Radiographic Stage Does Not Correlate With Symptom Severity in Thumb Basilar Joint Osteoarthritis

C. Edward Hoffler II, MD, PhD

Jonas L. Matzon, MD

Kevin F. Lutsky, MD

Nayoung Kim, BS

Pedro K. Beredjiklian, MD

Abstract

Background: We hypothesize that thumb basilar joint osteoarthritis (TBJA) radiographic stage does not correlate with patient-reported measures of symptom severity.

Methods: Patients with unilateral TBJA who completed the 11-item QuickDASH (Disabilities of the Arm, Shoulder, and Hand), Short-Form 12 Health Survey (SF-12) Mental Component and SF-12 Physical Component surveys were prospectively enrolled in the study. The Eaton-Littler radiographic stage was assigned for each patient. The correlation between the radiographic score and disease stage was calculated.

Results: Sixty-two patients (15 men, 47 women; average age, 62.3 years) formed the basis of this study. The average QuickDASH score (and standard deviation) for patients with stage 1 TBJA was 31.5 (11.4); for those with stage 2, it was 37.9 (17.4); with stage 3, it was 30.1 (13.0), and with stage 4, it was 39.4 (12.5). Eaton-Littler stage did not correlate significantly with QuickDASH scores (rho = -0.014, $P = 0.91$). Neither SF-12 Mental Component scores (MCS-12: rho = 0.019, $P = 0.89$) nor the SF-12 Physical Component scores (PCS-12: rho = 0.145, $P = 0.26$) correlated with TBJA stage.

Conclusion: Radiographic severity in TBJA does not correlate with validated patient-reported symptom scores. Metrics that link radiographic and subjective components of TBJA may improve surgical decision making and monitoring of treatment response.

Level of Evidence: Prognostic, level II

Thumb basilar joint osteoarthritis (TBJA) is a common problem affecting nearly 7% of men and 15% of women aged >30 years.[1] Symptoms vary in severity, and patients range from those who are asymptomatic to those with unremitting pain and significant disability. Weakness, loss of motion, and deformity are common. Radiographic severity of TBJA is assessed most commonly based on the staging system described by Eaton et al,[2] with higher stages indicating more advanced osteoarthritis. Previous studies have demonstrated that non-surgical treatment can be successful in improving symptoms even in patients with more advanced stages of arthritis.[3-5]

Conventional wisdom accepts that the radiographic stage of TBJA does not correlate with patient-reported symptoms; however, few if any objective data support this belief.[6,7] We evaluated the extent to which radiographic stage of TBJA correlates to patient symptoms. We hypothesize that radiographic stage

of TBJA does not correlate with patient-reported symptom severity as measured by validated outcome surveys.

Methods

Our Institutional Review Board approved this cross-sectional study protocol. We prospectively enrolled consecutive patients with a diagnosis of unilateral thumb TBJA who presented to the clinics of four fellowship-trained orthopaedic hand surgeons between June 2013 and March 2014. The diagnosis was based on clinical history, physical examination, and radiographic evaluation. Patients with bilateral disease, concomitant ipsilateral upper extremity diagnoses, or unwillingness to participate were excluded from the study.

Patient age, gender, and hand dominance were recorded. All patients completed self-reported health and disability scores, including the Quick Disabilities of the Arm, Shoulder, and Hand 11-item questionnaire (Quick-DASH) and the Medical Outcomes Study 12-Item Short-Form Survey (SF-12). The QuickDASH is scored from 0 to 100, with a higher score indicating greater disability. The SF-12 has a Physical Component score (PCS-12) and a Mental Component score (MCS-12). Both the PCS-12 and MCS-12 are scored from 0 to 100, with higher scores indicating better function. Both scores are structured to a national norm with a mean score of 50.0 and a standard deviation of 10.0.[8]

The Eaton-Littler radiographic stage[2] of osteoarthritis for each patient was determined by the treating physicians, who were blinded to the results of the questionnaires. The reliability of the staging system between the four treating hand surgeons was assessed with an interobserver agreement analysis of a subset of 10 randomly selected sets of radiographs. The intraclass correlation coefficient (ICC) ([ICC(3,1)]) and the Fleiss kappa statistic were used to evaluate agreement.

The Spearman rank analysis was used to correlate arthritis stage with demographic data and survey outcomes. A power analysis was performed to determine whether our samples were sufficient to detect a clinically meaningful correlation. Logistic regression was used to evaluate gender and Eaton-Littler stage. A two-way analysis of variance (ANOVA) with Eaton-Littler stage and dominant-hand involvement was used to determine whether patients with symptomatic dominant hands reported themselves as being less healthy. Significance was attributed to P values of $<5\%$. A post hoc power analysis indicated that our sample has 80% power to detect a correlation of rho = 0.31. A clinically significant correlation should have a rho of at least 0.3.[9]

Results

Sixty-two patients who presented for treatment of unilateral thumb TBJA formed the basis of this study. There were 15 men and 47 women, with an average age 62.3 years (range, 32 to 81 years). Based on radiographic stage, there were 12.9% stage 1 (n = 8), 32.3% stage 2 (n = 20), 45.2% stage 3 (n = 28), and 9.7% stage 4 (n = 6) patients. Age correlated weakly with Eaton-Littler stage (rho = 0.27; P = 0.03), whereas gender (P = 0.22) did not correlate with stage of disease. The dominant hand was involved in 25 patients (40.3%). Dominant-hand involvement did not significantly affect stage (P = 0.94) or QuickDASH (P = 0.61), PCS-12 (P = 0.73), or MCS-12 (P = 0.93) scores. Interobserver agreement for arthritis staging was significant with ICC 0.80 (95% confidence interval [CI], 0.63 to 0.92; P = $<$0.001). The Fleiss kappa statistic was 0.73 (95% CI, 0.54 to 0.93; P = 0.001).

The average QuickDASH, PCS-12, and MCS-12 scores, and their standard deviations are shown in Table 1. Eaton-Littler radiographic stage[2] did not correlate significantly with QuickDASH (rho = $-$0.014; P = 0.91), PCS-12 (rho = 0.145; P = 0.26), MCS-12 (rho = 0.019, P = 0.89) survey scores.

Discussion

This study assesses the relationship between the severity of radiographic TBJA and patient-reported symptoms. We present prospective evidence that radiographic stage does not correlate with patient-reported symptom severity. Eaton-Littler radiographic stages[2] did not correlate with QuickDASH or with the PCS-12 and MCS-12 components of the SF-12.

The QuickDASH provides a composite assessment of global upper extremity function and symptoms. Questions explicitly explore the patient's pain, neurologic symptoms, and sleep disturbance. The minimal important difference (MID) for the QuickDASH is 19 points.[10] The MID refers to a minimum change in score that indicates a meaningful change in health status. As such, the metric is most appropriate to longitudinal comparisons of the same patient or patients. However, the maximum difference between scores in this study, 7.9 points, compared with an MID of 19 points contextualizes the magnitude of the differences between stages and suggests that, in a longitudinal study, this magnitude would not be clinically meaningful.

The PCS-12 has been studied in patients with various hand and wrist diagnoses, including carpal tunnel

Table 1

Patient-reported Score Data for Thumb Basilar Joint Osteoarthritis Radiographic Stages[a]

Survey	Eaton-Littler Radiographic Stage[2] Mean Test Score (SD)				Spearman Rank Coefficient Rho	P Value
	1[b]	2[c]	3[d]	4[e]		
QuickDASH	31.5 (11.4)	37.9 (17.7)	30.1 (13.0)	39.4 (12.5)	−0.014	0.914
PCS-12	40.7 (9.3)	40.3 (8.7)	43.6 (7.9)	43.9 (4.9)	0.145	0.261
MCS-12	58.3 (6.8)	52.0 (10.8)	53.5 (9.6)	59.1 (2.4)	0.019	0.886

[a] No significant correlations were identified.
[b] Sample size = 8 (13%)
[c] Sample size = 20 (32%)
[d] Sample size = 28 (45%)
[e] Sample size = 6 (10%)
MCS-12 = Short-Form 12 Health Survey Mental Component, PCS-12 = Short-Form 12 Health Survey Physical Component, QuickDASH = 11-item Disabilities of the Arm, Shoulder, and Hand Survey

syndrome and wrist fractures. In this setting, the PCS-12 has an MID of 7.3 points.[11] Again, the MID provides a context for the maximum difference between stages of 3.6 points in this study. The difference would not be clinically meaningful in a longitudinal study.

The MCS-12 scores did not differ statistically between radiographic stages, which likely indicates there are not clinically significant differences, either. Schmitt et al[11] attempted to quantify MCS-12 responsiveness with respect to upper extremity diagnoses. During this 3-month study, PCS-12 and DASH demonstrated improving health and declining disability, while the MCS-12 score remained unchanged. This suggests the MCS-12 may not be sensitive to changes in upper extremity symptoms in individual patients. However, Hall et al[12] found MCS-12 to correlate with grip strength in a comparison between patients. These data collectively suggest that Eaton-Littler radiographic stages do not predict clinically significant differences that can be detected by the MCS-12 score.

The poor correlation between patient-reported symptoms and radiographic staging may be attributable to poor reliability of the staging system and a lack of agreement between observers. The interrater reliability of Eaton-Littler radiographic stage has been reported as moderate by most authors, with kappa ranging from 0.497 to 0.56,[13-16] but Hansen et al[17] found it to be poor with kappa = 0.17. Most recently, interrater reliability has been quantified with kappa 0.53 to 0.54 and an intraclass correlation of 0.56.[18] In our study, there was substantial agreement between the four hand surgeons with kappa = 0.7318 and ICC = 0.80. Whereas the agreement between observers in this study is higher than the values reported in the literature, the lower end of our CIs falls within range of literature values. Therefore, while our reliability values are higher, literature values are within our measurement range.[19]

The absence of a correlation between Eaton-Littler radiographic stage and the QuickDASH, PCS-12, and MCS-12 scores suggests two possibilities. First, Eaton-Littler staging is based on radiographic features, which are not predictive of patient-reported symptom severity. Trabecular connectivity, trabecular number,[20] dorsal subluxation,[21] and trapezial inclination[22] have been reported to change with TBJA. Some authors suggest that specific radiographic views may improve staging.[20] Also, different radiographic views may allow measurement of more relevant features of the osteoarthritic joint.[22]

An alternative conclusion is that these validated questionnaires are inadequate to assess patient-reported thumb basilar joint function and symptoms. Recently, investigators have designed a thumb basilar joint–specific self-administered questionnaire called the Nelson score.[23] Results correlated strongly with visual analog pain scores, had high test-retest reliability, and trended toward being more responsive than DASH scores in patients treated with thumb basilar joint surgery. MID and minimal detectable change were not reported. The Brief Michigan Hand Questionnaire (Brief MHQ) is a 12-point abbreviated survey tailored to distal radius fractures, carpal tunnel syndrome, rheumatoid arthritis, and TBJA.[24] Its sensitivity, as defined by the standardized response mean, was better than that of the original MHQ, although TBJA had the lowest sensitivity of the four diagnoses. Whether these newer surveys correlate with Eaton-Littler radiographic staging remains unknown.

Similar to the results of patient-reported disability, gender and hand dominance did not correlate with TBJA stage. However, age

correlated weakly with Eaton-Littler stage, suggesting that radiographic progression has a small component dictated by aging and senescence. Haara et al[1] have demonstrated an increasing prevalence of radiographic TBJA as people age. Nearly 27% of men and 37% of women aged >75 years had radiographic evidence of thumb base osteoarthritis. Similar to the results in the current study, the authors noted an increasing prevalence of more severe osteoarthritis (ie, Kellgren-Lawrence grades 3 and 4) with age.

This study is limited by the sample size and subgroup size. In particular, we have a small number of stage 1 and stage 4 patients, which limits our effective sample size. However, our data have adequate power to detect any correlation with rho of ≥0.31, that is, moderate to strong correlations.[9] A clinically significant correlation should have a rho of at least 0.3.[9] Any undetected correlation is very weak and likely not clinically significant.

Another limitation is that none of our survey metrics are specific to thumb function. Also, we did not exclude patients with concomitant upper extremity osteoarthritis that may not be discussed in a traditional hand surgery clinic. For example, a patient with significant osteoarthritis of the glenohumeral joint may find that his or her shoulder limits upper extremity function as much or more than TBJA does. However, the patient may not mention it specifically to a hand surgeon.

Summary

Eaton-Littler radiographic stage does not correlate with patient-reported symptom severity as measured by the QuickDASH and SF-12. Alternate radiographic characterizations of TBJA may allow us to correlate anatomic changes with symptom severity. Validated metrics that link radiographic and subjective components of TBJA may improve surgical decision making and monitoring of treatment response. Patients commonly want to know when looking at their radiographs how "bad" their arthritis is; based on the results of our study, the answer appears to be, "It's not how it looks but how it feels."

References

Evidence-based Medicine: Levels of evidence are described in the table of contents. In this article, reference 1 is a level II study.

References printed in **bold type** are those published within the past 5 years.

1. Haara MM, Heliövaara M, Kröger H, et al: Osteoarthritis in the carpometacarpal joint of the thumb. Prevalence and associations with disability and mortality. *J Bone Joint Surg Am* 2004;86(7):1452-1457.

2. Eaton RG, Lane LB, Littler JW, Keyser JJ: Ligament reconstruction for the painful thumb carpometacarpal joint: A long-term assessment. *J Hand Surg Am* 1984;9(5):692-699.

3. **Becker SJ, Bot AG, Curley SE, Jupiter JB, Ring D: A prospective randomized comparison of neoprene vs thermoplast hand-based thumb spica splinting for trapeziometacarpal arthrosis. *Osteoarthritis Cartilage* 2013;21(5):668-675.**

4. Heyworth BE, Lee JH, Kim PD, Lipton CB, Strauch RJ, Rosenwasser MP: Hylan versus corticosteroid versus placebo for treatment of basal joint arthritis: A prospective, randomized, double-blinded clinical trial. *J Hand Surg Am* 2008;33(1):40-48.

5. Swindells MG, Logan AJ, Armstrong DJ, Chan P, Burke FD, Lindau TR: The benefit of radiologically-guided steroid injections for trapeziometacarpal osteoarthritis. *Ann R Coll Surg Engl* 2010;92(8):680-684.

6. Barron OA, Glickel SZ, Eaton RG: Basal joint arthritis of the thumb. *J Am Acad Orthop Surg* 2000;8(5):314-323.

7. Van Heest AE, Kallemeier P: Thumb carpal metacarpal arthritis. *J Am Acad Orthop Surg* 2008;16(3):140-151.

8. Ware JE Jr, Kosinski M, Turner-Bowker DM, Sundaram M, Gandek B, Maruish ME. *User's Manual for the SF-12v2 Health Survey Second Edition.* QualityMetric, Lincoln, RI2009.

9. Cohen J: A power primer. *Psychol Bull* 1992;112(1):155-159.

10. Polson K, Reid D, McNair PJ, Larmer P: Responsiveness, minimal importance difference and minimal detectable change scores of the shortened disability arm shoulder hand (QuickDASH) questionnaire. *Man Ther* 2010;15(4):404-407.

11. Schmitt JS, Di Fabio RP: Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *J Clin Epidemiol* 2004;57(10):1008-1018.

12. **Hall SA, Chiu GR, Williams RE, Clark RV, Araujo AB: Physical function and health-related quality-of-life in a population-based sample. *Aging Male* 2011;14(2):119-126.**

13. **Berger AJ, Momeni A, Ladd AL: Intra- and interobserver reliability of the Eaton classification for trapeziometacarpal arthritis: A systematic review. *Clin Orthop Relat Res* 2014;472(4):1155-1159.**

14. Kubik NJ III, Lubahn JD: Intrarater and interrater reliability of the Eaton classification of basal joint arthritis. *J Hand Surg Am* 2002;27(5):882-885.

15. Dela Rosa TL, Vance MC, Stern PJ: Radiographic optimization of the Eaton classification. *J Hand Surg Br* 2004;29(2):173-177.

16. **Spaans AJ, van Laarhoven CM, Schuurman AH, van Minnen LP: Interobserver agreement of the Eaton-Littler classification system and treatment strategy of thumb carpometacarpal joint osteoarthritis. *J Hand Surg Am* 2011;36(9):1467-1470.**

17. **Hansen TB, Sørensen OG, Kirkeby L, Homilius M, Amstrup AL: Computed tomography improves intra-observer reliability, but not the inter-observer reliability of the Eaton-Glickel classification. *J Hand Surg Eur Vol* 2013;38(2):187-191.**

18. **Choa RM, Giele HP: Inter- and intrarater reliability of osteoarthritis classification at the trapeziometacarpal joint. *J Hand Surg Am* 2015;40(1):23-26.**

19. Pohl E, Gerokostopoulos A, Guo H: Determining the Right Sample Size for Your Test: Theory and Application. In: *Reliability and Maintainability Symposium, 2013 Proceedings Annual.* New York, NY, Institute of Electrical and Electronics Engineers, 2013, pp 1-12.

20. **Lee AT, Williams AA, Lee J, Cheng R, Lindsey DP, Ladd AL: Trapezium trabecular morphology in carpometacarpal arthritis. *J Hand Surg Am* 2013;38(2):309-315.**

21. **Kurosawa K, Tsuchiya I, Takagishi K: Trapezial-metacarpal joint arthritis:**

Radiographic correlation between first metacarpal articular tilt and dorsal subluxation. *J Hand Surg Am* 2013;38(2):302-308.

22. Ladd AL, Messana JM, Berger AJ, Weiss AP: Correlation of clinical disease severity to radiographic thumb osteoarthritis index. *J Hand Surg Am* 2015;40(3):474-482.

23. Bettinger PC, Linscheid RL, Cooney WP III, An KN: Trapezial tilt: A radiographic correlation with advanced trapeziometacarpal joint arthritis. *J Hand Surg Am* 2001;26(4):692-697.

24. Waljee JF, Kim HM, Burns PB, Chung KC: Development of a brief, 12-item version of the Michigan Hand Questionnaire. *Plast Reconstr Surg* 2011;128(1):208-220.