

Detection and characterisation of bone destruction in murine rheumatoid arthritis using statistical shape models

James M. Brown^{a1}, Ewan Ross^b, Guillaume Desanti^b, Atif Saghir^b, Andy Clark^b, Chris Buckley^b, Andrew Filer^b, Amy Naylor^b, Ela Claridge^{a2}

^aSchool of Computer Science, University of Birmingham, Birmingham B15 2TT, U.K.

^bInstitute of Inflammation and Ageing, University of Birmingham, Birmingham B15 2TT, U.K

Abstract

Rheumatoid arthritis (RA) is an autoimmune disease in which chronic inflammation of the synovial joints can lead to destruction of cartilage and bone. Pre-clinical studies attempt to uncover the underlying causes by emulating the disease in genetically different mouse strains and characterising the nature and severity of bone shape changes as indicators of pathology. This paper presents a fully automated method for obtaining quantitative measurements of bone destruction from volumetric micro-CT images of a mouse hind paw. A statistical model of normal bone morphology derived from a training set of healthy examples serves as a template against which a given pathological sample is compared. Abnormalities in bone shapes are identified as deviations from the model statistics, characterised in terms of type (erosion / formation) and quantified in terms of severity (percentage affected bone area). The colour-coded magnitudes of the deviations superimposed on a three-dimensional rendering of the paw show at a glance the severity of malformations for the individual bones and joints. With quantitative data it is possible to derive population statistics characterising differences in bone malformations for different mouse strains and in different anatomical regions. The method was applied to data acquired from three different mouse strains. The derived quantitative indicators of bone destruction have shown agreement both with the subjective visual scores and with the previous biological findings. This suggests that pathological bone shape changes can be usefully and objectively identified as deviations from the model statistics.

Keywords

Statistical shape models; Quantification of bone destruction; Rheumatoid arthritis; Murine models.

1 Introduction

1.1 Rheumatoid arthritis

Rheumatoid arthritis (RA) is an inflammatory disease of autoimmune origin. It is most common in the elderly population, but it can affect people of all ages. It is a chronic disease, and being one of the most common causes of disability, it constitutes a public health problem.

¹ Present address: MGH/HST Athinoula A. Martinos Center for Biomedical Imaging, Boston, U.S.A.

² Corresponding author.

The autoimmune response mounted by the body affected by RA gives rise to chronic inflammation of the synovial joints between the bones, which are the most common type of movable joints in the body. A proportion of sufferers will develop persistent inflammation of the synovial membrane (synovium) leading to the destruction of both cartilage and bone. Damage to the bone is thought to occur through imbalance of two processes: bone erosion (the breakdown of bone through secretion of enzymes that demineralise the bone matrix) and bone formation (deposition of new bone mineral into the underlying bone matrix). In healthy patients the two mechanisms are tightly regulated to ensure that on average, bone integrity remains constant. In patients with RA bone erosion dominates and over time the destruction of the bone surface can impair normal joint function and lead to structural deformities. In approximately 90% of cases where inflammation is persistent, patients will be clinically disabled within 20 years (Buckley, 1997).

1.2 Mouse models of the disease

The exact cause of rheumatoid arthritis is unknown, which limits the number of available treatment options. Pre-clinical studies attempt to uncover the underlying causes by emulating the disease in animals. Animal models of disease allow biomedical researchers to investigate medical conditions through experimentation that would otherwise be infeasible or unethical to perform on human subjects. Mice are particularly well-suited due to their high genetic homology with humans, speed of breeding and inexpensive housing costs. A number of laboratory mouse strains have been developed and bred to be near genetically identical to one another. This allows for the roles of different genes to be investigated in a way that eliminates genetic variation as a factor and helps to gain better understanding of disease mechanisms and of the efficacy of new treatments. Characterisation of bone destruction is one of the key factors in understanding RA pathogenesis.

In longitudinal studies mice may be examined for traits associated with inflammatory arthritis. These could be based either on measurements (e.g. paw and ankle thickness, body weight, grip strength) or on subjective qualitative scoring (e.g. limb deformity, colouration, gait). Typically a hind paw is examined. Based on such observations and measurements individual animals are assigned a physiological disease index score according to disease severity. Assessment of disease in this way is an effective method for determining the gross differences between individual animals, but it does not provide explicit information about how the disease affects bones and joints.

1.3 Objective characterisation of RA

The nature and severity of bone shape changes are important indicators of pathology in mouse models of rheumatoid arthritis. Researchers that rely on the assessments of the external signs of the disease, as described above, cannot characterise a specific nature of bone destruction in different models of RA. Observation of the pattern of joint involvement and the bias towards bone formation compared to bone destruction allows characterisation of the nature of the arthritis, for example spondyloarthropathy is characterised by bone formation at the sites of tendon insertion whilst rheumatoid arthritis shows predominantly bone erosion at sites close to the articular joints. The spatial pattern of malformations can help to uncover cellular mechanisms of bone destruction by, for example, identifying the sites where activated osteoclasts proliferate causing bone erosion, or studying regulatory factors that control osteoclast differentiation and activity at specific bone junctions (Goldring, 2003; Romas et al., 2000).

In order to assess the degree of damage to bone, x-ray based techniques such as micro-CT are employed to reveal changes in bone structure. The technique has been demonstrated early on in animal studies of RA as a suitable means for evaluating bone loss and osteophytosis (formation of bone spurs) through visual examination of a sequence of two-dimensional micro-CT slices (Pettit et al., 2001; Pine et al., 2007). The process of identifying bone abnormalities is carried out manually and relies on the experimenter's own experience of how such bones appear under normal (non-pathological) circumstances.

A number of methods have been proposed to automate the analysis of bone destruction in RA, both in animal models and in humans. Of particular interest to this paper are volumetric methods that consider three-dimensional bone surface data. Quantification of bone volume is one possible approach, where gross differences can be determined automatically from image data. This approach has been used to assess mice with collagen induced arthritis (CIA), which were shown to have significantly lower bone volume and density than wild-type (normal) controls (Yang et al., 2013). One of the main limitations of this approach is that it provides no data about bone morphology, or the nature of the bone destruction that has occurred. This problem has since been tackled using surface-based registration of bones (Joshi et al., 2013).

In a longitudinal study of treatment efficacy, three-dimensional CT images were acquired from patients with established RA. Individual bones were manually segmented, and surface representations generated as triangulated meshes. Surface registration of bones at different stages of treatment was followed by calculation of point-wise distances between them, and used to determine local differences in bone volume. These differences were visualised as colourised surface renderings, highlighting the presence of abnormalities. This approach is effective in assessing bone changes over time in individual patients, but is not well-suited to detecting abnormalities in multiple subjects imaged at a single time point. Subsequently this work was extended to analyse statistical shape differences in wrist bone shapes of different populations (Joshi et al., 2015). This involved constructing an atlas of bone shapes where each surface point is characterised by the mean and variance and the p-value of the locations in individual samples. Shape differences were then visualised on colour-coded surface maps of p-values. A similar approach was used to identify erosive changes of carpal bones in longitudinal patient studies. The detected erosions were quantified by measuring changes in their depth.

The osteoarthritic changes affecting the mandibular joint were analysed in a study aiming to correlate the nature of morphological changes of the bone shape and the amount of pain experienced by a patient (Cevidane et al., 2010). Individual bone instances were segmented, meshed and registered via spherical mapping using Procrustes alignment, followed by building a statistical shape model (SSM). Variations in the three aspects of morphological changes (flattening, erosions and osteophytes) for mild, moderate and severe conditions were visualised by generating the meshes from the SSM.

1.4 Objectives and contributions of this work

This paper presents an automated method for obtaining quantitative measurements of bone destruction in mouse models of RA from micro-CT images. A statistical model of normal bone morphology derived from a training set of healthy examples serves as a “template” against which a given sample is compared. Abnormalities in bone shapes for a given RA model are identified as

deviations from the model statistics and are then characterised in terms of type (erosion / formation) and quantified in terms of severity (affected bone area). The colour-coded magnitudes of the deviations superimposed on a three-dimensional rendering of the paw show at a glance the spatial distribution and severity of the bone erosions and formations and their association with specific joints. The method has been applied to investigate the nature of bone destruction in three different mouse models of inflammatory arthritis, providing an insight into the different ways that joints are affected by the disease.

1.5 Outline of the paper

The hypothesis underpinning this work is that shapes of bones affected by pathology depart from statistically normal bone shape variations. This suggests the three-step procedure: first, the development of a statistical shape model of a normal limb; second, the detection of bone regions that lie outside the statistically predicted boundaries; and third, qualitative and quantitative characterisation of the detected abnormal regions both globally and locally. Step 1 was the subject of an earlier paper (Brown et al., 2014). This paper focuses on the detection and quantification of the bone destruction.

The mouse models used in the experiments are discussed in section 2.1. Section 2.2 briefly describes data acquisition and sample preparation. The construction of the articulated statistical shape model (ASSM) of a normal mouse hind paw is outlined in section 2.3. Shape model fitting and abnormality detection are described in sections 2.4 and 2.5 respectively. Validation of the abnormality detection methods is presented in section 2.6 followed by results (section 3), discussion (section 4) and conclusions (section 5).

2 Materials and methods

2.1 Mouse models of rheumatoid arthritis used in the study

Mouse models of rheumatoid arthritis work by targeting the genes and proteins that are implicated in its pathogenesis. Although RA is predominantly an erosive disease, a number of mouse models show evidence of both erosion and formation of bone (Pettit et al., 2001) (Aya et al., 2005). Accurate assessment of these bone changes is required to identify those mouse models that best replicate the human disease progression and phenotype and to identify protective genetic mutations and drug efficacy in pre-clinical trials. This study used three different mouse models of RA.

The K/BxN serum transfer model: These mice develop a transient form of arthritis in response to a serum injection. They show evidence of pannus formation, synovial hyperplasia and erosion of cartilage and bone that is reminiscent of human RA (Monach et al., 2008) (Ditzel, 2004). Because the K/BxN serum transfer model has 100% penetrance and high reproducibility of clinical score, a strain of mice that are resistant to arthritis was used to test the sensitivity of the ASSM. These inflammation-resistant mice have a genetic mutation in the TTP gene which is described in detail in other publications (Ross et al., 2016); throughout this manuscript they will simply be referred to as “non-responders”.

Collagen antibody induced arthritis (CAIA) mice: As in the K/BxN serum transfer model, CAIA mice experience transient inflammatory arthritis, characterised by inflammation of the synovium and destruction of both cartilage and bone (Oestergaard et al., 2008).

TNF^{ΔARE} transgenic mice: These mice develop a progressive arthritis spontaneously. They experience chronic polyarthritis (affecting five or more joints) that is non-resolving and highly erosive (Kontoyiannis et al., 1999). Despite this, TNF^{ΔARE} mice lack the severe swelling of the joints observed in the other models.

Two of these models (KBxN and CAIA) are resolving models induced by injection of autoantibody and one (TNF^{ΔARE}) is a chronic model caused by a genetic modification in the mouse that results in overproduction of a pro-inflammatory protein (tumour necrosis factor α). Normal, untreated mice (wild-type mice) were used to develop the shape model of normal bones.

2.2 Sample preparation

All experiments were carried out at the University of Birmingham, UK (project licence number 40/3253 or 70/8003) following strict guidelines governed by the UK Animals (Scientific Procedures) Act 1986 and approved by the local ethics committee (BERSC: Birmingham Ethical Review Subcommittee). Mice were housed in individually ventilated cages in groups of 3-6 individuals on a 12 hour light-dark cycle with ad libitum access to standard laboratory mouse chow diet and water. Six mice were used for each experiment. Mice were sacrificed at 12 weeks of age. Both hind limbs were dissected above the hip and fixed in formalin over 24 hours at room temperature and then transferred to 70% ethanol in preparation for imaging. Hind limb samples were imaged in 70% ethanol using a Skyscan 1172 micro-CT scanner (Bruker). The x-ray beam was set to a source voltage of 60 kV and source current of 167 μ A. A 0.5mm aluminium filter was used to minimise soft x-ray emission and projections were taken every 0.45 degrees at 1000 ms exposure. Image volumes (2000 x 2000 x 1187 isotropic voxels of size 13.59 μ m) were reconstructed using the Feldkamp algorithm (Feldkamp et al., 1984)(in NRecon 1.6.1.5, Bruker) having applied flat field and beam hardening correction. A radiodensity range of -300.0 to 3000 Hounsfield units (HU) was chosen to isolate the bony structures from the imaging medium. CTAnalyser (v1.12) image analysis software (Kharitonov, 2003) was used to extract an isosurface mesh representation of the mouse limb from the reconstructed and thresholded micro-CT slices. Poisson surface reconstruction (Powell, 1964) (MeshLab 1.3.2) was then used to generate a smooth, uniformly sampled surface mesh that preserved the original surface topology. Internal structures of no interest to this study, such as trabecular bone, were removed using ambient occlusion (Bowden, 2000). Finally, the number of mesh vertices (originally of the order of 2×10^6) was reduced by 90% using quadric edge collapse decimation (Boisvert et al., 2008). Together these procedures generated a simplified surface representation of manageable size so that registration could be performed in a reasonable time frame and within feasible memory requirements.

2.3 Building articulated statistical model of the normal sample

The development of the articulated statistical model of the mouse hind paw was described in detail in an earlier publication (Brown et al., 2014). For completeness, this section provides the essential details of the methods used.

The model comprises shape statistics of the hind paw bones of ten wild-type mice. It is initialised by hand-labelling the constituent bones and joints of a single hind mouse paw (a reference) according to the hierarchy shown in Figure 1. A global rigid registration of the entire reference mesh to an unlabelled sample mesh uses the curvature of the meshes to bring the two into a coarse alignment. This is followed by consecutive registration of individual bones and their descendants down a

hierarchy using the iterative closest point (ICP) algorithm (Besl and McKay, 1992). Transformation parameters for subsequent bones are constrained to a subset of vertices within a frustum projecting from a terminal joint of an already registered parent bone. This articulated registration procedure provides correspondence between the reference and sample vertices. Using this information, labels are propagated from the reference to sample, effectively performing its segmentation.

Registration carried out in this step aligns the reference mesh with a sample mesh. This direction of transformations is chosen because in the reference set each individual bone can be manipulated individually whereas prior to segmentation a sample mesh is not differentiated into individual objects. In order to bring a sample into correct alignment with the reference, inverse transformations are performed in reverse hierarchical order. The result of this process is complete alignment of the sample with the reference in its original coordinate system. This process is applied to all of the samples used in model construction, registering them with the reference.

Description of shape variability uses a well known idea of the point distribution model (PDM) (Cootes et al., 1995), based on principal component analysis (PCA) . After establishing point correspondence between the individual shapes, the mean shape and the principal components of the shape variations are computed. In the *articulated* statistical shape model, ASSM, the shape statistics are derived for each bone individually (van de Giessen et al., 2010). Details of development and validation of ASSM can be found in an earlier paper (Brown et al., 2014).

2.4 Shape model fitting for pathological mice

The ASSM described above was trained exclusively with ten non-pathological samples to produce a model of “normal” bone shape. Any given instance of the model should therefore only consist of bone shapes that are likely to be found among normal mice. The objective of the shape model fitting step is to find an instance of the model that approximates as closely as possible a pathological sample being analysed. By constraining model instances to fall within three standard deviations of the mean shape, differences due to normal anatomical variability are discounted whilst highlighting those that might have arisen due to pathology.

The model fitting process proceeds in two steps. The first step, articulated registration, effectively initialises the fitting process and is identical to the one used for model construction as described in Section 2.3. The only difference here is that rather than registering the annotated reference to a normal sample (as used in model construction), the mean model shape is registered to a pathological sample, labelled, and inversely transformed into the model coordinate system. As in 2.3 this process also generates one-to-one correspondence between the sample and the model vertices. In the second step a non-rigid deformation is found for each individual bone that minimises the distances between the model points and corresponding sample points. The iterative algorithm implemented for this task is based on similar active shape model fitting algorithms, e.g. (Besl and McKay, 1992).

The shape of a sample bone is approximated by a non-rigid deformation of the mean bone shape, defined by a linear combination of the principal components computed by PCA during model construction:

$$x \approx \bar{x} + Pb \quad (1)$$

The vector b parametrises the deformation and is constrained to lie within ± 3 standard deviations from the mean, P is the eigenvector matrix, and x and \bar{x} are the mean model points and corresponding sample points, respectively. After each iteration the point correspondence between the model and the sample is re-computed and the process is repeated for a fixed number of iterations n . A value of $n = 100$ was used for this work, above which only negligible deformations were observed for any of the shape model experiments. The result of this iterative process is a model instance that best approximates the sample given the constraint on the vector b .

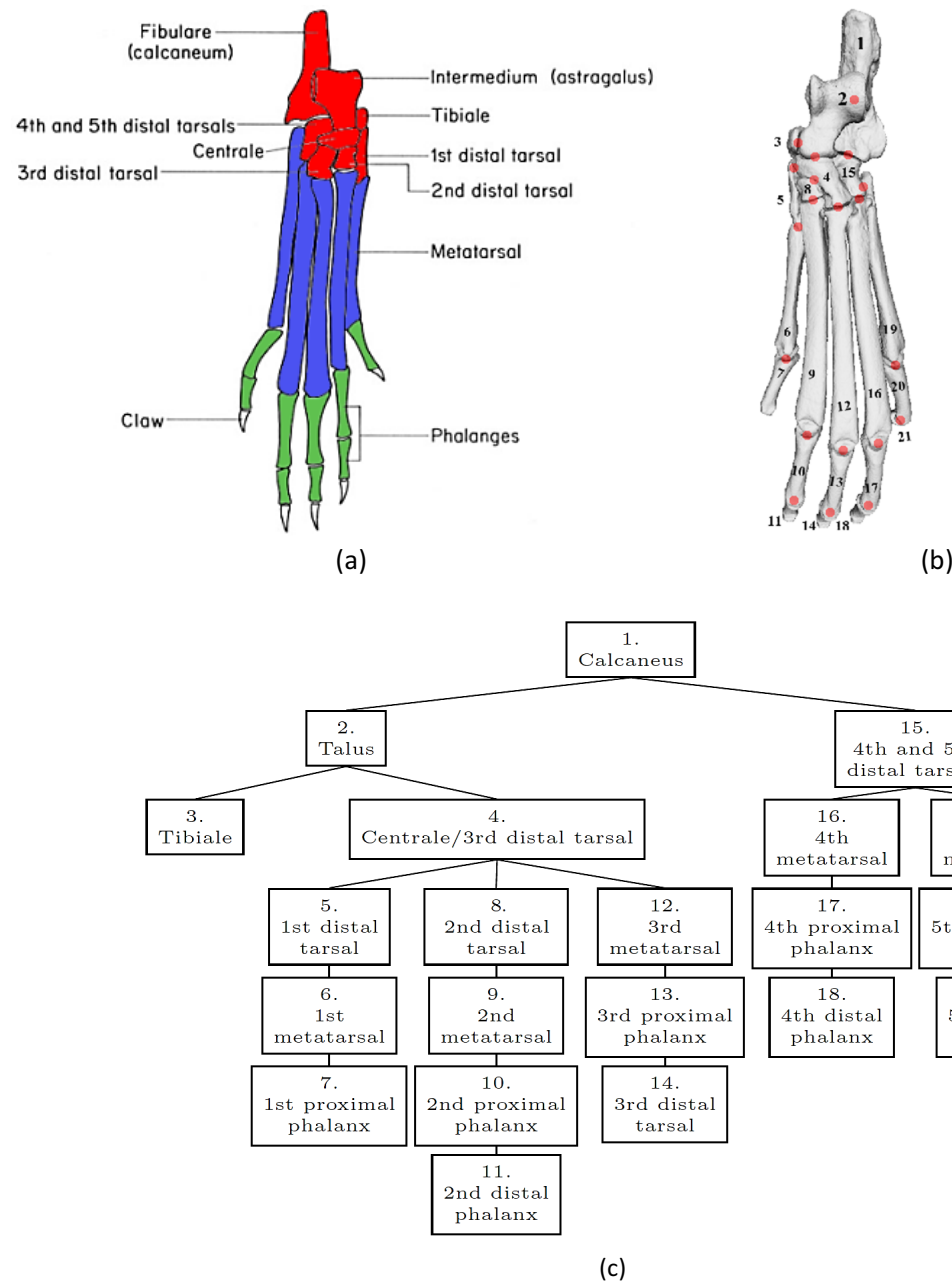


Figure 1. (a) Main anatomical regions: heel (red), metatarsals (blue), phalanges (green). (b) (c) The hierarchy of the mouse hind paw, with corresponding labels. The connections between nodes represent joints, shown as red circles. The heel (calcaneus) bone is the root node of the hierarchy, and is therefore the first bone to undergo alignment. Transformations applied to any parent node are inherited by the entire subtree, allowing for the connected bones to be initialised and registered accordingly. This particular traversal order was selected from a number of less successful approaches which produced unsatisfactory registration results in initial experiments.

2.5 Abnormality detection and analysis

This work aims to develop a tool with which to better understand the diversity in nature and severity of bone shape destruction in different arthritis mouse models. Of particular interest are the total bone loss (erosion) and gain (formation), the statistics of erosions and formations, and their magnitudes and locations for the individual bones and joints. By computing the signed distance between the corresponding points (vertices) in the sample and in the best-fitting model instance, the required quantities can be computed in an objective manner.

2.5.1 Abnormality detection, characterisation and visualisation

The process of fitting the model onto a sample allows for identification of abnormalities in terms of fitting error. Regions where the model and sample show good agreement are assumed to be non-pathological. Conversely, regions where the sample deviates from the model statistics are regarded as abnormal. Numerically, the Euclidean distance between corresponding points can be used as a measure of how closely the model matches the sample. Let $\vec{m} = \langle m_1, m_2, m_3 \rangle$ and $\vec{s} = \langle s_1, s_2, s_3 \rangle$ be a model point and its corresponding sample point, respectively. The unsigned Euclidean distance is calculated as:

$$E(\vec{m}, \vec{s}) = \sqrt{(m_1 - s_1)^2 + (m_2 - s_2)^2 + (m_3 - s_3)^2} \quad (2)$$

In order to differentiate between bone erosions and formations the signed Euclidean distance (SE) is calculated between the model and sample points. The model and sample meshes each consist of a set of vertices, connected to form triangular faces. For any given face its surface normal is computed as the cross-product of any two of its edges. For each vertex its normal \vec{N} is computed as the mean of the normals of the faces that are connected to it. The signed Euclidean distance

$$SE(\vec{m}, \vec{s}) = E(\vec{m}, \vec{s}) (\text{sgn}(\vec{N} \cdot (\vec{m} - \vec{s}))) \quad (3)$$

is then used to characterise and quantify the nature of deformations. Values close to zero are where the model and sample surfaces closely match one another. Positive values indicate the presence of elevations, whereas negative values are due to indentations. Physiologically, this translates into regions of deposition (gain, formation), and bone resorption (loss, erosion) respectively. This can be visually represented as a heatmap (see Figure 2 (d)). Figure 2 illustrates the consecutive steps of the process, starting from micro-CT data through registration, segmentation and labelling, through to heatmap visualisation.

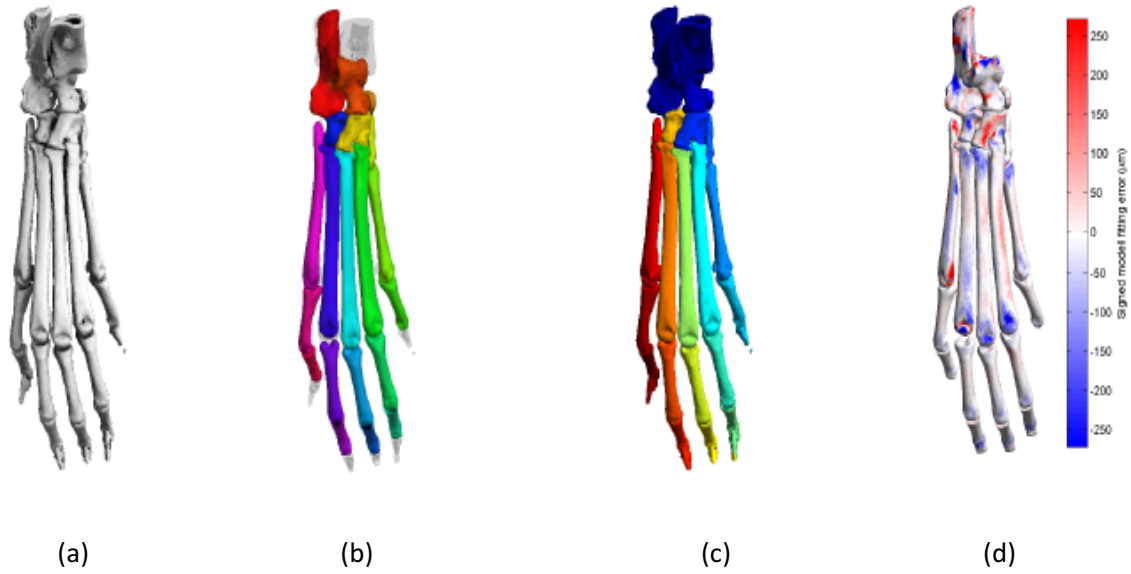


Figure 2. Visualisation of the consecutive steps of the process leading to quantification and visualisation of bone deformation. (a) Three-dimensional reconstruction of micro-CT data; (b) Results of registration of the articulated model (ASSM) to the sample shown in (a). Colours indicate the individual model bones; (c) Results of segmentation where different colours indicate labels assigned to the individual bones; (d) Heatmap showing the departure of the sample from the model of normal bone shapes. The red and blue regions of the heatmap correspond to bone formation and bone erosion, respectively.

2.5.2 Delineation of abnormal regions

In order to provide meaningful numerical indicators of bone deformations, abnormal regions are delineated and measured in terms of their surface area. A tolerance threshold is first defined to account for imperfections in registration and shape model fitting. The threshold is determined by calculating the mean Euclidean distance between model and sample vertices across the original training set. The statistics are computed for each bone independently, rather than for the limb as a whole, to account for the variability in registration accuracy between different bones. In the results presented, error values within three standard deviations of the mean error (per bone) are regarded as normal departures from the mean bone shape. Any vertices with error values above this threshold are therefore regarded as abnormal, and used in the abnormal surface area calculation.

In the next step the mesh faces adjacent to the abnormal vertices are grouped together to form labelled patches representing either erosion or formation. A patch is defined as a maximally adjacent set of mesh faces bearing the same label: negative (bone erosion), zero (normal) or positive (bone formation). The face labels are assigned by considering the sign of its vertices and the vertices of the adjacent faces. The most trivial case is where an individual face comprises three vertices with the same label, receiving the same label itself. Non-trivial cases emerge as a result of small surface irregularities where patches of faces are surrounded by a combination of negative, positive and zero-valued vertices. To remove these local irregularities a library of rules was defined to assign each mesh face the most consistent label. Labelling proceeds as follows. Each vertex is visited in turn, and the surrounding faces are assigned labels based on the current vertex label and the labels at neighbouring vertices. Given an input configuration of labels in clockwise order, and adjusted to be rotationally-invariant, the library provides the appropriate assignment of face labels. As each vertex is examined, the relevant rule is extracted from the library and the appropriate labels assigned to

the adjacent faces. Figure 3a shows an example of the library entry for positive vertices of degree three, their corresponding vertex labels and the face labels to be assigned. Figure 3b shows an illustrative example of the face label assignment for vertices of degree three and six (the maximum degree found in the surface mesh). This labelling system provides a segmentation of the mesh surface into three classes: normal, erosion and formation.

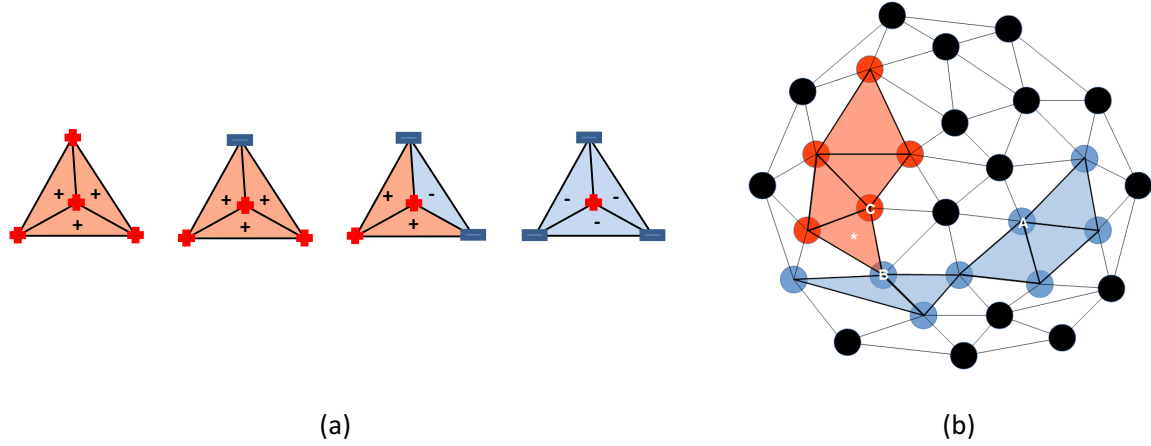


Figure 3. (a) A vertex library entry for positive vertices of degree three. Large red plus sign indicates a positive vertex, large blue minus sign indicates a negative vertex, red triangle face and small plus sign indicates that the face was labelled as formation, and blue face and small minus sign indicates the erosion label. (b) Labelling of faces as regions of bone erosion/formation. Each vertex is visited in turn, and the labels of surrounding faces assigned based on its own label. Vertex A shows a simple case, where its blue label is given to three of the surrounding faces, due to each of them having two blue vertices. At vertex B, despite there being a mixture of red and blue vertices, only the two blue faces are assigned. The face marked by an asterisk does not receive its red label until vertex C is visited.

Bone surface areas corresponding to the labelled patches are calculated by summing the areas of their individual mesh triangles. For any given triangle $\Delta = v_1v_2v_3$, its area is given by:

$$A(\Delta) = 1/2 |(v_2 - v_1) \times (v_3 - v_1)| \quad (4)$$

The overall degree of bone destruction measured for each bone is then expressed as the total percentage coverage, e.g. x% erosions, y% formations. In order to provide a more detailed characterisation of the abnormalities present, individual patches are considered. A queue-based connected component labelling algorithm (see Algorithm 1) assigns individual labels to maximally connected faces of the same abnormal type (erosion or formation). The patch surface area is calculated as a sum of the surface areas of the triangles forming the patch. The equivalent diameter D_{eq} is then calculated as the diameter of a circle with area equal to the patch size.

$$D_{eq} = 2 \sqrt{\frac{A(\Delta)}{\pi}} \quad (5)$$

Algorithm 1 A non-recursive algorithm for connected-component labelling of triangulated meshes

```
1: Initialisation:
2:  $BW \leftarrow$  binary labelling of mesh faces
3:  $N \leftarrow$  number of faces in the mesh
4:  $visited \leftarrow$  zero array of length  $N$ 
5:  $CC \leftarrow$  zero array of length  $N$ 
6:  $Q.push(0)$ 
7:  $currentLabel \leftarrow 1$ 
8:
9: Outer loop:
10: while  $Q$  is not empty do
11:
12:    $f \leftarrow Q[0]$ 
13:
14:   if  $BW[f] == 1$  then
15:      $W.push(f)$ 
16:
17:     Inner loop:
18:     while  $W$  is not empty do:
19:
20:        $g \leftarrow W[0]$ 
21:        $CC[g] \leftarrow currentLabel$ 
22:        $BW[g] \leftarrow 0$ 
23:
24:        $abnormalNeighbours \leftarrow$  indices of faces adjacent to  $g$  (where  $BW == 1$ )
25:        $BW[abnormalNeighbours] \leftarrow 0$ 
26:        $W.pop()$ 
27:        $W.append(abnormalNeighbours)$ 
28:
29:     end while
30:
31:      $currentLabel \leftarrow currentLabel + 1$ 
32:
33:   end if
34:
35:   Update:
36:    $visited[f] \leftarrow 1$ 
37:    $notVisited \leftarrow$  indices of faces adjacent to  $f$  (where  $visited == 0$ )
38:    $Q.pop()$ 
39:    $Q.append(notVisited)$ 
40:
41: end while
```

2.6 Method validation

The model's ability to characterise and quantify bone abnormalities was validated using digital phantoms. Three leave-one-out experiments were conducted using wild-type samples that had been computationally deformed to simulate the presence of bone erosions and formations. This was achieved by elevating or depressing the normal bone surface at a number of locations, with variations in their cross-sectional area and height/depth. To achieve ground truth that was biologically realistic the abnormalities were generated at locations which were observed in real data. Parameters characterising their size and shape were derived from the statistics of measured deformations in one of the diseased populations. An example of a sample with artificial bone

formations is shown in Figure 4. Parameters characterising their location and surface area were used as the “ground truth”, which was then compared to measurements generated by the ASSM.

Figure 5 shows the results of the validation for each sample, grouped by bone destruction type. This data provides a broad overview of the affected regions as the percentage of the total surface area affected. In all three samples the measured bone erosion is accurate to within 0.1% of total surface area. Its accuracy is relatively consistent across all three anatomical regions. For bone formation the results are more variable. The model is most accurate in the metatarsal region. In the heel region, the degree of bone formation is consistently underestimated. One possible explanation for this is a lack of variation in the model, preventing it from deforming inwards to fit the sample surface. This may have been caused by there being too few samples in the model, or poor point correspondence due to misregistration. In the phalanx region, bone formation is overestimated. Bone formation is rarely found in the phalanx regions of the mouse models studied, hence the absence of ground truth values in that region. Furthermore, the lowest articulated registration accuracy is generally found in the phalanx region, providing a poor initialisation for shape model fitting.

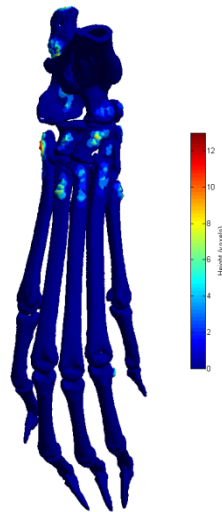


Figure 4. Sample with computationally deformed erosions and formations. The abnormalities were generated at locations which could realistically be observed in real data, and of a similar size and shape.

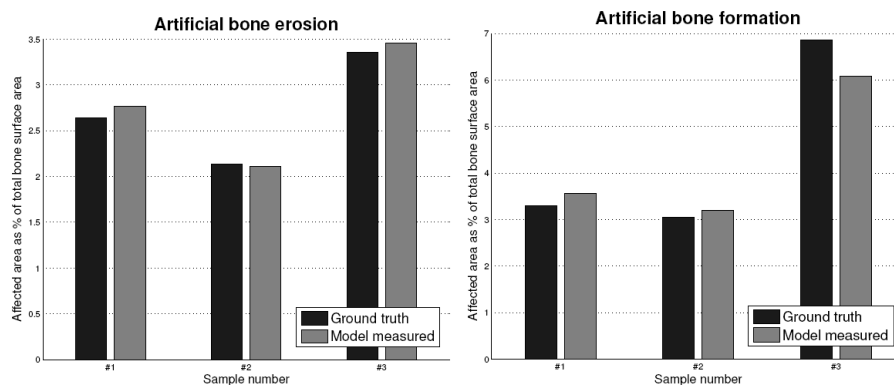


Figure 5. Results of three leave-one-out experiments using artificially deformed wildtype samples. The amount of overall bone erosion measured using the model is accurate to within 0.1 %. For bone formation, the model showed similar accuracy for samples 1 and 2 (~0.2 %) but produced a larger error for sample 3 (~1 %).

2.7 Comparison with bone destruction scores

A scoring system for evaluation of severity of bone destruction in micro-CT data was devised to enable comparison of the visual assessment with numerical scores obtained by computer analysis. Scores were assigned according to the criteria shown in Table 1 for each of the three main regions: heel, metatarsals and phalanges (see Figure 1). The scores were acquired by three independent observers and averaged to give a combined score for each hind paw sample. Measurements for abnormalities computed from the model are given as the percentage surface area affected, separately for erosion and formation. Although the data is not directly comparable, this test served to establish whether similar trends can be observed. Linear and rank correlation coefficients were computed as numerical indicators and scatterplots were produced to visualise relationships between the two scoring systems for the main anatomical regions (heel, metatarsals and phalanges), separately for erosions and formations.

Type	Visual Score	Observations
Bone erosion	0	Normal, no signs of erosion
	1	Roughness of bone surface
	2	Pitting / indentations
	3	Full thickness holes
Bone formation	0	Normal, no signs of formation / deformity
	1	Rough appearance, small osteophytes
	2	Spurs or signs of bone fusion
	3	Whole bone deformity / complete fusions

Table 1. Scoring system for bone destruction observed from micro-CT image data. Scores are assigned by independent observers.

Figure 6 (left) shows the scatterplot for visual scores of bone erosion against the percentage area affected by erosions, as measured by the model. Pearson correlation coefficient $r=0.62$ ($p<0.0015$) and Spearman rank correlation coefficient $\rho=0.63$ ($p<0.0015$) indicate strong correlation between the two. The visual scores reveal low levels of erosion in the phalanges as compared to the rest of the paw, which is recapitulated by the model measured results.

Figure 6 (right) shows the scatterplot for the visual scores of bone formation against the model measurements. Pearson correlation coefficient $r=0.69$ ($P<0.0002$) and Spearman rank correlation coefficient $\rho=0.76$ ($p<1.6 \times 10^{-5}$) indicate strong correlation between the two. As with the erosion data, relatively low levels of bone formation are found in phalanges by visual scoring. One sample, however, has significantly higher phalangeal bone formation scores than the other samples. In examining the image data, narrow strips can be observed along the lengths of several of the phalanges. These are likely to be image artefacts caused by incorrect misalignment compensation. Their narrow width also explains the model's inability to detect and measure their surface area.

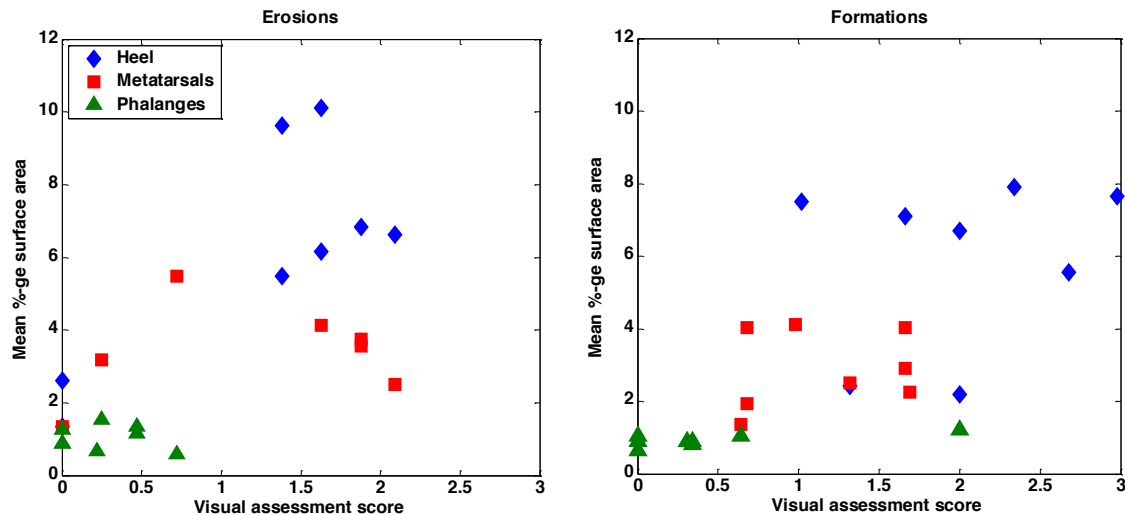


Figure 6. Comparison between visual erosion (N = 7) and formation scores (N = 8) and model measured erosions.

3 Results

This section demonstrates the use of the articulated statistical shape model (ASSM) as a tool for characterising the different ways in which bone destruction can take place in three phenotypically different mouse models.

3.1 K/BxN serum transfer-induced arthritis

Four K/BxN mice were analysed and compared. Two of the mice were “non-responders” which had no clinical signs of inflammation and received microCT visual scores of zero (see Figure 6). Micro-CT imaging studies indicate that normal bone integrity is maintained in the non-responders, showing no evidence of localised erosions. Application of the ASSM and the resulting heat maps are supportive of these observations, as shown in Figure 7 (a)-(b). The uniform baseline “elevation” that appears as a pink hue suggests that the non-responders may have thicker cortical bone than the wild type individuals. Figure 7(c)-(d) shows the two other samples that responded to the serum. Upon visual inspection the micro-CT data does not reveal any obvious signs of bone destruction. However, comparison with the normal model reveals both erosion and formation of bone, predominantly around the synovial joints. It is further clear that whilst both mice were responsive to the serum, there is a difference in severity between the two samples shown. The presence of bone formation along the metatarsals indicates that some abnormal thickening or perhaps bowing has taken place that is not accommodated by wild-type variation. Erosions can be found in common locations on both samples, such as the calcaneus, talus and metatarsals. Importantly, it is apparent that bone remodelling is not exclusively erosive. Evidence of bone formation can be found on the calcaneus, distal tarsals and metatarsals of both samples.

3.2 Collagen antibody-induced arthritis (CAIA)

Two CAIA mice were analysed using the ASSM, as shown in Figure 7(e)-(f). It can be seen that the two samples are very different from one another in terms of bone phenotype. In the first sample, a bulbous formation of bone on the distal region of the 4th metatarsal causes the registration to partially fail, resulting in subsequent misalignment of the connected phalanges. Elsewhere, the model reveals widespread bone remodelling (erosion / formation) of the calcaneus, talus,

metatarsals and phalanges. The second sample offers a very different bone phenotype consisting of full thickness bone erosions, and small studded patches of newly formed bone. Application of the model reveals speckled hotspots (shown in red) that are consistent with the sparse nature of the bone formation located on the calcaneus, distal tarsals and metatarsals. Full thickness erosions present on the metatarsal are not as well handled by the model, being detected either partially or not at all.

3.3 $TNF^{\Delta ARE}$ transgenic mice

Two $TNF^{\Delta ARE}$ mice were examined using the ASSM and the results are shown in Figure 7(g)-(h). The two samples are very similar in terms of the location and severity of the bone destruction present. Notable erosions can be observed on the 2nd - 4th metatarsals, as well as the flat portion of calcaneus. Full thickness erosions present on the metatarsal are not well handled by the model, being detected either partially or not at all. Interestingly, a small degree of bone formation can be observed along the bodies of the metatarsals which is indicative of bowing. This can be verified upon inspection of opposite side of the paw (not shown), where the apparent "bone loss" is in fact due to curvature of the bones.

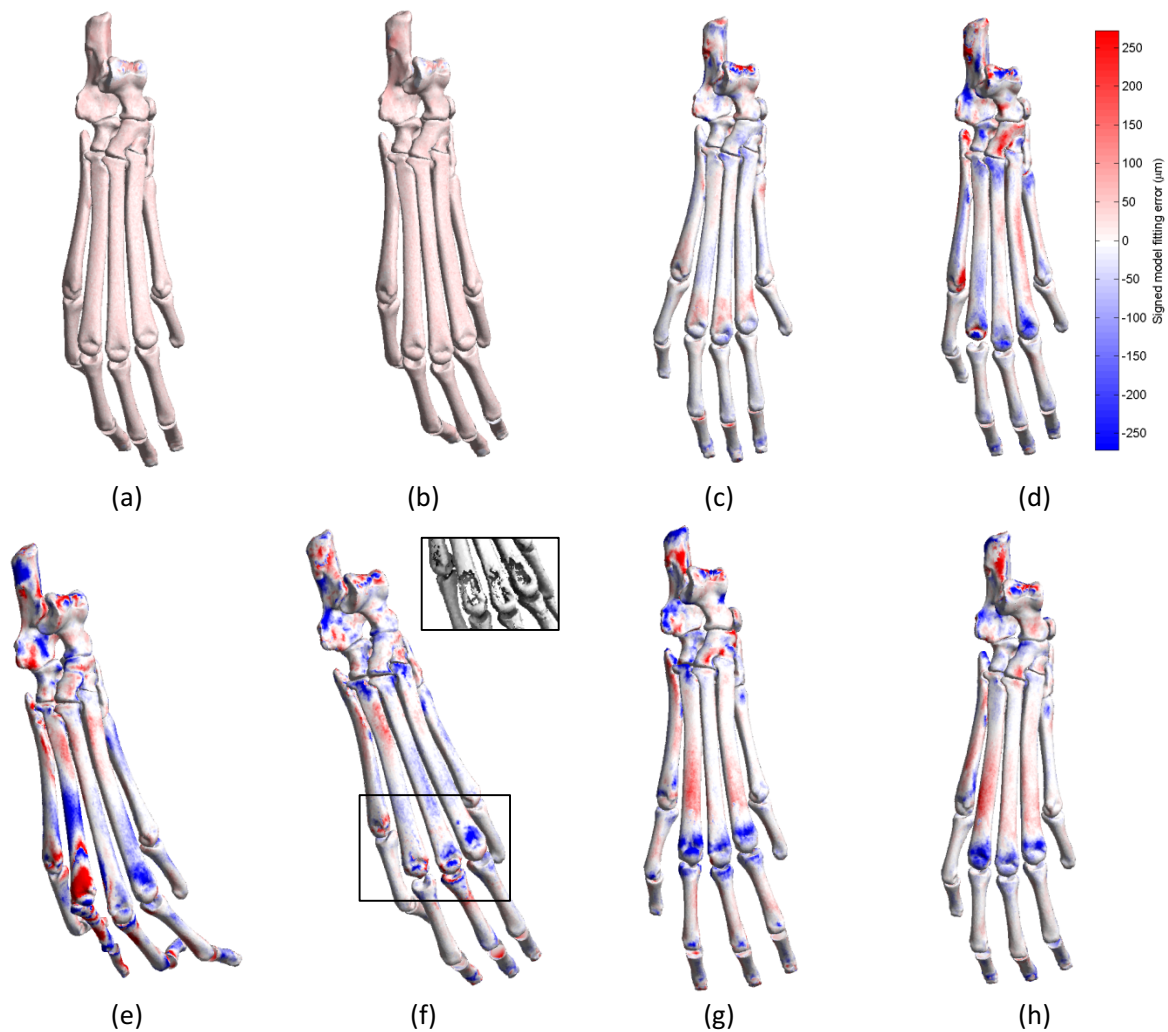


Figure 7. Heatmaps of three phenotypically different mouse models. The heatmaps show the departure of a sample from the model expressed as the signed distance (μm). The red and blue regions of the heatmap correspond to bone formation and bone erosion, respectively.

(a)-(b) Non-responsive mice after K/BxN serum-transfer. The micro-CT data does not show any obvious signs of bone destruction. With the exception of some uniform baseline error, the heatmaps show no evidence of localised bone erosion or formation having taken place.

(c)-(d) Responsive mice after K/BxN serum-transfer. In contrast to the non-responders above, it is immediately apparent from the heatmaps that bone remodelling has occurred. The degree of bone destruction differs between the two samples, with sample (c) showing smaller changes when compared with sample (d).

(e)-(f) Collagen antibody induced arthritis (CAIA) mice. The two samples present two very different bone phenotypes that reveal two important limitations of the proposed method; robustness to extreme bone formation, and sensitivity to full thickness erosions. The black arrow in (e) indicates a rarely observed, bulbous region of abnormal bone formation which caused the connected subtree to be misaligned. The resulting poor correspondence prevents the model from reliably identifying all the abnormalities present. The lack of sensitivity to full thickness erosions is most evident on the metatarsals in (f), where blue patches in the heatmap do not coincide with particularly deep erosions (see the inset showing micro-CT data where full thickness erosions are present).

(g)-(h) $\text{TNF}^{\Delta\text{ARE}}$ transgenic mice. The two samples are very similar in terms of the location and severity of the bone destruction present. Notable erosions can be observed on the 2nd - 4th metatarsals, as well as the at portion of calcaneus.

3.4 Quantitative comparisons

Figure 8 shows the mean scores for bone erosion and bone formation for the three mouse types. K/BxN mice are additionally subdivided into responders (RESP) and non-responders (NON): See methods (section 2.1) for detail of the non-responders.

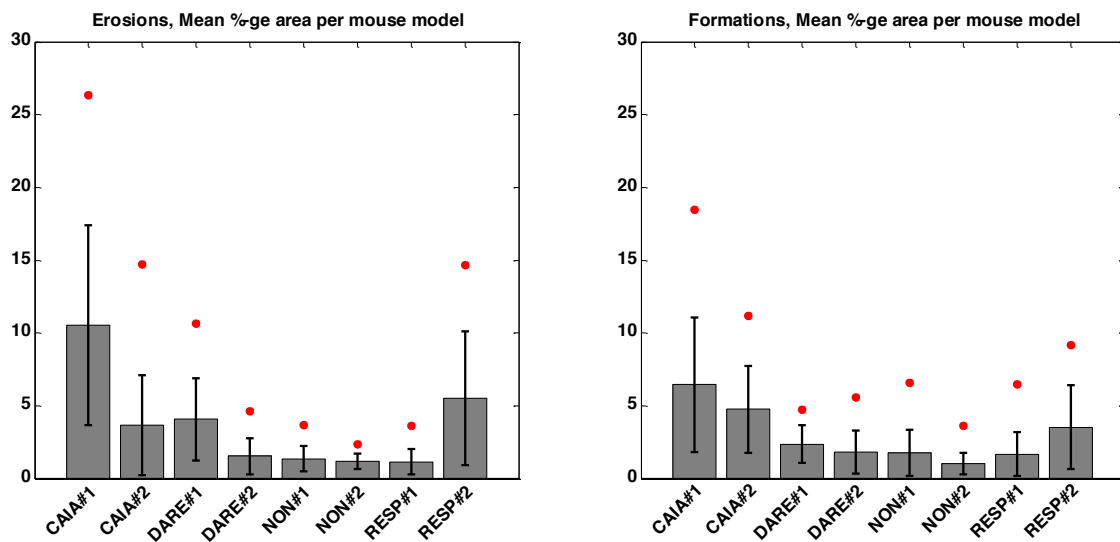


Figure 8. Mean scores for bone erosion and bone formation computed for the three mouse types. K/BxN mice are subdivided into responders (RESP) and non-responders (NON). CAIA and $\text{TNF}^{\Delta\text{ARE}}$ (DARE) are also shown. X-axis shows the mouse type for each individual mouse, Y-axis shows the total percentage area of erosions (left) and formations (right) per mouse. Bars show standard deviation. Red dots show the maximum percentage area.

Figure 9 illustrates differences in bone destruction for all the mouse types in the three gross anatomical regions: the heel, metatarsals and phalanges.

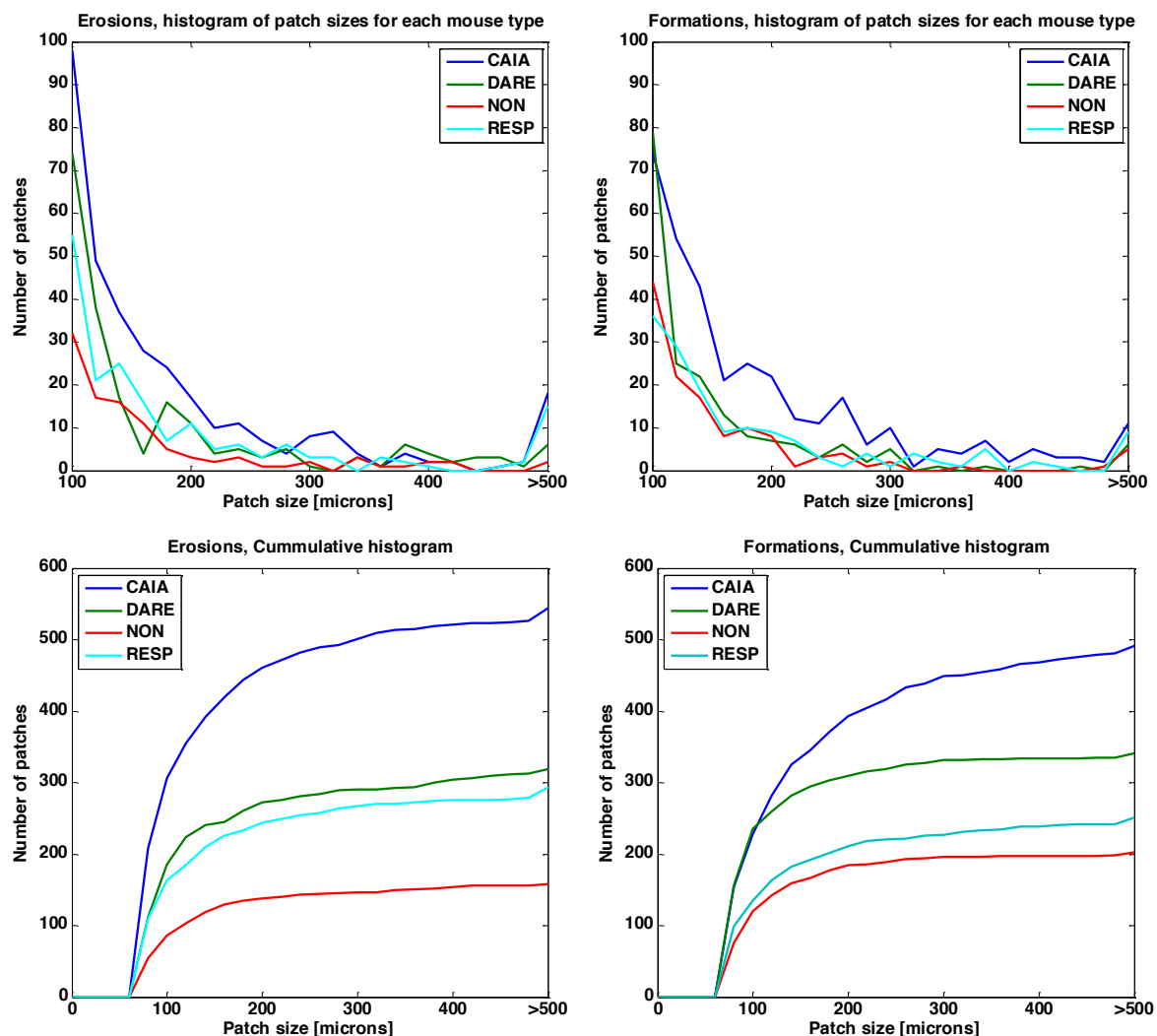


Figure 10. Histograms show the distribution of the patch sizes in the CAIA (dark blue), $TNF^{\Delta ARE}$ (DARE -green), K/BxN non-responder (NON - red) and K/BxN responder (RESP - light blue). Patch diameter (microns) was calculated for each mouse model for bone erosion (left) and formation (right). Top row: the distribution of all patch sizes; middle row: only patches above 100 microns to give better visual separation of the data; bottom row: cumulative patch sizes with base set at 80 microns, for better visualisation.

4 Discussion

The ultimate goal of this work is to provide the rheumatological research community with the necessary tools to characterise and quantify murine bone destruction in an entirely automated and unbiased fashion. Using computer software researchers would supply unprocessed micro-CT images and be provided with bone phenotype data that describes the type, severity and location of the bone destruction present. Such data would allow for reliable, detailed comparison of different disease models and treatments used in mouse studies. Furthermore, large volumes of image data could be assessed with minimal user interaction that would otherwise be prohibitively time consuming, and subject to inter-operator error. The work presented in this paper has proposed a number of measures, demonstrated the feasibility of their automatic derivation and applicability to answering questions of importance for RA research.

Although this study comprised a limited number of samples, it has served as a useful proof of concept by demonstrating the ability of the method to expose differences in pathology between several different mouse models with a range of severities and malformation types. In routine use it is envisaged that over 100 mice per annum will be examined. In such large studies quantification becomes increasingly important for conducting what would otherwise be labour-intensive assessments that are subject to operator biases and human error. Automation in this context is a significant advantage in terms of time saved and, more importantly, in standardising data output and comparability between samples.

4.1 Characterisation of bone destruction

4.1.1 Heatmaps

Heatmaps are three-dimensional representations of the signed distance between the model and vertices of a sample, mapped as red (formations) through to blue (erosions) on the sample surface. They provide an intuitive, semi-quantitative, representation of the bone deformations. Heatmaps were enthusiastically endorsed by the collaborating RA researchers for allowing them to perceive at a glance the size, location, nature and severity of areas affected by a disease. The location, magnitude and nature of deformations is not so clearly apparent in either pathology slides or 3-dimensional rendering of micro-CT slides.

4.1.2 Quantitative indicators

This paper has developed a number of quantitative indicators of bone destruction to aid the RA researchers in the objective characterisation of the disease models. They all derive from two assumptions: that normal bone shape anatomy can be sufficiently represented by a statistical shape model; and that pathological changes make bone shape statistics depart from the normal model. Consequently, the disease indicators are all based on differences between the model of normal bone anatomy and a given sample. A global measure, which however proved too crude, is a total distance between the model and the sample computed as the sum of differences between all the corresponding vertices of the model and the sample mesh representations. A less crude measure could distinguish between erosions and formations by taking into account the signature of the distance. However, averaged over all the vertices it proved too insensitive to local changes. The measures finally adopted consider individual erosion and formation patches on each bone separately. The percentage patch area as a fraction of the area of a given model bone provides a relative measure of destruction per bone, irrespective of its size. From this measure further statistics can be derived, for example erosions / formations for each individual bone (not shown), for all the bones at a given level of hierarchy (Figure 9), or for the entire sample. The equivalent patch diameter is an absolute measure (μm), based on which size distribution statistics can be computed (Figure 10). Another informative measure that could have been developed is a volume of each individual patch. This was easy to compute for bone formations, but proved difficult even to define for full thickness holes.

In the present implementation the bone destruction statistics are computed in relation to individual bones, rather than to joints which are predominantly affected in RA. The joints at which bone destruction is most visually discernible, and thus the most interesting for these studies, are the tarsometatarsal and metatarsophalangeal joints. Errors in segmentation due to narrowing of the joint space and the limited spatial resolution may in some cases result in inaccurate values for bone

loss and gain. In general, quantitative measures should be considered reliable for the periarticular surfaces, but less so for the articulating surfaces. Visualisations presented via heatmaps do not suffer so much in this respect as their main purpose is to indicate sites of abnormality, their general nature (erosions and formations) and relative magnitude. In mouse models of RA for histological assessment most researchers look at the ankle region including the joint between the calcaneum and distal tarsals. This is largely for pragmatic reasons as this large joint is most readily located in 2D and can be compared between samples. Heatmaps consistently depicted the pathology at the metatarsal joints as expected, in addition showing other areas around the tarsal and heel joints that were particularly affected.

4.1.3 Key biological findings revealed by the analysis

The analysis based on the articulated statistical shape model of a normal mouse hind paw has provided interesting insights into the nature of bone destruction in the three mouse models of RA. The initial observations are summarised below.

Although K/BxN and CAIA mice are very similar in terms of mechanism of arthritis induction and the transient inflammation produced, this work has identified apparent differences in the effects of this inflammation on bone structure. Application of the ASSM has revealed that whilst both models of arthritis are predominantly erosive, both also bear clear signs of bone formation having taken place. This is supported by previous studies of both models (Oestergaard et al., 2008; Ruiz-Heiland et al., 2012). The nature of the bone destruction is somewhat different as the CAIA serum-transfer model presents with more severe deformities such as the full thickness bone erosions and the presence of highly abnormal bone formations. This caused the ASSM model to fail occasionally, however it was still able to highlight several interesting bone shape differences including apparent “thinning” of cortical bone, as well as bone spurs on the calcaneus, located where the tendons attach to bone.

Histograms in Figure 10 show the distribution of patch sizes in the three mouse models, in addition making a distinction between responders and non-responders in the K/BxN mice. Consistently the ASSM detected the fewest patches of all sizes for the control mouse group (K/BxN non-responders) with the majority less than 100 micron in diameter.

The heatmaps generated by the model indicate that $\text{TNF}^{\Delta\text{ARE}}$ mice have a primarily erosive phenotype, which is consistent with reported findings (Kontoyiannis et al., 1999) (Jacques et al., 2013). The erosions are more pronounced than those observed in the other models, which follows given that $\text{TNF}^{\Delta\text{ARE}}$ mice are born with the mutation that produces chronic synovitis (joint inflammation) at an early age. Interestingly, this data also demonstrates relatively high rates of bone formation (osteophytosis) in contrast to the consensus in the published literature that bone formation in this model is minimal or absent (Jacques et al., 2013) (Caplazi et al., 2015). Further work is required to characterise the bone formation process fully but, given that the patches of bone formation are almost exclusively below 150 micron in size, it seems likely that they have previously escaped detection by eye. In comparison, the CAIA model has previously been noted as having high levels of bone formation (Jacques et al., 2013). This was confirmed by the ASSM and is particularly characterised by large (>200 micron) osteophytes (formations) as evident from the histograms (Figure 10) and by eye (Figure 7). Conversely, the K/BxN model had the lowest level of bone erosion of the three arthritis models and the cumulative levels of bone formation were equivalent to the control, indicating that this model triggers low levels of bone turnover. Given that the K/BxN and

CAIA models are conceptually relatively similar the large differences between them are perhaps surprising and, to our knowledge, have not been previously reported.

The data produced by the ASSM highlight the value of this approach to analysing micro-CT data from murine models of arthritis. These models are commonly used in rheumatology research and the micro-CT analysis is generally scored by eye. Our results demonstrate that the ASSM has the ability to detect bone changes that can discriminate between the different murine arthritis models commonly used in research. In addition, as identified by the $\text{TNF}^{\Delta\text{ARE}}$ bone formation results, this method has the capacity to capture information not readily visible by eye. The ability of the ASSM to detect small changes in bone resorption and formation will allow automated and unbiased analysis of the effect of genes and compounds on bone changes in in vivo models of arthritis. A comprehensive paper addressing the differences in the bone phenotypes resulting from various mouse models of RA is now in preparation.

4.2 Further work

To attain an ultimate goal of a fully automated method there are several areas requiring improvement. Construction of an ASSM based on a single reference sample (section 2.3) is inherently biased towards the geometry of the chosen sample. The effects of this bias were not investigated, but are considered an important issue, to be examined in future work. Large differences in pose observed in a number of diseased samples caused fully automatic model-to-sample registration fail. Similarly, failures occurred due to large departures from a statistically “normal” bone shapes at the high levels of the bone hierarchy. At present approximately one in five cases required some form of manual adjustment, usually applied to a single bone. Variations in relative bone position are handled to some extent by the kinematic constraints imposed during registration (Brown et al., 2014). However, in cases where mice have developed extreme structural deformities, registration of the model can fail. Occasionally problems occurred due to bone fusion at a joint and presence of severe abnormalities such as full thickness bone erosions. Some of these problems can be tackled by utilising non-rigid correspondence in order extract as much variation as possible from the training data. This would also improve the articulated registration accuracy at intermediate steps, minimising error propagation down the hierarchy. With a larger number of samples available in the future a statistical model of erosions and formations per disease model can be built by mapping erosions and formations to the template to show the likelihood of a deformation arising at a given anatomical location. Population-based statistical maps of this nature were previously reported to be useful in human studies (Cevdanes et al., 2010; Joshi et al., 2015; Joshi et al., 2013).

5 Conclusions

In this work an articulated statistical shape model has been demonstrated as a suitable approach for quantifying bone destruction in mouse models of RA. Validation experiments using *in silico* generated known deformations produced an error < 1% for most of the bones with the exception of a false positive error of about 4% for bone formations in the phalanx region. Application of the model to real data acquired from a number of mouse models has shown agreement both with the clinical scores and with the previous biological findings, suggesting that bone shape changes can be usefully identified as deviations from the model statistics. Quantitative data provided objective evidence for differences in bone deformation occurring in different mouse models of rheumatoid

arthritis. This quantitative data allows for unbiased and reproducible analysis of bone changes in arthritis models and, ultimately, it is hoped that the developed method may be employed in pre-clinical studies to assess the degree of bone protection conferred by potential therapies and to help answer questions about how different mechanisms contribute to RA pathogenesis.

Acknowledgements

This work was supported by Arthritis Research UK (Arthritis Research UK programme grant (19791) and an Arthritis Research UK Foundation Fellowship (AJN); and through the Engineering and Physical Sciences Research Council studentship (JMB) from the PSIBS Doctoral Training Centre (EP/F50053X/1). We are indebted to Prof. George Kollias, Fleming Institute, Greece for providing the TNF^{ΔARE} mouse (Kontoyiannis et al. 1999). We are grateful to Dr Iain Styles (School of Computer Science, University of Birmingham) for guidance and mentoring throughout this project.

The authors are grateful to anonymous reviewers for their helpful comments.

References

- Aya, K., Alhawagri, M., Hagen-Stapleton, A., Kitaura, H., Kanagawa, O., Novack, D.V., 2005. NF-κB–inducing kinase controls lymphocyte and osteoclast activities in inflammatory arthritis. *The Journal of Clinical Investigation* 115, 1848-1854.
- Besl, P.J., McKay, N.D., 1992. Method for registration of 3-D shapes, Robotics-DL tentative. *International Society for Optics and Photonics*, pp. 586-606.
- Boisvert, J., Cheriet, F., Pennec, X., Labelle, H., Ayache, N., 2008. Geometric variability of the scoliotic spine using statistics on articulated shape models. *IEEE Transactions on Medical Imaging* 27, 557-568.
- Bowden, R., 2000. Learning non-linear Models of Shape and Motion. Doctoral dissertation, Brunel University.
- Brown, J.M., Naylor, A., Buckley, C., Filer, A., Claridge, E., 2014. 3D articulated registration of the mouse hind limb for bone morphometric analysis in rheumatoid arthritis, In: Ourselin, S., Modat, M. (Eds.), *Biomedical Image Registration, 6th International Workshop, WBIR 2014, Proceedings*. Springer, pp. 41-50.
- Buckley, C., 1997. Science, medicine, and the future. Treatment of rheumatoid arthritis. *British Medical Journal* 315, 236.
- Caplazi, P., Baca, M., Barck, K., Carano, R., DeVoss, J., Lee, W., Bolon, B., Diehl, L., 2015. Mouse models of rheumatoid arthritis. *Veterinary Pathology*, 0300985815588612.
- Cevidanes, L., Hajati, A.-K., Paniagua, B., Lim, P., Walker, D., Palconet, G., Nackley, A., Styner, M., Ludlow, J., Zhu, H., 2010. Quantification of condylar resorption in temporomandibular joint osteoarthritis. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology* 110, 110-117.

- Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J., 1995. Active shape models-their training and application. *Computer Vision and Image Understanding* 61, 38-59.
- Ditzel, H.J., 2004. The K/BxN mouse: a model of human inflammatory arthritis. *Trends in Molecular Medicine* 10, 40-45.
- Feldkamp, L., Davis, L., Kress, J., 1984. Practical cone-beam algorithm. *Journal of the Optical Society of America A* 1, 612-619.
- Goldring, S., 2003. Pathogenesis of bone and cartilage destruction in rheumatoid arthritis. *Rheumatology* 42, 11-16.
- Jacques, P., Lambrecht, S., Verheugen, E., Pauwels, E., Kollias, G., Armaka, M., Verhoye, M., Van der Linden, A., Achten, R., Lories, R.J., 2013. Proof of concept: enthesitis and new bone formation in spondyloarthritis are driven by mechanical strain and stromal cells. *Annals of the Rheumatic Diseases*, annrheumdis-2013-203643.
- Joshi, A., Leahy, R., Badawi, R., Chaudhari, A., 2015. Registration-based bone morphometry for shape analysis of the bones of the human wrist. *IEEE Transactions on Medical Imaging* 35, 416 - 426.
- Joshi, A.A., Leahy, R.M., Badawi, R.D., Chaudhari, A.J., 2013. Morphometry for early monitoring of treatment response in rheumatoid arthritis, *IEEE 10th International Symposium on Biomedical Imaging (ISBI)*. IEEE, pp. 121-124.
- Kharitonov, V., 2003. Software CTAnalyser, version 1.10. 9.0. Kontich, Bélgica, DVD licenciado.
- Kontoyiannis, D., Pasparakis, M., Pizarro, T.T., Cominelli, F., Kollias, G., 1999. Impaired on/off regulation of TNF biosynthesis in mice lacking TNF AU-rich elements: implications for joint and gut-associated immunopathologies. *Immunity* 10, 387-398.
- Oestergaard, S., Rasmussen, K., Doyle, N., Varela, A., Chouinard, L., Smith, S., Qvist, P., Karsdal, M., 2008. Evaluation of cartilage and bone degradation in a murine collagen antibody-induced arthritis model. *Scandinavian Journal of Immunology* 67, 304-312.
- Pettit, A.R., Ji, H., von Stechow, D., Müller, R., Goldring, S.R., Choi, Y., Benoist, C., Gravallese, E.M., 2001. TRANCE/RANKL knockout mice are protected from bone erosion in a serum transfer model of arthritis. *The American Journal of Pathology* 159, 1689-1699.
- Pine, P.R., Chang, B., Schoettler, N., Banquerigo, M.L., Wang, S., Lau, A., Zhao, F., Grossbard, E.B., Payan, D.G., Brahn, E., 2007. Inflammation and bone erosion are suppressed in models of rheumatoid arthritis following treatment with a novel Syk inhibitor. *Clinical Immunology* 124, 244-257.
- Powell, M.J., 1964. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The Computer Journal* 7, 155-162.
- Romas, E., Bakharevski, O., Hards, D.K., Kartsogiannis, V., Quinn, J., Ryan, P., Martin, T.J., Gillespie, M.T., 2000. Expression of osteoclast differentiation factor at sites of bone erosion in collagen-induced arthritis. *Arthritis and Rheumatism* 43, 821.
- Ross, E., Naylor, A., O'Neil, J., Crowley, T., Ridley, M., Crowe, J., Smallie, T., Tang, T., Turner, J., Norling, L., 2016. Treatment of inflammatory arthritis via targeting of tristetraproline, a master

regulator of pro-inflammatory gene expression. *Annals of the Rheumatic Diseases*, annrheumdis-2016-209424.

Ruiz-Heiland, G., Horn, A., Zerr, P., Hofstetter, W., Baum, W., Stock, M., Distler, J.H., Nimmerjahn, F., Schett, G., Zwerina, J., 2012. Blockade of the hedgehog pathway inhibits osteophyte formation in arthritis. *Annals of the Rheumatic Diseases*, annrheumdis-2010-148262.

van de Giessen, M., Foumani, M., Streekstra, G.J., Strackee, S.D., Maas, M., van Vliet, L.J., Grimbergen, K.A., Vos, F.M., 2010. Statistical descriptions of scaphoid and lunate bone shapes. *Journal of Biomechanics* 43, 1463-1469.

Yang, S., Hollister, A.M., Orchard, E.A., Chaudhery, S.I., Ostanin, D.V., Lokitz, S.J., Mathis, J.M., 2013. Quantification of bone changes in a collagen-induced arthritis mouse model by reconstructed three dimensional micro-CT. *Biological Procedures Online* 15, 1.