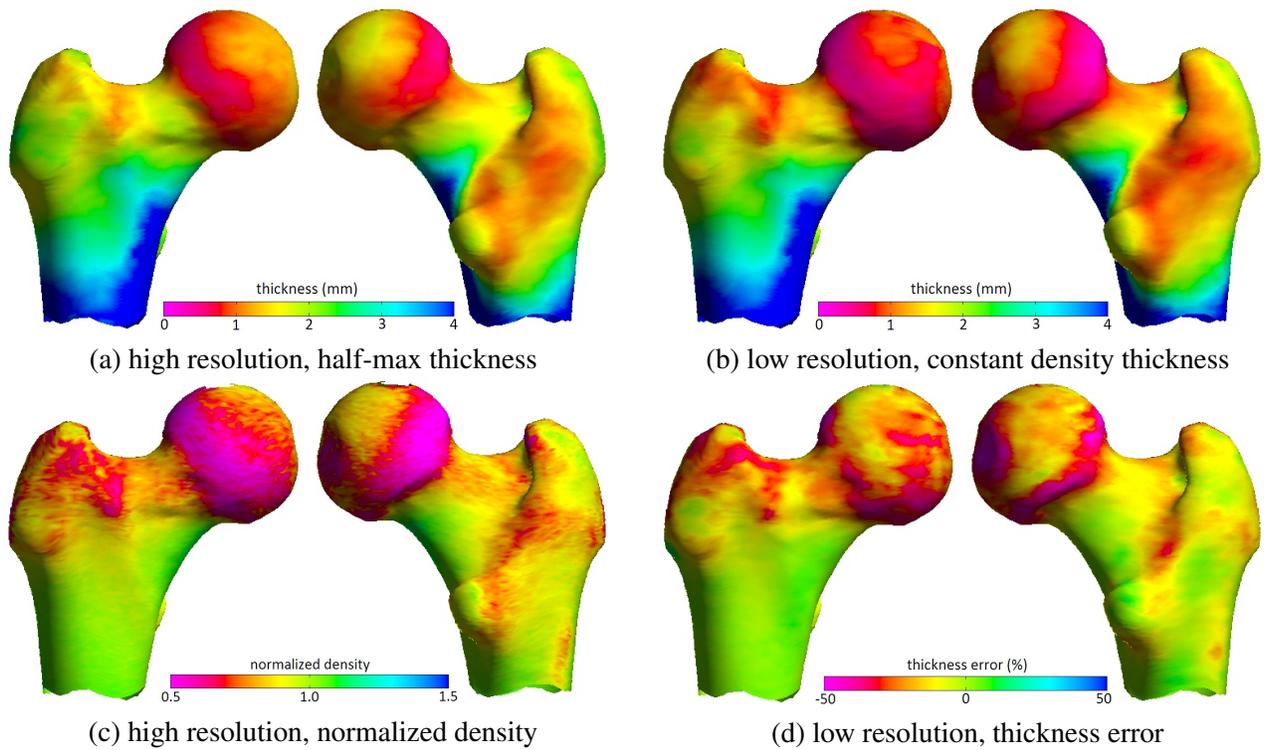


Figure 9: Illustrative results displayed on the canonical femur. Measurements on each of the sixteen femurs were expressed in the canonical morphology and then averaged. (a) Gold standard cortical thickness, calculated using the half-max method on the high resolution data. (b) Constant density thickness estimation from the low resolution data, with  $\hat{\gamma}_1$  estimated using the method described in Section 2.2. (c) Cortical density in the high resolution data, normalized by the high resolution density that corresponds to  $\hat{\gamma}_1$ . (d) The low resolution, constant density thickness estimation error, expressed as a percentage of the gold standard thickness.

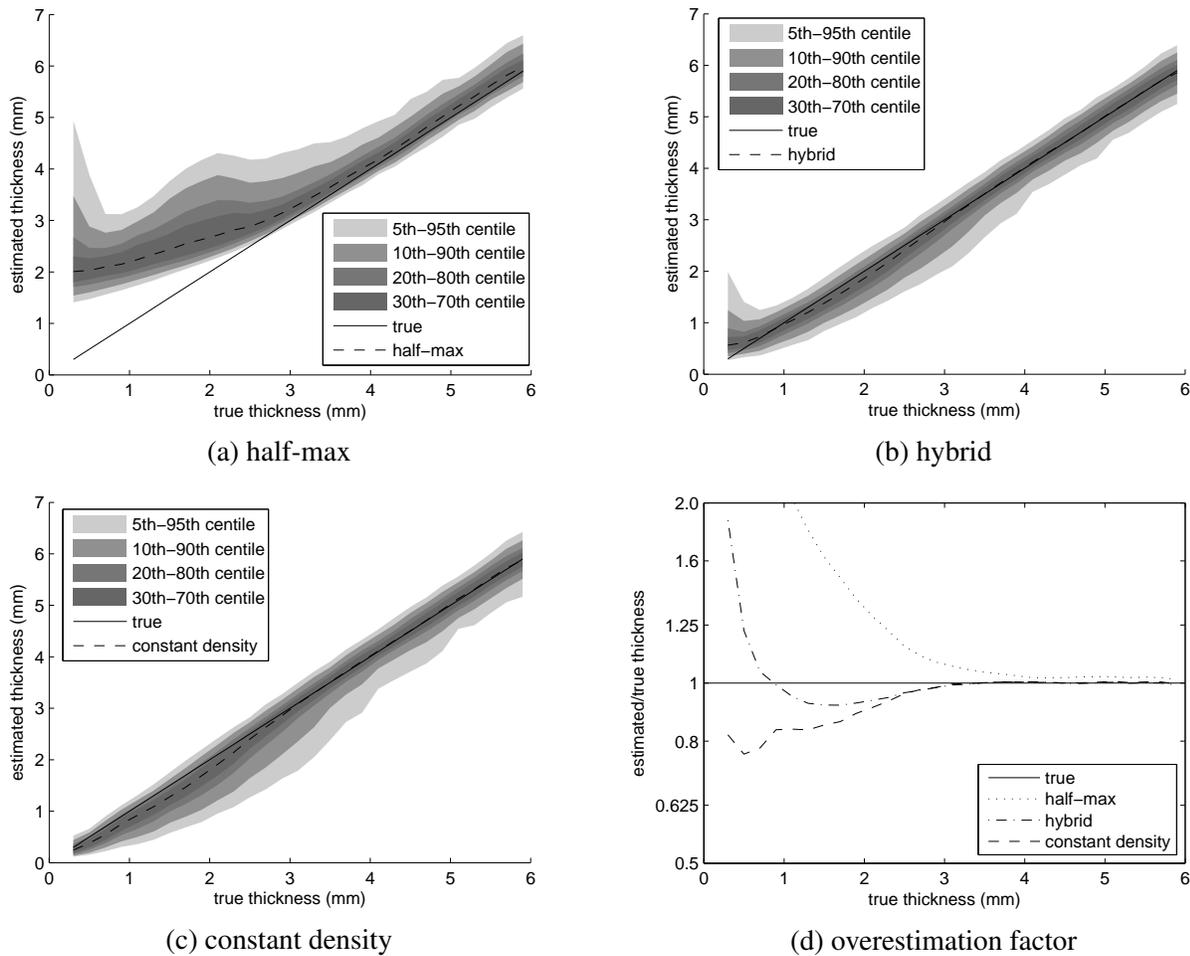


Figure 10: Cortical thickness estimation in low resolution data. Data from all sixteen femurs contributed to these graphs. True thickness was measured in the high resolution data using the half-max method.

corresponds to  $\hat{y}_1$ , it appears that  $\hat{y}_1$  is biased towards the *peak* cortical density, which is a good estimate of the actual density at many locations on the femur but an overestimate at others. Figure 9(d) confirms that overestimating  $y_1$  leads to underestimation of thickness by roughly the same factor. These errors are not very apparent in Figures 9(a) and (b) because they are confined to regions of thin cortex. An error of, say, 30% in 1 mm cortex is fairly inconsequential in the context of the 0–4 mm dynamic range.

Figure 10 compares the performance of the constant density approach with the half-max and hybrid alternatives. The shortcomings of the half-max method are clear in Figure 10(a), with increasingly inaccurate estimates below 2.5 mm. This is consistent with the performance anticipated in Figure 4(b), as is the asymptotic measurement of 2 mm for vanishingly thin cortex. In contrast, the constant density and hybrid methods perform relatively well. Figure 10(d) indicates very little estimation error until 3 mm, which is where we know the constant density assumption starts to break down. From 3 mm down to 0.5 mm there is no catastrophic failure: median errors peak at around 20%, mirroring the 20% density reduction apparent in Figure 7(a). Below 0.5 mm, the pixel dimensions of the raw CT data approach the limit for effective sampling of the blurred cortex, with unavoidable consequences for the estimation error. The apparently superior performance of the hybrid method in the 0.5–3 mm range owes much to chance. The cadaveric femurs were scanned in air, with

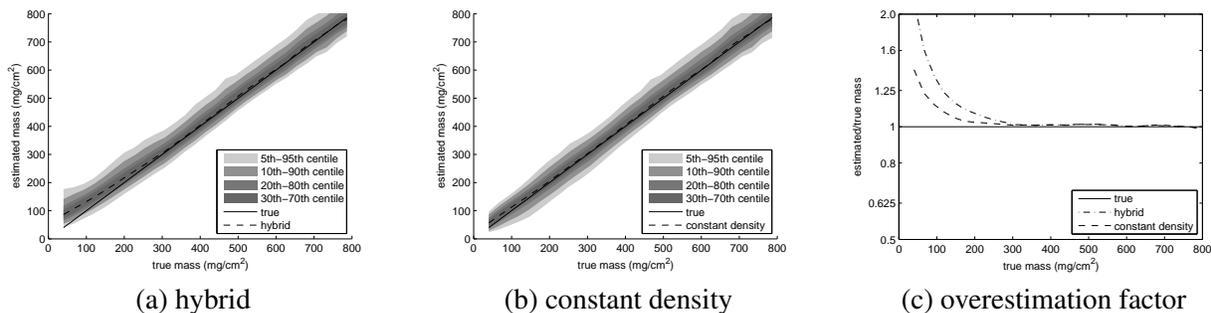


Figure 11: Cortical mass estimation in low resolution data. Data from all sixteen femurs contributed to these graphs. True mass was measured in the high resolution data using the half-max method.

$y_0 \ll y_2$ . We have seen in Figure 2 (bottom) how the hybrid method is bound to overestimate thickness in such circumstances, and this fortuitously ameliorates the underestimation caused by too high a value of  $\hat{y}_1$ .

Figure 11 shows how mass estimation is far less sensitive to any variation in the cortical density. Gaussian blurring is area preserving, so as long as the blurred  $y_0$ - $y_1$ - $y_2$  pulse fits the data well, the area under the pulse will be the same as the area under the data, providing a good mass estimate irrespective of the particular value of  $\hat{y}_1$ . When  $y_0 \neq y_2$ , there is the added requirement that the fitted pulse coincide with the true location of the cortex, otherwise the mass estimate is corrupted by a  $y_0$ - $y_2$  transition at the wrong location. All these observations are borne out in Figure 11, where we see mass estimation apparently unaffected by the variation in cortical density. The hybrid method overestimates mass for thin cortices, as predicted in Figure 2 (bottom,  $y_0$ - $y_2$  transition at the wrong location), and both techniques eventually fail when the cortex is undersampled.

To validate the preceding claims of cause and effect, we simulated idealised cortical sections with a thickness-density relationship based on that observed in Figure 7(a), and two sets of  $y_0/y_2$  values that simulate scanning in air and *in vivo*. Subsequent mass and thickness estimation results can be found in Figure 12. Consistent with the observations in Figure 2, the hybrid method produces larger estimates than the constant density method in air, but slightly smaller estimates *in vivo*. There is an excellent match between Figures 12(a,b) and the experimental results in Figures 10 and 11. Where there are discrepancies, they are easily explained. The low thickness sampling breakdown happens further to the left in the simulations, since these were conducted with a smaller pixel dimension. By increasing the pixel size we have observed the breakdown point moving to the right, and it is for this reason that we are confident that this is nothing more than a sampling effect. The actual hybrid results appear to lie somewhere between the air and *in vivo* simulations, but this is consistent with the cadaveric femurs not being perfectly clean. Any soft tissue attached to the periosteal surface will increase  $y_0$  above  $-1000$  HU, closer to the assumed *in vivo* value of 0 HU.

We therefore have good reason to commend Figures 12(c) and (d) as accurate predictors of likely *in vivo* performance. For cortical mass estimation, the constant density method outperforms the hybrid method since it is better able to model the observed data. *In vivo* the difference is relatively small, corresponding to the discrepancies between the red and blue curves in Figure 2 (top left). The story is much the same when estimating thickness, though both techniques are affected by an increasingly inappropriate value of  $\hat{y}_1$  below 3 mm. While it would be straightforward to correct the thickness estimates by exploiting prior knowledge of the thickness-density relationship, we have no evidence to suggest that the distribution observed in Figure 7(a) generalises beyond the sixteen specimens examined here. In the absence of any such evidence, we must accept thickness errors that are bounded by the extent to which the local density deviates from the peak density.

Although the hybrid method is slightly less accurate, it is significantly faster than the constant density approach. The former requires just one pass through the data to measure  $t^{\text{hm}}$ ,  $y_1^{\text{hm}}$  and  $y_b^{\text{hm}}$  at all points on the femoral surface, and then calculate  $\hat{y}_1$ , as described in Section 2.2. The half-max estimates can then be cor-

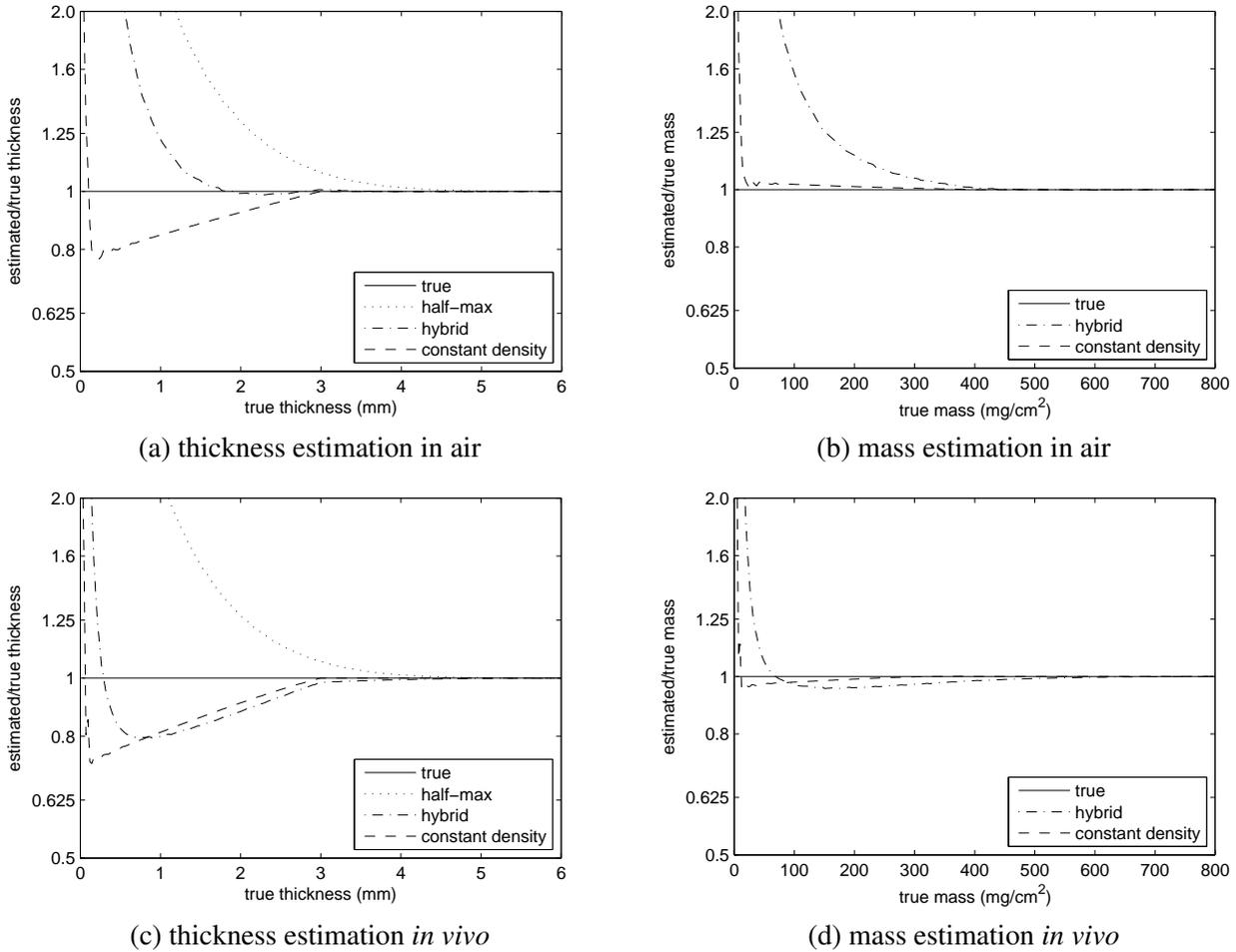


Figure 12: Simulated thickness and mass measurements. To facilitate comparison of these simulations with Figures 10(d) and 11(c), the blur was set to approximate that of the low resolution scans ( $\sigma = 0.85$  mm), as was the mapping from HU to  $\text{mg}/\text{cm}^3$ . In all cases, the true cortical density  $y_1$  was modelled as varying linearly from 1300 HU at zero thickness to 1750 HU at 3 mm, thereafter remaining at 1750 HU. The various thickness and mass estimation techniques assumed  $\hat{y}_1 = 1750$  HU. (a) and (b) simulate imaging in air, with  $y_0 = -1000$  HU and  $y_2 = 300$  HU. (c) and (d) simulate imaging *in vivo*, with  $y_0 = 0$  HU and  $y_2 = 300$  HU.

rected, immediately, using equations (6) and (7). In contrast, the constant density approach requires a second pass through the data to estimate  $t^{cd}$  and  $m^{cd}$  using the just calculated value of  $\hat{y}_1$ . To put this observation into perspective, it takes 10 seconds to map cortical thickness at 6500 locations (more than sufficient for a single femur) using a single core of a 2.67 GHz Intel Core i7 processor and the hybrid method, and twice as long for the constant density approach.

## 4 Discussion

It remains to discuss the relative merits of cortical mass and thickness estimation, and to comment on the benefits of estimating  $\hat{y}_1$  using the method described in Section 2.2, in comparison with our earlier approach of sampling regions of thick cortex below the lesser trochanter. Both of these themes are best addressed in the context of an example, so we here revisit our work on cortical thickening following two years' treatment with recombinant human parathyroid hormone, hPTH(1-34). In [Poole et al. \(2011\)](#), CT scans of 65 osteoporotic women were analysed at baseline and 24 months, with cortical thickness estimated using the constant density technique and  $\hat{y}_1$  sampled from regions of thick cortex. All the resulting measurements were mapped onto the canonical femur for subsequent cohort analysis, producing the images in Figure 13(c)<sup>3</sup>. Statistically significant regions of thickened cortex were identified by statistical parametric mapping (SurfStat, [Worsley et al. \(2009\)](#)). The resulting vertex  $p$ -maps (Figure 13, yellow/orange) are sensitive to focal effects, while the cluster  $p$ -maps (cyan/blue) are sensitive to distributed effects. The reader is referred to [Poole et al. \(2011\)](#) for a full description of this work. Here, we augment the study with constant density cortical thickness maps based on the new method of estimating  $\hat{y}_1$  (Figure 13(d)), as well as cortical mass maps using both density estimation methods (Figures 13(a) and (b)).

The first point to note is the insensitivity of the cortical mass maps to the density estimate  $\hat{y}_1$ . This confirms the findings in Section 3 and reinforces the message that measuring cortical mass, using the various techniques described in this paper, is a relatively safe undertaking. A good estimate of  $\hat{y}_1$  is not required, and it does not matter how much  $y_1$  varies across the femur in reality. However, there are some drawbacks. Firstly, it is necessary to know the calibration relating HU to mineral  $\text{mg}/\text{cm}^3$ . More significant, though, is the danger of disregarding potentially significant density effects. For example, for the 65 individuals analysed here, baseline cortical density was  $1185.1 \pm 67.9 \text{ mg}/\text{cm}^3$  (mean  $\pm$  one standard deviation), whereas at 24 months it was  $1133.6 \pm 78.1 \text{ mg}/\text{cm}^3$ . A paired t-test confirms that these densities are significantly different ( $p < 4.9 \times 10^{-13}$ ). This is not surprising: hPTH(1-34) functions by increasing bone remodeling and hence porosity ([Burr et al., 2001](#); [Sato et al., 2004](#); [Arlot et al., 2005](#)). Consequently, in those regions where the mass increases in Figures 13(a) and (b), it is reasonable to postulate a far more significant increase in thickness, since density has decreased. However, unlike the mass maps, the cortical thickness maps in Figures 13(c) and (d) must be approached with some caution. There is the underlying assumption of constant cortical density: might a regionally fluctuating density undermine the apparent statistically significant results? Perhaps the regional mass increases correlate with density, not thickness? This seems unlikely: we know that average density has decreased, so we would be hypothesising a density increase in the regions most affected by hPTH(1-34), counteracted by an even greater decrease in the unaffected regions, in direct contradiction to how the drug is known to work. Nevertheless, Figures 13(c) and (d) show changes in cortical thickness *assuming constant density in any one scan*, and must be interpreted as such.

A more subtle but equally important consideration is the possibility of measurement bias affecting the thickness maps. Figures 13(c) and (d) differ because the new density estimates, obtained by model fitting, indicate around 2% more density reduction than the old estimates, obtained by sampling regions of thick cortex

<sup>3</sup>Figure 13(c) is in fact slightly different from the corresponding figure in [Poole et al. \(2011\)](#), since we now use an improved smoothing algorithm when mapping onto the canonical femur.

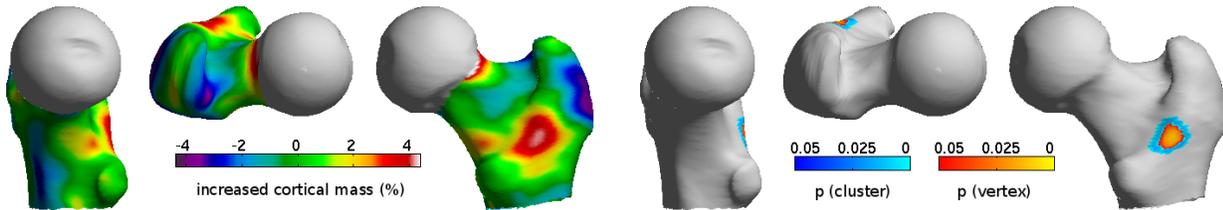
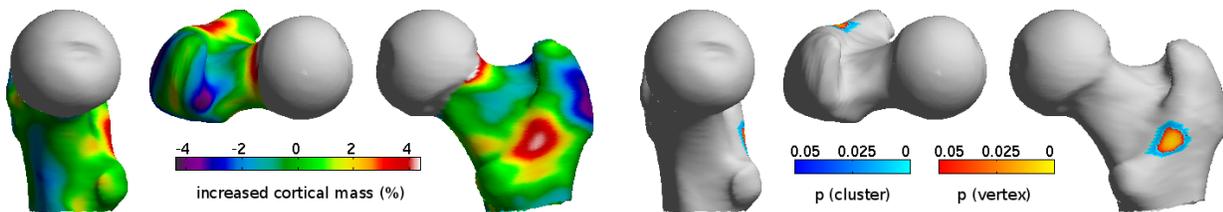
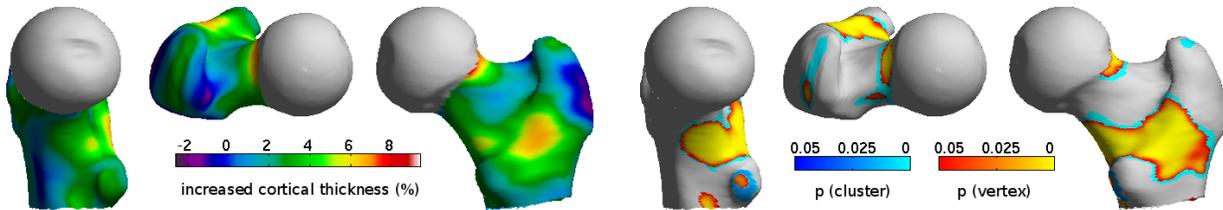
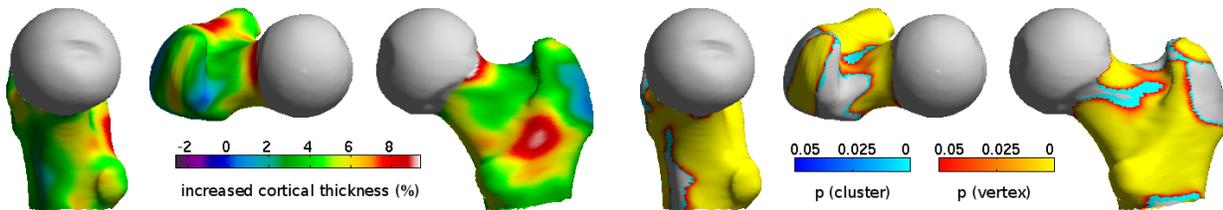
(a) Change in cortical mass,  $\hat{y}_1$  estimated from regions of thick cortex(b) Change in cortical mass,  $\hat{y}_1$  estimated as in Section 2.2(c) Change in cortical thickness,  $\hat{y}_1$  estimated from regions of thick cortex(d) Change in cortical thickness,  $\hat{y}_1$  estimated as in Section 2.2

Figure 13: Cortical mass and thickness effects following two years' treatment with hPTH(1-34). Measurements from 65 subjects were mapped onto the canonical femur and averaged. The statistical parametric maps (right) indicate significant regions of mass increase (a,b) or thickening (c,d). Vertex effects are based on the magnitude of peaks corrected for multiple comparisons, while the cyan/blue areas are connected clusters of vertices exceeding an uncorrected  $p$ -value of 0.001.

below the lesser trochanter. Might either or both of these methods be biased? Any systematic measurement bias that affects the baseline and 24 month density estimates to different degrees will manifest themselves as false thickness effects. So extreme care must be taken when interpreting any small, apparent thickness differences: bear in mind that a 10% thickness increase in 1 mm cortex is only  $100\ \mu\text{m}$ , and we are inferring this information from clinical CT scans with a point spread function of around 3 mm. In this instance, there is a known source of bias: the 24 month scans extended, on average, 4.5 mm lower down the femur than the baseline scans. This means that the baseline scans are more susceptible to imprecise  $\hat{y}_1$  estimates (caused by a lack of sufficiently thick cortical bone) than the 24 month scans. Referring to Figure 3(b), it is clear that sampling the thickest available cortex will tend to *underestimate* the true, peak cortical density: had there been more data, and the samples extended further to the right, we would have encountered higher density values. In contrast, the model-fitting approach tends to *overestimate*  $\hat{y}_1$  when there is marginally sufficient data, as demonstrated clearly in Figure 8(c).

So Figures 13(c) and (d) most likely represent lower and upper bounds on the true cortical thickness effect, and we are sanguine that the difference is only around 2%, despite the difficulties posed by this challenging study. Although all the scans extended below the lesser trochanter – some only just so – these were severely osteoporotic women with pathologically thin cortices. Furthermore, the point spread function of the CT system was around 3 mm, compared with 2 mm for the cadaveric studies. Nevertheless, the mean confidence interval for  $\hat{y}_1$  was  $\pm 14.8$  HU and the worst case was  $\pm 38.0$  HU: we therefore decided to include all 65 subjects in the cortical thickness analysis.

In summary, the merits of the new density estimation method are twofold. Compared with our previous approach of sampling the density in regions of supposedly thick cortex, the new technique makes use of all the available data. It can therefore produce estimates of  $\hat{y}_1$  that correctly exceed the peak density observed anywhere in the imaged cortex. When there is insufficient data to fully constrain the model, it does tend to overestimate density just as the old technique tends to underestimate. This brings us to the second advantage: the new method furnishes a confidence interval for  $\hat{y}_1$ , allowing objective inclusion or exclusion of particular scans. Our experience with the hPTH(1-34) study suggests that exclusion should rarely be necessary.

## 5 Conclusions

Cortical mass and thickness can be mapped across the proximal femur using either a constant density approach or a hybrid half-max/constant density approach. The former is more accurate but requires two passes through the data compared with a single pass for the latter. When a suitable calibration phantom is included in the scan, cortical mass can be measured in  $\text{mg}/\text{cm}^2$  to high accuracy. Cortical thickness estimation requires a good estimate of the cortical density, which is assumed to be constant at all points on the femur. While the model-fitting method presented in this paper estimates this density in a manner that can be considered optimal, in that it exploits all the available data, thickness estimates are nevertheless biased when there is any variation in the actual cortical density. In many studies, it may therefore be advisable to base any firm deductions on observations of cortical mass alone, unless the study involves comparison of two groups with different densities, in which case cortical thickness analysis, when applied with care, may well reveal more significant group-dependent effects.

## Acknowledgments

Ken Poole is supported by the Arthritis Research Campaign, the Evelyn Trust and the NIHR Cambridge Biomedical Research Centre. The cadaveric femurs were from the *Melbourne Femur Collection Research Tissue Bank of the Victorian Institute of Forensic Medicine*, with kind permission of Professor John Clement.

## References

- Arlot, M., Meunier, P. J., Boivin, G., Haddock, L., Tamayo, J., Correa-Rotter, R., Jasqui, S., Donley, D. W., Dalsky, G. P., Martin, J. S., Eriksen, E. F., 2005. Differential effects of teriparatide and alendronate on bone remodeling in postmenopausal women assessed by histomorphometric parameters. *Journal of Bone and Mineral Research* 20 (7), 1244–1253.
- Bouxsein, M. L., Delmas, P. D., Aug. 2008. Considerations for development of surrogate endpoints for anti-fracture efficacy of new treatments in osteoporosis: A perspective. *Journal of Bone and Mineral Research* 23 (8), 1155–1167.
- Buie, H. R., Campbell, G. M., Klinck, R. J., MacNeil, J. A., Boyd, S. K., Oct. 2007. Automatic segmentation of cortical and trabecular compartments based on a dual threshold technique for in vivo micro-CT bone analysis. *Bone* 41 (4), 505–515.
- Burr, D. B., Hirano, T., Turner, C. H., Hotchkiss, C., Brommage, R., Hock, J. M., 2001. Intermittently administered human parathyroid hormone (1-34) treatment increases intracortical bone turnover and porosity without reducing bone strength in the humerus of ovariectomized cynomolgus monkeys. *Journal of Bone and Mineral Research* 16 (1), 157–165.
- Cann, C. E., Feb. 1988. Quantitative CT for determination of bone mineral density: a review. *Radiology* 166 (2), 509–522.
- Carpenter, R. D., Beaupre, G. S., Lang, T. F., Orwoll, E. S., Carter, D. R., 2005. New QCT analysis approach shows the importance of fall orientation on femoral neck strength. *Journal of Bone and Mineral Research* 20 (9), 1533–1542.
- de Bakker, P. M., Manske, S. L., Ebacher, V., Oxland, T. R., Cripton, P. A., Guy, P., Aug. 2009. During sideways falls proximal femur fractures initiate in the superolateral cortex: Evidence from high-speed video of simulated fractures. *Journal of Biomechanics* 42 (12), 1917–1925.
- Dougherty, G., Newman, D., Jul. 1999. Measurement of thickness and density of thin structures by computed tomography. *Medical Physics* 26 (7), 1341–1348.
- Hangartner, T. N., Mar. 2007. Thresholding technique for accurate analysis of density and geometry in QCT, PQCT and  $\mu$ CT images. *Journal of Musculoskeletal and Neuronal Interactions* 7 (1), 9–16.
- Hangartner, T. N., Gilsanz, V., 1996. Evaluation of cortical bone by computed tomography. *Journal of Bone and Mineral Research* 11 (10), 1518–1525.
- Holzer, G., von Skrbensky, G., Holzer, L. A., Pichl, W., 2009. Hip fractures and the contribution of cortical versus trabecular bone to femoral neck strength. *Journal of Bone and Mineral Research* 24 (3), 468–474.
- Johnell, O., Kanis, J. A., Oden, A., Johansson, H., Laet, C. D., Delmas, P., Eisman, J. A., Fujiwara, S., Kroger, H., Mellstrom, D., Meunier, P. J., 3rd, L. J. M., O'Neill, T., Pols, H., Reeve, J., Silman, A., Tenenhouse, A., 2005. Predictive value of BMD for hip and other fractures. *Journal of Bone and Mineral Research* 20 (7), 1185–1194.
- Kanis, J. A., Burlet, N., Cooper, C., Delmas, P. D., Reginster, J. Y., Borgstrom, F., Rizzoli, R., 2008. European guidance for the diagnosis and management of osteoporosis in postmenopausal women. *Osteoporosis International* 19 (4), 399–428.

- Kaptoge, S., Beck, T. J., Reeve, J., Stone, K. L., Hillier, T. A., Cauley, J. A., Cummings, S. R., 2008. Prediction of incident hip fracture risk by femur geometry variables measured by hip structural analysis in the study of osteoporotic fractures. *Journal of Bone and Mineral Research* 23 (12), 1892–1904.
- Mayhew, P. M., Thomas, C. D., Clement, J. G., Loveridge, N., Beck, T. J., Bonfield, W., Burgoyne, C. J., Reeve, J., 2005. Relation between age, femoral neck cortical stability, and hip fracture risk. *The Lancet* 366 (9480), 129–135.
- More, J. J., 1977. The Levenberg-Marquardt algorithm: Implementation and theory. In: Watson, A. (Ed.), *Numerical Analysis. Lecture Notes in Mathematics* 630, Springer-Verlag, pp. 105–116.
- Parker, M., Johansen, A., 2006. Hip fracture. *British Medical Journal* 333 (7557), 27–30.
- Poole, K. E., Mayhew, P. M., Rose, C. M., Brown, J. K., Bearcroft, P. J., Loveridge, N., Reeve, J., 2010. Changing structure of the femoral neck across the adult female lifespan. *Journal of Bone and Mineral Research* 25 (3), 482–491.
- Poole, K. E. S., Treece, G. M., Ridgway, G. R., Mayhew, P. M., Borggrefe, J., Gee, A. H., 2011. Targeted regeneration of bone in the osteoporotic human femur. *PLoS ONE* 6 (1), e16190.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., 2002. *Numerical recipes in C: The art of scientific computing*, 2nd Edition. Cambridge University Press.
- Prevrhal, S., Engelke, K., Kalander, W. A., Mar. 1999. Accuracy limits for the determination of cortical width and density: the influence of object size and CT imaging parameters. *Physics in Medicine and Biology* 44 (3), 751–764.
- Prevrhal, S., Fox, J. C., Shepherd, J. A., Genant, H. K., Jan. 2003. Accuracy of CT-based thickness measurement of thin structures: Modeling of limited spatial resolution in all three dimensions. *Medical Physics* 30 (1), 1–8.
- Sambrook, P., Cooper, C., 2006. Osteoporosis. *The Lancet* 367 (9527), 2010–2018.
- Sanders, K. M., Nicholson, G. C., Watts, J. J., Pasco, J. A., Henry, M. J., Kotowicz, M. A., Seeman, E., 2006. Half the burden of fragility fractures in the community occur in women without osteoporosis. when is fracture prevention cost-effective? *Bone* 38 (5), 694–700.
- Sato, M., Westmore, M., Ma, Y. L., Schmidt, A., Zeng, Q. Q., Glass, E. V., Vahle, J., Brommage, R., Jerome, C. P., Turner, C. H., Apr. 2004. Teriparatide [PTH(1-34)] strengthens the proximal femur of ovariectomized nonhuman primates despite increasing porosity. *Journal of Bone and Mineral Research* 19 (4), 623–629.
- Streekstra, G. J., Strackee, S. D., Maas, M., ter Wee, R., Venema, H. W., Sep. 2007. Model-based cartilage thickness measurement in the submillimeter range. *Medical Physics* 34 (9), 3562–3570.
- Treece, G. M., Gee, A. H., Mayhew, P. M., Poole, K. E. S., Jun. 2010. High resolution cortical bone thickness measurement from clinical CT data. *Medical Image Analysis* 14 (3), 276–290.
- Verhulp, E., van Rietbergen, B., Huiskes, R., 2008. Load distribution in the healthy and osteoporotic human proximal femur during a fall to the side. *Bone* 42 (1), 30–35.
- Worsley, K., Taylor, J., Carbonell, F., Chung, M., Duerden, E., Bernhardt, B., Lyttelton, O., Boucher, M., Evans, A., 2009. Surfstat: A Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. *NeuroImage* 47 (Supplement 1), S102–S102, Organization for Human Brain Mapping, 2009 Annual Meeting.