HHS Public Access

Author manuscript

Osteoarthritis Cartilage. Author manuscript; available in PMC 2020 January 01.

Published in final edited form as:

Osteoarthritis Cartilage. 2019 January ; 27(1): 114–117. doi:10.1016/j.joca.2018.09.005.

A Graphic User Interface for the Evaluation of Knee Osteoarthritis (GEKO): An open-source tool for histological grading

Heidi E. Kloefkorn[1], **Brittany Y. Jacobs**[1], **Danny F. Xie**[1], and **Kyle D. Allen**[1]

[1]J Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL

Abstract

Objective: In osteoarthritis (OA) models, histology is commonly used to evaluate the severity of joint damage. Unfortunately, semi-quantitative histological grading systems include some level of subjectivity, and quantitative grading systems can be tedious to implement. The objective of this work is to introduce an open source, graphic user interface (GUI) for quantitative grading of knee OA.

Methods: Inspired by the 2010 OARSI histopathology recommendations for the rat, our laboratory has developed a GUI for the evaluation of knee OA, nicknamed GEKO. In this work, descriptions of the quantitative measures acquired by GEKO are presented and measured in 42 histological images from a rat knee OA model. Using these images, across-session and within-session reproducibility for individual graders is evaluated, and inter-grader reliability across different levels of OA severity is also assessed.

Results: GEKO allowed histological images to be quantitatively scored in less than 1 min per image. In addition, intra-class coefficients (ICCs) were largely above 0.8 for across-session reproducibility, within- session reproducibility, and inter-grader reliability. These data indicate GEKO aided in the reproducibility and repeatability of quantitative OA grading across graders and grading sessions.

Conclusions: Our data demonstrate GEKO is a reliable and efficient method to calculate quantitative histological measures of knee OA in a rat model. GEKO reduced quantitative grading times relative to manual grading systems and allowed grader reproducibility and repeatability to be

Corresponding Author: Kyle D Allen, Ph.D., Associate Professor, J Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, 1275 Center Drive, Biomedical Sciences Building, JG56, Gainesville, Florida 32610, (352) 273-9337, kyle.allen@bme.ufl.edu.

Contributions

HEK and KDA conceived and designed the experiment. HEK, BYJ, and DFX wrote the software. HEK acquired the histological data. HEK and KDA analyzed the data and drafted this manuscript. All authors have edited the manuscript and approved the final submission.

Competing Interests

The authors have no competing interests to disclose.

easily assessed within a grading session and across time. Moreover, GEKO is being provided as a free, open-source tool for the OA research community.

Introduction

Preclinical models of osteoarthritis (OA) represent a critical link in the translational pipeline. In these models, OA-related damage is commonly evaluated using histological assessments, including the Mankin scheme[1] (or one of its modified versions[2–4]) and the 2006 Osteoarthritis Research Society International (OARSI) score[5]. In 2010, OARSI tasked OA experts to identify "consensus of scoring systems for the most important species used in OA animal model research[6]." Moreover, within the guiding principles of this initiative, Aigner and colleagues wrote:

> *"Clearly, there will never be a perfect scoring system fulfilling all need in all respects: but the basic requirement is simplicity such that the scoring system should be easy to follow and reproducible for sinsle observers as well as multiple observers*[7]."

These remain key goals for OA histopathology, ultimately seeking to improve robustness and repeatability of OA assessments across studies.

To build field consensus, key nomenclature were defined in the 2010 OARSI histopathology initiative's guiding principles[7]. First, "staging" was defined as an overall disease assessment, whereas "grading" was defined as assessments at a specific site or region. While grading provides relatively more detail on OA features than staging, grading is more time-consuming. Furthermore, the 2010 OARSI guiding principles defined "scoring" as a general term for semi-quantitative and quantitative evaluations, whereas "measuring" was defined as specifically evaluating an OA feature in a quantitative manner. The semiquantitative nature of staging, grading, and scoring systems includes some level of subjectivity, and thus, can be relatively difficult to replicate across experiments and labs. Moreover, OA histopathology can be tedious, leading to challenges in throughput and repeatability.

To address throughput and repeatability, our laboratory has developed a graphic user interface (GUI) for the evaluation of knee OA, nicknamed GEKO. Inspired by quantitative measures in the 2010 OARSI recommendations for the rat[8], GEKO loads a series of histological images, guides users through the measurement of several OA features, calculates measures of joint damage, and returns these quantitative measures in a comma delimited file. GEKO is introduced here, beginning with descriptions of quantitative measures acquired by our software and method of use. In addition, across-session and within- session reproducibility for individual graders using GEKO is evaluated, as well as inter-grader reliability across different levels of OA severity for both GEKO and manual grading. Our data demonstrate GEKO can reliably and efficiently calculate quantitative histological measures of rat knee OA, reducing grading times and allowing grader reproducibility and repeatability to be easily assessed within a study. Finally, in the spirit of the 2010 OARSI histopathological initiative, GEKO is being provided as a free, open-source method for the OA research community. An executable program and MATLAB-based scripts are available at https://www.orthobme.com/resources.html.

Methods

A GUI for the Evaluation of Knee OA (GEKO)

GEKO is a MATLAB-based program designed to help graders measure histological features of knee OA. GEKO was inspired by the 2010 OARSI histopathology recommendations for the rat[8]; as such, GEKO is specifically designed to grade geometric changes in frontal plane histological images of rat knee OA. While GEKO could be applied to frontal plane histological images from other species, GEKO users should note that, to comply with the 2010 OARSI histopathology recommendations, histopathology scoring for the specific species used should also be reported, thereby allowing for comparisons between studies. In this way, GEKO serves as a supplement, but not replacement, of current OARSI histopathology recommendations for species other than rats.

GEKO loads a series of histological images, presents images one at a time, and provides instructions to assist the grader in identifying specific OA features (Supplemental Figure 1). Please note, while toluidine blue is typically used for histopathology in our lab, GEKO can used for any stain that allows the features in Supplemental Figure 1 to be identified. Within GEKO, the grader marks 6 features per image, including tibial plateau width, medial synovial capsule thickness, osteochondral interface, affected cartilage surface width, lost cartilage, and osteophyte diameter. These marks are then used to calculate quantitative measures of knee OA, including surface, middle, and deep cartilage matrix loss widths; total cartilage degeneration width; osteophyte size; and, joint capsule thickness (Supplemental Figure 2 and Supplemental Table 1).

Histological Images of Rat Knee OA

To evaluate GEKO, images of post-traumatic knee OA in the rat were acquired from a past experiment [9]. All prior methods and testing were performed with University of Florida Institutional Animal Care and Use Committee (IACUC) approval; no additional animals were used for this study.

In this prior work[9], post-traumatic OA was modeled in 250 g male Lewis rats by surgically transecting the medial collateral ligament and medial meniscus (MCLT+MMT). Sham surgery consisted of medial collateral ligament transection alone, while naïve animals received no surgical manipulation. Animals were euthanized at 1, 2, 4, and 6 weeks post-operation. For the evaluation of GEKO, a single section representing evidence of knee OA was selected for grading from 42 animals (6 MMT and 3 sham per time point, 6 naïve control animals); the set of histological images was selected to provide a range of OA severity. Please note, the purpose of this study was to evaluate grader reproducibility using GEKO, not to evaluate histological differences between groups; differences between MCLT +MMT, sham, and naïve animals have been previously reported[9].

Grading Reproducibility

Four blinded graders independently evaluated histological images in four separate GEKO grading sessions, with each grading session separated by one week. While grading, graders did not communicate with other graders.

In each grading session, graders were presented 48 randomized histological images (42 unique histological images, plus two replicates of three images from OA-affected knees). Prior to grading, the image set was independently randomized for each grader, with the criteria that repeated images be separated by at least 1 different image. Each week, the set of repeated images was changed. As a follow- up experiment, three graders evaluated a set of 48 images both manually and using GEKO.

To assess within-session reproducibility, repeated image grades (n=3 per grader with 3 repeated measures) were used to calculate alpha model within-session intra-class correlation coefficients (ICCs) with two-way random compensation (SPSS). Similarly, alpha model ICCs with two-way random compensation were used to assess across-session reproducibility for each grader and inter-grader reliability within a session. Finally, alpha model ICCs with two-way random compensation were used to evaluate inter-grader reliability for manual and GEKO measures. All reported ICCs represent average consistency agreement. Statistical significance between manual and GEKO ICCs was determined using Student's t-test (paired, two-tailed).

Grading Time

To assess grading time, GEKO tracked the time spent on each image. GEKO grading times in the first session were then compared to manual grading times for one experienced grader using the same 48-image set. Statistical significance was determined using Student's t-test.

Results

GEKO reduced session grading time from 377±193 seconds per image to 48±19 seconds per image ($p<0.0001$, Student's t-test).

Average within-session ICCs were above 0.85, with average within-session ICCs for surface cartilage matrix loss width, deep cartilage matrix loss width, total cartilage degeneration width, osteophyte size, and joint capsule thickness above 0.9 (Table 1). Average across-session ICCs dropped slightly, but remained above 0.75 (Table 1). Deep cartilage matrix loss width was the least reproducible, with across- session ICCs ranging from 0.714 to 0.811.

GEKO and manual grading had inter-grader ICCs were above 0.9 for tibial plateau width, total cartilage degeneration width, osteophyte size, and joint capsule thickness (Table 1). However, manual inter-grader ICCs were higher for surface and deep cartilage matrix loss width, while GEKO inter-grader ICCs were higher for middle depth cartilage matrix loss width ($p<0.05$). GEKO inter-grader ICCs were above 0.7 for all measures except deep cartilage matrix loss width, while manual inter-grader ICCs were above 0.7 for all measures except middle cartilage matrix loss width.

GEKO and manual measures did not statistically differ for any measure (Table 2, Student's t-test).

Discussion

GEKO markedly reduced grading times, achieved reasonably high inter-grader ICCs, and enabled testing of within-session and across-session reproducibility. In particular, measuring within-session and across- session reproducibility may allow assessment of unknown sources of error, such as grader skill and fatigue.

GEKO was inspired by the 2010 OARSI recommendations for the rat[8], which focused on grading focal medial tibial plateau damage in post-traumatic OA models. GEKO can be extended to grading lateral compartment tibial cartilage, though reproducibility of those grades have not been assessed (our histological images lacked lateral compartment damage). Similarly, GEKO principles could also be extended to femoral cartilage or sagittal sections; however, the code would need to be updated to account for a rounded osteochondral interface. As OA histopathology assessments evolve, we plan to expand GEKO to include other assessments, like femoral cartilage damage and sagittal section grading.

Other software is available for histological grading; however, GEKO is designed to fill a niche for preclinical OA models. For example, ImageJ and FIJI are free and offer tools capable of collecting GEKO-like measures, but these packages require some data transcription and calculation after image analysis. Commercial software, such as OsteoMetrics, OsteoMeasure, and Bioquant Osteom, offer more detailed image assessments, but these products are neither free nor open source. As such, GEKO aims to make rapid, quantitative histological OA grading broadly available to the OA research community.

A previous publication reports manual inter-grader ICCs for rat knee OA[8]. In that study, all cartilage matrix loss widths, total cartilage degeneration width, and osteophyte size produced inter-grader ICCs above 0.9. Our manual inter-grader ICCs were comparable for all measures except middle depth cartilage matrix loss width, and our GEKO inter-grader ICCs were comparable for all measures except of deep cartilage matrix loss width. Moreover, direct comparison of manual and GEKO grading show higher GEKO inter-grader ICCs for middle depth cartilage matrix loss width, and higher manual inter-grader ICCs for surface and deep cartilage matrix loss width.

Low GEKO inter-grader ICCs for deep cartilage may be due to low variance in the parameter. In GEKO, deep cartilage is defined as the bottom 8% of cartilage depth. Because lesion width is small at this depth, missing by a few pixels can have a relatively large effect on the measured ICC (see large 95% confidence interval in Table 1). Also, GEKO has strict rules for calculating deep cartilage matrix loss width, while manual graders tend to measure this width at the bottom of the lesion regardless of lesion depth. While GEKO's approach may be less biased, it may also be less consistent.

Table 2 demonstrates some interesting trends on how graders evaluate histological slides during manual and GEKO grading. In GEKO, graders outline the lesion; then, lesion traces are mathematically converted into surface, middle, and deep cartilage matrix loss width (Supplemental Figure 2). While not statistically significant, surface and deep cartilage matrix loss width tends to be lower in GEKO, while middle depth cartilage matrix loss tends to be higher. Inspection of graded images indicated lesion traces in GEKO tended to start

and stop at the tips of fibrillated cartilage; during manual grading, graders tended to measure loss widths from the bottom of fibrillated cartilage. Also, GEKO determines the depth of middle and deep cartilage mathematically; in manual grading, these locations are determined visually. For middle depth cartilage, this may have resulted in some inconsistency during manual grading. For deep cartilage, manual graders tended to measure the deep cartilage matrix loss width at the bottom of the lesion, regardless of depth. This may have been consistent, but not necessarily accurate.

GEKO can be expanded to yield additional measures. For example, our group recently published quantitative subchondral bone and subintima measures, which we aim to add to GEKO[10]. In addition, a better approach to cartilage measures may be continuously defining the relationships between the cartilage surface, osteochondral interface, and potentially the tidemark, allowing for new measures of cartilage thickening and the spatial location and orientation of cartilage changes.

In conclusion, GEKO reduced overall grading time for histological images of knee OA. In addition, repeatability controls were easily introduced during grading. These controls allow for a more thorough exploration of grader variability. Overall, GEKO is a robust tool to improve quantitative histological grading.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) of the National Institutes of Health under award numbers K99/R00AR057426 and R01AR071431. The histological grading data could not have been collected without contributions from Emily Lakes, Ph.D. and Yash Shah. Their contributions are greatly appreciated.

References

1. Mankin HJ, Dorfman H, Lippiello L, Zarins A. Biochemical and metabolic abnormalities in articular cartilage from osteo-arthritic human hips. II. Correlation of morphology with biochemical and metabolic data. Journal of Bone and Joint Surgery 1971; 53: 523–537. [PubMed: 5580011]
2. Neo H, Ishimaru JI, Kurita K, Goss AN. The effect of hyaluronic acid on experimental temporomandibular joint osteoarthrosis in the sheep. Journal of Oral and Maxillofacial Surgery, 1997; 55: 1114–1119. [PubMed: 9331236]
3. Hayami T, Pickarski M, Wesolowski GA, McLane J, Bone A, Destefano J, et al. The role of subchondral bone remodeling in osteoarthritis: reduction of cartilage degeneration and prevention of osteophyte formation by alendronate in the rat anterior cruciate ligament transection model. Arthritis & Rheumatism 2004; 50: 1193–1206. [PubMed: 15077302]
4. Furman BD, Strand J, Hembree WC, Ward BD, Guilak F, Olson SA. Joint degeneration following closed intraarticular fracture in the mouse knee: a model of posttraumatic arthritis. Journal of Orthopedic Research 2007; 25: 578–592.
5. Pritzker KP, Gay S, Jimenez SA, Ostergaard K, Pelletier JP, Revell PA, et al. Osteoarthritis cartilage histopathology: grading and staging. Osteoarthritis & Cartilage 2006; 14: 13–29. [PubMed: 16242352]
6. Berenbaum F The OARSI histopathology initiative - the tasks and limitations. Osteoarthritis & Cartilage 2010; 18 Suppl 3: S1.

7. Aigner T, Cook JL, Gerwin N, Glasson SS, Laverty S, Little CB, et al. Histopathology atlas of animal model systems - overview of guiding principles. Osteoarthritis & Cartilage 2010; 18 Suppl 3: S2–6 [PubMed: 20864020]

8. Gerwin N, Bendele AM, Glasson S, Carlson CS. The OARSI histopathology initiative - recommendations for histological assessments of osteoarthritis in the rat. Osteoarthritis & Cartilage 2010; 18 Suppl 3: S24–34.

9. Kloefkorn HE, Jacobs BY, Loye AM, Allen KD. Spatiotemporal gait compensations following medial collateral ligament and medial meniscus injury in the rat: correlating gait patterns to joint damage. Arthritis Research & Therapy 2015; 17: 287. [PubMed: 26462474]

10. Kloefkorn HE, Allen KD. Histological changes in the subchondral bone and synovium correlate to rodent behavior in a model of post-traumatic knee OA. Connective Tissue Research 2017; 58:373–385. [PubMed: 27797605]

Table 1:

Within-session and Across-session ICCs as Measured via GEKO, and Inter-grader ICCs as Measured via GEKO and Manual Grading.

Data are presented as mean +/− 95% confidence intervals. A * indicates a significant difference between manual and GEKO grading ($p<0.05$). Graders 2 and 4 were unavailable for manual grading due to a change of employment; Grader 5 was added to assure a minimum of three graders for this follow-up experiment.

	Within-Session Reproducibility ICCs (Images repeated within a session)					**Across-Session Reproducibility ICCs (42 image set)**					**Inter-Grader Reliability ICCs (42 image set)**	
	Grader				**All Graders Mean (95% CI)**	**Grader**				**All Graders Mean (95% CI)**	**GEKO (Graders 1, 3, & 5)**	**Manual (Graders 1, 3, & 5)**
Histological Gradess	**1**	**2**	**3**	**4**		**1**	**2**	**3**	**4**			
Tibial Plateau Width (μm)	**0.907**	0.814	**0.900**	0.826	0.862 (0.785–0.939)	**0.927**	0.874	**0.949**	0.845	0.899 (0.823–0.975)	***0.916 (0.859–0.952)***	0.873 (0.736–0.910)
Surface Cartilage Matrix Loss Width (μm)	**0.982**	0.844	0.864	**0.983**	***0.918 (0.799–1.000)***	**0.944**	0.845	0.676	**0.958**	0.856 (0.649–1.000)	0.820 (0.654–0.900)	***0.966 (0.944–0.981)****
Middle Cartilage Matrix Loss Width (μm)	**0.965**	0.714	0.893	**0.995**	0.892 (0.692–1.000)	**0.958**	**0.931**	0.643	**0.972**	0.876 (0.627–1.000)	***0.944 (0.905–0.969)****	0.645 (0.405–0.799)
Deep Cartilage Matrix Loss Width (μm)	**0.943**	NV	0.862	**0.992**	***0.932 (0.769–1.000)***	0.714	0.760	0.733	0.811	**0.755 (0.687–0.822)**	0.535 (0.208–0.740)	0.867 (0.777–0.924)*
Total Cartilage Degeneration Width (% of Tibial Plateau)	**0.988**	**0.995**	**0.996**	**0.947**	***0.982 (0.944–1.000)***	**0.965**	**0.995**	**0.994**	0.849	***0.951 (0.841–1.000)***	***0.983 (0.971–0.990)***	***0.977 (0.965–0.988)***
Osteophyte Size (μm)	**0.995**	**0.998**	**0.961**	**0.986**	***0.985 (0.958–1.000)***	**0.985**	**0.995**	**0.984**	**0.981**	***0.986 (0.977–0.996)***	***0.971 (0.952–0.984)***	***0.982 (0.977–0.990)***
Medial Joint Capsule Repair (μm)	**0.957**	**0.976**	**0.986**	**0.965**	***0.971 (0.951–0.991)***	**0.944**	**0.971**	**0.977**	**0.963**	***0.964 (0.941–0.987)***	***0.939 (0.899–0.965)***	***0.985 (0.825–0.940)***

Table 2:

Histological Grades Using Manual Grading and GEKO.

Data are presented as mean +/− 95% confidence intervals.

Histological Grades		Naive	Sham	MMT – Week 1	MMT – Week 2	MMT – Week 4	MMT – Week 6
Tibial Plateau Width (μm)	*Manual*	2298 (2175–2421)	2409 (2304–2514)	2418 (2296–2540)	2417 (2276–2557)	2609 (2464–2753)	2602 (2405- 2798)
	GEKO	2310 (2194–2425)	2401 (2292–2522)	2417 (2274–2559)	2429 (2278–2580)	2700 (2540–2859)	2590 (2379- 2801)
Surface Cartilage Matrix Loss Width (μm)	*Manual*	0 (0–0)	109 (0–338)	744 (560–928)	789 (376–1202)	1041 (832–1250)	987 (649–1326)
	GEKO	0 (0–0)	78 (0–241)	467 (319–615)	360 (167–553)	780 (566–994)	905 (417–1393)
Middle Cartilage Matrix Loss Width (μm)	*Manual*	0 (0–0)	40 (0–125)	128 (59–196)	133 (25–242)	183 (11–356)	208 (75–342)
	GEKO	0 (0–0)	93 (0–289)	336 (166–506)	477 (111–842)	426 (204–647)	472 (145–800)
Deep Cartilage Matrix Loss Width (μm)	*Manual*	0 (0–0)	37 (0–115)	75 (30–120)	69 (7–132)	103 (0–217)	101 (0–209)
	GEKO	0 (0–0)	16 (0–48)	13 (0–43)	66 (0–147)	19 (0–48)	16 (0–33)
Total Cartilage Degeneration Width (% of Tibial Plateau)	*Manual*	0 (0–1)	5 (0–15)	54 (47–61)	49 (29–70)	55 (50–59)	54 (48–61)
	GEKO	0 (0–0)	5 (0–15)	47 (41–53)	44 (26–63)	50 (47–54)	48 (36–61)
Osteophyte Size (μm)	*Manual*	0 (0–0)	53 (0–163)	0 (0–0)	152 (0–326)	538 (376–700)	480 (307–652)
	GEKO	0 (0–0)	55 (0–156)	0 (0–0)	183 (41–324)	512 (368–656)	481 (364–598)
Medial Joint Capsule Repair (μm)	*Manual*	386 (351–420)	501 (407–595)	688 (523–854)	519 (353–686)	517 (424–609)	557 (351–762)
	GEKO	383 (322–445)	507(407–606)	713 (551–876)	552 (365–739)	535 (410–659)	529 (363–694)

Osteoarthritis Cartilage. Author manuscript; available in PMC 2020 January 01.