Journal of Biomechanics 43 (2010) 792–795

Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech
www.JBiomech.com

Short Communication

A computational technique to measure fracture callus in radiographs

Trevor J. Lujan [a,*], Steven M. Madey [a], Dan C. Fitzpatrick [b], Gregory D. Byrd [a], Jason M. Sanderson [a], Michael Bottlang [a]

[a] Biomechanics Laboratory, Legacy Research and Technology Center, Portland, OR 97232, USA
[b] Slocum Center for Orthopedics and Sports Medicine, Eugene, OR 97408, USA

ARTICLE INFO

Article history:
Accepted 5 October 2009

Keywords:
Bone
Fracture
Callus
Image processing
Validation

ABSTRACT

Callus formation occurs in the presence of secondary bone healing and has relevance to the fracture's mechanical environment. An objective image processing algorithm was developed to standardize the quantitative measurement of periosteal callus area in plain radiographs of long bone fractures. Algorithm accuracy and sensitivity were evaluated using surrogate models. For algorithm validation, callus formation on clinical radiographs was measured manually by orthopaedic surgeons and compared to non-clinicians using the algorithm. The algorithm measured the projected area of surrogate calluses with less than 5% error. However, error will increase when analyzing very small areas of callus and when using radiographs with low image resolution (i.e. 100 pixels per inch). The callus size extracted by the algorithm correlated well to the callus size outlined by the surgeons ($R^2=0.94$, $p<0.001$). Furthermore, compared to clinician results, the algorithm yielded results with five times less inter-observer variance. This computational technique provides a reliable and efficient method to quantify secondary bone healing response.

1. Introduction

The formation of fracture callus occurs in the presence of secondary bone healing and has relevance to the mechanical stability at the fracture. For instance, the amount of callus formation in long bone fractures is predictive of bending stiffness (Eastaugh-Waring et al., 2009; Marsh, 1998; Tiedeman et al., 1990). Unfortunately, callus measurement from plain radiographs is a subjective process, with inter-physician variability of 20–25% (Bhandari et al., 2002; Whelan et al., 2002). This subjectivity has implicitly hindered efforts to predict fracture mechanics and treatment efficacy from radiographs (Corrales et al., 2008; Martin et al., 2000).

Image processing algorithms have potential to render callus measurement objective and thereby reduce observer error. However, previous studies which measured callus with image processing protocols did not document the accuracy and objectivity of their methods (Augat et al., 1997; Morcuende et al., 2004). To establish credibility, it is critical that computational techniques be verified and validated (Anderson et al., 2007). This research endeavors to verify and validate a novel computational method to objectify callus measurement from plain radiographs.

2. Methods

The proposed algorithm analyzes fracture callus in digital radiographs using MATLAB (Mathworks; Natwick, MA). User-defined boundary conditions assist the demarcation of cortical fragments, and the callus is automatically segmented. To convert to metric area, a hardware feature of known dimension is chosen as a length standard. Numerical error was analyzed with two and three-dimensional surrogates. For clinical validation, callus formation in distal femur fractures was assessed manually by orthopaedic surgeons and compared to non-clinicians using the algorithm.

2.1. Cortex segmentation

Segmentation of the cortex was guided by user inputs and defined by two-dimensional intensity gradients (∇I_{xy}). First, the user selects a region of interest (ROI) that encompasses the fracture site and the periosteal callus (Fig. 1A). The user then specifies rectangular regions that enclose the ends of the anterior and posterior bone fragments (Fig. 1B). The final user input is to connect a user-defined finite sequence of straight line segments along the external cortices (Fig. 1B). The external cortex is defined for each row as the local ∇I_{xy} maxima, nearest to this line (Fig. 1C). The inner cortex is automatically defined when the I_{xy} magnitude reaches a local maxima ($\nabla I_{xy}=0$), deep to the external cortex (Fig. 1C). The cortex delineations originate at the ROI boundary and expand by permitting only a one pixel shift per row. Cortex endpoints are defined in the user specified regions (Fig. 1B) by a composite ∇I_{xy} score, taken tangential and normal to the cortical bone (Fig. 1D). Finally, the fracture was bridged by the shortest line connecting the tip of the outermost cortex with the periosteal surface of the adjacent cortex (Fig. 2A).

2.2. Callus segmentation

Automatic segmentation of callus involved noise reduction, edge detection and edge localization. To reduce salt-and-pepper noise while maintaining edge

* Corresponding author: Legacy Biomechanics Laboratory, 1225 NE 2nd Avenue, Portland, OR 97215, USA. Tel.: +1 503 413 5487; Fax: +1 503 413 4942.
E-mail address: tlujan@biomechresearch.org (T.J. Lujan).

0021-9290/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jbiomech.2009.10.013

contrast, a median filter was applied over the ROI (Southard and Southard, 1995). The callus edge was initially approximated by selecting pixels in the periosteal region that had a ∇I_{xy} magnitude greater than an empirical threshold of 1.5 (Fig. 2B). After a flood-fill operation, pixel clusters in the periosteal region not adjacent to the external cortices were eliminated. Morphological opening was applied and the callus region was approximated (Figs. 2C and D).

For further refinement, pixels in the callus region were deselected if their intensity was not greater than the sum of a two components threshold (Fig. 2E). The first component was two standard deviations above the local background intensity. The second component was five percent of the difference between the local background intensity and local cortex intensity. Locality was determined with Euclidean distance. Finally, morphological opening and closing operations were applied.

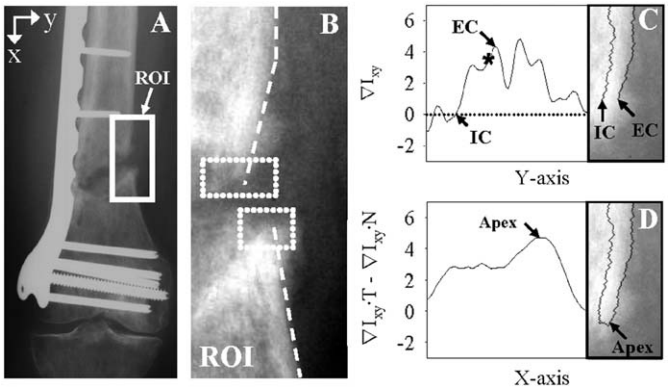

Fig. 1. Cortex segmentation. (A) User selects ROI. (B) User defines regions enclosing fragment ends (dotted) and snaps line segments (dashed) along external cortex. (C) External cortex (EC) at the local max ∇I_{xy} nearest the point where the user approximated the EC (∗). Inner cortex (IC) deep to EC where ∇I_{xy} equals zero. (D) Apex of bone fragment defined at max difference between the intensity gradients tangent (T) and normal (N) to the external cortex.

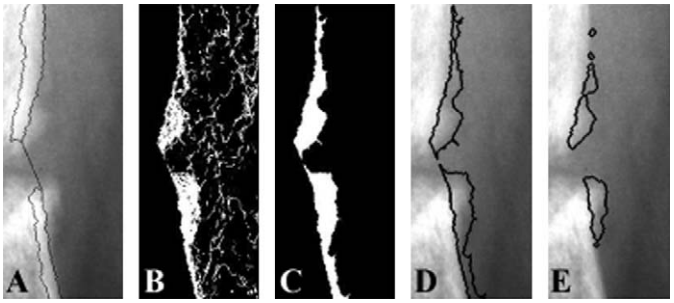

Fig. 2. Callus segmentation. (A) Bone fragments connected. (B) Pixels selected with ∇I_{xy} greater than 1.5. (C) Opening and flood-fill operations executed on objects adjacent to bone fragments. (D) Callus perimeter estimated. (E) Callus perimeter refined by deselecting pixels with intensity values below local threshold.

2.3. Verification of algorithm accuracy

To quantify the inherent error caused by filtering, morphological operations and discretization, the algorithm analyzed sets of manufactured binary images of semi-circular callus surrogates. Pixels per inch (PPI), length standard and callus size were varied by two orders of magnitude. To test algorithm sensitivity to callus maturation and image plane rotation, surrogate models of a fracture callus were developed. The development of these surrogates was guided by quantitative data measured from twenty radiographs of femur fractures taken 6 and 24 weeks post-fracture (Table 1). Surrogates were assembled using callus and cortex components. The callus components were composed of epoxy and fiberglass (profile area $\approx 180\ mm^2$). To produce low and high density surrogate callus, corresponding to actual callus at 6 and 24 weeks post-fracture, the weight concentration of fiberglass beads was set to 12.5% and 50%, respectively (Fig. 3A). Fiberglass mat was added to create the radiographic effect of callus heterogeneity (Fig. 3B). The cortical components (Sawbones, Vashon WA) were then press fit into the callus lumen. Once radiographs of the surrogates were captured (70 kV, 0.06 s) and digitized, a soft tissue envelope was superimposed upon the image and an isochoric or volume preserving transformation was applied to give curvature. These final image processing steps improved the surrogates' resemblance to clinical fractures (Fig. 3C, Table 1), while preserving the known size of the callus surrogate for determination of algorithm accuracy. Radiographs were analyzed by three different operators using a 5 mm length standard.

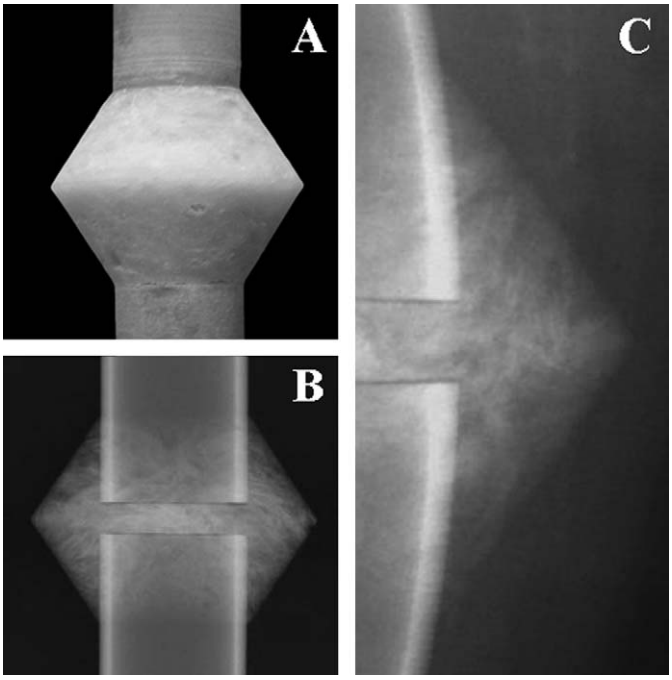

Fig. 3. Callus surrogate. (A) Callus fabricated with fiberglass and epoxy. (B) Radiograph of low density callus. (C) Soft tissue envelope and curvature applied to digital image.

Table 1
Comparison of the surrogate models to actual fracture callus.

Radiograph	Callus		Background		Cortex		
	Mean intensity	Mean intensity variation	Mean intensity	Mean intensity variation	Mean intensity	Mean intensity variation	Mean curvature (deg.)
Low density							
Clinical	0.35	0.09	0.14	0.08	0.71	0.15	7
Surrogate	0.36	0.11	0.13	0.07	0.72	0.13	7
High density							
Clinical	0.48	0.12	0.16	0.09	0.72	0.13	8
Surrogate	0.55	0.16	0.19	0.10	0.70	0.10	7

Clinical radiographs in the low and high density groups were from 6 weeks post-fracture ($n=10$) and 24 weeks post-fracture ($n=10$), respectively. Intensity values were normalized by the maximum intensity bin of 255.

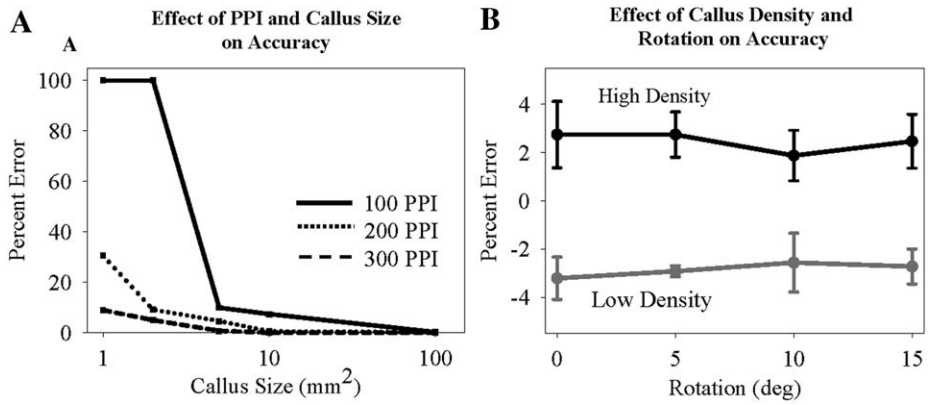

Fig. 4. Accuracy and sensitivity of the algorithm. (A) Numerical error of algorithm when using a 5 mm length standard. Callus size and PPI significantly affect error ($p < 0.001$). (B) Error when analyzing radiographic images of callus surrogates ($n=3$). Accuracy was slightly affected by callus density ($p=0.01$), but not image plane rotation ($p=0.5$).

2.4. Clinical validation of algorithm

Ten antero-posterior and lateral radiographs of distal femur fractures treated with locked plating were analyzed immediately after surgery and at a post-operative time point. Five of the fractures were multifragmentary.

Three clinicians manually outlined the area of callus formation. Two clinicians (DCF, SMM) were senior trauma surgeons, the other was an orthopaedic resident (GDB). Using a pen tablet, clinicians independently outlined the cortex and periosteal callus in each digital radiograph (300 PPI). Film radiographs of the digital images were available for reference. Callus area was computed from the clinicians' digitally marked outlines.

Three non-clinicians evaluated the same radiographs using the proposed algorithm with the following standard operating procedure: (1) orient images to obtain vertical bone alignment with external cortex facing right; (2) ensure ROI encompasses the internal cortex and callus; and (3) reference the initial post-operative radiographs when tracing the external cortex.

2.5. Statistics

The effect of independent factors on accuracy was assessed with one-way and two-way ANOVAs. If significance was detected ($p \leq 0.05$), Tukey post-hoc tests determined significance between factor levels. Differences in callus area between the clinician and algorithm were tested with a Pearson r correlation, while case-by-case comparisons were examined with paired *t*-tests.

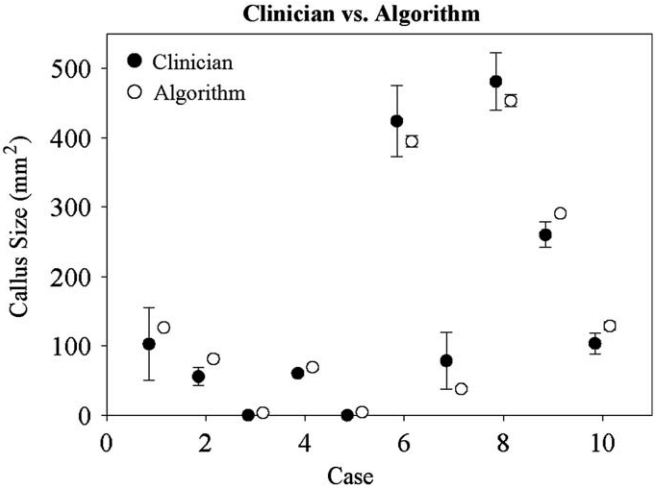

Fig. 5. Fracture callus outlined manually by clinicians ($n=3$) and outlined with the algorithm by non-clinicians ($n=3$). There was no statistical difference between the groups. The algorithm had five times less inter-observer variance. Bars=standard deviation.

3. Results

The algorithm's inherent error in measuring callus area was a function of PPI ($p < 0.001$) and callus size ($p < 0.001$) (Fig. 4A). When image resolution was upgraded from 100 to 200 PPI, the average error was reduced by 31% ($p < 0.001$). Callus areas of 1 mm^2 were detected on average with 30% larger error than areas of 5 mm^2 ($p=0.01$). Although the size of length standard did not have an overall effect on error ($p=0.78$), a notable 13% improvement in error was observed when increasing the length standard from 1 to 5 mm. Further increase to the length standard only reduced error by an additional 0.3%.

The low and high density surrogate models were a good approximation of actual callus at 6 and 24 weeks post-fracture, respectively (Table 1). Relative to the adjacent cortex, the radiodensities of the low and high density callus surrogates were 47% and 81%, respectively. At low and high densities, the algorithm was accurate within 5% of the actual surrogate area (Fig. 4B). The algorithm predicted a significantly greater area for the high density callus than the low density callus ($p < 0.001$). Rotation of the radiograph in the image plane did not affect accuracy ($p=0.9$). The average inter-observer standard deviation for the low and high density surrogate was 0.8% and 1.1%, respectively.

Overall, the clinician and algorithm groups had a high positive correlation when analyzing area of fracture callus ($R^2=0.94$, $p < 0.001$). There was no significant difference between the two groups in analyzing any specific case (Fig. 5). For all ten cases, the average standard deviation or inter-observer variation in the algorithm group was 4% and in the clinician group was 22%. At 200 PPI, the algorithm on average analyzed images in 60 s on a standard computer (2.80 GHz processor).

4. Discussion

The algorithm estimated callus area in long bone fractures with good precision and accuracy. Area was over-predicted by ~3% in surrogate callus with a high radiodensity, while area was under-predicted by ~3% in surrogate callus with low radiodensity. To minimize error implicit in the image processing technique, the length standard and image resolution should be at least 5 mm (roughly the diameter of plate screws) and 200 PPI (typical for digital radiographs), respectively.

The callus size extracted by the algorithm correlated extremely well with an independent analysis by orthopaedic surgeons. By limiting subjective user interaction, the computational method

had five times less inter-observer variance relative to the surgeons. This technique provides a reliable measure of callus size without requiring time-consuming and prohibitive consultations with physicians.

Limitations exist with this technique. Accuracy results are based on surrogates that model a simplified fracture callus and should be interpreted accordingly. Also, the segmentation of the cortex required some user interaction. Although user input did not result in large variations between observers, it nevertheless inhibits objectivity and processing speed. To further improve objectivity, a region growing procedure could automate cortex segmentation (Gelaude et al., 2006). User interaction was not required to segment the callus, since constant thresholds were automatically applied. To make more robust, the algorithm could be iterated to optimize these thresholds. The algorithm requires the external cortex to be vertically aligned. However, since the functions used for callus segmentation were invariant to rotation, moderate variations in vertical alignment did not affect accuracy. Finally, two-dimensional projections are not ideal in measuring a three-dimensional biological process. Nevertheless, radiographs do adequately approximate callus growth at the fracture site (Augat et al., 1997) and require 45 times less radiation exposure than equivalent computed tomography scans (Chauhan et al., 2004).

In summary, the proposed algorithm presents a reliable and valid technique to measure a radiographic feature with biomechanical and clinical relevance. By reducing the inter-observer variation in callus measurement, this technique may improve the utility of radiographs to predict bending stiffness at fracture sites (McClelland et al., 2007). Furthermore, the influence of fracture fixation constructs on secondary bone healing can now be efficiently analyzed in large prospective and retrospective cohort studies.

Conflict of interest

The authors have no conflict of interest.

Acknowledgements

We wish to thank Dr. Larry Marsh for his clinical insight and recommendations. This project was supported by a grant from the Legacy Research Advisory Committee.

References

Anderson, A.E., Ellis, B.J., Weiss, J.A., 2007. Verification, validation and sensitivity studies in computational biomechanics. Comput. Methods Biomech. Biomed. Eng. 10, 171–184.

Augat, P., Merk, J., Genant, H.K., Claes, L., 1997. Quantitative assessment of experimental fracture repair by peripheral computed tomography. Calcif. Tissue Int. 60, 194–199.

Bhandari, M., Guyatt, G.H., Swiontkowski, M.F., Tornetta 3rd, P., Sprague, S., Schemitsch, E.H., 2002. A lack of consensus in the assessment of fracture healing among orthopaedic surgeons. J. Orthop. Trauma 16, 562–566.

Chauhan, S.K., Clark, G.W., Lloyd, S., Scott, R.G., Breidahl, W., Sikorski, J.M., 2004. Computer-assisted total knee replacement. A controlled cadaver study using a multi-parameter quantitative CT assessment of alignment (the Perth CT Protocol). J. Bone Jt. Surg. Br. 86, 818–823.

Corrales, L.A., Morshed, S., Bhandari, M., Miclau III, T., 2008. Variability in the assessment of fracture-healing in orthopaedic trauma studies. J. Bone Jt. Surg. Am. 90, 1862–1868.

Eastaugh-Waring, S.J., Joslin, C.C., Hardy, J.R., Cunningham, J.L., 2009. Quantification of fracture healing from radiographs using the maximum callus index. Clin. Orthop. Relat. Res. 467, 1986–1991.

Gelaude, F., Vander Sloten, J., Lauwers, B., 2006. Semi-automated segmentation and visualisation of outer bone cortex from medical images. Comput. Methods Biomech. Biomed. Eng. 9, 65–77.

Marsh, D., 1998. Concepts of fracture union, delayed union, and nonunion. Clin. Orthop. Relat. Res., S22–S30.

Martin, J., Marsh, J.L., Nepola, J.V., Dirschl, D.R., Hurwitz, S., DeCoster, T.A., 2000. Radiographic fracture assessments: which ones can we reliably make? J. Orthop. Trauma 14, 379–385.

McClelland, D., Thomas, P.B., Bancroft, G., Moorcraft, C.I., 2007. Fracture healing assessment comparing stiffness measurements using radiographs. Clin. Orthop. Relat. Res. 457, 214–219.

Morcuende, J.A., Gomez, P., Stack, J., Oji, G., Martin, J., Fredericks, D.C., et al., 2004. Effect of chemotherapy on segmental bone healing enhanced by rhBMP-2. Iowa Orthop. J. 24, 36–42.

Southard, T.E., Southard, K.A., 1995. Performance of filters for noise reduction in maxillary alveolar bone imaging. IEEE Trans. Biomed. Eng. 42, 13–20.

Tiedeman, J.J., Lippiello, L., Connolly, J.F., Strates, B.S., 1990. Quantitative roentgenographic densitometry for assessing fracture healing. Clin. Orthop. Relat. Res., 279–286.

Whelan, D.B., Bhandari, M., McKee, M.D., Guyatt, G.H., Kreder, H.J., Stephen, D., et al., 2002. Interobserver and intraobserver variation in the assessment of the healing of tibial fractures after intramedullary fixation. J. Bone Jt. Surg. Br. 84, 15–18.