Cell Tissue Bank
DOI 10.1007/s10561-012-9307-3

ORIGINAL PAPER

Validity of an automatic measure protocol in distal femur for allograft selection from a three-dimensional virtual bone bank system

Lucas Eduardo Ritacco · Christof Seiler · German Luis Farfalli · Lutz Nolte · Mauricio Reyes · Domingo Luis Muscolo · Luis Aponte Tinao

Received: 23 January 2012 / Accepted: 16 March 2012

Abstract Osteoarticular allograft is one possible treatment in wide surgical resections with large defects. Performing best osteoarticular allograft selection is of great relevance for optimal exploitation of the bone databank, good surgery outcome and patient's recovery. Current approaches are, however, very time consuming hindering these points in practice. We present a validation study of a software able to perform automatic bone measurements used to automatically assess the distal femur sizes across a databank. 170 distal femur surfaces were reconstructed from CT data and measured manually using a size measure protocol taking into account the transepicondyler distance (A), anterior-posterior distance in medial condyle (B) and anterior-posterior distance in lateral condyle (C). Intra- and inter-observer studies were conducted and regarded as ground truth measurements. Manual and automatic measures were compared. For the automatic measurements, the correlation coefficients between observer one and automatic method, were of 0.99 for A measure and 0.96 for B and C measures. The average time needed to perform the measurements was of 16 h for both manual measurements, and of 3 min for the automatic method. Results demonstrate the high reliability and, most importantly, high repeatability of the proposed approach, and considerable speed-up on the planning.

Keywords Bone bank system · Lograft selection · 3D surgical planning

Introduction

For locally aggressive or recurrent benign bone tumors and bone sarcomas, primary tumoral resection with wide margins is considered as a first treatment. This requires a wide surgical resection that entails a residual bone defect, which can be reconstructed with a fresh frozen bone allotransplantation.

The improvements in diagnostic and therapeutic techniques have produced an increase of the patient survival as well as decreased complications and side effects, and an improved quality of life (Enneking et al. 1993; Enneking and Campanacci 2001). Indeed, it has been emphasized on the better performance of more biologically oriented bone reconstructions,

L. E. Ritacco
Department of Medical Informatics, Virtual Planning and Navigation Unit, Italian Hospital of Buenos Aires, Buenos Aires, Argentina

L. E. Ritacco (✉) · G. L. Farfalli · D. L. Muscolo · L. A. Tinao
Institute of Orthopedics "Carlos E. Ottolenghi", CINEOT, Italian Hospital of Buenos Aires, Potosí 4247, 1199 Buenos Aires, Argentina
e-mail: lucasritacco@hotmail.com

C. Seiler · L. Nolte · M. Reyes
Institute for Surgical Technology, University of Bern, Bern, Switzerland

Published online: 08 April 2012

especially in young and physically active patients, due to the normal biomechanics of the limb in question when using a bone allotransplantation. For instance, it has been shown that a poor selection of the bone to be transplanted with respect to its shape and size may derive in changes in the joint mobility and load distribution, resulting in joint fractures and early joint degeneration (Muscolo et al. 1992; Muscolo et al. 2000).

Determining the size and shape of the allograft is a critical issue in the correct selection of it. However, current approaches are very time-consuming, mostly based on manual measurements performed directly on the bones present in the bone databank or on their corresponding three-dimensional virtual models, reconstructed from computerized tomography (CT) images (Ritacco 2010).

Ideally, an automatic allograft selection system could be based on shape-matching, retrieved from CT images, between the patient anatomy and the complete bone databank. Recently, such approach was demonstrated for pelvic allograft selection (Paul et al. 2010). Nevertheless, the practicality of this method is conditioned by the exhaustive computations performed on a large bank of donors. More importantly, for each patient it would be necessary to allocate extra time to the clinical workflow, in order to segment the anatomy and have a workable model.

A more straightforward approach using simpler shape descriptors was proposed to determine possible donor bones for a given patient. The method consists of defining geometrical features characterizing the distal femur size (Ritacco 2010). These features include the transepicondyle distance (A), the anterior-posterior distance in the medial condyle (B), and the anterior-posterior distance in the lateral condyle (C). The so-called ABC protocol can then be regarded as a screening of the bone databank, therefore speeding further computations and most importantly, not requiring the pre-processing mentioned previously, since such measures can be extracted either directly from the image or if tools available, from a rough bone segmentation characterizing the location of the points used by the ABC protocol. The method provides a small subset of possible candidates for allograft selection, on which further analysis is required to select the best donor match. The advantage of such approach relies on not requiring segmentation of the bone, considerably reducing pre-operative planning time.

An automatic bone morphology measurement software, it allows fast computations of bone morphology based on a non-rigid image registration technique tailored for orthopedic research (Seiler et al. 2009). This approach seems ideal for retrospective scenarios to automatize the use of the ABC protocol on existing bone databank, as well as to providing with a flexible tool allowing extension of this and other population-based bone morphology studies. For instance, this technique was employed for the purpose of quantifying changes in the spatial distribution of bone after mechanical unloading (Li et al. 2007).

The objective of this work is to introduce and validate an automatic bone morphology measuring tool as a key element towards fast and accurate, yet clinically oriented, bone allograft selection.

Methods

A total of 170 complete left femora were selected from our database (age range: 23–83, 62.57 ± 15; (67 males and 93 females) were scanned on a Toshiba Aquilion CT scanner, with a resolution of 0.877 mm and slice increments of 1 mm. Three-dimensional reconstructions of all specimens were created from CT images. Two operators referred here as "Observer 1" and "Observer 2" were requested to measure the following distal femur morphometric parameters using a specialized 3D software (Mimics, Materialise, Leuven, Belgium) on a plane perpendicular to the long axis of the bone: (1) Transepicondylar axis (A): the distance between the most medial point in the medial epicondyle and the most lateral point in the lateral epicondyle. (2) Medial condyle distance (B), determined as the distance between the most anterior and most posterior points, respectively, in the anterior-posterior direction. (3) Lastly, the length of the lateral condyle (C) determined with the same method used for the medial condyle. The resulting distances were exported for statistical analysis. To assess intra-observer repeatability, Observer 1 was requested to perform a second run of measurements 1 week after.

For the automatic measurements, Observer 1 was requested to select pairs of points corresponding to A, B and C parameters on the reference bone. In this study manual measurements were used as a control.

Intra- and inter-observer reliability of this protocol was assessed by manually measuring 170 femora, and was evaluated using an intra-class correlation coefficient (ICC). Figure 1 illustrates the ABC-protocol.

Image pre-processing

The set of CT scans were initially pre-processed by semi-automatic segmentation using Amira software, and cropping of each image to separate left and right femur bones. Image segmentation refers to the task of assigning a label to each pixel in an image in order to identify different structure or structures. In the case of segmenting the bone structure as a whole (i.e. no differentiation of different bone tissue types), only two labels are necessary for background and foreground. The resulting binary images (see Fig. 2c) were consequently used as image masks on the original CT images. This resulted on images only containing bone tissue intensities, while being zero-valued outside (see Fig. 2b). Finally, three-dimensional models were created from this image through iso-contouring (Lorensen and Cline 1987) in order to produce a three-dimensional representation of the bone surface describing the bone morphology. The open-source library insight toolkit (Ibanez et al. 2003) was used to perform all post-segmentation tasks. Figure 2 shows the entire pre-processing pipeline.

After image pre-processing, the key task of this work, called image registration, was performed. The next section explains in more details this task.

Automatic morphology assessment through image registration

Image registration refers to the task of finding a spatial transformation that aligns a so-called moving image (also called target image) into an image fixed (also

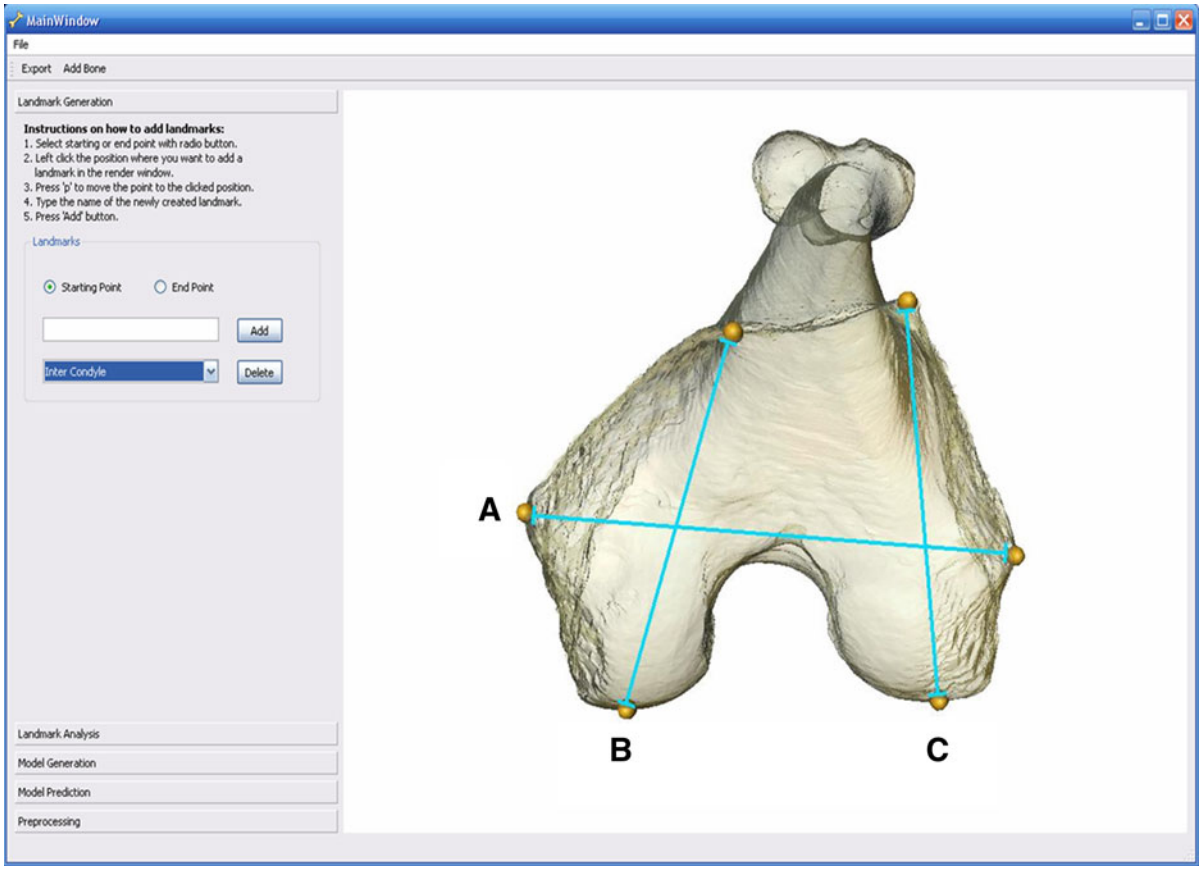

Fig. 1 The ABC protocol. A set of three distances is measured corresponding to 1. Transepicondylar axis (**a**). Medial condyle distance (**b**), and length of the lateral condyle (**c**). Graphical user interface for the automatic bone morphology tool. A observer was requested to set the corresponding ABC landmarks (shown in example as *yellow spheres*)

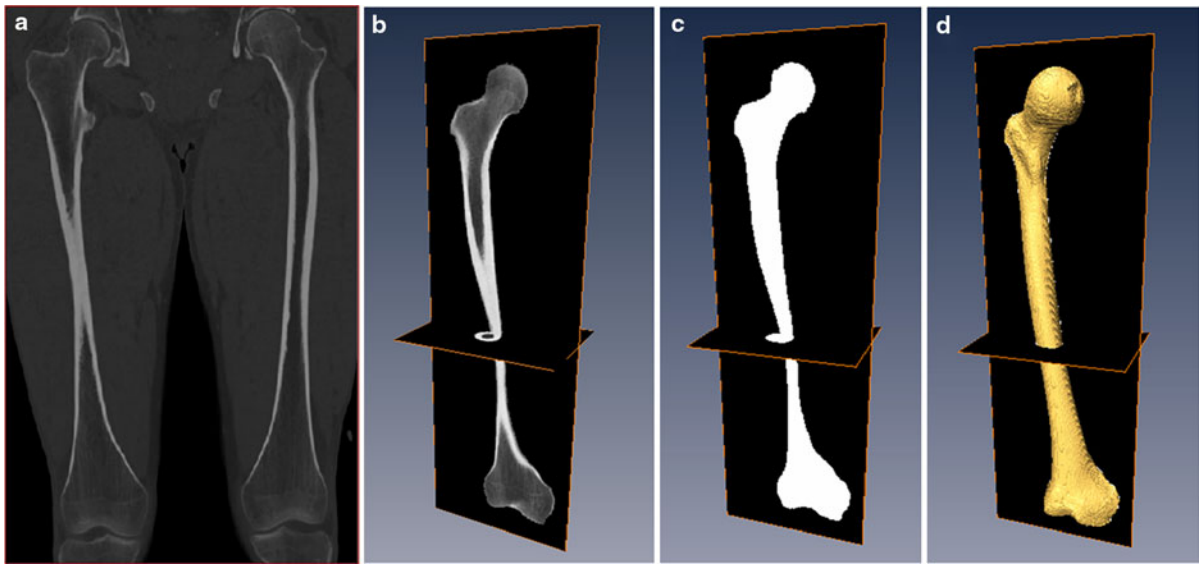

Fig. 2 Pre-processing pipeline. The original CT image (**a**), is cropped to separate *left* and *right* femur bones, which are segmented semi-automatically (**b**). The resulted segmentation is used to mask the original CT image (**c**). A surface reconstruction is additionally performed (**d**)

called source image). The transformation is found through optimizing a given numerical criteria, called similarity measure, which can be based on image intensity information, geometrical features, etc. Some common similarity measures include the square of image intensity sum differences (SSD), correlation coefficient (CC), and mutual information between image intensity patterns, amongst others. In addition, the type of transformation also has to be considered depending on the structure of interest and the problem at hand. For instance, some common types of transformation include rigid transformation, which considers rotation and translation of the structure; affine transformation, which is a geometrical transformation considering rotation, translation and anisotropic (i.e. different in each direction) scaling. The ability of the transformation to cover more complex type of image deformations defines the degree of freedom of the transformation. At the end of the spectrum of degrees of freedom, one finds non-rigid image registration, providing the maximum level of complexity and details achievable in medical image registration. Indeed, in non-rigid image registration, the goal is to find for each voxel in the source image, its anatomically correspondent in the target image.

Image registration has been extensively studied and used in many medical applications due to its ability to provide a way of comparing different image datasets on a common spatial reference system. Some examples include image fusion of pre- and post-operative images, multimodality image registration (e.g. PET/CT), computer-assisted surgery, atlas construction, etc. For a comprehensive review of medical image registration the reader is referred to Hill et al. (2001) and Fischer and Modersitzki (2008).

In this study, image registration was used to provide with an automatic and efficient way to realize high-throughput bone morphology assessment. The image registration pipeline is as follows. First, a common reference image is chosen from the selected cohort. Based on this selection, affine-based image registration is performed between the reference image and the remaining datasets (Ourselin et al. 2000). The aim of the affine-based image registration is to provide the non-rigid registration with an initialization for the final image warping. The algorithm parameters were chosen empirically as three multi-resolution levels, 10 maximum number of iterations per level, normalized correlation coefficient metric, and tri-linear interpolation. The non-rigid registration we adopted is based on the pioneer work of (Thirion 1998), which was later on improved to account for higher robustness and speed (Vercauteren et al.). We adopted this algorithm based on its high accuracy as well as the low computation

times compared to other developed techniques (Klein et al. 2009; Ourselin et al. 2000), which makes it appealing for use in clinical routine. Similarly to the affine-based image registration, the selection of parameters for non-rigid registration algorithm was chosen empirically as three multi-resolution levels, ten iterations per level, Gaussian regularization standard deviation of 1 pixel, sum of square intensity differences as metric, and tri-linear interpolation. After non-rigid registration, the output of the algorithm consists of displacement vector fields. This means that the algorithm assigns for each voxel in the source image a three-dimensional vector, describing the position of its corresponding point in the target image. In this way a complete mapping of anatomical corresponding points can be established. Such information represents the core of the proposed methodology. Figure 3 exemplifies the steps involved in the image registration pipeline. Figure 3d shows the resulting deformed image being morphed to match the reference image. In this way by repeating this process for each image in the database, anatomical point correspondences are established for every pixel in each 3D volume. This rich map of anatomical correspondences can then be exploited to perform automatic morphology assessment over a population. To illustrate the quality of the image morphing, Fig. 3e shows a checkerboard visualization of the reference and the resulting deformed image. In this type of visualization the images are jointly displayed in alternated blocks. This visualization allows the user to qualitatively evaluate the accuracy of the image registration process, specifically for non-rigid transformations.

Similarly to the pre-processing tasks, all developments regarding image registration were conducted using the open-source library, insight toolkit (ITK).

Clinical processing workflow

The processing pipeline previously discussed describes the relevant technical details and methods required to establish point correspondences between

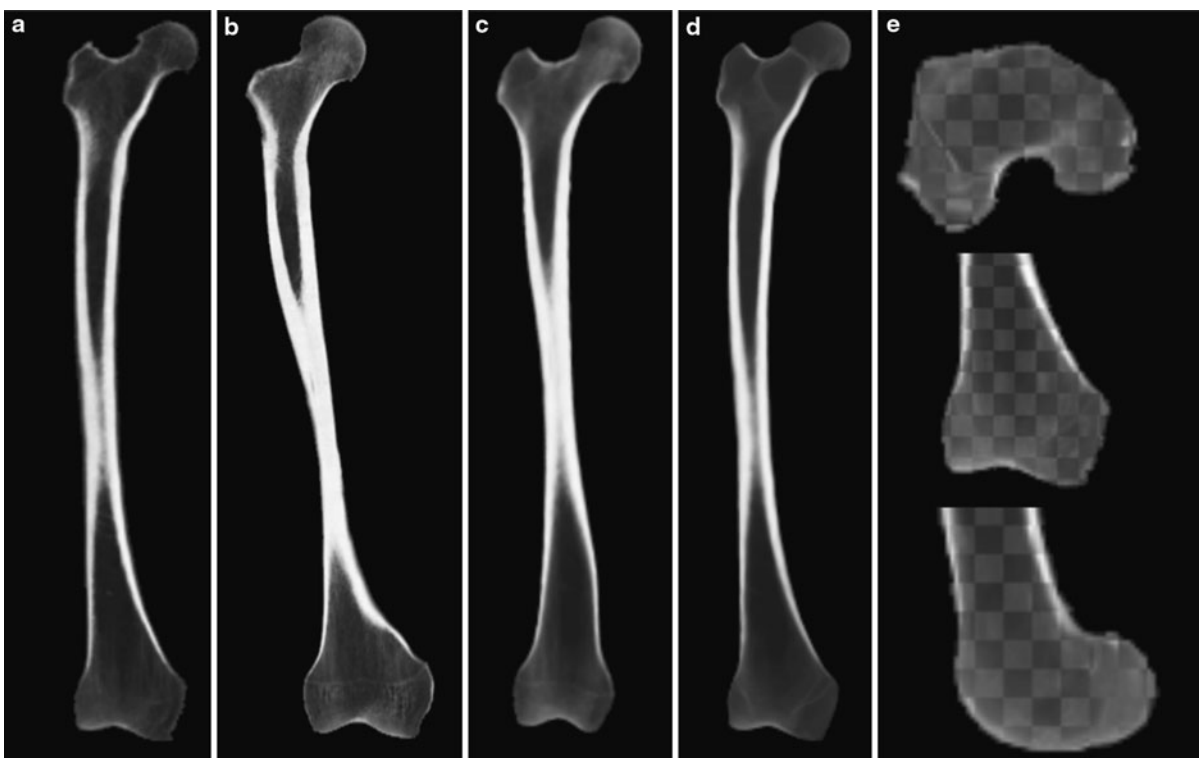

Fig. 3 Image registration pipeline. Original reference (**a**) and target pre-processed images (**b**). Target image after affine transformation (**c**). Target image after non-rigid transformation (**d**). Zoomed axial, coronal and sagittal checkerboard views between reference and final deformed images (**e**). Note: reference and original target plane ((**a**) and (**b**), respectively), are not anatomically corresponding due to position of the bone in the scanner

images. This pipeline needs to be encapsulated into a user-friendly software in order to make it clinically relevant. In this direction, a complete graphical user interface was developed to assist on this task. Depending on the stage of processing, two clearly scenarios can be distinguished. Namely, bone data bank preparation and storage, and automatic bone morphology assessment and outputs retrieval.

For the former, the user needs to perform the following steps:

1. Add cohort group to softwares' database. This involves selecting a folder on disk, where pre-process images are stored.
2. Select reference image. This involves selection of one image dataset as reference.
3. Proceed with image registration pipeline. The entire registration pipeline, described in the previous section, is performed on the selected cohort and reference image. The resulting displacement vector fields are stored, and the database is updated with this information.

For the latter, a typical scenario considers:

1. Select cohort. The software interface allows the user to define a subset cohort based on matching criteria such as age range, gender, ethnic group, etc. This information is collected from the original dicom stacks or from manually added input.
2. Define 3D landmarks on the reference model. A three-dimensional model of the reference bone is displayed and the user can set landmarks by clicking on the surface of the model. These landmarks can be optionally stored. Based on selected landmarks, the software can compute automatically distances on the entire selected cohort. The results can be exported into a text file for later analysis.

Results

Manual and automatic measurements

Intraobserver (Observer 1 vs Observer 1). Manual measures

A single operator was considered for testing intraobserver repeatability while using the ABC protocol twice on one hundred and seventy one-hundred and seventy (170) distal femoral. For A measurements the maximal, mean and standard deviation (SD) were: maximal difference 1.67 mm, mean of 0.34 mm, and standard deviation (SD) 0.30 mm. Similarly for B measurements the results were: maximal difference was 2.04 mm, mean of 0.40 mm, and SD of 0.37 mm. Finally, for C measurements the results were: maximal difference of 1.57 mm, mean of 0.35 mm, and SD of 0.28 mm. The obtained intraclass correlation coefficient for all measurements was of 0.99 in the range of 0.993–0.997 (95 % CI).

Interobserver (Observer 1 vs Observer 2). Manual measures

Interobserver variability between two separate observers was quantified when they measured the ABC parameters of 170 distal femoral allografts leading to a maximal difference of 1.82 mm for A measure and a mean of 0.41 mm (SD 0.36 mm), for B measures maximal difference was 5.27 mm, a mean of 0.46 mm (SD 0.51 mm). For C measures, maximal difference was 4.58 mm, a mean of 0.43 mm (SD 0.56 mm). Intraclass correlation coefficient was of 0.99 in A measure, in the range 0.99–0.99 (95 % CI), 0.98 for B in a range of 0.98–0.99 (95 % CI) and 0.98 for C measures in a range of 0.97–0.98 (95 % CI).

Interobserver (automatic vs Observer 1). Automatic vs manual

Finally, for 145 bones the automatic and manual measurements were compared. For A measure the maximal difference was of 3.29 mm, a mean of 0.68 mm (SD 0.55 mm), for B measure the maximal difference was of 3.31 mm, a mean of 0.96 mm (SD 0.73). For C measure the maximal difference was of 5.56 mm, a mean of 1.04 mm (SD 0.82). The intraclass correlation coefficient for measure A was of 0.99 in a range of 0.98–0.99 (95 % CI), 0.96 for measure B in a range of 0.95–0.97 (95 % CI) and 0.96 for measure C in a range of 0.94–0.97 (95 % CI).

A whisker plot describing the accuracy of the intra- and inter-observer as well as the automatic measurements is presented in Fig. 4. In addition, considering that all intra-observer differences were below 2.0 mm, an analysis of percentage of number of measurements below this value was performed.

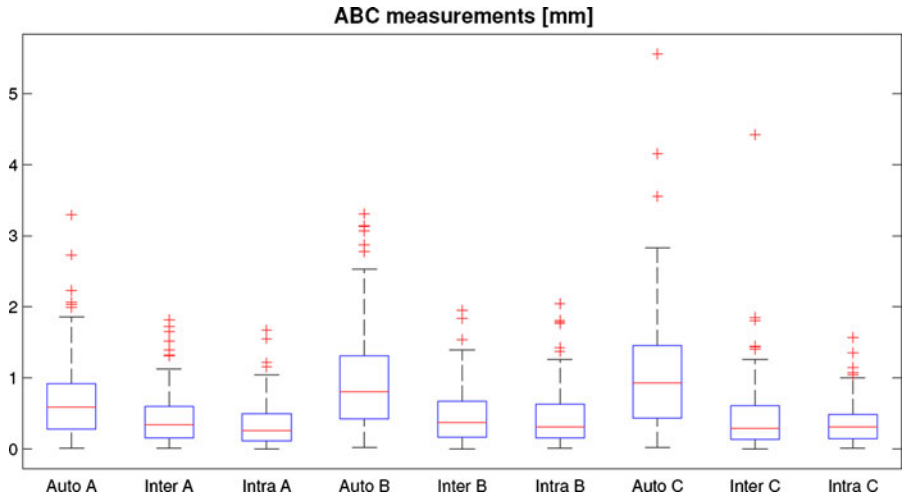

Fig. 4 Whisker plots for differences between manual vs manual measurements (Intra A/B/C, Inter A/B/C) and differences between automatic vs manual (Auto A/B/C) measurements. N = 145 for automatic and N = 170 for manual measurements. This diagram is a convenient way of graphically depicting groups of numerical data through their five-number summaries: the *bottom* and *top* of the *box* are always the 25th and 75th percentile (the *lower* and *upper* quartiles, respectively), and the band near the *middle* of the *box* is always the 50th percentile (the median)

Discussion

The evaluation of the automatic ABC measurements showed a good agreement with the manual measurements. Indeed, all intra-observer differences were below 2.0 mm, such value could then be regarded as the possible maximum error produced by a human operator. Under this consideration the automatic method presented a robustness of 97.9 % for measure A, 91 % for B and 90.3 % for C. In terms of computation speed, the average time needed to perform the measurements was of 16 h for both observers, and of 3 min for the automatic method. This clearly highlights the potential of the proposed automatic bone morphology tool for large databases. With respect to the construction of the database, the average time involved for pre-processing of the data was of 5 min per bone. This does not include the time required for image segmentation, since it is dependent on the user's experience and employed segmentation method. However, for our segmentation setup using commercially available software, a typical segmentation required approximately 30 min per bone. Such task is commonly performed offline from the surgical planning workflow and it is necessary for both manual and automatic approaches. The major advantage of the automatic method relies on the capacity of redefining the measurement protocol (e.g. different from the ABC protocol, herein presented) by simply selecting points on the reference bone model, and having the complete map of measurements for the database at no additional cost.

Up to our knowledge, this is the first work dealing with automatic measurements in a virtual bone bank system.

These kind of tools show to be promising for large datasets due to the automatic and fast features (high-throughput) as well as exhaustive statistical studies on bone morphology.

One limitation of this work is that such measures were applied on virtual models rather than the real ones, which implies a possible bias introduced by the 3D reconstruction step. However, automatic and manual method measurements were applied on same data and with same ABC protocol. Other important point is that the ABC protocol do not take into account biomechanical factors (Brin et al. 2010; Eckhoff et al. 2005; Eckhoff 2001; (Eckhoff et al. 2003). However, the same algorithm can be used to recognize other patterns including the ones presented on that work. The ABC protocol used in this work serves as screening in order to speed-up and ease the process of allograft selection, through, for instance, visual inspections in the same virtual environment, of the possible candidates and the patient anatomy. The main advantage of the proposed automatic method on

existing bone databanks, is the minimal work needed to perform accurate and robust quantification bone morphology compared to manual measurements, which are time consuming and error prone. In conclusion this work demonstrates the usefulness of three-dimensional models when searching and selecting the best similar host-donor allograft match. The results suggest that a robust technique which provides, reliability, and most importantly, repeatability, has been established. This tool is suitable for such bone banks with huge amount of long bones (more than 50). This method can then be used to match a candidate from the bone bank as a screening search to the patient's femur. On the other hand, the results stemming from the use of this measurement protocol enable accurate selection of allografts achieving the best possible match size considering the geometry of available allograft candidate femur specimens in an automatic manner.

Acknowledgments Lucas Eduardo Ritacco, the main author, has been supported by Swiss National Science Foundation, for applying a short fellowship during January, February and March 2011 in which developed the present study at Institute for Surgical Technology, University of Bern, Switzerland. Thanks to Maria del Carmen Ianella, Statistics Project Management.

References

Brin YS, Livshetz I, Antoniou J, Greenberg-Dotan S, Zukor DJ (2010) Precise landmarking in computer assisted total knee arthroplasty is critical to final alignment. J Orthop Res 28(10):1355–1359

Eckhoff DG, Dwyer TF, Bach JM, Spitzer VM, Reinig KD (2001) Three-dimensional morphology of the distal part of the femur viewed in virtual reality. J Bone Jt Surg Am 83(Suppl 2(Pt 1)):43–50

Eckhoff DG, Bach JM, Spitzer VM, Reinig KD, Bagur MM, Baldini TH, Rubenstein D, Humphries S (2003) Three-dimensional morphology and kinematics of the distal part of the femur viewed in virtual reality. J Bone Jt Surg Am 85(Suppl 4):97–104

Eckhoff DG, Bach JM, Spitzer VM et al (2005) Three-dimensional mechanics, kinematics, and morphology of the knee viewed in virtual reality. J Bone Jt Surg Am 87((Suppl 2)):71–80

Enneking WF, Campanacci DA (2001) Retrieved human allografts: a clinicopathological study. J Bone Jt Surg Am 83:971–986

Enneking WF, Dunham W, Gebhardt MC, Malawer M, Pritchard DJ (1993) A system for the functional evaluation of reconstructive procedures after surgical treatment of tumors of the musculoskeletal system. Clin Orthop Relat Res 286:241–246

Fischer B, Modersitzki J (2008) Ill-posed medicine—an introduction to image registration. Inverse Probl 24:1–19

Hill DL, Batchelor PG, Holden M, Hawkes DJ (2001) Medical image registration. Phys Med Biol 46(3):R1-45 (Epub 2001/03/30)

Ibanez L, Schroeder W, Ng L, Cates J et al (2003) The ITK, Software Guide. Citeseer

Klein A, Andersson J, Ardekani B, Ashburner J, Avants B, Chiang MC, Christensen G, Collins D, Gee J, Hellier P, Song J, Jenkinson M, Lepage C, Rueckert D, Thompson P, Vercauteren T, Woods R, Mann J, Parsey R (2009) Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage 46(3):786–802

Li W, Kezele I, Collins DL, Zijdenbos A, Keyak J, Kornak J, Koyama A, Saeed I, LeBlanc A, Harris T, Lu Y, Lang T (2007) Voxel-based modelling and quantification of the proximal femur using inter-subject registration of quantitative CT images. Bone 41(5):888–895

Lorensen W, Cline H (1987) Marching cubes: a high resolution 3D surface construction algorithm. In proceedings of the 14th annual conference on computer graphics and interactive techniques, SIGGRAPH'87, New York, NY, pp 163–169

Muscolo DL, Petracchi LJ, Ayerza MA, Calabrese ME (1992) Massive femoral allografts followed for 22–36 years. Report of six cases. J Bone Jt Surg Br 74:887–892

Muscolo DL, Ayerza MA, Aponte-Tinao LA (2000) Survivorship and radiographic análisis of knee osteoarticular allografts. Clin Orthop Relat Res 373:73–79

Ourselin S, Roche A, Prima S, Ayache N. Block Matching (2000) A General framework to improve robustness of rigid registration of medical images. in: DiGioia AM, Delp S (eds) Third international conference on medical robotics, imaging and computer assisted surgery (MICCAI 2000) volume 1935 of lectures notes in computer science, Pittsburgh, Penn, USA,October 11–14 2000, pp 557–566

Paul Laurent, Docquier Pierre-Louis, Cartiaux Olivier, Cornu Olivier, Delloye Christian, Banse Xavier (2010) Selection of massive bone allografts using shape-matching 3-dimensional registration. Acta Orthop 81(2):252–257

Ritacco LE, Espinoza Orías AA, Aponte-Tinao L, Muscolo DL, de Quirós FG, Nozomu I (2010) Three-dimensional morphometric analysis of the distal femur: a validity method for allograft selection using a virtual bone bank. Stud Health Technol Inform 160((Pt 2)):1287–1290

Seiler C, Weber S., Schmidt W, Fischer F, Reimers N, Reyes M (2009) Automatic propagation for left and right symmetry assessment of tibia and femur. A computational anatomy based approach. Presented at the 9th annual meeting of CAOS (computer assisted orthopaedic surgery); final programme, Boston, MA, USA, Jun 17–20 2009, pp 195–198

Thirion JP (1998) Image matching as a diffusion process: an analogy with Maxwell's demons. Med Image Anal 2(3): 243–260

Vercauteren T, Pennec X, Perchant A, Ayache N (2009) Diffeomorphic demons: efficient non-parametric image registration. Neuroimage 45(1 Suppl):S61–S72