

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jmbbm

Research Paper

An experimental approach to determining fatigue crack size in polyethylene tibial inserts

Carly A. Lockard^a, Anthony P. Sanders^{a,b}, Bart Raeymaekers^{a,*}^aDepartment of Mechanical Engineering, University of Utah, Salt Lake City, UT 84112, USA^bOrtho Development Corporation, Draper, UT 84020, USA

ARTICLE INFO

Article history:

Received 2 May 2015

Received in revised form

16 September 2015

Accepted 21 September 2015

Available online 30 September 2015

Keywords:

Fatigue

UHMWPE

Prosthetic knee joint

ABSTRACT

A major limiting factor to the longevity of prosthetic knee joints is fatigue crack damage of the polyethylene tibial insert. Existing methods to quantify fatigue crack damage have several shortcomings, including limited resolution, destructive testing approach, and high cost. We propose an alternative fatigue crack damage visualization and measurement method that addresses the shortcomings of existing methods. This new method is based on trans-illumination and differs from previously described methods in its ability to non-destructively measure subsurface fatigue crack damage while using a simple and cost-effective bench-top set-up. We have evaluated this method to measure fatigue crack damage in two tibial inserts. This new method improves on existing image-based techniques due to its usability for subsurface damage measurement and its decreased reliance on subjective damage identification and measurement.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Each year, approximately 700,000 primary total knee arthroplasty (TKA) surgeries are performed in the United States to treat degenerative joint diseases by replacing the damaged natural cartilage with a prosthetic knee joint (OECD, 2013). A prosthetic knee joint typically comprises a metal femoral component that is attached to the femur and articulates with a polyethylene tibial insert, which is constrained in a metal tibial plateau that is anchored in the tibia. It is well documented that the statistical survivorship of prosthetic knee joints declines dramatically after 15–20 years of use (Stiehl et al., 2006; Heyse et al., 2014), primarily as a result of loosening at the implant-bone interface, instability, infection,

polyethylene wear, arthrofibrosis, and malalignment (Schroer et al., 2013). Polyethylene wear and aseptic loosening, which results from polyethylene wear, account for approximately 40% of the total failure incidence (Schroer et al., 2013; Hossain et al., 2010). Tibial inserts are particularly susceptible to polyethylene wear resulting from fatigue because of cyclic loading caused by reciprocating rolling/sliding of the joint. In addition, the limited congruency between the articulating surfaces of the prosthetic knee joint results in a small area of contact and, thus, high contact stress directly beneath the surface of the tibial insert condyles (Simis et al., 2006). Fatigue wear is typically observed as fatigue crack damage, defined as progressive and localized damage resulting from cyclic loading (Blunn et al., 1997).

*Corresponding author.

E-mail address: bart.raeymaekers@utah.edu (B. Raeymaekers).

In contrast to adhesive/abrasive wear, which is measured during knee simulator wear testing, fatigue wear is only classified as occurring or not occurring during testing (ISO/IEC, 2009). The absence of a fatigue wear measurement standard for knee simulator wear testing is due in part to the lack of adequate, quantitative subsurface fatigue crack damage measurement methods, and prevents detailed analysis of the fatigue crack damage that occurs during knee simulator wear testing. In addition, it limits the quantitative analysis of retrieval implants.

Few methods have been developed to identify fatigue crack damage in tibial inserts. Cornwall et al. (1995) estimate the surface damage area based on observation of damage within 10 regions on the tibial insert, whereas Harman et al. (2001) provide a more precise measurement by hand-tracing the damage on digital images. However, both of these measurement methods focus exclusively on surface rather than subsurface fatigue crack damage.

Several sub-surface fatigue crack detection techniques have also been documented in the literature. Muratoglu et al. (2004) use backlighting to enhance the visualization of subsurface fatigue crack damage in knee simulator wear-tested tibial inserts, which they classify only as occurring or not occurring. Thin-slice microscopy provides the ability to reveal and measure subsurface cracks through the depth of the material, via sectioning of the insert in ultra-thin slices using a microtome (Medel et al., 2011), after the specimen is cooled to its glass transition temperature to limit deformation during slicing (Petzow, 1999). Fig. 1A shows a schematic of the thin-slice microscopy technique applied to a tibial insert with a subsurface crack. Thin slices of the tibial insert are analyzed with an optical microscope. Although thin-slice microscopy allows subsurface crack measurement, it is destructive and the possibility of creating artifacts during the slicing process exists. Scanning acoustic tomography (SAT) is based on reflection of high-frequency acoustic waves to map local differences in acoustic impedance in a specimen, which indicates the presence of cracks (Todo et al., 1999). Fig. 1B shows a schematic of the SAT technique, along with a resulting processed image of the subsurface fatigue cracks. The resolution of SAT decreases with increasing specimen thickness and depends on the material and the surface roughness of the specimen (Brand et al., 2011). Microscale computed-tomography (Micro-CT) uses x-rays to sequentially scan thin slices of the specimen with a resolution of 1–100 μm (Rasesh, 2010). The information about sub-surface cracks contained in the thin slice images is assembled into a 3D representation of the specimen, enabling quantitative evaluation of subsurface crack in three dimensions (Rasesh, 2010; Teeter et al., 2010). Fig. 1C shows a schematic of the technique. Optical microscopy using indirect illumination allows non-destructive examination of subsurface damage using indirect or oblique angle illumination to enhance visualization of surface features by means of shadows and highlights (Furmanski et al., 2011). Fig. 1D shows a schematic of a specimen lit at an oblique angle under a microscope camera and the resulting view of the highlights and shadows created by surface damage features.

The existing methods to quantify fatigue crack damage have several shortcomings, including limited resolution, destructive approach, and/or high cost. No method seems to exist that enables quantitative measurement of subsurface

fatigue crack damage in a non-destructive way, without the need for expensive equipment (such as Micro-CT). Hence, this paper attempts to fill this gap, and proposes an alternative fatigue crack damage visualization and measurement method that addresses the shortcomings of existing methods. This new method is based on trans-illumination and differs from previously described methods in its ability to measure subsurface fatigue crack damage non-destructively while using a simple and cost-effective bench-top set-up.

2. Materials and methods

2.1 Tibial insert specimens

We have analyzed two tibial insert specimens used in a left knee application. Tibial insert 1 is a 7 mm thick posterior-

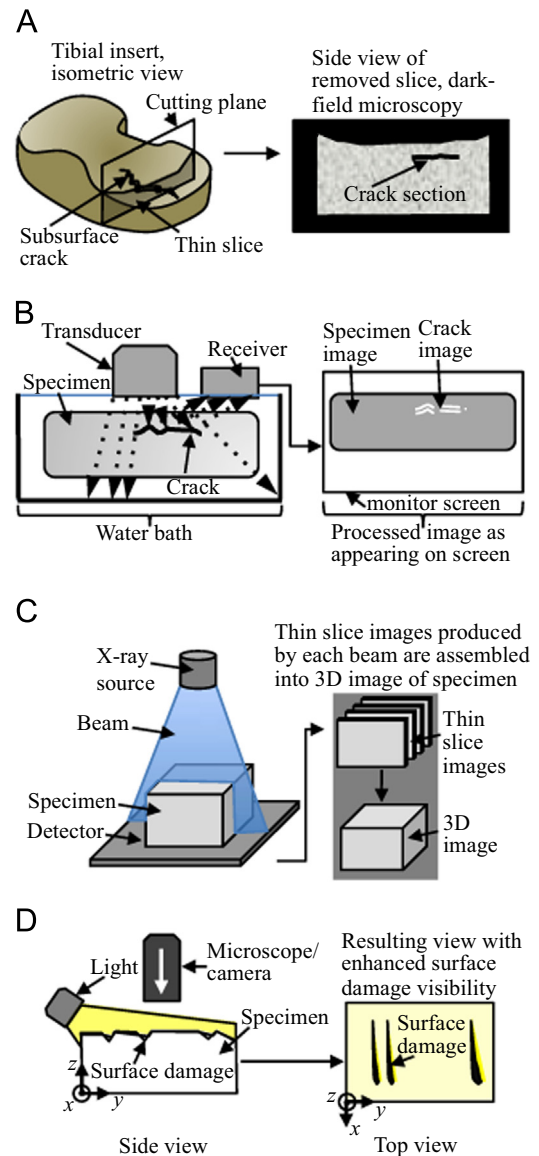


Fig. 1 – Existing methods to quantify fatigue crack damage, including (A) thin-slice microscopy, (B) scanning acoustic tomography, (C) micro-CT, and (D) optical microscopy using indirect illumination.

stabilized/cruciate sacrificing type that was used in knee simulator wear testing. Tibial insert 2 is a 14 mm thick cruciate retaining type, retrieved after 9 years use in vivo. Both tibial inserts are machined from compression molded Ticona GUR 1050 UHMWPE, packaged with nitrogen flush and vacuum, and sterilized using 25–40 kGy gamma irradiation. Prior to knee simulator wear testing, tibial insert 1 was unpackaged and aged for two weeks in accordance with the standard accelerated aging procedure for gamma irradiated UHMWPE (ASTM International, 2008). Fig. 2A shows tibial insert 1 after knee simulator testing, and Fig. 2B shows tibial insert 2 after retrieval. A dashed ellipse locates the fatigue crack damage region. Each specimen is divided in two to remove a slice of material for oxidation analysis (not discussed in this manuscript).

Tibial insert 1 is wear tested at the University of Nebraska Medical Center Orthopaedic Biomechanics Laboratory, Omaha, Nebraska, using a modified Instron-Stanmore Model KC knee simulator (Instron Corp., Norwood, MA.) in accordance with ISO standard 14243-1 (ISO/IEC, 2009), and with the following applied loading: axial load (178–2600 N), anterior–posterior load (0.6–230 N), internal external torque (0–6 Nm). Bovine calf serum diluted with deionized water to a protein mass concentration of 20 g/l is used as lubricant (ISO/IEC, 2009). Simulator testing is conducted at 1 Hz for 75,177 simulated gait cycles, and aborted prior to the standard test duration of five million cycles due to visible fatigue crack damage. The tibial inserts are cleaned before fatigue crack damage measurement with a soft cloth to remove

any surface debris and are stored at -18°C to retard further oxidation.

2.2 Fatigue crack damage measurement

The fatigue crack damage area is quantified by measuring regions of decreased light transmission in images of the trans-illuminated specimen. Fig. 3 shows light transmission through trans-illuminated fatigue crack damaged (Fig. 3A) and pristine (Fig. 3B) tibial inserts. The fatigue crack damage appears as a darkened crescent on the medial condyle of the fatigue crack damaged tibial insert.

Fig. 4 shows a schematic of the experimental set-up for image capture. The tibial insert is backlit using a uniform light source created by a fiber optic illuminator (Fiber-Lite, Bausch & Lomb, USA) and diffusing screen (0.007 mm matte two-sided film, Inventables, USA). The tibial insert is secured in a vise and is surrounded with opaque plastic sheeting to prevent light from passing around the sides of the tibial insert. Sheeting is also used to block ambient light from the entire setup between the fiber optic illuminator and the digital camera (Powershot ELPH 300 HS, 12 Megapixels, Canon, USA). A scale bar with printed lines at 1 mm intervals is attached to the tibial insert flush with the photographed surface. To prevent dimensional distortion created by an oblique viewing angle, the images are captured with the camera lens parallel to the inferior surface of the tibial insert. The distance between the camera lens and tibial insert is adjusted to ensure that the fatigue crack damage is in focus.

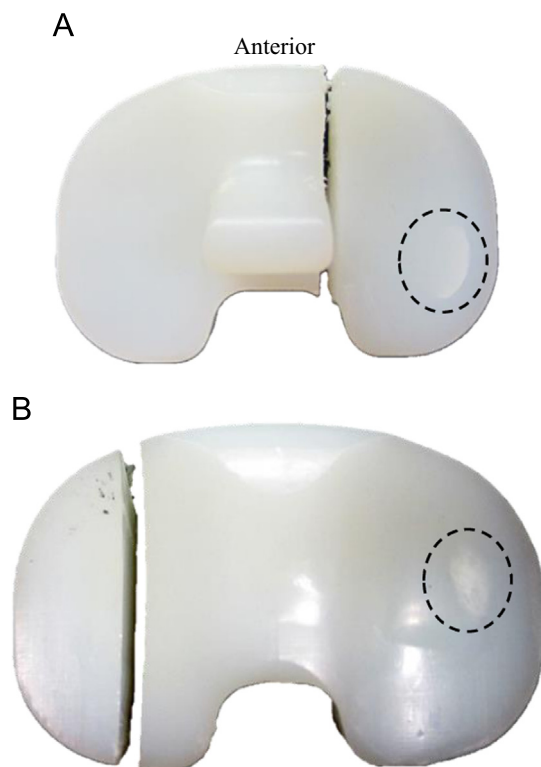


Fig. 2 – Top view of (A) tibial insert 1 after knee simulator testing, and (B) tibial insert 2 after retrieval. The dashed ellipse encloses the fatigue crack damage region. A slice of material has been removed from both specimens for oxidation analysis.

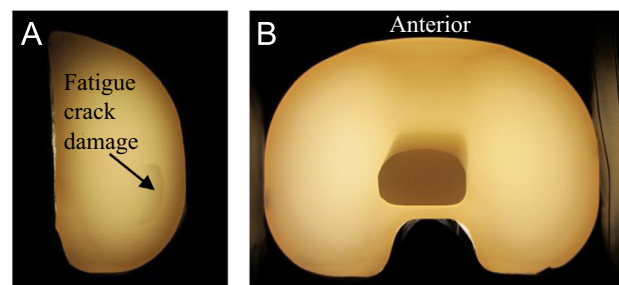


Fig. 3 – Light transmission through trans-illuminated left knee tibial inserts (A) with fatigue crack damage and (B) without fatigue crack damage.

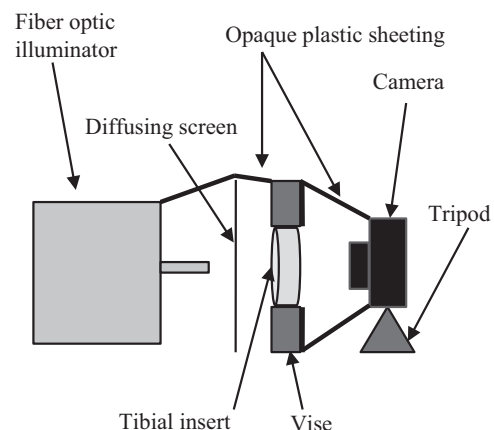


Fig. 4 – Experimental set-up for image capture.

Fig. 5 shows a detailed overview of the image processing algorithm (we use tibial insert 2 as an example). Trans-illuminated images are imported into Matlab, converted to grayscale, and filtered using a Wiener filter to remove high-frequency noise that degrades the effectiveness of the edge detection algorithm (Fig. 5A). Edge detection is performed using the Canny edge detection algorithm, resulting in a black and white image of edge and non-edge pixels, respectively (Canny, 1986) (Fig. 5B). The low and high Canny algorithm threshold for each tibial insert fatigue crack damage image is selected by the user for high sensitivity to edges in the fatigue crack damage region. Under-detection of edges is undesirable due to the potential for under-estimating the fatigue crack damage area. Over-detection of edges, in which noise is incorrectly classified as edges, is corrected through the edge pixel sorting process described in Section 2.3. The selected low and high thresholds for the fatigue crack damage image for tibial insert 1 and 2 are 0.0225 and 0.2500, and 0.0300 and 0.0740. The edge pixels are sorted based on

the pixel density into two categories: (1) edge pixels caused by fatigue crack damage, and (2) non-damage edge pixels caused by image noise (see Section 2.3). The non-damage edge pixels are removed, leaving only the fatigue crack damage edge pixels (Fig. 5C). Then all pixels except the outer contour of the fatigue crack damage are removed (Fig. 5D). Fig. 5E shows the outer edge pixels superimposed on the original fatigue crack damage image for comparison. The outermost fatigue crack damage edge pixels are used as interpolation points for a piecewise cubic spline. Fig. 5F shows the cubic spline fit to the outer fatigue crack damage edge pixels, whereas Fig. 5G shows the cubic spline superimposed on the original fatigue crack damage image. The area within the cubic spline is calculated to estimate the size of the fatigue crack damage. Fig. 5H shows the same cubic spline superimposed on an image of the entire tibial insert medial condyle to illustrate the size and position of the cubic spline relative to the tibial insert condyle. A magnified inset shows a detailed view of the fatigue crack damage and superimposed cubic spline.

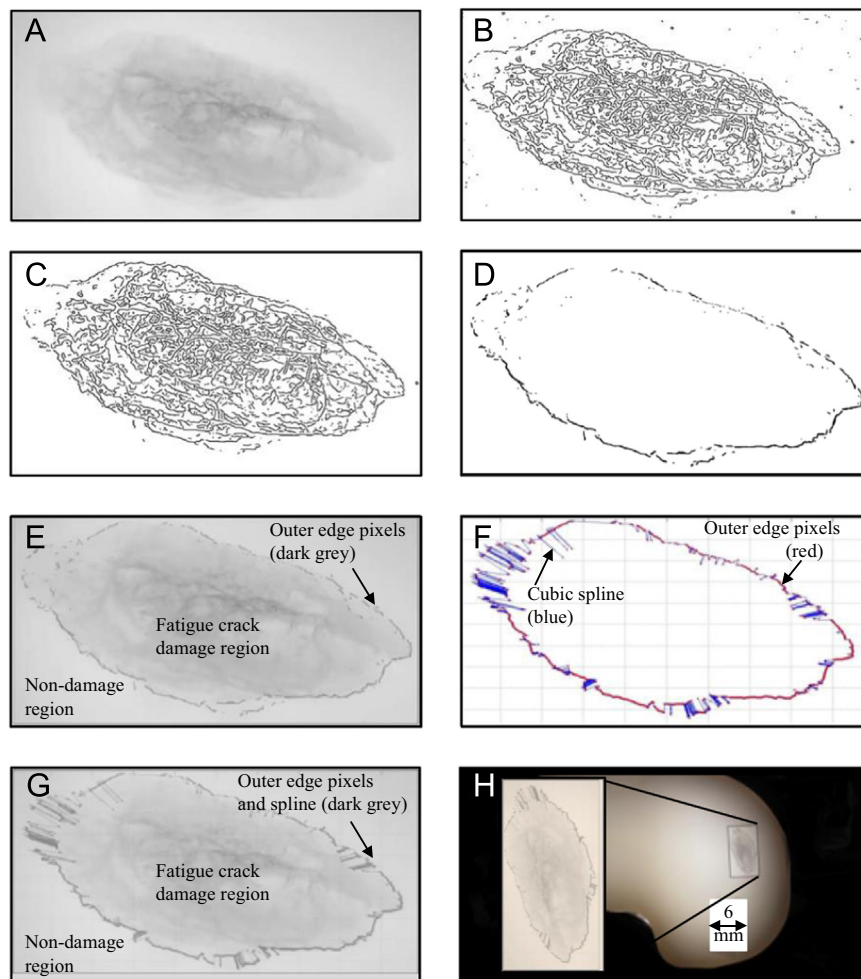


Fig. 5 – Image processing and fatigue crack damage area measurement process, illustrated for tibial insert 2. (A) The original image is converted to grayscale and filtered with a Wiener filter. (B) A Canny edge finding algorithm is used to detect edges within the image. (C) The edge pixels identified by the Canny algorithm are sorted into true fatigue crack damage edges and non-damage edges. Non-damage edges are eliminated. (D) All but the outermost edge pixels are removed. (E) The outer edge image superimposed on the original fatigue crack damage image for comparison. (F) A cubic spline is fit to the outer fatigue crack damage pixels. (G) The cubic spline superimposed on the fatigue crack damage image. (H) View of the entire tibial insert with cubic spline superimposed and with the fatigue crack damage magnified (inset).

2.3 Distinguishing between fatigue crack damage and non-damage

The edge pixels identified by the Canny algorithm are sorted into true fatigue crack damage edges and false edges by considering the edge pixel density. A false edge delineates an image region that is free of fatigue crack damage but contains contrasting light and dark pixels, which resemble damage edges, due to corruption by noise and light saturation. The edge pixel density is defined, for a specific pixel of interest, as the number of edge pixels within a surrounding neighborhood divided by the total number of pixels in that neighborhood. The fatigue crack damage regions of the tibial insert image display a higher edge pixel density than non-damage regions of the image. This difference in edge pixel density is exploited to categorize each edge pixel as fatigue crack damage or as a false edge. Two sorting parameters must be determined: (1) the optimal neighborhood size for computing the edge pixel density surrounding the pixel of interest and (2) the edge pixel density threshold that distinguishes fatigue crack damage edge pixels from false edge pixels.

The optimal neighborhood size is determined as follows. First, two rectangular clusters are manually selected within the regions that exhibit visible fatigue crack damage and no damage, respectively. Fig. 6 shows a 414 by 234 pixel rectangular cluster within the fatigue crack damage (Fig. 6A) and non-damage region (Fig. 6B) of tibial insert 2 as an example.

Next, we divide each of the two rectangular clusters in non-overlapping, square neighborhoods, and calculate their respective intra-cluster correlation coefficient $0 \leq \rho \leq 1$ for 12 different neighborhood sizes, ranging from 25 pixels (5 by 5 pixels) up to half of the width of the cluster. This measure quantifies the relatedness of the pixel density of the clusters by comparing the variance within clusters with the variance between clusters (Killip et al., 2004) and is defined as

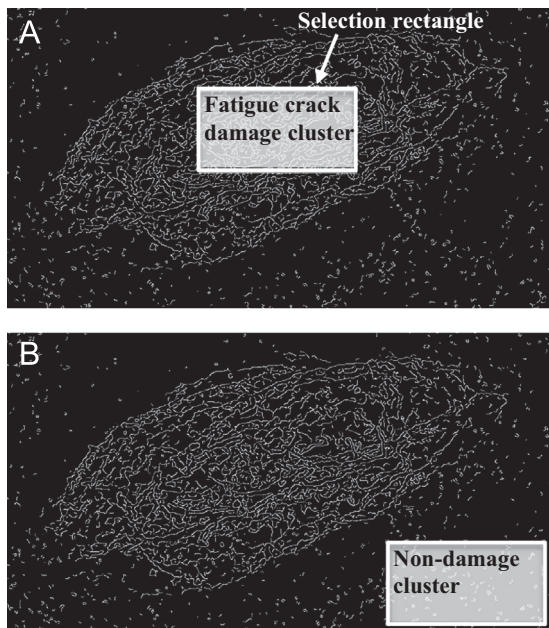


Fig. 6 – (A) Fatigue crack damage and (B) non-damage cluster selection.

$$\rho = \frac{S_b^2}{S_b^2 + S_w^2} \quad (1)$$

Here, S_w^2 is the within-cluster variance, and S_b^2 is the between-cluster variance, both referring to edge pixel densities. We use the subscript D and ND to refer to the fatigue crack damage and non-damage clusters, respectively. Thus, maximizing the intra-cluster correlation coefficient for the fatigue crack damage ρ_D and non-damage ρ_{ND} clusters via the optimal selection of neighborhood size will provide the best capability to distinguish true edge pixels from false ones. S_w^2 is the variance in edge pixel density occurring within a single cluster (for a given neighborhood size) due to the differences between individual neighborhood samples in that cluster, i.e.

$$S_w^2 = \frac{\sum_{i=1}^n (x(i) - \bar{x})^2}{n-1}, \quad (2)$$

where n is the total number of non-overlapping neighborhood samples that fit within the cluster for a given neighborhood size ($4 \leq n \leq 2438$ in our study), $x(i)$ represents the edge pixel density for each neighborhood sample i , and \bar{x} represents the overall edge pixel density for the entire cluster. We have calculated S_w^2 for each neighborhood size, for both the fatigue crack damage ($S_{w,D}^2$) and non-damage clusters ($S_{w,ND}^2$). The between-cluster variance S_b^2 is the variance due to different edge pixel density between the two clusters, calculated as

$$S_b^2 = \frac{n_D (\bar{x}_D - \bar{x})^2 + n_{ND} (\bar{x}_{ND} - \bar{x})^2}{(n_D + n_{ND}) - 1} \quad (3)$$

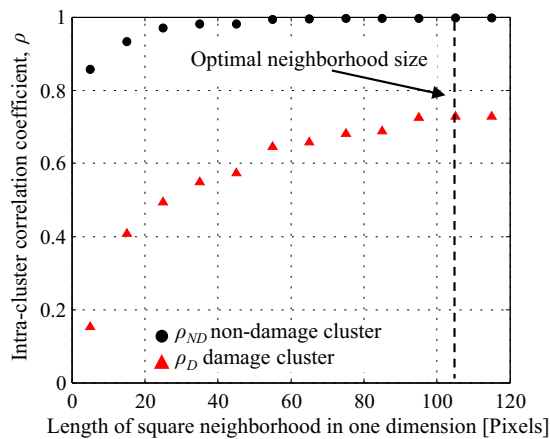
where n_D and n_{ND} represent the number of non-overlapping neighborhood samples in the rectangular fatigue crack damage and non-damage selections for a given neighborhood size, \bar{x} represents the overall average edge pixel density of the image under analysis, and \bar{x}_D and \bar{x}_{ND} are the edge pixel densities of the rectangular fatigue crack damage and non-damage clusters. S_b^2 is calculated for each neighborhood size.

An intra-cluster correlation coefficient ρ approaching 1.0 indicates that the between-cluster variance is much larger than the within-cluster variance. Physically, this means that the variance of the edge pixel densities considering both clusters (fatigue crack damage and non-damage) is greater than the variance of the edge pixel densities within the cluster of interest. The within-cluster variance provides information about the similarity in edge pixel density within the cluster of interest, while the between-cluster variance provides information specifically about how distinguishable the fatigue crack damage cluster is from the non-damage cluster.

The between-cluster variance, within-cluster variance, and intra-cluster correlation coefficient are all dependent on the neighborhood size used to calculate the edge pixel density. Table 1 summarizes the results for the twelve neighborhood sizes tested on the image of tibial insert 2. From Table 1 we observe that the between-cluster variance increases with increasing neighborhood size, reaching its maximum value of $2.231 \cdot 10^{-3}$ for the second largest neighborhood size used in this study, i.e., 105 by 105 pixels. In contrast, the within-cluster variance decreases with increasing neighborhood size. The combination of an increasing between-cluster variance and decreasing within-cluster variance results in an increase in intra-cluster correlation coefficient with increasing

Table 1 – Between-cluster variance S_b^2 , within-cluster variance S_w^2 , and fatigue crack damage intra-cluster correlation coefficient ρ_D of each tested neighborhood size.

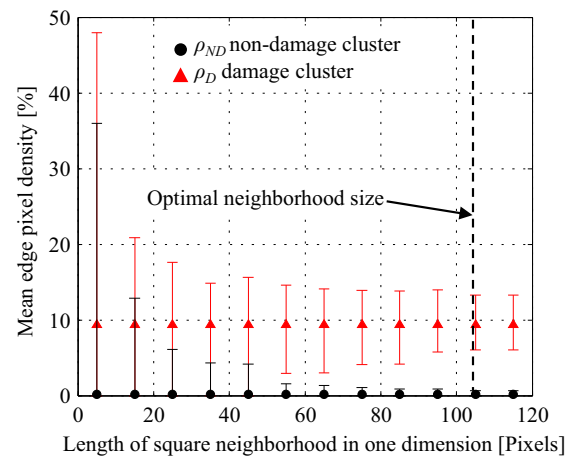
Neighborhood size in one dimension [pixels]	S_b^2	$S_{w,D}^2$	$S_{w,ND}^2$	ρ_D	ρ_{ND}
5	$2.107 \cdot 10^{-3}$	$1.176 \cdot 10^{-2}$	$3.516 \cdot 10^{-4}$	0.152	0.857
15	$2.109 \cdot 10^{-3}$	$3.062 \cdot 10^{-3}$	$1.515 \cdot 10^{-4}$	0.408	0.933
25	$2.113 \cdot 10^{-3}$	$2.167 \cdot 10^{-3}$	$6.311 \cdot 10^{-5}$	0.494	0.971
35	$2.119 \cdot 10^{-3}$	$1.749 \cdot 10^{-3}$	$4.104 \cdot 10^{-5}$	0.548	0.981
45	$2.126 \cdot 10^{-3}$	$1.583 \cdot 10^{-3}$	$3.897 \cdot 10^{-5}$	0.573	0.982
55	$2.143 \cdot 10^{-3}$	$1.180 \cdot 10^{-3}$	$1.294 \cdot 10^{-5}$	0.645	0.994
65	$2.150 \cdot 10^{-3}$	$1.118 \cdot 10^{-3}$	$1.080 \cdot 10^{-5}$	0.658	0.995
75	$2.161 \cdot 10^{-3}$	$1.016 \cdot 10^{-3}$	$8.679 \cdot 10^{-6}$	0.680	0.996
85	$2.185 \cdot 10^{-3}$	$9.898 \cdot 10^{-4}$	$6.575 \cdot 10^{-6}$	0.688	0.997
95	$2.198 \cdot 10^{-3}$	$8.340 \cdot 10^{-4}$	$6.614 \cdot 10^{-6}$	0.725	0.997
105	$2.231 \cdot 10^{-3}$	$8.376 \cdot 10^{-4}$	$4.471 \cdot 10^{-6}$	0.727	0.998
115	$2.231 \cdot 10^{-3}$	$8.376 \cdot 10^{-4}$	$4.471 \cdot 10^{-6}$	0.727	0.998

**Fig. 7 – Intra-cluster correlation coefficient for the fatigue crack damage ρ_D and non-damage ρ_{ND} clusters for the range of tested neighborhood sizes.**

neighborhood size, with a maximum value of $\rho_D=0.727$ in our study. Note that the variance here is dimensionless, since it is expressed as edge pixels over total pixels.

Fig. 7 shows the intra-cluster correlation coefficients for the fatigue crack damage ρ_D (triangular marker) and non-damage ρ_{ND} (circular marker) clusters, as a function of neighborhood size. The optimal neighborhood size is selected as the smallest neighborhood at which the intra-cluster correlation coefficient for the fatigue crack damage cluster ρ_D reaches a plateau. In the case of tibial insert 2, we found that $\rho_D=0.727$ for a 105 by 105 pixels neighborhood size.

Fig. 8 shows the edge pixel density range for the fatigue crack damage (triangular marker) and non-damage clusters (circular marker) versus neighborhood size. The error bars indicate the range of the edge pixel densities. For the smallest neighborhood size, the edge pixel densities calculated for the fatigue crack damage and non-damage clusters overlap significantly, which corresponds with low between-cluster variance, and they vary widely for each cluster, signifying high within-cluster variance. This results in a low intra-cluster correlation coefficient (0.152) and limited ability to distinguish between the two clusters at the 5 by 5 pixel

**Fig. 8 – Mean edge pixel density of fatigue crack damage and non-damage clusters for each tested neighborhood sample size. The error bars indicate the range of the edge pixel densities.**

neighborhood size. In contrast, the optimal neighborhood size, 105 by 105 pixels, corresponds to no overlap between the density ranges, indicating increased between-cluster variance, and a small density range for each cluster, signifying increased within-cluster variance and a high fatigue crack damage intra-cluster correlation coefficient (0.727).

The second parameter in addition to the neighborhood size that is used to distinguish fatigue crack damage edge pixels from false edge pixels is the edge pixel density threshold, defined as the maximum edge pixel density in the manually selected non-damage cluster. Edge pixels are eliminated when the calculated edge pixel density falls below this threshold. An example of the result of this elimination is the difference in image noise between Fig. 5B and C. When evaluating the fatigue crack damage, the edge pixel density is calculated for each edge pixel in the image of the tibial insert using the optimum neighborhood size. If the pixel density exceeds the threshold it is classified as a damage pixel, and vice versa.

3. Results and discussion

We have measured fatigue crack damage of the two tibial inserts described in Section 2. Table 2 lists the image size, optimal neighborhood size, edge pixel density threshold, and the fatigue crack damage intra-cluster correlation coefficient ρ_D for the two tested tibial insert images.

Note that ρ_D is much smaller for tibial insert 1 than 2. This is because the fatigue crack damage was visually fainter in tibial insert 1 than 2 and, thus, when selecting the damage cluster, it contained a significant portion of undamaged material, thereby making it more difficult to distinguish between true damage and non-damage edges. Fig. 9A shows the visually observed damage of tibial insert 1, and Fig. 9B shows a magnified view of the damage. Fig. 9C shows the cubic spline fit to the damage region, which closely fits with the visible edge on the left edge of the fatigue crack damage. In Fig. 9C, the orientations of x- and y-axes of the plotted cubic spline correspond respectively to the medial-lateral and anterior-posterior axes of the tibial insert. The total measured fatigue crack damage area is 22 mm². On the right side, where no distinct edge exists, the edge of the fatigue crack damage region is approximated by a straight section of cubic spline connecting the two points at the far ends of the visible edge line. We observe jagged cubic spline regions in Fig. 9C, resulting from image noise remaining after filtering. This is most prominent in regions where the visible edge of the

fatigue crack damage appears faint or discontinuous and the Canny edge detection algorithm is less effective. Incomplete fatigue cracks with sections of intact material may contribute to faint and discontinuous fatigue crack damage edges.

Fig. 10A shows the visually observed damage of tibial insert 2, and Fig. 10B shows a magnified view of the damage. Fig. 10C shows the cubic spline fit to the damage region. The total measured fatigue crack damage area is 38 mm². The orientations of x- and y-axes of the plotted cubic spline correspond to the medial-lateral and anterior-posterior axes of the tibial insert, respectively. We observe jagged cubic spline regions in Fig. 10C, resulting from image noise remaining after filtering.

This new method to quantify fatigue crack damage, based on trans-illumination, overcomes the reliance on subjective identification of fatigue crack damage edges (Harman et al., 2001; Furmanski et al., 2011) and manual tracing techniques (Harman et al., 2001). In addition, the fatigue crack area is quantified, which has not been done for existing techniques including SAT (Todo et al., 1999) and trans-illumination for qualitative evaluation (Muratoglu et al., 2004). The measured fatigue crack damage area allows quantitatively comparing fatigue crack damage severity between tibial insert specimens.

The measurement resolution is dependent on the fatigue crack damage image scale. The fatigue crack damage images for tibial insert 1 and tibial insert 2 scale as 841 pixels per

Table 2 – Image size, optimal neighborhood size, edge pixel density threshold, and fatigue crack damage intra-cluster correlation coefficient for the tested images of tibial inserts 1 and 2.

Image	Image size, width by height [pixels]	Rectangular cluster size, width by height [pixels]	Optimal neighborhood size [pixels]	Edge pixel density threshold	ρ_D
Tibial insert 1	393 by 537	295 by 196	85 by 85	0.0030	0.154
Tibial insert 2	744 by 1400	414 by 234	105 by 105	0.0053	0.727

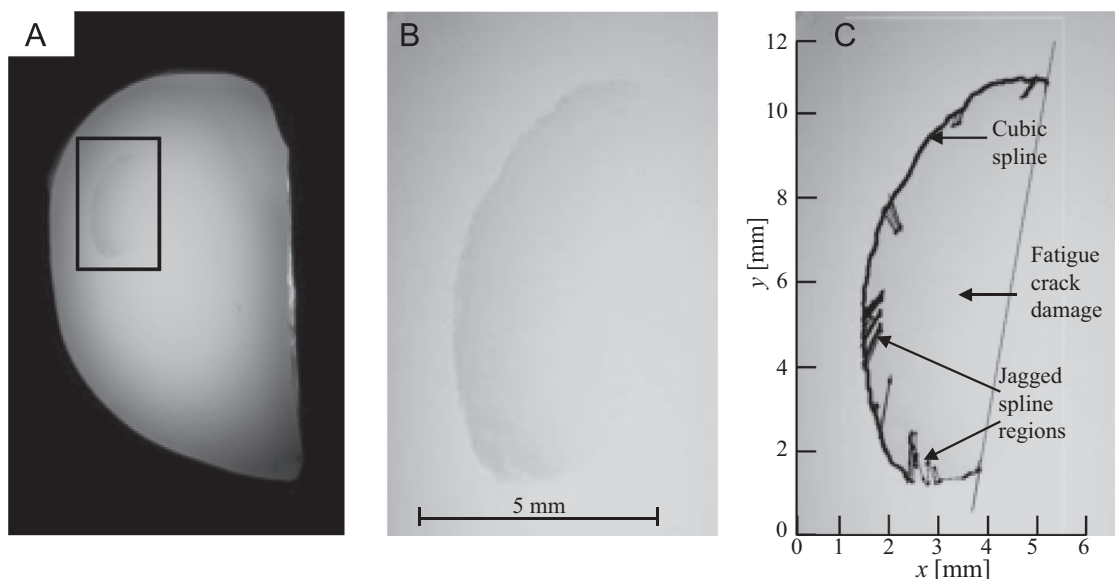


Fig. 9 – The cubic spline superimposed on the fatigue crack damage image for tibial insert 1. (A) The fatigue crack damage area (enclosed in black rectangle) relative to tibial insert 1. (B) Magnified view of fatigue crack damage region. (C) The cubic spline superimposed on the fatigue crack damage image.

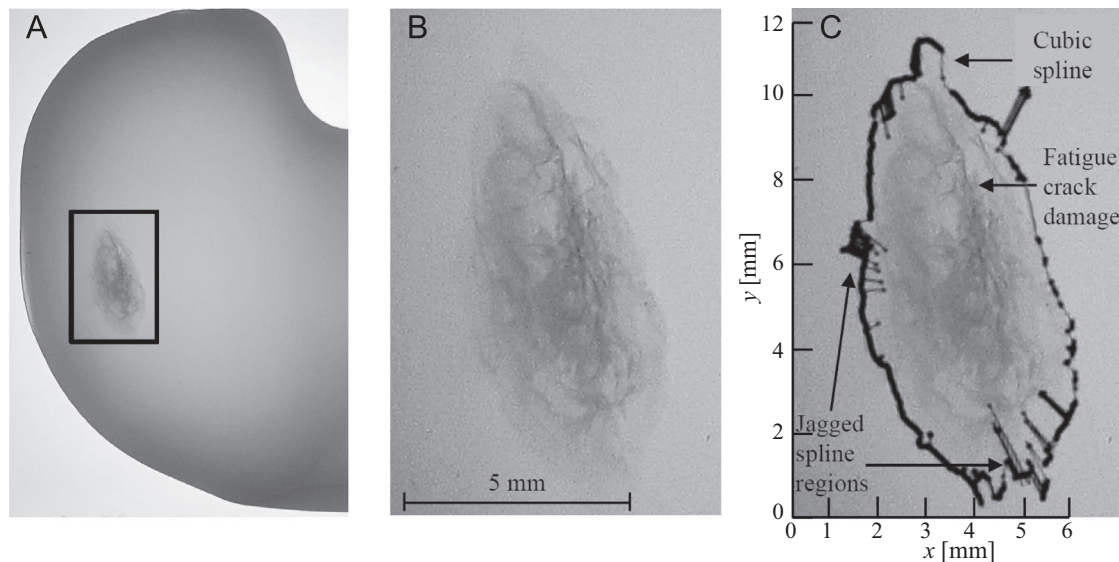


Fig. 10 – Cubic spline superimposed on the fatigue crack damage image for tibial insert 2. (A) The fatigue crack damage area (enclosed in black rectangle) relative to the tibial insert. (B) A magnified view of the fatigue crack damage region. (C) The cubic spline superimposed on the fatigue crack damage image.

mm² and 784 pixels per mm², respectively. The measurement accuracy when using this technique is dependent on three factors. First, if the fatigue crack damage plane is not parallel to the camera's image plane during image capture, the resulting area projection will underestimate the actual fatigue crack damage area. The orientation of the fatigue crack damage plane relative to the image plane in the two tibial inserts that we have tested is unknown and warrants further investigation. Second, the fatigue crack damage area measurement is dependent on accurate detection of the outermost edges of the fatigue crack damage. As demonstrated by the two tibial inserts examined in this work, distinct edges do not necessarily enclose the entire fatigue crack damage area. Fig. 9A and B show tibial insert 1, which exhibits a crescent-shaped fatigue crack damage region with an indistinct edge on the concave side of the crescent. Tibial insert 2 exhibits greater edge continuity, but still displays regions where the outer edge is not detected. This effect is visible in Fig. 10C where the spline forms a jagged edge. This lack of a distinct enclosing edge, or detection of the fatigue crack front presents potential for measurement inaccuracies, similar to other non-destructive methods. Noise within the image, which has the potential to be incorrectly categorized as damage, can also contribute to inaccuracies in the fatigue crack damage area measurement. The latter effect is observed as sections of jagged cubic spline in Figs. 9C and 10C where the spline jumps between the true fatigue crack damage edge pixels and misidentified non-damage edge pixels. This problem is reduced, but not completely eliminated, by the use of the Wiener filter during image processing. In the two specimens evaluated in this study, the area added or removed by the jagged spline regions due to noise represented only a small portion of the total enclosed area, thus the jagged spline regions minimally affected the accuracy of the technique. To quantify the effect of removing the jagged regions, the contributing edge pixels are identified and

removed. The area within the resulting piecewise cubic spline increases by 0.01 mm² for tibial insert 1 and by 0.14 mm² for tibial insert 2. This represents a 0.05% and 0.41% change in fatigue crack damage area for tibial insert 1 and tibial insert 2, respectively.

The trans-illumination technique has not yet been validated for various fatigue crack damage depths and sizes. The concern regarding the effect of depth on the visibility of fatigue crack damage is partially addressed by the tendency of fatigue crack damage to occur between 1 and 2 mm subsurface, where the shear stress resulting from Hertz contact is maximum, rather than deep within the tibial insert (Bell et al., 1996; Bartel et al., 1986). However, further validation with a variety of fatigue crack damage depths is desirable for confirming the robustness of the trans-illumination based method. Lastly, this method relies on the assumption that increased tibial insert opacity represents fatigue crack damage. However, deformation of the material may result in increases in UHMWPE opacity prior to the formation of macroscopic fatigue cracks.

Despite these limitations the trans-illumination and fatigue crack damage edge detection technique represents a simple and economical method for quantifying fatigue crack damage area and location within tibial insert specimens, with greater precision and accuracy than existing subsurface fatigue crack damage measurement methods. Note also that the method is valid for quantifying any subsurface damage, irrespective of its origin.

4. Conclusions

We have demonstrated a non-destructive, cost-effective method for quantifying fatigue crack damage in polyethylene tibial inserts, which overcomes the problems associated with existing fatigue crack damage assessment methods. This

novel fatigue crack damage measurement method is based on trans-illumination to highlight fatigue crack damage and image processing to identify and measure fatigue crack damage areas. We have successfully tested this technique on two tibial inserts.

This new technique presents a precise, objective method for measuring subsurface fatigue crack damage and improves on existing image-based techniques due to its usability for subsurface damage measurement and its reduced reliance on manual damage identification and measurement. The precision of this measurement process is within 1 mm² based on the scale bar used in capturing the images. The use of an edge detection and cubic spline-fitting algorithm to measure the fatigue crack damage area reduces the subjectivity present in other image-based wear identification methods. It is determined that this technique allows fatigue crack damage to be successfully identified and measured for the range of fatigue crack damage present in the two tested tibial inserts.

REFERENCES

- ASTM International, 2008. ASTM F2003-02. Standard Practice for Accelerated Aging of Ultra-High Molecular Weight Polyethylene After Gamma Irradiation in Air.
- Bartel, D.L., Bicknell, V.L., Wright, T.M., 1986. The effect of conformity, thickness, and material on stresses in ultra-high molecular weight components for a total joint replacement. *J. Bone Jt. Surg. Am.* 68, 1041–1051.
- Bell, C., Walker, P.S., Abeysondera, M.R., Simmons, J.M.H., King, P. M., Blunn, G.W., 1996. Effect of oxidation on the delamination of ultrahigh-molecular-weight polyethylene tibial components. *J. Arthroplast.* 13 (3), 280–290.
- Blunn, G.W., Joshi, A., Minns, R.J., Lidgren, L., Lilley, P., Ryd, L., Engelbrecht, E., Walker, P.S., 1997. Wear in retrieved condylar knee arthroplasties: a comparison of wear in different designs of 280 retrieved condylar knee prostheses. *J. Arthroplast.* 12 (3), 281–290.
- Brand, S., Czurratis, P., Hoffrogge, P., Temple, D., Malta, D., Reed, J., Petzold, M., 2011. Extending acoustic microscopy for comprehensive failure analysis applications. *J. Mater. Sci. Mater. Electron.* 22, 1580–1593.
- Canny, J., 1986. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell. (PAMI)* 8 (6), 679–698.
- Cornwall, B.G., Bryant, T.J., Hansson, C.M., Rudan, J., Kennedy, L. A., Cooke, T.D.V., 1995. A quantitative technique for reporting surface degradation of UHMWPE components of retrieved total knee replacements. *J. Appl. Biomater.* 6, 9–18.
- Furmanski, J., Kraay, M.J., Rimnac, C.M., 2011. Crack initiation in retrieved cross-linked highly cross-linked ultra-high-molecular-weight polyethylene acetabular liners: an investigation of 9 cases. *J. Arthroplast.* 26 (5), 796–801.
- Harman, M.K., Banks, S.A., Hodge, W.A., 2001. Polyethylene damage and knee kinematics after total knee arthroplasty. *Clin. Orthop.* 392, 383–393.
- Heyse, T.J., Ries, M.D., Bellemans, J., Goodman, S.B., Scott, R.D., Wright, T.M., Lipman, J.D., Schwarzkopf, R., Figgie, M.P., 2014. Total knee arthroplasty in patients with juvenile idiopathic arthritis. *Clin. Orthop.* 472 (1), 147–154.
- Hossain, F., Patel, S., Haddad, F.S., 2010. Midterm assessment of causes and results of revision total knee arthroplasty. *Clin. Orthop. Relat. Res.* 468, 1221–1228.
- ISO/IEC, 2009. ISO 14243-1: Implants for Surgery-Wear of Total Knee-joint Prostheses – Part 1: Loading and Displacement Parameters for Wear-Testing Machines with Load Control and Corresponding Environmental Conditions for Test.
- Killip, S., Mahfoud, Z., Pearce, K., 2004. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Ann. Fam. Med.* 2 (3), 204–208.
- Medel, F.J., Kurtz, S.M., Parvizi, J., Klein, G.R., Kraay, M.J., Rimnac, C.M., 2011. In vivo oxidation contributes to delamination but not pitting in polyethylene components for total knee arthroplasty. *J. Arthroplast.* 26 (5), 802–810.
- Muratoglu, O.K., Bragdon, C.R., Jasty, M., O'Connor, D.O., Von Knoch, R.S., Harris, W.H., 2004. Knee-simulator testing of conventional and cross-linked polyethylene tibial inserts. *J. Arthroplast.* 19 (7), 887–897.
- OECD, 2013. Health at a Glance 2013: OECD Indicators.
- Petzow, G., 1999. Metallographic Etching: Techniques for Metallography, Ceramography, Plastography, ASM International, United States of America.
- Rasesh, K., 2010. New developments in x-ray micro-computed tomography technology and its applications to study bone micro-architecture. *J. Acoust. Soc. Am.* 127, 2006.
- Schroer, W.C., Berend, K.R., Lombardi, A.V., Barnes, C.L., Bolognesi, M.P., Berend, M.E., Ritter, M.A., Nunley, R.M., 2013. Why are total knees failing today? Etiology of total knee revision in 2010 and 2011. *J. Arthroplast.* 28 (Suppl. 1), S116–S119.
- Simis, K.S., Bistolfi, A., Bellare, A., Pruitt, L.A., 2006. The combined effects of crosslinking and high crystallinity on the microstructural and mechanical properties of ultra high molecular weight polyethylene. *Biomaterials* 27, 1688–1694.
- Stiehl, J.B., Hamelynck, K.J., Voorhorst, P.E., 2006. International multi-centre survivorship analysis of mobile bearing total knee arthroplasty. *Int. Orthop.* 30 (3), 190–199.
- Teeter, M.G., Yuan, X., Naudie, D.D.R., Holdsworth, D.W., 2010. Technique to quantify subsurface cracks in retrieved polyethylene components using micro-CT. *J. Long Term Eff. Med. Implants* 20 (1), 27–34.
- Todo, S., Tomita, N., Kitakura, T., Yamano, Y., 1999. Effect of sliding locus on subsurface crack formation in ultra-high-molecular-weight polyethylene knee component. *Biomed. Mater. Eng.* 9, 13–20.