This is a repository copy of *Strain uncertainties from two digital volume correlation approaches in prophylactically augmented vertebrae: local analysis on bone and cement-bone microstructures*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/109235/

Version: Accepted Version

Strain uncertainties from two digital volume correlation approaches in prophylactically augmented vertebrae: local analysis on bone and cement-bone microstructures

Gianluca Tozzi[a], Enrico Dall'Ara[b,c*], Marco Palanca[d], Marco Curto[a], Federica Innocente[d], Luca Cristofolini[d]

[a]School of Engineering, University of Portsmouth, UK
[b]Department of Oncology and Metabolism, University of Sheffield, UK
[c]INSIGNEO Institute for In Silico Medicine, University of Sheffield, UK
[d]School of Engineering and Architecture, Alma Mater Studiorum - Università di Bologna, Italy

For correspondence:

Dr Enrico Dall'Ara
Department of Oncology and Metabolism
INSIGNEO Institute for In Silico Medicine
Sir Frederick Mappin Building, Pam Liversidge building
Sheffield
S1 3JD
United Kingdom

Tel: +44 (0)114 2226175
Email: e.dallara@sheffield.ac.uk

ABSTRACT

Combination of micro-focus computed tomography (micro-CT) in conjunction with *in situ* mechanical testing and digital volume correlation (DVC) can be used to access the internal deformation of materials and structures. DVC has been exploited over the past decade to measure complex deformation fields within biological tissues and bone-biomaterial systems. However, before adopting it in a clinically-relevant context (i.e. bone augmentation in vertebroplasty), the research community should focus on understanding the reliability of such method in different orthopaedic applications involving the use of biomaterials. The aim of this study was to evaluate systematic and random errors affecting the strain computed with two different DVC approaches (a global one, "ShIRT-FE", and a local one, "DaVis-DC") in different microstructures within augmented vertebrae, such as trabecular bone, cortical bone and cement-bone interdigitation. The results showed that systematic error was insensitive to the size of the computation sub-volume used for the DVC correlation. Conversely, the random error (which was generally the largest component of error) was lower for a 48-voxel (1872 micrometers) sub-volume (64-221 microstrain for ShIRT-FE, 88-274 microstrain for DaVis-DC), than for a 16-voxel (624 micrometers) sub-volume (359-1203 microstrain for ShIRT-FE, 960-1771 microstrain for DaVis-DC) for the trabecular and cement regions. Overall, the local random error did not appear to be influenced by either bone microarchitecture or presence of biomaterial. For the 48-voxel sub-volume the global approach was less sensitive to the gradients in grey-values at the cortical surface (random error below 200 microstrain), while the local approach showed errors up to 770 microstrain. Mean absolute error (MAER) and standard deviation of error (SDER) were also calculated and substantially improved when compared to recent literature for the cement-bone interface. The

2

multipass approach for DaVis-DC further reduced the random error for the largest volume of interest. The random error did not follow any recognizable pattern with the six strain components and only ShiRT-FE seemed to produce lower random errors in the normal strains. In conclusion this study has provided, for the first time, a preliminary indication of the reliability and limitations for the application of DVC in estimating the micromechanics of bone and cement-bone interface in augmented vertebrae.

1. Introduction

The efficacy of prophylactic augmentation with injectable biomaterials (i.e. poly-methyl-methacrylate (PMMA)-based cements) in improving the mechanical stability of vertebrae is still a matter of debate (Kamano et al., 2011; Cristofolini et al., 2016). In particular, a deep understanding of internal microdamage in the bone tissue and at the cement-bone interface, which could potentially promote further damage to treated vertebrae, is currently missing.

This is probably due to the intrinsic limitations in most experimental techniques like digital image correlation (DIC) (Palanca et al., 2016) in not being able to capture and quantify internal microdamage evolution under load. In this perspective, digital volume correlation (DVC) is ideal to investigate the local internal damage in treated vertebrae. In fact, with the rapid progress of micro-focus computed tomography (micro-CT) in conjunction with *in situ* mechanical testing (Nazarian & Muller, 2004; Tozzi et al., 2012, 2013), DVC has become a powerful tool to examine full-field internal deformations in trabecular bone (Liu & Morgan, 2007; Gillard et al., 2014; Dall'Ara et al., 2014, Roberts et al., 2014), cortical bone (Christen et al., 2012; Dall'Ara et al., 2014), whole bones (Hussein et al., 2012, 2013; Danesi et al., 2016; Tozzi et al., 2016), cellular scaffolds (Madi et al., 2013) and cement-bone interface (Tozzi et al., 2014).

In order to expand the applications of DVC to biological tissues, including investigation of clinically-relevant issues such as bone augmentation, it is important to understand what is the error associated to the DVC measurement for specific sets of images, scanning protocols and correlation strategies. To this extent, the uncertainties of DVC in calculating strain in bone tissue have been quantified (Roberts et al., 2014). Moreover, the strain uncertainties in relation to a virtual

4

displacement applied to one single micro-CT image was also evaluated (Madi et al., 2013). However, it is recommended that strain uncertainties of any specific DVC approach are quantified on repeated scans (i.e. in a known deformation field such as zero-strain) to account for the intrinsic noise of the input images. This repeated scans methodology has been already adopted to quantify strain errors associated to trabecular bone (Liu & Morgan, 2007; Gillard et al., 2014; Dall'Ara et al., 2014), cortical bone (Dall'Ara et al., 2014), whole bones (Hussein et al., 2012) and cement-bone interface (Zhu et al., 2015). However, as DVC typically exploits different correlation and strain calculation strategies to compute strains (i.e. local vs global approaches, different registration metrics, etc.), it is important to quantify the level of uncertainty in the strain determination, by comparing two or more DVC methodologies using the same original image dataset. Palanca et al. (2015) compared the output of three different DVC approaches (a global and two local ones) applied on the same micro-CT biopsies of trabecular and cortical bone, where accuracy and precision in strain fields for both virtual displacements and repeated scans were investigated. Moreover, the presence of preferential components (normal or shear) for strain measurement in the different correlation approaches was also evaluated (Palanca et al., 2015).

Given a specific pattern/texture inside the bone specimen, DVC uncertainties are affected by the features that can be recognized in the sequence of images, which in turn depends on the spatial resolution of the image, and on the number of voxels included in the computation window (sub-volume) (Roberts et al., 2014). This pattern distribution can be related to the intrinsic natural features of the material (i.e. trabeculae in trabecular bone) or to radiopacifier particles usually incorporated in bone cements (i.e. ZrO_2 and $BaSO_4$) (Lewis et al., 1997). Thus, the DVC-computed

strain errors can be affected by the presence of biomaterials within the bone. Zhu et al. (2015) proposed a first attempt to investigate the strain uncertainties in specimens including cement and bone. They focused on images with voxel size of 22 micrometers, with smallest computation sub-volume of 32 voxels. The noise affecting computed strains was lowest within the cement (~500 microstrain), slightly higher in the bone regions partially interdigitated with cement (~700 microstrain), and more than doubled in the trabecular bone (~1400 microstrain). Zhu et al. (2015) used a single local DVC approach based on Fast Fourier Transform (described as DaVis-FFT in Palanca et al., 2015) with multipass and overlaps up to 75%, on one single cement-bone specimen in dry conditions, focusing on a single component of strain (the axial one, e_{zz}). However, recent literature in the DVC computation of bone tissue (Palanca et al., 2015) clearly indicated how DVC strain uncertainties obtained for the same local approach (DaVis-FFT) used in Zhu et al. (2015) are very much reduced if a direct correlation (described as DaVis-DC) is used instead of a FFT-based one (DaVis-FFT), and no overlap is used in multipass strategy. Furthermore, it is known (Gillard et al., 2014; Palanca et al., 2015) that looking at one single strain component (i.e. e_{zz}) is not sufficient for a complete understanding of the error pattern, as variability of strain error among the six components could be quite large. Very recently, uncertainty analyses of local and global DVC approaches applied to the whole natural and augmented porcine vertebrae were performed (Palanca et al., 2016b). In that study it was found that, despite the strain error produced similar trends in function of the computation sub-volumes for both groups, in the augmented vertebrae the random error of the strain components computed with the two DVC methods were different, especially for higher spatial resolution. In particular, the augmentation increased the error for the global approach, while reducing it for the

6

local. It is not clear yet how the DVC errors are influenced by the tissue microstructure and by the biomaterial distribution.

The main aim of this study was to evaluate and quantify strain measurement uncertainties at tissue level in five specific locations within different augmented vertebrae. This was done in order to better understand how the bone microstructure (trabecular and cortical), the presence of biomaterial and its integration with bone (cement-bone interface) could explain differences in performance of the two DVC approaches.

2. Methods

2.1 Specimens

Five thoracic vertebrae (T1-T3) were harvested from fresh porcine thoracic spines. All the surrounding soft tissues were removed, as well as the growth plates. The endplate areas of the vertebrae were potted in poly-methyl-methacrylate (PMMA) similar to Danesi et al. (2014). The spinous process was used to center the specimen in the transverse plane and align it about its vertical axis. The posterior arch was subsequently removed. Cement routinely used for vertebroplasty (Mendec Spine, Tecres, Italy) was then injected in the vertebral bodies by means of a proprietary device, following the instructions of the manufacturer. This is an acrylic-based cement, containing pellets of $BaSO_4$ (~300 micrometers) as a radiopacifier. The vertebrae were heated before and after augmentation in a circulating bath at 40°C, to allow optimal flow and consolidation of the cement.

2.2 Experimental procedures and volumes of interest (VOIs)

All the specimens (n=5) were placed in a loading device (CT5000, Deben Ltd, UK) equipped with a custom-designed environmental chamber, in order to closely simulate *in situ* loading conditions that are typically being applied to such vertebral bodies (Danesi et al., 2016; Tozzi et al., 2016). The specimens were immersed in saline solution and constrained against rotation inside the loading device with sandpaper disks glued to the bottom compressive platen. Each unloaded specimen was micro-CT imaged (XTH225, Nikon Metrology, UK) twice without repositioning, in order to reproduce a zero-strain condition. Prior to each imaging session a full conditioning of the micro-CT (up to 225 kV) was performed to stabilize x-rays and reduce at minimum fluctuations in the selected settings (i.e. kV, microA), throughout the duration of test. The micro-CT scanner was set to a voltage of 88 kV and a current of 110-115 microA. With an isotropic voxel size of 39 micrometers and exposure of 2 s, the image acquisition was performed with a rotational step of 0.23°, over 360° for a total scanning time of approximately 90 min.

In order to investigate the performance of the DVC approaches for the different bone tissues (cortical and trabecular), for the cement, and for the interdigitated regions, five volumes of interest (VOIs) were identified within each vertebral body. The five VOIs were cropped using MeVisLab (MeVis Medical Solution AG, Germany) and consisted in parallelepipeds of 300*300*432 voxels for the largest possible area that could be inscribed in all vertebrae (VOI-1, data presented in Palanca et al. (2016b) and reported here for completeness and for comparison) and of 152*152*432 voxels for smaller VOIs including areas of: fully cement-augmented trabecular bone (VOI-2), interface between augmented and non-augmented trabecular bone (VOI-3),

trabecular bone (VOI-4), and regions containing both trabecular and cortical bone, with surrounding saline solution (VOI-5) (Fig.1). The VOI-5 region was selected to understand how inappropriate (or ineffective) image masking could influence the DVC performance in the two approaches, particularly for the local DVC. The dimensions for the smallest VOIs (2-5) were able to include the different regions of interest within the augmented vertebra. To allow for the most standardized and less operator-dependent workflow, and investigate the worst-case scenario, no beam hardening and noise artifacts were corrected in the images. In order to allow comparison between the results obtained from different DVC approaches, the image datasets used in the present study will be made available to the scientific community at (*images will be shared through https://figshare.com/ upon acceptance and a doi will be added here*) or by contacting the corresponding author.

For each VOI in each specimen, the solid volume fraction (SV/TV) was computed as the sum of the volume of cement and/or bone, divided by the total volume of the VOI (Table 2). In VOI-4 and VOI-5 the SV/TV is equivalent to the bone volume fraction (BV/TV). The values of SV/TV were obtained via a manual thresholding of the grey-scale histograms with ImageJ (NIH) software, using its BoneJ plugin (Doube et al., 2010). The images were segmented by using a single level threshold, chosen in the valley between the first and second peak of the frequency distribution of the greyscale (histograms). The threshold value was adapted by visual comparison of the segmented and greyscale image in order to separate bone and cement from the background values. The SV/TV value, computed as average ± standard deviation between specimens for each VOI, was used to assess possible correlations with the DVC strain errors.
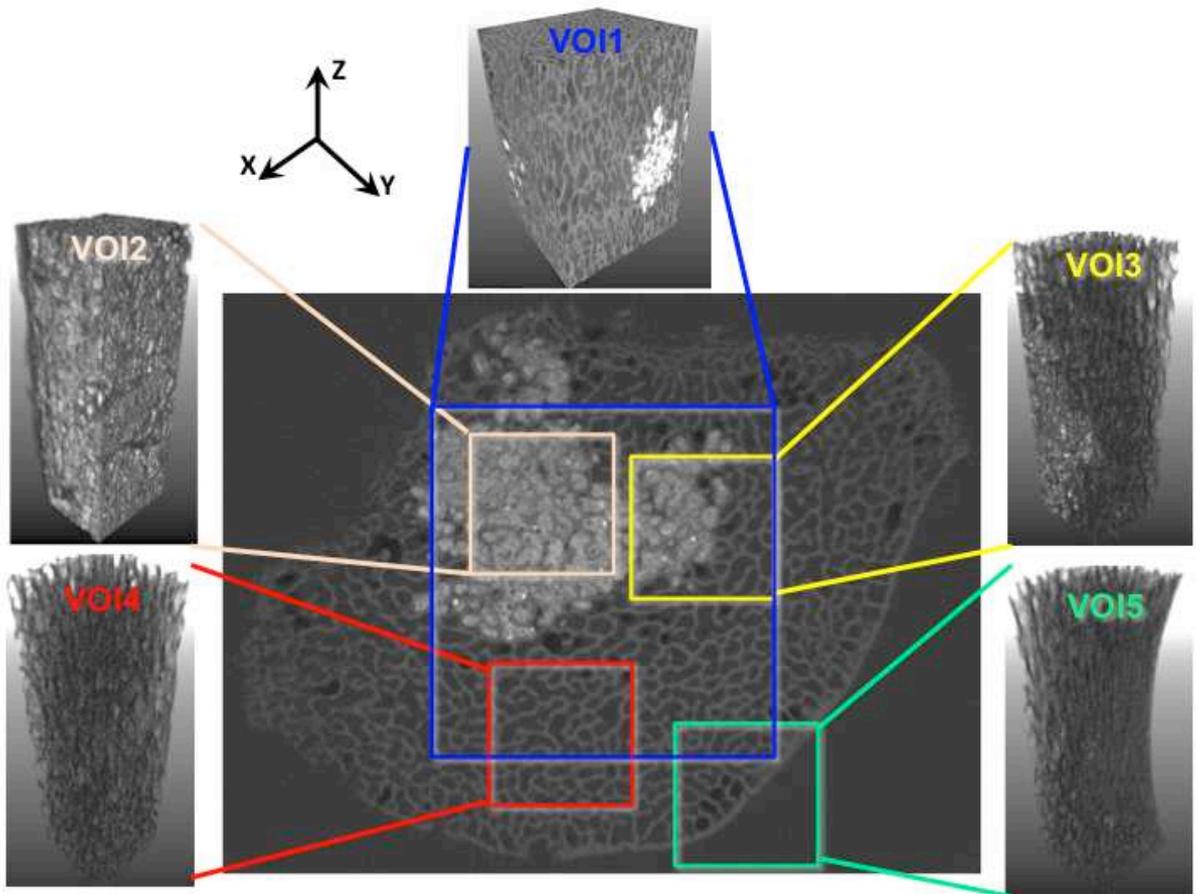
Fig 1: Transverse section of a vertebra, showing the five different volumes of interest (VOIs) selected for the DVC computation. Specifically, VOI-1 was the largest volume that could be inscribed in all vertebrae, VOI-2 a region of full cement-bone augmentation, VOI-3 a region of partial cement-bone augmentation, VOI-4 a region of trabecular bone, and VOI-5 a region of trabecular and cortical mixture surrounded by saline solution. Data related to VOI-1 has been reported for completeness and adapted from Palanca et al. (2016b).

Table 2: SDER and solid volume fraction (SV/TV) for DaVis-DC and ShIRT-FE calculated for a sub-volume size of 48 voxels in the five specimens and for each VOI. SDER is reported as median and standard deviation, whereas SV/TV as average and standard deviation. Data related to VOI-1 has been reported for completeness and adapted from Palanca et al. (2016b).

VOI	SDER DaVis-DC (microstrain)	SDER ShIRT-FE (microstrain)	SV/TV (%)
VOI-1 (300*300*432 voxels)	66 ± 52	35 ± 52	57.5 ± 10.9
VOI-2 (152*152*432 voxels)	45 ± 69	75 ± 48	84.1 ± 10.9
VOI-3 (152*152*432 voxels)	63 ± 47	52 ± 42	54.5 ± 6.4
VOI-4 (152*152*432 voxels)	61 ± 46	83 ± 48	32.9 ± 3.6
VOI-5 (152*152*432 voxels)	159 ± 406	51 ± 41	31.4 ± 5.2

2.3 Digital volume correlation (DVC) approaches

Two different DVC approaches were compared in this work, namely a 'local correlation' and 'global correlation'. The operating principles of the two DVC methods have been detailed elsewhere (Palanca et al., 2015, 2016b). Briefly, the local approach (DaVis-DC) is implemented in the DaVis software (v8.2.1, LaVision, Germany). DaVis-DC sub-divides the 3D images into smaller sub-volumes that can be correlated independently as a discrete function of grey-levels. The matching between the sub-volumes is achieved via a direct cross-correlation function (DC). Additionally, a piece-wise linear shape function and a third-order spline interpolation

in the image reconstruction are employed to correlate the pattern information contained in the reference and deformed images. The displacement field vector is obtained at the center of each sub-volume and the strain field is subsequently computed using a centered finite differences (CFD) scheme. The employed global approach (ShIRT-FE) is a combination of an home-written elastic registration software known as Sheffield Image Registration Toolkit (ShIRT) (Barber et al., 2007) and a Finite Element (FE) software package (Ansys v.14.0, ANSYS, US) as reported in Dall'Ara et al. (2014). In ShIRT the recognition of corresponding features in the subsequent 3D images is obtained by superimposing a grid with selectable nodal spacing (or sub-volume) to the entire volume of interest. ShIRT solves elastic equations at the nodes of the selected grid to evaluate the nodal displacements. The grid is then converted into an eight-node hexahedral mesh and the displacements computed by ShIRT at each node are imposed as boundary conditions in the FE model, where the strain field is then computed.

In order to evaluate the random errors associated to the displacement and the systematic and random errors associated to the strain for both DVC methods, two sub-volume sizes of 16 and 48 voxels were investigated for the five VOIs in each specimen. The larger sub-volume (48 voxels – 1872 micrometers) was chosen in order to obtain sufficient measurement points in the VOIs, and while it showed acceptable uncertainties of the strain components averaged over the whole organ for augmented vertebrae, it also revealed different behavior for the two DVC methods (Palanca et al., 2016b). The lower sub-volume (16 voxels – 624 micrometers) was chosen in order to evaluate the error for smaller registration regions, which could be beneficial especially for the boundary between the cement and bone. Moreover, both sub-volume sizes produced a 100% of correlated volume (defined as in Palanca et

al. (2015, 2016b)) for both local and global approaches. Finally, two different multipass schemes (available only on DaVis-DC) with decreasing sub-volume size of 128-112-96-80-64-48 voxels for VOI1 and 48-32-16 voxels for VOI2-VOI5 were tested with 0% overlap, In particular, the multipass was pushed to a final size of 16 voxels in the local VOIs (2-5), to explore the potential improvements for the local DVC approach, but still producing a larger number of measurement points (spatial resolution) when compared to 48 voxels.

2.4 Evaluation of errors as a consequence of the computation sub-volume

To quantify the errors, different indicators were computed:

- Ideally, the displacements were null; in the real experiment the actual displacements were affected by the inevitable unknown micro-movements of the moving parts of the scanner. To quantify the random error of the displacements, their variability was computed within each specimen. The systematic error for the displacements could not be quantified.

- As the test was based on a zero-strain condition, any non-zero values of strain were considered as error. Systematic and random errors for each specimen were computed as the average and standard deviation, separately, for each component of strain. For each VOI and sub-volume size, the median of the values of the errors obtained for the five specimens was then reported for each strain component.

- The mean absolute error (MAER) and standard deviation of error (SDER) were computed as:

$$\text{MAER} = \frac{1}{N}\sum_{k=1}^{N}\left(\frac{1}{6}\sum_{c=1}^{6}\left|\varepsilon_{c,k}\right|\right) \qquad \text{(Eq. 1)}$$

$$\text{SDER} = \sqrt{\frac{1}{N} \sum_{k=1}^{N} \left(\frac{1}{6} \sum_{c=1}^{6} \left| \varepsilon_{c,k} \right| - \text{MAER} \right)^2} \qquad \text{(Eq. 2)}$$

where "ε" represents the strain; "c" represents the six independent strain components; "k" represents the measurement point; N is the number of measurement points. MAER and SDER correspond to the indicators formerly called as "accuracy" and "precision" (Liu & Morgan, 2007).

- Linear correlations between the SV/TV and the random error computed for each component of the strain, or the SDER, were computed for each VOI and for both DVC approaches (Mann-Withney U test, α=0.05, Minitab 17, UK).

3. Results

The random errors affecting the displacements ranged between 0.01 and 1.61 of the voxel size (from 0.66 to 63.08 micrometers) for DaVis-DC, and from 0.01 to 0.04 voxels (from 0.50 to 1.53 micrometers) for ShIRT-FE. Random errors were typically larger for smaller sub-volume sizes (Table 1) and this difference was more pronounced for DaVis-DC than ShIRT-FE. The multipass scheme available for DaVis-DC notably improved the performance in VOI-1 (sub-volume output of 48 voxels), VOI-4 (sub-volume output of 16 voxels) and partially in VOI-2 and VOI-3 (sub-volume output of 16 voxels) when compared to the results obtained with sub-volume of 16 voxels, obtaining values comparable to the case when a sub-volume of 48 voxels was used. Multipass in VOI-5 also (sub-volume output of 16 voxels) produced an improvement compared to the case in which a sub-volume of 16 voxels was used, but less relevant when compared to the results obtained with a sub-volume of 48 voxels.

Both the local (DaVis-DC) DVC and the global (ShIRT-FE) approaches did not show a clear trend in the systematic (Fig. 2) and random (Fig. 3) errors affecting the specific components of strain. Moreover, the sub-volume size (16- or 48-voxel) did not seem to generally affect the order of magnitude of the systematic error. DaVis-DC experienced absolute systematic errors mostly lower than 100 microstrain, with a maximum peak of 350 microstrain (e_{xx} in VOI-3) for VOI-1, VOI-2, VOI-3 and VOI-4. The main exception was observed for DaVis-DC in relation to VOI-5 (Fig. 2), where considerably higher systematic errors (up to ~6000 microstrain) where found with the 16-voxel sub-volume size. However, the use of a 48-voxel sub-volume size produced errors ranging from -223 to 428 microstrain for e_{xz} and e_{xx}, respectively. The multipass strategy for DaVis-DC did not drastically reduce the strain uncertainties for all VOIs, but only in few cases such as e_{xx} in VOI-1 and e_{xx}, in VOI-3. In some other cases the multipass had a rather detrimental effect and considerably increased the strain error, particularly when compared with the 48-voxel sub-volume size (i.e. e_{xy} in VOI-1, e_{yy} in VOI-5 and e_{xy} in VOI-5). In ShIRT-FE, for the six components, absolute strain values were always lower than 100 microstrain (for all VOIs).

Once again the random error evaluation did not indicate any preferential direction in the six strain components for the different VOIs, but more regular patterns could be identified (Fig. 3, all values for sub-volume 48 in supplementary material). For the sub-volume size of 48 voxels in VOI-1, VOI-2, VOI-3 and VOI-4, DaVis-DC computed errors that were generally lower than 200 microstrain with a maximum value of 274 microstrain for e_{zz} in VOI-1. The sub-volume size of 16 voxels increased the random error to thousands of microstrain in DaVis-DC as well in VOI-1, VOI-2, VOI-3 and VOI-4, with a maximum of 1771 microstrain for e_{zz} in VOI-4. VOI-5 still presented the worst case with very large errors (several-thousands

microstrain) for the 16-voxel sub-volume size, and up to 770 microstrain for the 48-voxel size. The multipass for DaVis-DC was only able to reduce the uncertainties for VOI-1, when a final sub-volume size of 48 voxels was used. For VOI-2, VOI-3, VOI-4 and VOI-5 the multipass, with a final sub-volume pushed at 16 voxels, could only mitigate the errors relative to the 16-voxel sub-volume alone, without any considerable improvements. In ShIRT-FE the strain uncertainties for all the components with a sub-volume size of 48 voxels were consistently lower or close to 200 microstrain. For a sub-volume size of 16 voxels ShIRT-FE produced a maximum strain error of ~1200 microstrain. Interestingly, this global approach seems to produce lower random errors for the normal strains, rather than the shear ones for all VOIs.

The strain values obtained in DaVis-DC (local approach) for VOI-5 were clearly influenced by the presence of the saline solution in the micro-CT images as shown in Fig. 4 (a, b, d, e). Conversely, ShIRT-FE (global approach) seemed to be less sensitive to saline region, and the major strain uncertainty was related to boundary effect (Fig. 4a, c, d, f).

In order to facilitate the comparison with published literature the MAER and SDER were also computed as scalar values similar to (Liu & Morgan, 2007), so as to have a single value to be associated with each VOI (Fig. 5). Both MAER and SDER followed a decreasing trend with the increase of sub-volume size from 16 to 48 voxels.

In particular, for the 48-voxel sub-volume the MAER and SDER in VOI-1, VOI-2, VOI-3 and VOI-4 for both DaVis-DC and ShIRT-FE were consistently better or close to 200 microstrain and 100 microstrain, respectively. In VOI-5, DaVis-DC produced MAER and SDER (48 voxels) better than 400 microstrain and 200

microstrain, respectively. The multipass scheme was only able to reduce the error in VOI-1 (48-voxel final sub-volume), but not in the smaller VOIs (2-5) when the final sub-volume output was pushed to 16 voxels. ShIRT-FE confirmed the same trend as for the other VOIs with strain errors better than 150 microstrain.

The effect of SV/TV was not clearly associated with the strain uncertainties. In terms of SDER (48-voxel sub-volume, Table 2) the outputs of ShIRT-FE and DaVis-DC did not show any linear correlation with the SV/TV ($p > 0.21$). The random error for each strain component (not reported here for brevity) showed inverse linear correlation with SV/TV only for e_{xx} ($p = 0.012$, $R^2 = 0.61$) and e_{xz} ($p = 0.036$, $R^2 = 0.45$) computed in VOI-2 (48 voxels sub-volume) with DaVis-DC.
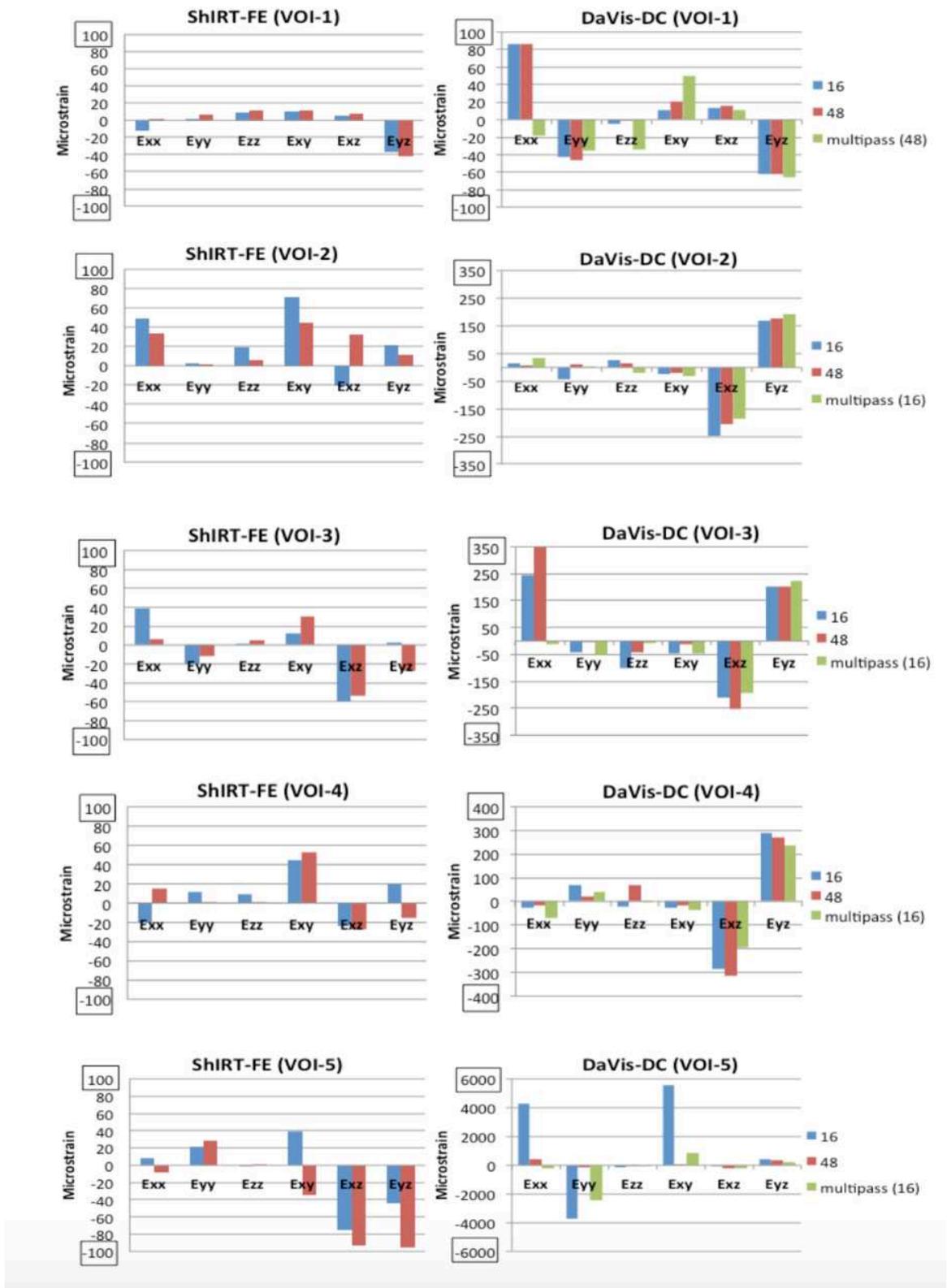
Fig 2: Systematic error with ShIRT-FE (left) and DaVis-DC (right) in the five VOIs (1-5): median between five specimens. Data related to VOI-1 has been reported for completeness and adapted from Palanca et al. (2016b).
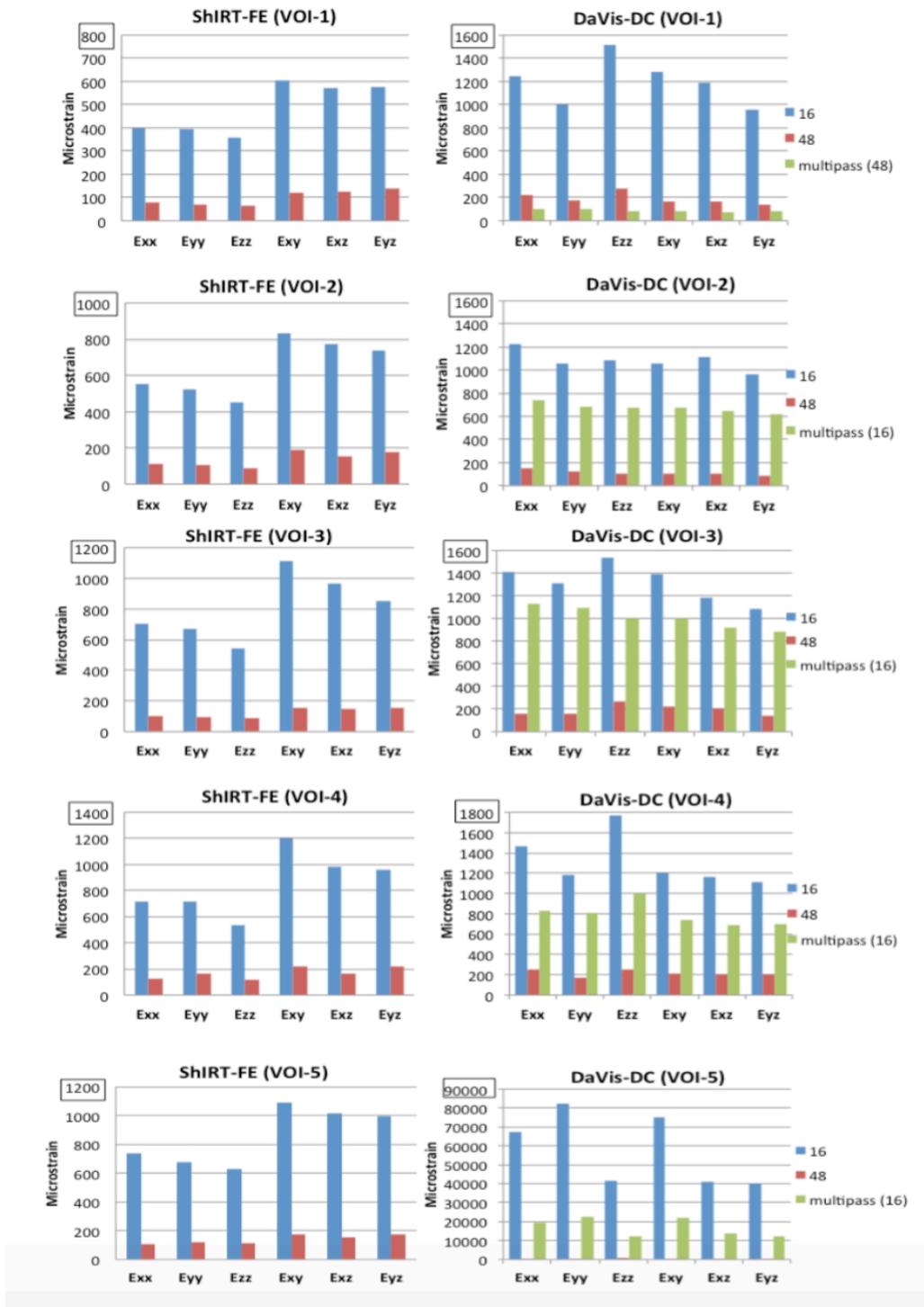
18

Fig 3: Random error with ShIRT-FE (left) and DaVis-DC (right) in the five VOIs (1-5): median between five specimens. Data related to VOI-1 has been reported for completeness and adapted from Palanca et al. (2016b).
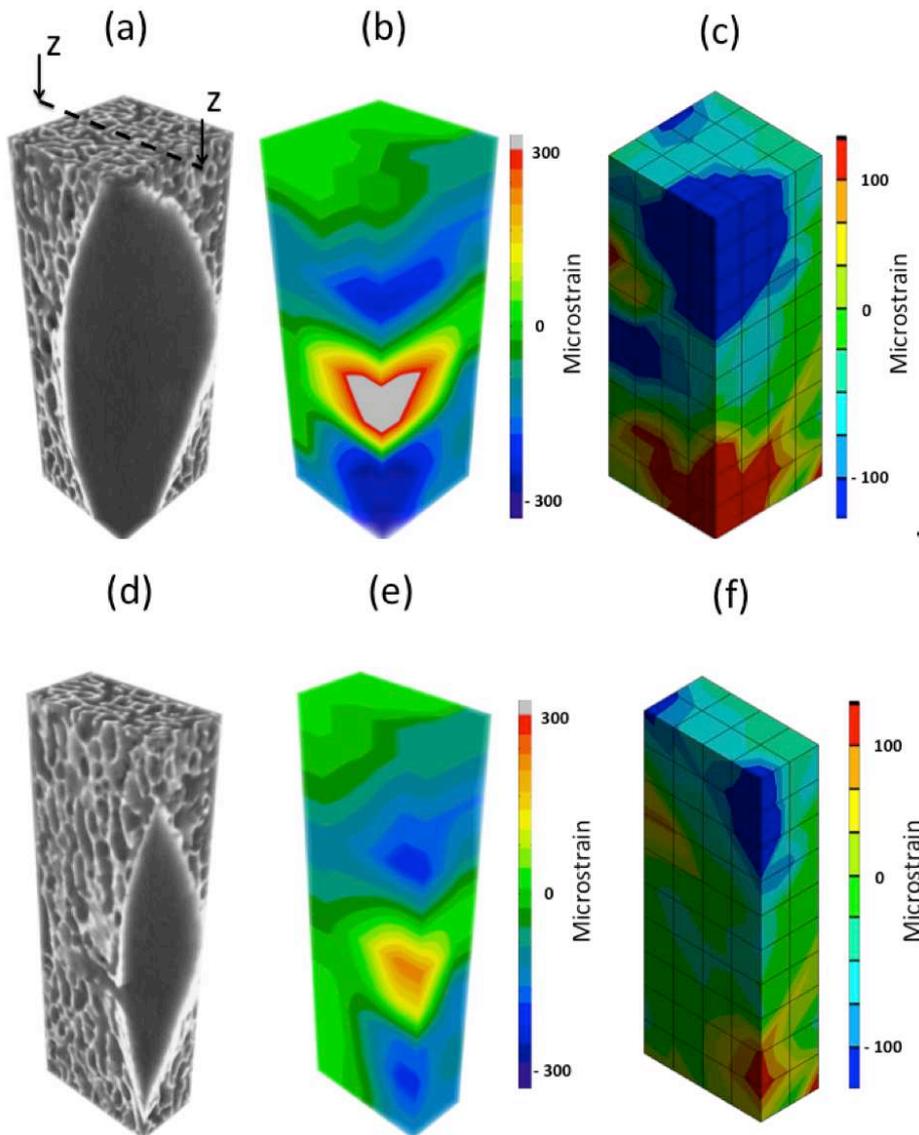
Fig 4: The first row reports the volumetric view of VOI-5 for micro-CT (a), DVC strain maps computed with DaVis-DC (b) and ShIRT-FE (c) with sub-volume size of 48 voxels. The second row reports the z-z planar section for micro-CT (d), DVC strain maps computed with DaVis-DC (e) and ShIRT-FE (f). For DaVis-DC the largest random errors mainly corresponded to the region of saline solution and negatively influenced the result in the trabecular/cortical region, whereas strain error in ShIRT-FE are localized mainly in the boundaries of the image.
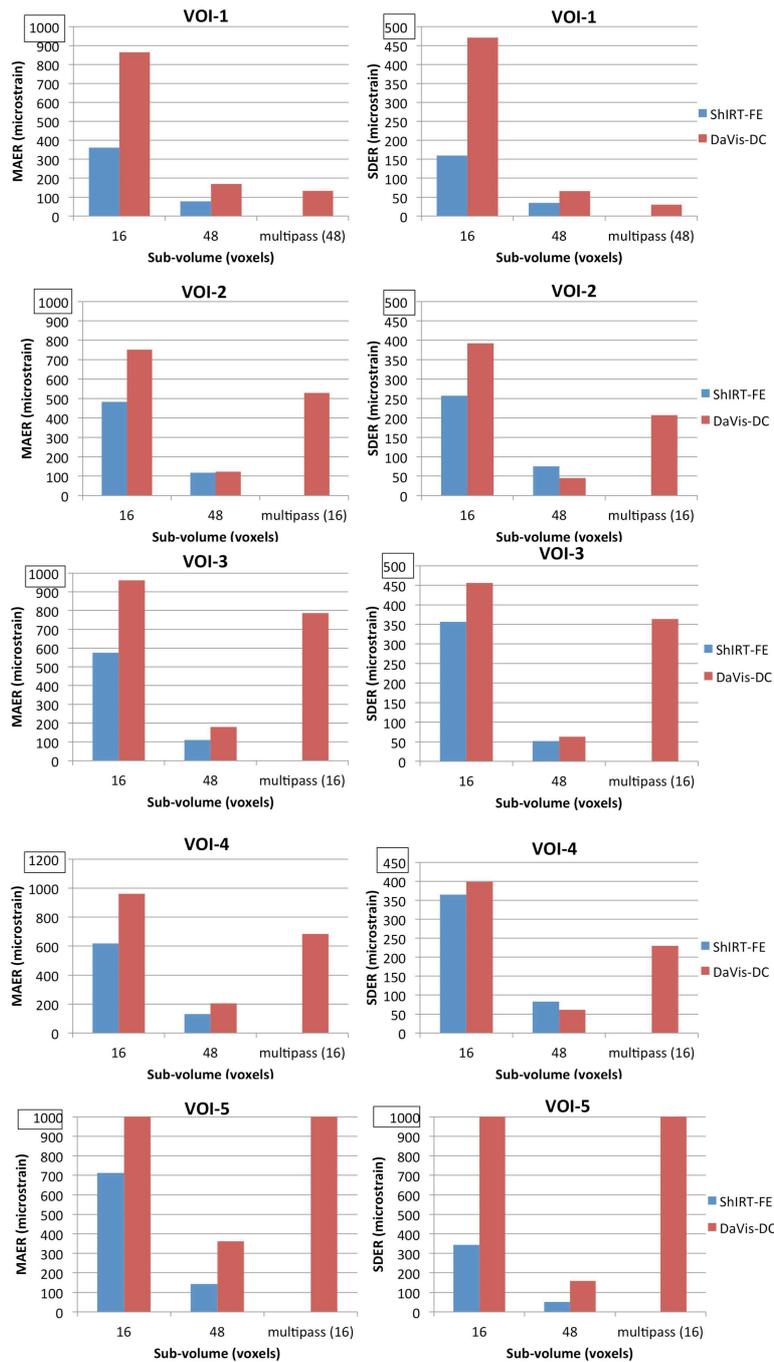
Fig 5: DVC strain uncertainties reported as 'MAER' (left) and 'SDER' (right) (formerly known as 'accuracy' and 'precision' respectively in Liu & Morgan, 2007) for both ShIRT-FE and DaVis-DC with sub-volume sizes of 16 and 48 voxels, and for the multipass scheme (only DaVis-DC, final sub-volume size equal to 48 for VOI-1 and 16 for the other VOIs). Data related to VOI-1 has been reported for completeness and adapted from Palanca et al. (2016b).

4. Discussion

The main aim of this work was to evaluate the effect of bone microstructure, biomaterial and its integration with bone (cement-bone interface) on the systematic and random strain error distributions within prophylactically augmented vertebrae, when two different DVC approaches are used. For VOI-1, which was intended as an organ-level investigation, DaVis-DC and ShIRT-FE were similar in terms of magnitude of systematic and random errors. For more details please refer to Palanca et al. (2016b). For all VOIs the effect of sub-volume size on the systematic error seemed negligible as well as the multipass (16-voxel final sub-volume). Most of the strain components in VOI-2, VOI-3, and VOI-4 were included in the range ±50 microstrain and absolute maximum strains of ~70 microstrain for ShIRT-FE (in VOI-2) and ~350 microstrain for DaVis-DC (in VOI-3). However, in VOI-5 there was a visible difference between the two DVC approaches. ShIRT-FE reported values comparable to the other VOIs, with absolute strains always lower than 100 microstrain, whereas DaVis-DC produced absolute errors up to ~5600 microstrain. Moreover, the effect of sub-volume size was remarkable in some strain components: some components of error (i.e. e_{xy}) for a 16-voxel sub-volume were ~200 times higher than for the 48-voxel sub-volume. This was expected for the local DVC approach due to the absence of trackable features outside the bone (Fig. 4), which becomes critical for the local DaVis-DC when computing smaller sub-volumes (higher spatial resolution).

Not surprisingly random errors for both approaches were largely influenced by the sub-volume size in all VOIs, where errors for the 16-voxel sub-volume were

much higher than those for the 48-voxels, and a more repeatable trend in the strain components was observed (Fig. 3). In all VOIs except VOI-5, DaVis-DC produced strain errors up to in the order or thousand microstrain (maximum of ~1800 microstrain in VOI-4) for the 16-voxel sub-volume and errors in the order of hundred microstrain for the 48-voxel sub-volume (maximum of ~250 microstrain in VOI-4). The multipass was able to further reduce the error in VOI-1 (close or below 100 microstrain for all components) only when the final sub-volume was 48 voxels. In the smaller VOIs (2-4) where the final sub-volume was pushed at 16 voxels, the multipass was only able to reduce the error for the same sub-volume size without multipass to a minimum of ~600 microstrain in VOI-2. This can be related to a lack of convergence of the different steps in the multipass due to the reduced number of features with the smallest final sub-volume (16 voxels). In VOI-5, ShIRT-FE reported the same trend shown in the other VOIs with errors constantly lower than 200 microstrain for the 48-voxel sub-volume and close or lower than 1000 microstrain for the 16-voxel sub-volume. Similarly to the systematic error, DaVis-DC showed high sensitivity to the saline solution surrounding the tissue also for random errors as documented in Fig. 4. While errors outside the bone are in most cases acceptable, care should be taken when interpreting the results on the border of the specimen, where local approaches are affected by the absence of reference features. A possible solution to the problem for local DVC could rely in the use of appropriate overlap strategies to ensure a higher degree of continuity during correlation. However, current overlap scheme implemented in DaVis did not produce any improvements in strain error (Palanca et al., 2015) and further work is needed on that side. Also, an appropriate and controlled masking is suggested when local DVC approaches are used. However, it must be noted that for sub-volume size of 48

voxels DaVis-DC generated errors in the range 209 – 770 microstrain, which suggests how even the minimal variation in the image gray-scale intensity for the individual sub-volume could result in an important improvement of the local correlation strategy. There was no evidence of a clear directionality associated to strain error in all VOIs for both DVC approaches. Only ShIRT-FE seemed to indicate lower errors for the normal strains when compared to the shear components (Fig. 3), but no clear trend could be observed. The random error reported for the displacements (Table 1) is in line with the strain results. This is important as different strain calculation strategies could affect the final outcome, starting from comparable displacements. However, particularly for VOI-4 the multipass produced better displacements even when compared to the 48 voxels. This opens up discussion on how strain is actually computed. In fact, in this study only the centered finite differences (CFD) scheme available in DaVis software was used, but the influence of different strain computation of primary DVC output (displacement) surely requires further investigation. Overall, for both sub-volume size and DVC approaches (excluding the VOI-5 for DaVis-DC), both systematic and random errors resulted not particularly related to the bone microarchitecture and or the presence of biomaterial. Therefore, it seems that local material heterogeneities should not affect the precision of the DVC calculation, provided that enough recognizable patterns are available in the images.

Table 1: Random errors affecting the displacements (in micrometers) for DaVis-DC and ShIRT-FE, for a sub-volume size of 16 and 48 voxels for each VOI. The median

over the five specimens is reported. Data related to VOI-1 has been reported for completeness and adapted from Palanca et al. (2016b).

DISPLACEMENT RANDOM ERROR (MICROMETERS)							
VOI	Sub-Volume	DaVis-DC			ShIRT-FE		
		X	Y	Z	X	Y	Z
1	16	1.87	1.49	2.18	1.22	1.27	1.13
	48	1.56	1.10	1.16	1.25	1.35	1.24
	Multipass (48)	1.05	1.08	0.92	NOT AVAILABLE		
2	16	1.55	1.66	1.05	1.24	1.35	0.68
	48	1.11	0.89	0.47	1.24	1.23	0.63
	Multipass (16)	1.58	1.40	0.66	NOT AVAILABLE		
3	16	2.02	1.71	1.97	1.22	1.34	0.67
	48	1.47	1.19	0.76	1.15	1.30	0.54
	Multipass (16)	1.23	1.35	1.04	NOT AVAILABLE		
4	16	2.20	2.04	2.41	1.31	1.40	0.77
	48	1.73	1.57	1.36	1.25	1.35	0.54
	Multipass (16)	1.37	1.31	1.04	NOT AVAILABLE		
5	16	54.77	63.08	48.57	1.40	1.53	0.80
	48	2.59	2.04	2.18	1.18	1.55	0.50
	Multipass (16)	15.64	17.49	17.49	NOT AVAILABLE		

The MAER and SDER, reported as "accuracy" and "precision" in Liu & Morgan (2007) (Fig. 5), showed a clear reduction for both error indicators with a larger sub-volume, consistently with previous literature (Dall'Ara et al., 2014; Palanca et al., 2015). The multipass was still able to produce improvements for VOI-1, but not for the remaining VOIs (2-5), where the final sub-volume was 16 voxels. For VOI-4, containing only trabecular bone, both MAER and SDER were worse than those extrapolated via power law in Dall'Ara et al. (2014) for sub-volume size with physical dimension equal to 1872 micrometers and equivalent to the 48-voxel sub-volume in this study (MAER: ~200 microstrain in this study vs extrapolated 21 microstrain; SDER: ~50 microstrain in this study vs extrapolated 13 microstrain). This difference is probably due to the higher spatial resolution of the images used in the study of Dall'Ara et al. (2014) with respect to the one of the images used in this study (voxel size ~10 micrometers vs 39 micrometers).

The SV/TV was calculated for each VOI in order to take into account the effect of both bone and cement on the SDER, for the two DVC approaches. This choice was preferred to the BV/TV involving only bone tissue (Roberts et al., 2014), as the influence of bone cement with pellets of $BaSO_4$ (~300 micrometers) or other radiopacifiers could strongly modify the material texture and, therefore, influence the DVC analysis. It was found that there is no linear correlation between the SDER calculated in DaVis-DC and ShIRT-FE with SV/TV for all VOIs (p > 0.21). With regards to the single components, the random strain error produced a weak inverse correlation with SV/TV in VOI-2 (48-voxel sub-volume) for DaVis-DC only for e_{xx} (p = 0.012, R^2 = 0.61) and e_{xz} (p = 0.036, R^2 = 0.45). This could be related to the intrinsic nature of local DVC approaches, where small interrogation volumes in two scans are registered independently to map local grey-scale intensities in the images (if enough

features are available). Hence, the presence of radiopacifiers in the cement may have helped the local approach to produce lower errors, compared to areas with gradient of materials. However, this correlation is insufficient to justify the effect of microstructure or biomaterial in the strain error. The application of DVC to cement-bone composites was firstly introduced by Tozzi et al. (2014). In that study it was noted that the presence of radiopacifiers with suitable particle size in the cement helped the correlation process, producing better correlation in such areas. However, a detailed investigation of the effects of cement in the DVC strain uncertainties was not performed. Zhu et al. (2015) reported a first attempt to investigate this effect. They evaluated the DVC uncertainties with a local approach (DaVis-FFT) in zero-strain (repeated scans) on one cement-bone specimen in dry conditions (22 micrometers voxel size, smallest sub-volume of 32 voxels). They reported the MAER (referred to as "accuracy") and SDER (referred to as "precision") for only one strain component (e_{zz}). Thus, if results have to be compared with the current study, the SDER with DaVis-DC multipass for the 16-voxel sub-volume (our 624 micrometers vs their 704 micrometers) on the e_{zz} would be more appropriate and represent the worst case in both studies. The current results did not show the same decreasing trend from trabecular to cement as in Zhu et al. (2015). In fact, the SDER from trabecular bone regions (current VOI-4), to partially interdigitated (current VOI-3), to cement (current VOI-2) in the current study remains pretty constant (<100 microstrain for both approaches). However, our SDER was consistently better than that reported in Zhu et al. (2015) for e_{zz} in the bone region (230 microstrain in this study vs ~1400 microstrain in that study), partially interdigitated (364 microstrain in this study vs ~700 microstrain in that study) and cement region (207 microstrain in this study vs ~500 microstrain in that study). This is surely due to the specific choice

27

of FFT-based local DVC as well as extensive overlap (up to 75%) in Zhu et al. (2015), which were found to be less accurate when compared to a direct correlation approach for the same software (DaVis-DC) without any overlaps (Palanca et al., 2015). The present findings show that the local errors to be expected in the cement, bone, and in the interdigitated regions may not be so different, but that different sub-volume sizes may be required to minimize such errors in the different regions.

The current study has some limitations. Firstly, only two sub-volume sizes (16- and 48-voxel) were chosen in the present study. However, a more comprehensive trend for augmented vertebrae with sub-volume sizes up to 128 voxels has been recently reported in Palanca et al. (2016b). That study showed how random error could be reduced well below 100 microstrain in both DVC approaches for VOI-1. Thus, it is expected that also smaller VOIs could follow a similar trend. Secondly, the use of five specimens could not provide a statistical relevance, but only a trend that may be sufficient to have reliable information on strain uncertainties location and distribution. Thirdly, the strain error is only calculated in a zero-strain condition for repeated scans. This type of analysis should be expanded in order to take into account strain errors under load. Finally, the use of animal tissue is justified by easier handling and availability compared to human. This decision was taken for ethical reasons in this preliminary methodological work. Future work on DVC strain uncertainties from clinical CT images will expand our knowledge of the tool for a potential implementation in clinical practice.

Conclusions

The results obtained in this study aimed at better understanding the complexity of DVC strain uncertainties in prophylactically augmented vertebrae, and

of how the bone microstructure and the presence of injectable biomaterial could influence the strain error. Two different DVC approaches were tested (global ShIRT-FE and local DaVis-DC) and strain errors were evaluated for two sub-volume sizes (16- and 48-voxel). It was found that systematic error was insensitive to sub-volume changes, whereas the random errors were lower for the 48-voxel sub-volume (all values around or lower than 200 microstrain) in volumes of interest with larger amount of solid volume fraction, for both DVC approaches. The bone microstructure as well as the presence of biomaterial did not seem to have an important affect on DVC computation for both approaches. When the liquid (uniform material) was included in the image, DaVis-DC experienced higher errors (770 microstrain in the best case) than ShIRT-FE. MAER and particularly SDER were substantially improved when compared to recent literature in cement-bone interface. The multipass approach for DaVis-DC further reduced the minimum random error for the largest volume of interest (48-voxel final sub-volume) and reduced the maximum random error (16-voxel final sub-volume) in the other volumes. Finally, no anisotropy was found for the errors affecting the different components of strain, where only ShiRT-FE seemed to produce lower random errors in the normal strain components.

Conflict of interest statement

None.

Acknowledgements

REFERENCES

Christen, D., Levchuk, A., Schori, S., Schneider, P., Boyd, S. K., Muller, R., 2012. Deformable Image Registration and 3D Strain Mapping for the Quantitative Assessment of Cortical Bone Microdamage. J. Mech. Behav. Biomed. Mater. 8, 184–193.

Cristofolini, L., Brandolini, N., Danesi, V., Erani, P., Viceconti, M., Ferguson, S. J., 2016. A preliminary in vitro biomechanical evaluation of prophylactic cement augmentation of the thoracolumbar vertebrae. J. Mech. Med. & Biol. In press.

Dall'Ara, E., Barber, D., Viceconti, M., 2014. About the inevitable compromise between spatial resolution and accuracy of strain measurement for bone tissue: A 3D zero-strain study. J. Biomech. 47(12), 2956-63.

Danesi, V., Zani, L., Scheele, A., Berra, F., Cristofolini, L., 2014. Reproducible reference frame for in vitro testing of the human vertebrae. J. Biomech. 47, 313-8.

Danesi, V., Tozzi, G., Cristofolini, L., 2016. Application of digital volume correlation to study the efficacy of prophylactic vertebral augmentation. Clinical Biomechanics. 39, 14-24.

Doube, M., Kłosowski, M.M., Arganda-Carreras, I., Cordeliéres, F., Dougherty, R.P., Jackson, J., Schmid, B., Hutchinson, J.R., Shefelbine, S.J., 2010. BoneJ: free and extensible bone image analysis in ImageJ. Bone, 47, 1076-9. doi: 10.1016/j.bone.2010.08.023

Gillard, F., Boardman, R., Mavrogordato, M., Hollis, D., Sinclair, I., Pierron, F., Browne, M., 2014. The application of digital volume correlation (DVC) to study the microstructural behaviour of trabecular bone during compression. Journal of the Mechanical Behavior of Biomedical Materials. 29, 480-99.

Hussein, A.I., Barbone, P.E., Morgan, E.F., 2012. Digital volume correlation for study of the mechanics of whole bones. Procedia IUTAM: Full Field Measurements and Identication in Solid Mechanics. 4, 116-25.

Hussein, A.I., Mason, Z.D., Morgan, E.F., 2013. Presence of intervertebral discs alters observed stiffness and failure mechanisms in the vertebra. J. Biomech. 46, 1683-88.

Kamano, H., Hiwatashi, A., Kobayashi, N., Fuwa, S., Takahashi, O., Saida, Y., Honda, H., Numaguchi, Y., 2011. New vertebral compression fractures after prophylactic vertebroplasty in osteoporotic patients. AJR Am. J. Roentgenol. 197(2), 451-6.

Lewis, G., 1997. Properties of acrylic bone cement: State of the art review. Journal of Biomedical Materials Research. 38(2), 155-182.

Liu, L., Morgan, E.F., 2007. Accuracy and precision of digital volume correlation in quantifying displacements and strains in trabecular bone. J. Biomech. 40(15), 3516–20.

Madi, K., Tozzi, G., Zhang, Q.-H., Tong, J., Cossey, A., Au, A., Hollis, D., Hild, F., 2013. Computation of full-field displacement in a scaffold implant using Digital Volume Correlation and Finite Element Analysis. Med. Eng. & Phys. 35(9), 1298-1312.

Nazarian, A., Muller, R., 2004. Time-lapsed microstructural imaging of bone failure behavior. J. Biomech. 37, 55-65.

Palanca, M., Tozzi, G., Cristofolini, L., Viceconti, M., Dall'Ara, E., 2015. 3D Local Measurements of Bone Strain and Displacement: Comparison of Three Digital Volume Correlation Approaches. J. Biomech. Eng. 137(7). doi: 10.1115/1.4030174.

Palanca, M., Brugo, T.M., Cristofolini, L., 2015b. Use of digital image correlation to investigate the biomechanics of the vertebra. J. Mech. Med. Bio. doi: 10.1142/S0219519415400047.

Palanca, M., Tozzi, G., Cristofolini, L., 2016. The use of digital image correlation in the biomechanical area: a review. International Biomechanics doi: 10.1080/23335432.2015.1117395

Palanca, M., Tozzi, G., Dall'Ara, E., Curto, M., Innocente, F., Danesi, V., Cristofolini, L., 2016b. Digital volume correlation can be used to estimate local strains in natural and augmented vertebrae: An organ-level study. J. Biomech. In press.

Roberts, B.C., Perilli, E., Reynolds, K.J., 2014. Application of the digital volume correlation technique for the measurement of displacement and strain fields in bone: A literature review. J. Biomech. 47, 923-34.

Schreier, H.W., Sutton. M.A., 2002. Systematic errors in digital image correlation due to undermatched subset shape functions. Exp. Mech. 42, 303-10.

Tozzi, G., Zhang, Q.-H., Tong, J., 2012. 3D real-time micromechanical compressive behaviour of bone-cement interface: experimental and finite element studies. J. Biomech. 45, 356-363.

Tozzi, G., Zhang, Q.-H., Tong, J., Guillen, T., Ohrndorf, A., Christ, H.-J., 2013. Mechanical characterisation of a metallic foam-cement composite under selected loading conditions. J. Mater. Sci. Mater. Med. 24(11), 2509-2518.

Tozzi, G., Zhang, Q.-H., Tong, J., 2014. Microdamage assessment of bone-cement interfaces under monotonic and cyclic compression. J. Biomech. 47, 3466-74.

Tozzi, G., Danesi, V., Palanca, M., Cristofolini, L., 2016. Elastic full-field strain analysis and microdamage progression in the vertebral body from digital volume correlation. Strain. doi: 10.1111/str.12202

Zhu, M.L., Zhang, Q.H., Lupton, C., Tong, J., 2015. Spatial resolution and measurement uncertainty of strains in bone and bone-cement interface using digital volume correlation. J. Mech. Behav. Biomed. Mater. 57, 269-79.