Finite element analysis of the cranium: Validity, sensitivity and future directions

Analyse par éléments finis du crâne: validité, sensibilité et directions futures

Ricardo Miguel Godinho\textsuperscript{a,}\textsuperscript{*}, Viviana Toro-Ibacache\textsuperscript{b}, Laura C. Fitton\textsuperscript{a}, Paul O'Higgins\textsuperscript{a}

\textsuperscript{a} Hull York Medical School and Department of Archaeology of the University of York, York, UK
\textsuperscript{b} Centro de Análisis Cuantitativo en Antropología Dental and Instituto de Ciencias Odontológicas, Facultad de Odontología Universidad de Chile, Santiago, Chile

Keywords: Validation study; Sensitivity analysis; Finite Element Analysis; Geometric Morphometrics
Mots clés: Étude de validité; Analyse de sensibilité; Analyse par éléments finis; Morphométrie géométrique

* Corresponding author.

E-mail address: ricardomiguelgodinho@gmail.com (R.M. Godinho).
ABSTRACT

Finite Element Analysis (FEA) is increasingly applied in skeletal biomechanical research in general, and in fossil studies in particular. Underlying such studies is the principle that FEA provides results that approximate reality. This paper provides further understanding of the reliability of FEA by presenting a validation study in which the deformations experienced by a real cadaveric human cranium are compared to those of an FE model of that cranium under equivalent simulated loading. Furthermore, model sensitivity to simplifications in segmentation and material properties is also assessed. Our results show that absolute deformations are not accurately predicted, but the distribution of the regions of relatively high and low strains, and so the modes of global deformation, are reasonably approximated.

Résumé

La méthode des éléments finis (FEA) est de plus en plus appliquée dans la recherche biomécanique du squelette en général, et dans les études de fossiles en particulier. Ces études sont fondées sur le principe selon lequel les FEA fournissent des résultats qui se rapprochent de la réalité. Ce article fournit une meilleure compréhension de la fiabilité de la méthode des FEA en présentant une étude de validité dans laquelle les déformations subies par un vrai crâne de cadavre humain sont comparées à celles d'un modèle par éléments finis de ce crâne sous un chargement simulé équivalent. En outre, la sensibilité du modèle vis-à-vis de simplifications dans la segmentation et des propriétés des matériaux est également évaluée. Nos résultats montrent que les déformations absolues ne sont pas prédites avec précision, mais la répartition des régions de relativement hautes et basses contraintes, et par conséquent les modes de déformation globale, sont raisonnablement estimés.
1. Introduction

Over the last twenty years, Finite Element Analysis (FEA) has increasingly been applied in functional morphology and the biomechanics of extinct and living vertebrates (Rayfield, 2007; Richmond, 2007; Ross, 2005). It has been used to predict the mechanical behaviour of fossil material and so to infer function and ecology. However, the legitimacy of FEA is grounded on the premise that Finite Element Models (FEM) reflect the mechanical behaviour of the real structures they represent, that they produce valid results. The present paper extends previous work that assessed the validity of a model of the human cranium when simulating incisor bites (Toro-Ibacache et al., 2016). In this prior study, FEA approximated but did not exactly replicate experimentally measured strains. In particular, while the mode of deformation was reasonably approximated, the magnitudes were not. Further sensitivity analyses showed that the mode of deformation was altered little by varying the segmentation approach (i.e. what parts of the model segmentation are assigned to and given the material properties of teeth, cortical or cancellous bone), but the magnitudes of deformation were reduced approximately proportionately as the model was made more dense by replacing cancellous with cortical bone. Thus, in this model when simulating an incisor bite, a simple segmentation that preserves geometry (i.e., external form and major internal cavities) but effectively treats the whole of the rest of the cranium as if it were solid, and made of cortical bone and teeth approximated the mode of deformation found in experimentation with the real cranium. In this paper we extend this work to a simulated molar bite, to assess validity and sensitivity using the same cranium.

To assess validity, results from a finite element analysis (i.e., stress, strain and deformation) should match those obtained from the real specimen when the same loading regimen is applied (Bright and Rayfield, 2011; Grine et al., 2010; Kupczik, 2008; Rayfield, 2007; Richmond et al., 2005). While some studies have measured strains experienced by skeletal elements in vivo (Ross, 2001; Rubin and Lanyon, 1982) carrying out physical loading experiments on living individuals carries with it both ethical and practical difficulties which limit its use, as such, usually measurements are taken from loading experiments ex vivo with postmortem material (Bright and Rayfield, 2011; Groning et al., 2009; Kupczik et al. 2007; Richmond et al., 2005; Strait et al.,
As such, validation studies have typically compared the strain magnitudes and/or vector orientations experienced by a specimen \textit{ex vivo} with those predicted by the virtual finite element model following the simulation of a load (Bright and Rayfield, 2011; Szwedowski et al., 2011).

Until recently the most viable option to measure deformations was to affix strain gauges to bone (Bright and Rayfield, 2011; Daegling and Hylander, 2000; Ichim et al., 2007; Kupczik et al., 2007; Richmond et al., 2005; Strait et al., 2005; Szwedowski et al., 2011; Vollmer et al., 2000). However, these present some limitations (Richmond et al., 2005), which include technical difficulties associated with fixing gauges (Groning et al., 2009) and limits to the number of gauges that can be applied leading to the impossibility of dense measurement of strains being collected over regions of interest (Bright and Rayfield, 2011; Evans et al., 2012; Groning et al., 2009, 2012; Yang et al., 2007). To overcome these limitations Digital Speckle Pattern Interferometry (DSPI) has been applied to measurement of bone surface strains in validation studies (Bright and Groning, 2011; Groning et al., 2009, 2012; Toro-Ibacache et al., 2016). This is an optical full-field strain measurement technique that allows strains to be directly measured over a small area (several cm$^2$), determined by the field of view of the device.

Most validation studies have reported a degree of success in predicting skeletal behaviour using FEA. This said, models often fail to accurately reproduce absolute strain magnitudes while relative strain magnitudes between different regions of the model are generally consistent with relative strains from experimental loadings of the real specimen (Bright and Rayfield, 2011; Kupczik et al., 2007; Strait et al.; 2005; Toro-Ibacache et al., 2016). With regard to the cranium, it has been suggested that differences between the performance of \textit{in silico} models and actual skeletal material may be related to regional differences in material properties, to the presence of complex patterns of heterogeneity and orthotropy, and the difficulty in correctly reproducing variations in cortical bone thickness and cancellous bone architecture, given the constraints of imaging and model building (Bright and Rayfield, 2011; Ross, 2005; Strait et al., 2005; Szwedowski et al., 2011). This suggestion arises from studies that report regional differences in material properties of cortical bone in the human craniofacial skeleton (Dechow et al., 2010; Peterson and Dechow, 2002, 2003; Peterson et al., 2006). On the other hand, strain orientation in validations is commonly consistent between virtual and
physical specimens (Bright and Rayfield, 2011; Cuff et al., 2015; Porro et al., 2013; Toro-Ibacache et al., 2016). With regard to issues with resolution, and so segmentation, of cortical and trabecular bone, sensitivity analyses suggest that these mainly affect strain magnitudes (overall model stiffness), but less so strain vector orientations and relative strains (Fitton et al., 2015; O'Higgins and Milne, 2013; Parr et al., 2012; Toro-Ibacache et al., 2016).

While it would be desirable to have virtual models that exactly reproduce the performance of the real specimens, this can only be known for certain if experimental data for the full surface are available to guide model building. In most cases this is not possible, yet FE models are still useful to predict relative strains within and among models which renders FEA a useful approach with respect to many questions (Bright and Rayfield, 2011; Milne and O'Higgins 2012; O'Higgins and Milne, 2013; Strait et al., 2005), including the comparative cranial and post-cranial biomechanics of hominoids (Richmond, 2007; Wroe et al., 2007, 2010), hominins (Strait et al., 2009, 2010; Wroe et al., 2010) and recent human populations (Püschel, 2013).

Unknown input parameters and model simplifications are inevitable, particularly for studies of fossil hominins. As such, in order to improve accuracy of FE strain prediction and better understand the modelling process, sensitivity analyses should be carried out to assess how differences in model building approaches impact predicted strains and forces. Previous studies have examined the effects of variations in: model geometry (Strait et al., 2005), model constraints (Cox et al., 2011; Marinescu et al., 2005), material properties (Bright and Rayfield, 2011; Marinescu et al., 2005; Szwedowski et al., 2011), muscle loading magnitudes and vector directions (Fitton et al., 2012; Marinescu et al., 2005; Ross, 2005; Toro-Ibacache and O'Higgins, 2016), PDL segmentation (Groning et al., 2011; McCormack et al., 2014), modelling of cranial sutures (Kupczik et al., 2007; Wang et al., 2008, 2010, 2011) and static vs. dynamic loading (Wang et al., 2011) on model performance.

The answer to the question “is a model valid?” is dependent on the future application of the model and the questions to be addressed (Bright and Rayfield, 2011; Strait et al., 2005). Validations of FE models of fossil hominins are not possible, however if models of modern Homo sapiens can produce valid results, and the
researcher is aware of the models' sensitivity to input parameters, then FEA of fossil
hominins opens up in a meaningful way.

The present validation study and sensitivity analysis follows on from a previous
study by Toro-Ibacache et al. (2016) that investigated the validity of predicted strains in
a human cranium arising during loading and how modelling decisions impacted on
model performance. In this study, as in that of Toro-Ibacache et al. (2016), full field
strains are measured over the infra-orbital regions using DSPI technology. Validity of
model behaviour is investigated under an approximation to a molar bite whereas Toro-
Ibacache et al. (2016) approximated an incisor bite (see details of loading and
constraints below). The impact of simplifications of segmentation, and thus material
properties, is also investigated. These validation studies and sensitivity analyses are
relevant for fossil FE studies, in which validation is not possible, because they assess if
FE modelling decisions provide reliable relative strain magnitudes.

It is hypothesized that: (H\textsubscript{0}1) there are no differences in the magnitudes and
orientations of strains between the actual cranium under controlled experimental loading
conditions and the FE model under identical simulated loading; (H\textsubscript{0}2) there are no
differences in the distribution of regions of relatively low and high strains (here referred
to as strain pattern), irrespective of overall magnitude between predicted and
experimentally measured strains; (H\textsubscript{0}3) simplifications in segmentation do not impact
on strain magnitudes; (H\textsubscript{0}4) simplifications in segmentation do not impact on strain
pattern. It is predicted that H\textsubscript{0}1 and H\textsubscript{0}3 will be falsified and that H\textsubscript{0}2 and H\textsubscript{0}4 will not
be falsified.

2. Materials and Methods

This study compares the strains experienced by the infra-orbital region of a real
human cranium under mechanical loading with those of an FE model under equivalent
virtual mechanical loading. Moreover, the impact of simplification in segmentation, and
thus material properties, was also assessed. The loads and constraints approximate a
molar bite but were not designed to replicate physiological loading (see details below).

The cranium is that of a 74 year old male from Hull, England. Use for research
was approved under the Human Tissues Act 2004 (available at www.hta.gov.uk) and by
the relevant Hull York Medical School Ethics committee while it was cared for and handled during all experimental stages in accordance with the act and regulations. The same cadaveric head was previously used by Toro-Ibacache et al. (2016) for validation of masticatory function in an approximation to an incisor bite. In that study the head was dissected, segmented, modelled as an FE model and mechanically loaded to approximate the simulation of an incisor bite. The cranium was extremely well preserved, with no missing skeletal elements (with the exception of a congenitally absent premolar in each quadrant; see Fig. 1), and bone lightly embalmed and still wet (kept so by regular spray application of water between experiments). Therefore this is a suitable cranium for simulation of masticatory function in a living individual.

2.1. In vitro experiment

It is not possible, or necessary for the purpose of this investigation, to load and constrain the cranium to replicate muscle action, so simplified boundary conditions were applied and physically loaded onto the skull. The cranium was constrained by resting it on the two mastoid processes and the meso-lingual cusp of the first left molar (Fig. 1). A total compressive vertical force of 750 N was applied to the frontal bone, 15 mm anterior to bregma in increments of 75 N using a universal material testing machine (Lloyd's EZ50, Ametek-Lloyd Instruments Inc., UK). The mastoids acted as a fulcrum at which the skull rotated upon application of the downward compressive force to the frontal bone. The tooth acted as the biting point, transferring the load onto a 5 kN load cell (Omega DP25B-S; Omega Engineering Inc., Stamford, CT, USA) previously calibrated and repeatedly checked and recalibrated between experimental runs, by applying known compressive loads with the Lloyd’s testing machine described above. The cell was used to measure the reaction force at the tooth during loading and was used to check it scaled linearly with the compressive force applied by the testing machine.

A Q100 DSPI measuring system (DANTEC Dynamics GmbH, Ulm, Germany), which provides a maximum field of view of 25 X 33 mm, was used to measure the strains experienced by the infra-orbital plate during mechanical loading. That region was covered with a thin layer of white spray (DIFFU-THERM developer BAB-BCB; Technische Chemie KG, Herten, Germany) to prevent light reflection. The DSPI sensor
was then attached to a tripod that was glued to the infra-orbital plate using an acrylic-based adhesive (X60; HBM Inc., Darmstadt, Germany). The software Istra Q1002.7 (DANTEC Dynamics GmbH, Ulm, Germany) was used to characterize the surface of the region of interest before loading (Fig. 1B) and create phase maps prior to and during loading (Fig. 1A) that were used to estimate maximum ($\varepsilon_1$) and minimum ($\varepsilon_3$) principal strain magnitudes and directions. The software then converts the strains into 2D (Fig. 1D) strain contour plots. A 3D contour plot surface of the measured maximum and minimum principal strains is also created that may be superimposed onto the surface of the virtual model (Fig. 1E). In each experimental round the cranium was loaded with a total of 750 N in 10 increments of 75 N, as previously described. In between rounds, the load was removed and the cranium was checked to ensure it did not move and that the load was completely removed. From a total of eight experimental rounds four were inconsistent in strain magnitudes and/or directions, attributable to experimental errors (in reading consistent strain magnitudes, possibly due to undetected interferences/slippage; and, eventually, bone failure due to the high number of experimental rounds, after which no further experimental loading was performed). Thus, these four rounds were discarded and only the consistent rounds are presented and used as reference for the FEA simulations.

2.2. **FE experiment**

The FE model was based on a CT scan of the cadaveric head, performed before dissection and mechanical testing, using a Siemens 16-channel multidetector CT scanner equipped with a STRATON tube (Siemens Somatom Sensation 16; Siemens Healthcare, Erlangen, Germany) at 120 kV and 320 mA with an H60s edge-enhancing kernel, with an original voxel size of 0.48 X 0.48 X 0.7 mm that was later resampled to 0.349 X 0.349 X 0.349 mm. Three materials were originally segmented by Toro Ibacache et al. (2016), cortical bone, trabecular bone and teeth. Further segmentation was performed by thinning cortical bone in regions that presented cancellous bone. The cortical bone segmented from the internal surface of the cortical shell was subsequently allocated to cancellous bone, which was segmented as a bulk material. This additional segmentation was performed to make the model less stiff because preliminary FE
simulations yielded strains that were much lower than, but proportional to, those from the experiment. Moreover, because the resolution of the scan is relatively low, partial volume averaging artificially thickens cortical bone, thus requiring further segmentation in the cortical-cancellous bone boundaries. The refined model was later exported as a BMP image stack and directly converted into a voxel based finite element mesh using the bespoke software vox2vec. Pre and post-processing used the custom FEA program VoxFE (Fagan et al., 2007; Liu et al., 2012). Isotropic material properties were allocated to all materials. Young’s modulus of cortical bone was calculated by nanoindentation (cortical bone 16.3±3.7 GPa – 21.9±2.7 GPa) as previously reported by Toro-Ibacache et al. (2016) and within the range of previously reported values (Dechow et al., 1993; Schwartz-Dabney and Dechow, 2003). Trabecular bone was allocated an elastic modulus of 56 MPa, according to Misch et al. (1999). Teeth were modeled as single structures with a modulus of 50 GPa, within the published values for enamel (for a review, see He and Swain, 2008; He et al., 2006). All materials were allocated a Poisson’s ratio of 0.3. Model segmentation and material properties of this reference model (model 3) were then simplified by creating a two materials (cortical bone and teeth; model 2) and a one material model (cortical bone; model 1). In model 2 trabecular bone was combined with cortical bone. In model 1, trabecular bone and teeth were integrated into and represented as material with the properties of cortical bone.

The model was virtually positioned similarly to the real cranium so that the force vectors of the compressive applied loads matched between the real and virtual structures. Constraining of the model replicated the constraints applied to the real cranium. Thus, constraints were applied to both mastoid processes and to the meso-lingual cusp of the first molar. The mastoid processes were contrained at 32 nodes in all directions (left-right, anteroposterior and superiorinferior). The first molar was constrained at 21 nodes in the superiorinferior direction.

2.3. **In vitro vs. FE experiment**

Comparison between real and predicted strains is only meaningful if the same loading regimen is applied because differences in force magnitude and/or direction would potentially result in different strain distributions, magnitudes and directions.
Thus it was necessary to verify the force vectors had the same orientation in the real and virtual experiments. To that end reaction forces registered by the force transducer placed under the tooth were compared to the reaction forces calculated at the tooth of the FEM. Congruence of results can be interpreted as indicative of equivalent orientations of applied force vectors.

After comparison of tooth reaction forces, the deformations of the loaded crania were compared. In vitro and predicted strain directions and magnitudes were qualitatively assessed by visually comparing strain contour plots of maximum and minimum principal strains and strain vector directions. Quantitative assessment of differences between measured and predicted strain magnitudes relied on plotting magnitudes extracted along two equivalent lines traced in the DSPI and FEA outputs (Figs. 1B and 2A). This approach was chosen because the DSPI and FEA outputs differ in number of dimensions (2D in the DSPI and 3D in FEA) and in resolution, thus precluding direct (point to point) comparison between results. Line 1 was placed using vertices of the field of view of the DSPI to maximize reproducibility and crossing the whole FOV to maximize the number points at which strains were sampled. It extends from the alveolus above the lateral incisor to approximately above zygomaxillare. Line 2 was placed so it would sample a different region and extends from close to alare (one of the vertices of the FOV) to approximately the maxillary root of the zygoma. The lines traced in the DSPI were reproduced in the virtual cranium by manually best fitting the 3D strain contour plot created by the Istra Q 1002.7 program onto the virtual cranium and using it as a reference to trace the equivalent lines. The strain magnitudes extracted along each of the two lines from the experimental rounds were arithmetically averaged and ± two standard deviations were computed. This provided a range against which FEA results were compared.

The impacts of varying segmentations (material properties) of models were first assessed by visual inspection of the strain contour plots. Quantitative assessment of sensitivity is based on maximum ($\varepsilon_1$; predominantly tensile) and minimum ($\varepsilon_3$; predominantly compressive) principal strain magnitudes along the above described lines (Fig. 2A) and at thirty two points over the face (Fig. 2B). These magnitudes are compared to quantify the impacts of simplifying material properties on principal strains over the face.
To assess the sensitivity of global modes of deformation to segmentation simplifications, a principal component analysis (PCA) of changes in model sizes and shapes due to loading was performed using 51 landmarks placed over the whole cranium (Fig. 2C). An additional ‘incisor bite’ was simulated to allow the magnitudes of ‘error’ due to varying modelling parameters to be calibrated in relation to the magnitude of differences resulting from application of a different bite point. The ‘incisor bite’ used the reference model 3 and differs only in that the constraint at the molar was replaced by a similar constraint at the left central incisor.

3. Results

As previously described, comparison of deformations uses strain contour plots of $\varepsilon_{1}$ and $\varepsilon_{3}$ and plots of strain magnitudes extracted from two lines located in the infraorbital region that are equivalent between the DSPI and FEA outputs. In the comparison, directions, magnitudes and patterns are considered. Directions refer to the orientations of strain vectors. Magnitudes in the strain contour plots are colour coded. Thus, similar colours indicate similar magnitudes. ‘Strain pattern’ here refers to the spatial distribution of regions where high and low strains are experienced. It is compared using strain contour plots and plots of strain magnitudes. In the strain contour plots strain pattern is similar if the spatial distribution of regions of relatively high and low strains is equivalent, even if magnitude differs. In the plots of strain magnitudes, strain pattern is similar to the extent that the plotted curves follow similar courses, even though absolute strain magnitudes differ.

3.1. Real vs. predicted strains

Reaction forces measured during the consistent experimental rounds at the tooth were averaged to be used as reference for the FEA models. The mean reaction force is 424 N and that of the FEA is 425.61 N. Thus, the difference is negligible (0.38%) and the loading regimens considered equivalent.

When comparing the strain contour plot of the DSPI with the FE model it is apparent that the virtual models are stiffer, despite presenting comparable strain patterns
Moreover, there is a slight mismatch in the anatomical regions where the highest/lowest strains are experienced. This is visible, e.g., on the alveolar bone overlying the root of the left canine. In the 3D contour plot of the DSPI the alveolar bone experiences high strain in the root and low strain in the interradicular region. In the FEA strain contour plot the alveolar bone over the root of the canine shows low strain and high strains are located in the inter-radicular regions. Despite this difference, comparison between the DSPI output within the corresponding area of the FEA results shows some similarities (Figs. 3 and 4A). In ε1, these consist of a region of strain peaks that generally envelops the left canine and premolar, rising from the inter-radicular area between the lateral incisor/canine and the premolar/first molar and converging medially to the infra-orbital foramen. This high strain drops in the region overlying the roots of the premolar and, especially, of the canine. Above the first molar, at the root of the zygomatic process of the maxilla, there is a low strain area, more distinct in the real specimen, from which strains increase both anteriorly and posteriorly. In ε3 the main difference between the DSPI output and the FE model is in the lower relative strains displayed by the real specimen around the first molar and above the canine. In the remaining areas there are similarities between the models, with a relatively highly strained region that extends from the root of the zygomatic process to the inferior lateral border of the nasal aperture. In this region there are also foci of low strains, both in the real and virtual specimens, such as in the maxillary zygomatic process, the area next to the infra-orbital foramen and over the premolar roots (Figs. 3 and 4A). The numbers in Fig. 4A indicate the anatomical points at which strains peak in the real cranium (top row) and the FE model (bottom row); they are not exactly anatomically equivalent and so indicate a mismatch between the model and real specimen.

Fig. 4B shows some similarities but not consistency in detail between DSPI and FEA strain vector directions (note the vectors from FEA are automatically scaled by the software according to magnitude while those from DSPI are not). Similarities in ε1 are generally found in the bottom half of the strain vector and contour plots, with differences being more pronounced in the top half of each. In ε3 similarities are found in the left half of the plot, where supero-inferior compression (vectors oriented obliquely) predominates. Differences in vector directions are found over the root of the canine (bottom centre), both in ε1 and ε3.
In order to quantitatively compare real and predicted strain magnitudes these were extracted along lines 1 and 2 (Fig. 2). Fig. 5 presents a plot of the extracted strain magnitudes with peak strains, labelled according to Fig. 4, where the anatomical locations of these in the model and experiment are shown. In line 1 (Fig. 5A-B) there are differences in the absolute values of $\varepsilon_1$ between the predicted and the real strains. These differences are most marked throughout approximately the first third of the line. Despite these absolute differences, the general strain pattern is comparable. Thus, DSPI and FE results present the same strain peaks around the canine (points 1 and 2), despite a slight mismatch in the exact location of the peaks. When considering the second and third peaks (points 2 and 3) both DSPI and FEM outputs almost match in absolute values, with the main difference being that there is a significant drop in the experimental strains between points 2 and 3 in Fig. 5A, which is absent in the predicted strains. In $\varepsilon_3$ (Fig. 5B), there are also differences in the absolute values, with predicted strains being lower than real strains. Despite these differences the overall pattern of $\varepsilon_3$ (Fig. 5B) between real and predicted strains is similar, with a generally flat and regular line. The main differences between DSPI and FEA outputs are in the peak of predicted strains next to the premolar ($\approx 15$ mm), which is absent in the DSPI output, and the flatness of the first third of the line of the predicted strains (over the canine root), which is more irregular in the real strains.

In the scatter plot of $\varepsilon_1$ from line 2 (Fig. 5C) there are two main differences between real and predicted strains. The first is the significant drop in the real strains at $\approx 7$-8 mm, that anatomically corresponds to the area above the apex of the canine. In the FE model these low strains are directly over the tooth root, not extending upwards. The second difference is the drop in the predicted strains at about the apex of the premolar ($\approx 20$ mm, just before point 5 of Fig. 5C). Despite these two local differences the predicted absolute strains are mostly within the range of the real ones, with a generally similar pattern between the two. In $\varepsilon_3$ (Fig. 5D), both DSPI and FE extracted strains present a remarkable match. The only difference is found in the higher relative strains experienced by the FE model just next to the bite point (point 7, Fig. 5D).

The effects of simplifying the segmentation, and so reducing the numbers of different materials, are most marked near the teeth. This effect is, however, much smaller with increasing distance from the teeth. This is apparent is Figs. 5 and 6. The
former samples points directly over or close to teeth and shows increasing discrepancies between the FEA models as points come to lie directly over or very close to the teeth. In Fig. 6 this effect is even clearer, with almost complete overlap between models at points throughout the face that are distant from teeth and divergence between models at points over or nearby teeth (points three, eighteen and especially ten). These findings indicate that differences in material properties allocated to teeth and bone have marked local effects on predicted strains.

3.2. **Global modes of deformation**

Analysis of large scale deformations (Fig. 7) shows that simplification of model 3 (to models 2 and 1; two then one material) impacts on magnitude (distance from unloaded model) and a little less so on mode of deformation (divergence between vectors from unloaded to loaded models). Models 1 and 2 plot directly over each other. Thus, the effect of reducing materials from 2 to 1 is minimal but reducing from 3 to 2 impacts the mode and magnitude of deformation. Anatomically, the difference in mode of global deformation between model 3 and the two simplified models corresponds to a greater degree of torsion of the face (see inset warpings and transformation grids) towards the working side, which is seen as a relatively more superior deflection of the teeth and palate and greater supero-inferior shortening of the left orbit in model 3. These results are consistent with the graphs of facial strains (Figs. 5 and 6) that show higher strains in model 3 and greater differences at points at or near the alveolus. The impact of model simplification on mode and magnitude of deformation is small compared to the impact of varying bite point (molar to incisor).

4. **Discussion**

This study aimed to assess the validity of an FE model of a cadaveric cranium approximating a molar bite. To this end the real specimen and the FE model were loaded with similar boundary conditions and the resulting strains were compared.

To compare results it was first important to verify whether the loading was consistent between real and virtual loads. This is relevant because differences in
direction or magnitude of applied force would result in different loading regimens and hence impact on strains. To that end bite forces were compared between real and virtual experiments. The mean bite force of the experimental rounds (424 N) was similar to the bite force of the FEA (425.61 N) with a difference of 1.61 N (0.38%).

When comparing the distribution of strains between the 3D DSPI output and the FE results there is a degree of mismatch between the two (Figs. 3 and 4). This is evident, e.g., on the bone overlying the root of the canine. In the DSPI 3D plot this region presents high strains in the mesial half and low strains in the lateral half. In the FE strain contour plot there are low strains over the whole bone overlying the root of the canine. This mismatch is also found in the plots of extracted strain magnitudes (Fig. 5). Errors are inevitable in the FE model creation process because of lack of resolution and lack of full data on material property variation throughout the cranium. In particular, it is unclear if it is necessary, and if so, how, to incorporate material representing periodontal ligaments between teeth and their sockets in FE models. This would have a dramatic effect locally on strain magnitudes (Groning et al., 2011; Wood et al., 2011) and it is not known what effect applying one material property to both dentine and enamel may have on results. Beyond this, errors in the physical loading trials are evident in the range of results obtained among different runs of the experiment. Additionally FEA and DSPI results may disagree because of errors matching topographies and so sampling points when making strain contour plots. The most likely major sources of error are in model building and loading because there is greater consistency within DSPI and FEA results than between DSPI and FEA.

Despite this, the 3D contour plot of the real strains presents similarities to those from FEA. In both, $\varepsilon_1$ magnitudes peak in the region of the canine and premolar alveolar bone, rising through the facial skeleton and converging towards the nasomaxillary buttress (next to the infra-orbital foramen), with strains dropping over and close to the root of the canine and less so over the premolar root. Similarly, the strains are also lower in the zygomatic root, although clearly lower in the real specimen than in the FE model. The main difference between the two is that the low strain field over/near the canine is larger, extending more superiorly in the real specimen. In $\varepsilon_3$ both the real and virtual experiments yielded a high strain area that extends from the border of the nasal aperture to the root of the zygomatic process with “islands” of lower strains.
located close to the infra-orbital foramen and body of the maxillary zygomatic process. The main difference between the two lies in the higher relative strains experienced by the virtual model, particularly in the apex of the first molar, above the canines’ apex, and the inter-radicular spaces. The directions of strain vectors experienced by the real cranium and the FE model present similarities but are not fully consistent throughout the whole infra-orbital region (Fig. 4B).

Quantification of strain differences relied on plotting the strain magnitudes extracted from two lines (Fig. 5). Consistent with the strain contour plots (Fig. 4), similarities in strain pattern were found, despite discrepancies in magnitudes that are mainly located closer to the teeth, as noted in the results section.

Simplification of segmentation, and thus allocation of material properties to different regions, impacted over and near teeth, with a decrease in strains experienced by models 2 and 1. This effect was negligible to non-existent in the remaining facial skeleton. The change in strain magnitudes in the alveolus causes a change in strain pattern, when considering that region, which is reflected in the global modes of deformation (Fig. 7). Model 3 deforms more than models 2 and 1 (larger distance from the unloaded model) and presents a somewhat different global mode of deformation that anatomically corresponds to more torsion of the face towards the working side than models 1 and 2. This may be because converting cancellous bone to have the same material properties as cortical bone, as is done between models 3 and 2/1, reduces stiffness in a way that affects both mode and magnitude of deformation. Further simplification between models 2 and 1 has a much less marked effect and, in any case, the impact of all simplifications is small relative to the impact of changes in bite point from incisor to molar.

Hence, $H_0^1$ (magnitudes and directions of strains predicted by FEA simulation do not differ from the real ones) is falsified. Strain magnitudes in line 1 present clear differences, despite similarities in line 2. Strain directions present similarities but not full congruence between real and FEA results. $H_0^2$ (the pattern of strains does not differ between real and predicted strains) is partially falsified because, despite some differences, the pattern of strain distributions in the experiment and FE simulation is generally similar. $H_0^3$ (simplifications in segmentation, and hence material properties, do not impact on strain magnitudes) was falsified because model simplification
decreased the sampled strains. $H_0^4$ (simplifications in segmentation do not impact on strain pattern) was falsified, particularly because of strain magnitudes predicted at point 10 in models 1 and 2. If point 10 is ignored, the hypothesis is not falsified.

The results presented here are therefore somewhat consistent with those of Toro-Ibacache et al. (2016), who used the same cranium and a similar experimental setting (although the constrained tooth was the first left incisor and not the first molar) and reported successful prediction of where strains are relatively high or low, but not of precise absolute magnitudes. Despite that it should be noted that Toro-Ibacache et al. (2016) found a better match between real and predicted strain magnitudes than are found in the present study. This might be because the force applied in this study is greater than that applied by Toro-Ibacache et al. (2016) (750 N vs. 550 N), which may magnify the differences between experimental and predicted strains in this study. Furthermore, the triangle of support formed by the two mastoid processes and the first molar in this study is less stable than that in the study of Toro-Ibacache et al. (2016) (two mastoids and central incisor), which may cause additional error in this study. Nonetheless, our results and those of Toro-Ibacache et al. (2016) are consistent with those of other studies that report successful prediction of the spatial distribution of regions high and low strains but not of precise absolute strain magnitudes and this is therefore also likely to obtain in modelling fossil material. It should also be noted that the FE models used in this study and in that of Toro-Ibacache et al. (2016) were refined to proportionally decrease model stiffness and minimize absolute strains differences between real and predicted strains. Such refinement is simply not possible in extinct taxa where no experimental data can be obtained in vivo or from fresh skeletal material.

These findings therefore suggest that craniofacial FEA is useful to predict the general spatial distribution of regions of high and low strains. However, at present, without validation data to guide model refinement, it lacks the ability to predict fine local variations in strain and accurate absolute strain magnitudes and directions. This is consistent with the findings of several prior studies (Bright and Rayfield, 2011; Kupczik et al., 2007; Strait et al., 2005; Toro-Ibacache et al., 2016).

These limitations are related to the resolution of medical CT imaging (which is typically used in FE model building) and the extreme complexity of cranial material properties, as well the difficulty in precisely replicating loading where physiological
loadings are simulated. Medical CTs are unable to image fine anatomical details due to voxel size. Indeed some bony structures (e.g., trabeculae in cancellous bone and bones such, ethmoidal cells, vomer) are much thinner than individual voxels. This leads to imprecision in the segmentation of details of cranial external form and loss (or thickening) of fine detail of the extremely complex internal anatomy of the cranium. Material properties of cranial bone are heterogeneous, with modulus of elasticity varying between individuals and dependent on the direction of applied force, hence material properties are also anisotropic, within crania and within the same bone. Teeth also present extremely complex material properties, with clear differences between tissues (i.e., enamel, dentin, cement) and within the same tissue (for a review of the elastic modulus of different dental structures, see Naveh et al., 2012). Despite efforts to reproduce the complexity of cranial material properties (Strait et al., 2005), FE modelling still inevitably simplifies the full complexity of the real structures, as in this validation study. Simplifications used here include modelling of trabecular bone as a bulk material, allocation of isotropic properties to all materials and modelling of teeth as single material structures with Young's modulus of enamel. Moreover, absence of the periodontal ligament (PDL) also impacts on model performance and congruence of results. This effect is likely more marked in the alveolar region and other anatomically close structures, impacting less with increasing distance from the teeth.

Improvement in FEA for investigating craniofacial mechanics is therefore dependent to a considerable degree on advances in scanning technology that increase resolution. Thus, higher resolution scans would enable imaging of fine anatomical details, and facilitate allocation of more precise, heterogeneous, material properties (which have been shown to be individual specific, and can be estimated from voxel grey level – Hounsfield number). Such advances in imaging demand increased computational power and suitable software to handle substantially larger scans during segmentation and to solve much larger FE models. Even with such advances FEA will still demand many approximations, albeit at finer levels of detail and, as such, validity will continue to be an important question when considering the results of FEA studies. Despite this, present FEA studies based on medical CTs are useful to address broad questions about cranial strain distributions but lack detail that allows them to focus on finer questions and predict absolute strains. These findings should be borne in mind, especially when
attempting to predict the performance of skeletal structures under loads where no experimental data can be obtained to assess model validity, such as is the case when working with fossils.

Acknowledgements

Ricardo Miguel Godinho is funded by the Portuguese Foundation for Science and Technology (FCT, reference SFRH/BD/76375/2011) and Viviana Toro-Ibacache by Comisión Nacional de Investigación Científica y Tecnológica (Chile). We thank Sue Taft and Michael Fagan (University of Hull) for assistance during the experiments. We also thank the reviewers for comments and suggestions, which improved the manuscript. We are also grateful to the Editors, Roberto Macchiarelli and Clément Zanolli, for the invitation to submit this paper to this special volume commemorating the life and work of Laurent Puymerail. This work is particularly pertinent in this regard since Laurent preceded Ricardo Miguel Godinho in the Centre for Anatomical and Human Sciences at the University of York, where he held a Marie Curie Postdoctoral fellowship. While he was with Paul O’Higgins he made inroads into the development of the techniques we later applied in this work. We will miss him greatly, both on a human level as a great friend and colleague and, on a scientific level we are forever thankful for his contributions to this field.
References


**Fig. 1.** A: experimental set-up of the skull used in the validation study and data from DSPI. B: camera view with border and overlaid lines used to extract strain magnitudes for comparison of experimental and FEA results (see Fig. 2A and section 2.3). C: phase map with border. D: 2D strain contour plot of $\varepsilon_1$. E: 3D strain contour plot of $\varepsilon_1$, with scale which is also applicable to Fig 1D.

**Fig. 1.** A: Montage expérimental du crâne utilisé dans l’étude de validité et pour les données de DSPI. B: vue de l’appareil photo avec les limites et les lignes superposées utilisées pour extraire les ordres de grandeur des contraintes pour la comparaison des résultats expérimentaux et FEA (voir la Fig. 2A et la section 2.3.). C: carte de phase avec les limites. D: graphique 2D des niveaux de contraintes de $\varepsilon_1$. E: graphique 3D des niveaux de contraintes de $\varepsilon_1$, avec échelle identique à celle de la Fig. 1D.
Fig. 2. A: lines used to extract strain magnitudes and compare results between measured and predicted strains. B: 32 points over the facial skeleton used to quantify the impacts of simplifications of segmentation. C: 51 craniofacial landmarks used to assess global modes of deformation.

Fig. 2. A: les lignes utilisées pour extraire les grandeurs de contrainte et comparer les résultats entre les contraintes mesurées et prédites. B: 32 points sur le squelette facial utilisés pour quantifier les impacts des simplifications de la segmentation. C: 51 points de repère cranio-faciaux utilisés pour évaluer les modes de déformation globaux.
Fig. 3. DSPI 3D strain contour plots (superimposed on cranium; first column of crania) and FEA strain contour plots (columns two, three and four of crania) of $\varepsilon_1$ (first row of strain contour plots) and $\varepsilon_3$ (second row of strain contour plots). The top row shows cross sections of the several FE models, in which different colours in each cross section represent different segmented materials (three colours in model 3, two colours in model 2 and one colour in model 1). Location of the cross sections is depicted in the first image of the top row.

Fig. 3. Graphique 3D des niveaux de contrainte DSPI (superposé sur le crâne; première colonne de crânes) et graphiques des niveaux de contrainte par FEA (colonnes de crânes deux, trois et quatre) de $\varepsilon_1$ (première rangée des graphiques de niveaux de contraintes) et $\varepsilon_3$ (deuxième rangée des graphiques de niveaux de contraintes). La rangée supérieure montre les sections transversales des plusieurs modèles par éléments finis, dans lequel des couleurs différentes dans chaque section transversale représentent différents matériaux segmentés (trois couleurs dans le modèle 3, deux couleurs dans le modèle 2 et une couleur dans le modèle 1). La localisation des sections est représentée dans la première image de la rangée supérieure.
Fig. 4. DSPI (top) and FEA (bottom) strain contour plots of $\varepsilon_1$ (left) and $\varepsilon_3$ (right). Strains have been scaled among figures to emphasise differences in contours rather than magnitudes. Absolute strain magnitudes are shown in Fig. 3. Numbered points correspond to matching peak strain areas between measured (numbers with no apostrophe) and predicted (numbers with apostrophe) strains rather than to exact spatial location because of the slight spatial mismatch between the two (see discussion). The ellipse next to point 3' depicts a diffuse area of higher strains whereas in point 3 the area is more discrete. This matches results of extracted strain magnitudes (see Fig. 5A).

Fig. 4. Graphiques de niveaux de contrainte DSPI (en haut) et FEA (en bas) de $\varepsilon_1$ (à gauche) et $\varepsilon_3$ (à droite). Les contraintes ont été standardisées entre les figures pour souligner les différences relatives des niveaux plutôt que des grandeurs de contrainte absolues qui sont présentés dans la Fig. 3. Les points numérotés correspondent aux zones de déformation maximales entre les valeurs mesurées (numéros sans apostrophe) et prédites (numéros avec apostrophes) plutôt que la localisation spatiale exacte en raison de la légère disparité spatiale entre les deux (voir la discussion). L'ellipse à côté du point 3' représente une zone diffuse des contraintes plus élevées alors qu’au niveau du point 3, la zone est plus réduite. Cela correspond aux résultats des grandeurs de contrainte extraites (voir la Fig. 5A).
Fig. 5. DSPI and FEA output plots of line 1 (left) and 2 (right). Top plots depict $\varepsilon_1$ and bottom plots $\varepsilon_3$. Numbers above and below the lines represent peak strains that anatomically correspond to the numbers in Fig. 4.

Fig. 5. Graphiques de sortie de DSPI et de FEA de la ligne 1 (à gauche) et 2 (à droite). Les graphiques du haut représentent $\varepsilon_1$ et ceux en bas représentent $\varepsilon_3$. Les chiffres au-dessus et en dessous des lignes représentent les contraintes maximales qui correspondent anatomiquement aux numéros indiqués dans la Fig. 4.
Fig. 6. Plot of facial ε1 and ε3 values in the different FE models.

Fig. 6. Graphique des valeurs faciales de ε1 et ε3 dans les différents modèles FE.
Fig. 7. PCA of large scale deformations of FEA results. Models 1 and 2 overlap, thus the symbol for each was sectioned diagonally to allow visualisation of both. Deformations are magnified by a factor of 500 to facilitate visualisation.

Fig. 7. Analyse en composantes principales (ACP) des grandes déformations des résultats de la FEA. Les modèles 1 et 2 se chevauchent, donc le symbole de chaque diagonale est sectionné pour permettre la visualisation des deux. Les déformations sont amplifiées par un facteur de 500 pour faciliter la visualisation.