Longitudinal Stability of Common TMJ Structural Disorders

E.L. Schiffman[1], M. Ahmad[1], L. Hollender[2], K. Kartha[3,4], R. Ohrbach[4], E.L. Truelove[2], L. Zhang[5], J.S. Hodges[6], E. Sommers[2], G.C. Anderson[7], Y.M. Gonzalez[4], X. Guo[8], and J.O. Look[1]

Abstract

The longitudinal course of temporomandibular joint (TMJ) disc displacement (DD) and degenerative joint disease (DJD) has never been conclusively described with magnetic resonance imaging and computed tomography, respectively. This 8-y observational study's objective was to assess the longitudinal stability of DD and DJD among 401 subjects. The Validation Project provided baseline measures; follow-up was performed in the TMJ Impact Project. With magnetic resonance imaging, 2 radiologists rendered a consensus diagnosis of normal/indeterminate, DD with reduction, or DD without reduction. Computed tomography consensus diagnoses included normal/indeterminate, grade 1 DJD, or grade 2 DJD. Radiologist reliability was assessed by kappa; a Hui-Walter model was used to estimate, after accounting for diagnostic disagreement, the frequency of diagnostic progression and reversal. Permutation tests were used to test the statistical influence of concurrent baseline diagnoses on diagnostic changes at follow-up. Of 789 baseline joint-specific soft tissue diagnoses of DD, 598 (76%) joints showed no change; 109 (14%) demonstrated progression; and 82 (10%) had reversal. Of 794 joints with baseline joint-specific hard tissue diagnoses of DJD, progression was observed in 122 (15%) joints, no change in 564 (71%), and reversal in 108 (14%). Radiologist reliability (kappa) was 0.73 (95% CI, 0.64 to 0.83) for DD and 0.76 (95% CI, 0.68 to 0.83) for DJD. After accounting for the influence of diagnostic disagreement, progression of hard tissue diagnoses in the right TMJ occurred in 15.2% of subjects (95% CI, 10.5% to 20.8%) and reversal in 8.3% (95% CI, 4.9% to 12.3%); results were similar for soft tissue diagnoses and the left TMJ. Concurrent baseline soft tissue diagnoses were associated with hard tissue diagnostic changes at follow-up ($P < 0.0001$). Baseline hard tissue diagnoses showed no statistical association with soft tissue changes at follow-up ($P = 0.11$). Longitudinally, 76% of baseline TMJ soft tissue diagnoses and 71% of the baseline hard tissue diagnoses remained stable. Diagnostic reversal and progression were confirmed for both soft and hard tissues.

Keywords: biostatistics, imaging, joint disease, computed tomography, magnetic resonance imaging, orofacial pain/TMD

Introduction

Temporomandibular joint (TMJ) structural disorders include soft tissue disc displacement (DD) and hard tissue degenerative joint disease (DJD). These disorders were classically described as a longitudinal progression from normal disc position to DD with reduction (DDwR) to DD without reduction (DDw/oR) to DJD (Rasmussen 1981; Wilkes 1989). However, the longitudinal course of TMJ DD and DJD has never been conclusively described with magnetic resonance imaging (MRI) for soft tissue diagnoses and computed tomography (CT) for hard tissue diagnoses.

This report is an exploratory evaluation of the course of TMJ structural disorders with longitudinal data to 1) assess the classical model of progression, 2) estimate the contribution of radiologic error (disagreement in radiograph interpretation) to observed TMJ diagnostic changes, and 3) explain the influence of baseline joint soft tissue diagnoses on hard tissue diagnostic changes at follow-up and baseline hard tissue diagnoses on soft tissue diagnostic changes at follow-up.

[1]Department of Diagnostic and Biological Sciences, School of Dentistry, University of Minnesota, Minneapolis, MN, USA
[2]Department of Oral Medicine, School of Dentistry, University of Washington, Seattle, WA, USA
[3]Western New York MRI, Buffalo, NY, USA
[4]Department of Oral Diagnostic Sciences, School of Dental Medicine, University at Buffalo, Buffalo, NY, USA
[5]Biostatistical Design and Analysis Center, Clinical and Translational Science Institute, University of Minnesota, Minneapolis, MN, USA
[6]Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN, USA
[7]Department of Developmental and Surgical Sciences, School of Dentistry, University of Minnesota, Minneapolis, MN, USA
[8]Statistical Center for HIV/AIDS Research and Prevention, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

A supplemental appendix to this article is available online.

Corresponding Author:
E. Schiffman, School of Dentistry, University of Minnesota, 6-320 Moos Tower, Minneapolis, MN 55455, USA.
Email: schif001@umn.edu

Methods

Overview

This prospective observational study conformed with Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for human observational investigations (von Elm et al. 2014). Research methods received Institutional Review Board approval at the 3 study sites, and informed consent was obtained from all follow-up subjects.

Sample. In the baseline Validation Project, done at the University of Minnesota, University of Washington, and University of Buffalo, 724 subjects had comprehensive clinical assessments and TMJ imaging. Nineteen subjects were excluded from the validation analysis, resulting in 614 subjects with temporomandibular disorders and 91 with normal TMJs. The 19 exclusions included 5 subjects who could not be classified as cases or controls and 14 with comorbid conditions (Schiffman et al. 2010). Funding for follow-up assessments in the TMJ Impact Project was approved for a convenience sample of 400 participants drawn from the baseline Validation Project for follow-up in the TMJ Impact Project. This sample size addressed tests of hypothesis regarding jaw pain, jaw function, and disability (reported separately). Total subject accrual over the 3 study sites reached 401 before recruitment closure.

Time Course. Subjects were recalled about 8 y after their original assessment in the Validation Project.

Subject Assessment. The primary assessment at baseline and follow-up used bilateral TMJ MRI and CT to establish joint-specific soft and hard tissue diagnoses, respectively. Subject demographics, 2 subscales of the Graded Chronic Pain Scale (VonKorff et al. 1992)—Characteristic Pain Intensity (CPI) and Disability Points—and the Jaw Functional Limitation Scale (JFLS; Ohrbach et al. 2008) were assessed by questionnaire. Baseline questionnaire data were used to compare subjects from the Validation Project who were recalled for the TMJ Impact Project ($n = 401$) versus subjects who were not recalled ($n = 323$).

Comparison of Follow-up to Baseline. Rates of diagnostic progression (change per the classic model), no change, and diagnostic reversal (change counter to classic model) were determined.

Recall of Study Subjects

Over the 3 study sites, 620 subjects were potentially eligible for follow-up, that is, subjects who had participated in the Validation Project and were among those who agreed to be recalled for a future study. A third inclusion requirement was that a structured follow-up interview found no contraindication for imaging. Follow-up study aims required no exclusions of the 724 Validation Project participants. The Appendix gives details of the recall methods and the recruitment flow chart.

TMJ Imaging Acquisition

Acquisition protocols for the baseline TMJ MRI and CT have been reported (Ahmad et al. 2009). Bilateral TMJ MRI and CT at follow-up employed standardized acquisition protocols at the 3 sites updated to current radiologic methods. The baseline MRI technique, with 1.5T and a special surface coil, produced images that compared well with follow-up 3T images for assessment of the disc. Multidetector CT (MDCT) was used for hard tissue imaging at baseline and cone beam CT (CBCT) at follow-up. CBCT requires a substantially lower radiation dose than MDCT. CBCT has been shown to provide diagnostic information equivalent to MDCT and is considered the preferred TMJ imaging method (Zain-Alabdeen and Alsadhan 2012). The Appendix gives further details.

Consensus Protocol for Primary Study Diagnoses

Each of the 3 sites had 1 study radiologist. They used TMJ MRI to render diagnoses of normal/indeterminate, DDwR, or DDw/oR and CT to render diagnoses of normal/indeterminate, grade 1 DJD, or grade 2 DJD (Ahmad et al. 2009; Ahmad and Schiffman 2016). See the Appendix for diagnostic criteria for soft and hard tissue. Regarding the disc, radiologists gave an indeterminate diagnosis when the disc structure did not meet all criteria for classification as normal but also did not fulfill the criteria for DD (Ahmad et al. 2009). Hard tissue findings were considered indeterminate when only flattening or subcortical sclerosis was present (Ahmad et al. 2009). All radiologists were blinded to each subject's identity and clinical findings. To avoid detection bias, radiologists were not allowed to directly compare subjects' baseline and follow-up images. A consensus protocol was used for all imaging diagnoses as follows: Each radiologist interpreted the follow-up images obtained at his site. A second radiologist was then randomly selected from the remaining 2 sites to independently interpret the same images, so the second interpretations were equally divided between the other site radiologists. If the 2 independent readings of a subject's follow-up imaging were discrepant, the radiologists were required to come to diagnostic consensus. One week or more after the follow-up consensus-based diagnostic interpretations had been established and entered into the study database, the paired radiologists performed the same protocol using the subject's baseline images. This consensus reread of the baseline films was required for comparability between baseline and follow-up diagnoses because, for the Validation Project, only a single radiologist at each site performed radiographic interpretations.

Radiologist Reliability

The radiologists' reliability was assessed at calibration sessions in 2011 (baseline), 2012, and 2013 (end of study). These calibration sessions were identical to those reported for the Validation Project (Ahmad et al. 2009). Data from the 3 sessions were combined, so reliability of hard tissue interpretations was assessed in totals of 160 MDCT images and 160

CBCT images. For soft tissue interpretations, 160 magnetic resonance images were employed for disc position and 60 for joint effusions.

Statistical Methods

Comparing Subjects Included vs. Not Included in the TMJ Impact Project. The entire baseline sample included 724 subjects. We compared the portion of the baseline sample that did not return for follow-up ($n = 323$) with the follow-up sample ($n = 401$), using baseline measures. Two-sample t tests were used for age, jaw pain intensity, jaw physical function, and pain-related disability and Pearson's chi-square test for sex, education, income, and subject source (clinic cases, community cases, and controls).

Estimating the Fraction of Subjects Having True Changes in Diagnosis between Baseline and Follow-up. The brief description here is expanded in the Appendix, which includes computer code. Hard and soft-tissue diagnoses as well as left and right TMJs were analyzed separately. The following describes the analysis for 1 tissue type and 1 TMJ. The analysis data set included hard and soft tissue diagnoses for 386 participants from both the Validation Project and the TMJ Impact Project.

The analysis applied the Hui-Walter method of estimating the chance of diagnostic disagreement in the absence of a gold-standard diagnosis (Hui and Walter 1980) to the independent diagnoses made by study examiners before the consensus procedure took place. The 3 possible diagnoses for soft tissue were normal, DDwR, and DDw/oR and for hard tissue, normal, grade 1, and grade 2 DJD. Hui-Walter's key assumption is that, conditional on the true disease state, joint diagnoses rendered by the radiologists are subject to statistically independent errors. With enough examiners and enough subjects, probability (examiner m gives diagnosis j | true diagnosis is i) is identified without further assumptions; in particular, different examiners can have different disagreement probabilities. This identification is possible when right and left joints are analyzed separately because the number of possible latent true states is small.

The analysis has 2 further key assumptions. First, probability (examiner m gives diagnosis j | true diagnosis is i) is the same for the calibration exercises, the Validation Study diagnoses, and the TMJ Impact Project diagnoses. Second, the Hui-Walter analysis is subject to an identification failure called "label switching." For binary diagnoses, this is customarily avoided by assuming that each examiner's sensitivity is at least 0.5. Here, with 3-valued diagnoses, we assume analogously that probability (examiner m gives diagnosis i | true diagnosis is i) ≥ 0.5.

The analysis is Bayesian and computed with Markov chain Monte Carlo (MCMC), which produces a correlated sequence of draws from the posterior distributions of all unknowns, permitting inferences about them. Because each analysis included only the right joint or only the left joint for each person, the unknowns reported here are the fractions in this study population—that is, a joint has each possible pair of diagnoses: 1

from the Validation Project and 1 from the TMJ Impact Project. For example, for soft tissue, pairs of diagnoses showing no change were normal-normal, DDwR-DDwR, and DDw/oR-DDw/oR; pairs of diagnoses showing progression were normal-DDwR, normal-DDw/oR, and DDwR-DDw/oR; and pairs of diagnoses showing reversal were DDwR-normal, DDw/oR-normal, and DDw/oR-DDwR. The fraction progressing is the sum of the fractions with diagnosis pairs showing progression; the fraction reversing is the sum of the fractions with diagnosis pairs showing reversal. The estimated fraction is the median of the MCMC draws; 95% equal-tailed posterior intervals are the 2.5th and 97.5th percentiles of the MCMC draws. Computing used JAGS (Plummer 2013) implemented with the rjags package (Plummer et al. 2015) in the R system (R Development Core Team 2015). The Appendix gives further details.

Do Changes in Hard Tissue Diagnoses from Baseline to Follow-up Depend on Baseline Soft Tissue Diagnoses and, Analogously, Hard and Soft Tissue Role Reversal? For the first question, the test statistic is the sum, across baseline hard tissue diagnoses, of Pearson's chi-square testing the association between baseline soft tissue diagnosis and follow-up hard tissue diagnosis. To see this, consider the subset of subjects with a normal baseline hard tissue diagnosis. We want to test whether their follow-up hard tissue diagnosis is associated with their baseline soft tissue diagnosis. The usual test is Pearson's chi-square for association in a cross-tabulation. Now, do likewise for 2 other subsets of subjects: those with baseline hard tissue diagnosis of grade 1 DJD and those with baseline grade 2 DJD. We now have 3 Pearson chi-square statistics; add them to give the test statistic that we used. We computed a P value using a permutation test, which simulates the test statistic's distribution under the null hypothesis of no association by permuting among persons the person-specific pairs (left side and right side) of baseline soft tissue diagnoses. This permutation scheme preserves all features of the actual data set except the association being tested. We created 1,000 artificial data sets using this permutation scheme; the P value is the fraction of those artificial data sets in which the test statistic is at least as large as the test statistic from the actual data. Reversing the roles of hard and soft tissue diagnoses in the foregoing gives a test of the association between the joint-specific baseline hard tissue diagnosis and observed change in follow-up soft tissue diagnosis.

Results

The mean follow-up time for TMJ Impact Project subjects was 7.9 y (SD, 0.77; range, 5.8 to 10 y). The mean age at follow-up was 45.8 y (SD, 12.9); 83% of follow-up subjects were female. Table 1 gives demographic characteristics of these subjects and those from the original group of 724 baseline subjects who were not followed up. Baseline subjects who were and were not followed up did not differ significantly in sex, age, education, or income ($P \geq 0.18$). However, these 2 samples differed significantly ($P \leq 0.05$) on other prognostic factors: Baseline

Table 1. Demographic and Baseline Differences between Validation Project and TMJ Impact Project Subjects.

Variable/Scale	Validation Project, n = 323	Validation and TMJ Impact Projects, n = 401	P Value[a]	TMJ Impact Project at Follow-up, n = 401
Age, y	36.1 ± 13.2	37.4 ± 12.9	0.19	45.8 ± 12.9
Sex				
Male	61 (19)	68 (17)	0.50	68 (17)
Female	262 (81)	333 (83)		333 (83)
Education, y				
≤12	52 (16.1)	62 (15.5)	0.83	27 (6.75)
>12	271 (83.9)	338 (84.5)		373 (93.25)
Income				
<$40,000/y	161 (50.31)	189 (47.73)	0.49	88 (26.19)
≥$40,000/y	159 (49.69)	207 (52.27)		248 (73.81)
Jaw Functional Limitation Scale	1.50 ± 1.50	1.19 ± 1.38	0.004	0.73 ± 1.01
Graded Chronic Pain Scale				
Characteristic Pain Intensity	40.3 ± 28.3	34.7 ± 28.3	0.009	21.3 ± 21.1
Disability Points	2.50 ± 1.66	2.23 ± 1.73	0.034	1.62 ± 1.55
Distribution of cases and controls				
Clinic cases	78 (24.1)	78 (19.5)	0.27	
Community cases	208 (64.4)	269 (67.0)		
Community controls	37 (11.5)	54 (13.5)		

Values presented as n (%) or mean ± SD.
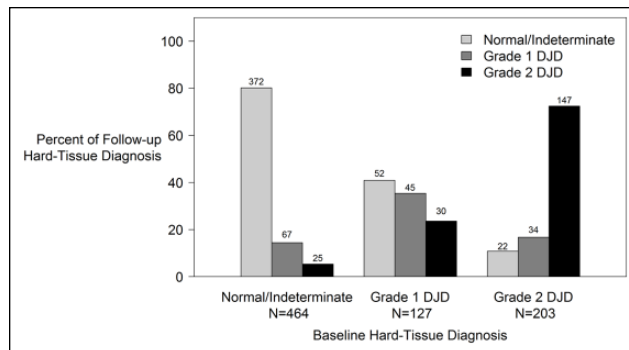TMJ, temporomandibular joint.
[a]For baseline difference.
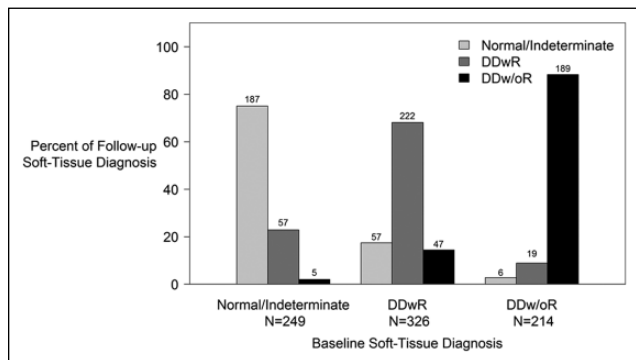
Figure 1. The longitudinal stability of soft tissue diagnoses per findings from temporomandibular joint magnetic resonance imaging: Of 789 joints with baseline diagnoses of disc displacement, on follow-up 598 (76%) joints showed no change; 109 (14%) demonstrated progression; and 82 (10%) had reversal. DDwR, disc displacement with reduction; DDw/oR, disc displacement without reduction.

Figure 2. The longitudinal stability of hard tissue diagnoses per findings from temporomandibular joint computed tomography: In 794 joints with baseline joint-specific hard tissue diagnoses of DJD, progression was observed in 122 (15%), no change in 564 (71%), and reversal in 108 (14%). DJD, degenerative joint disease.

CPI differed with a mean 40.3 (SD, 28.3) for subjects participating only at baseline versus 34.7 (SD, 28.3) for those who were followed up (P = 0.009). Likewise, the study samples differed in JFLS scores, with means of 1.50 and 1.19 (P = 0.004), and in Disability Points scores, with means of 2.50 and 2.23 (P = 0.034). Subjects included and not included in the TMJ Impact Study follow-up did not differ according to source (clinic vs. community) at baseline (P = 0.27).

Figure 1 depicts joint-specific baseline and follow-up soft tissue diagnoses. Of the 802 magnetic resonance images done, 789 (98%) could be interpreted and were used for analyses. Diagnostic reversal and progression of soft tissue diagnoses were both observed, with 109 joints (14%) showing progression, 598 (76%) showing no diagnostic change, and 82 (10%) showing reversal; images of 13 joints were nondiagnostic.

Figure 2 shows joint-specific baseline and follow-up hard tissue diagnoses. Of the 802 CT images done, 794 (99%) could be interpreted and were used for analyses. Both reversal and progression of hard tissue diagnoses were reported. Progression was observed in 122 joints (15%); 564 joints (71%) showed no diagnostic change from baseline; and reversal was reported in 108 joints (14%). No diagnosis could be rendered for images of 8 joints.

Figure 3 shows the distribution of joint-specific soft tissue diagnoses for 324 TMJs that had a concurrent diagnosis of DJD. Soft tissue diagnoses of normal/indeterminate and DDwR were found to coexist with a diagnosis of DJD. The distribution of joint-specific changes at follow-up in the absence of baseline DJD is very similar to Figure 3 (data not shown), suggesting that a concurrent diagnosis of DJD at baseline did not significantly influence soft tissue changes at follow-up.
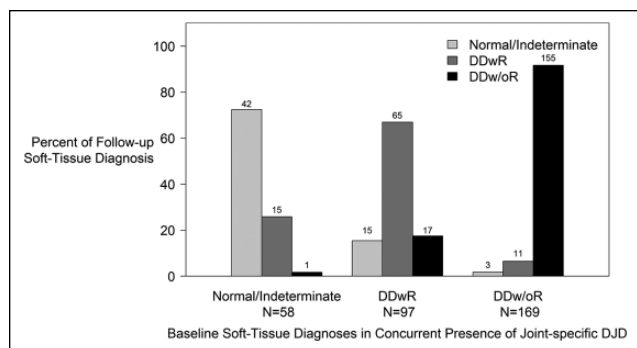
Figure 3. The stages of disc displacement with concurrent diagnosis of degenerative joint disease in 324 joints: Concurrent joint-specific baseline hard tissue diagnoses were not associated with soft tissue diagnostic change at follow-up (test statistic = 11.39, *P* = 0.11). DDwR, disc displacement with reduction; DDw/oR, disc displacement without reduction.

Concurrent joint-specific baseline hard tissue diagnoses were not associated with soft tissue diagnostic change at follow-up (test statistic = 11.39, *P* = 0.11). In contrast, joint-specific baseline soft tissue diagnoses significantly influenced change in hard tissue diagnoses at follow-up (test statistic = 53.35, *P* < 0.0001).

Effect of Radiologic Reliability (Diagnostic Agreement) on Reported Diagnostic Changes between Baseline and Follow-up

Interrater reliability for soft tissue diagnoses was good (kappa = 0.73; 95% CI, 0.64 to 0.83). For hard tissue diagnoses, agreement was similar (kappa = 0.76; 95% CI, 0.68 to 0.83). Although these measures of agreement are good, they leave open the possibility that some of the changes in diagnosis from baseline to follow-up could be explained by radiologic disagreement.

Table 2 reports estimates from the Hui-Walter analysis of the fractions of subjects having true progression and reversal of hard and soft tissue diagnoses (i.e., diagnostic changes that are not research artifacts attributable to error in image interpretation). Separate estimates are given for right and left jaw joints. Hard tissue progression was estimated to occur for 10% to 15% of subjects (depending on the joint), while hard tissue reversal occurred for about 8%. Estimated rates of progression and reversal were somewhat lower for soft tissue.

Discussion

This study used current radiologic methods to assess longitudinal stability of DD and DJD. More than 70% of baseline soft and hard tissue diagnoses remained stable long-term, except for grade 1 DJD. When diagnostic change did occur, we observed both reversal and progression of the baseline diagnoses. Our joint-wise comparison of radiographic diagnoses at baseline with corresponding diagnoses at follow-up revealed an unexpected frequency of diagnostic reversal to a less

Table 2. Estimated Fractions of Study Population with Diagnostic Progression or Reversal.

Temporomandibular Joint	Jaw Joint, % (95% Posterior Interval)	
	Right	Left
Hard tissue		
Progression	15.2 (10.5 to 20.8)	10.4 (6.5 to 15.1)
Reversal	8.2 (4.9 to 12.3)	7.8 (4.4 to 12.2)
Soft tissue		
Progression	9.0 (5.9 to 12.8)	11.0 (7.5 to 15.2)
Reversal	7.2 (4.4 to 10.8)	8.2 (5.1 to 12.1)

advanced stage. In contrast, joint progression to a more advanced disorder was anticipated except for diagnoses such as DDw/oR or grade 2 DJD, which were considered end-stage categories for soft and hard tissues, respectively.

The study sample included in the TMJ Impact Project (*n* = 401) was similar to the sample not recalled from the Validation Project for baseline age, sex, education, income, and recruitment from clinic versus community sources. CPI, JFLS, and Disability Points showed differences that were statistically but not clinically significant in that they represented ≤6% differences on the respective measurement scales (Dworkin et al. 2008). In contrast, CPI, JFLS, and Disability Points in subjects seen at baseline and follow-up decreased over time. This improvement was clinically significant, with average reductions ranging from >27% (Disability Points) to >38% (CPI and JFLS) on their measurement scales (Table 1).

Literature Review

A summary of relevant studies follows. First, a prospective 15-y MRI-based study of 47 subjects reported stable disc position (91%) and stable osseous status (96%) with minimal progression and no reversal (Sale et al. 2013). Baseline and follow-up images were assessed "side-by-side," creating a potential for detection bias especially if no reversal was expected a priori. Also, using MRI to assess hard tissues likely has a high rate of false negatives (Ahmad et al. 2009). Second, an MRI-based retrospective study of 29 joints over an average 89.3 mo reported progression (6 joints) and reversal (5 joints) for hard tissue diagnosis (Kurita et al. 2006). These 2 studies suggest that soft and hard tissue in the TMJ may progress diagnostically and that hard tissues may also undergo diagnostic reversal. Third, Alkhader and colleagues' (2010) cross-sectional study based on MRI and CT reported that 75% of the TMJs with osseous changes had DD. Fourth, a cross-sectional MRI-based study by Emshoff et al. (2001) found TMJ DJD in 20% of joints with no DD, 7% with DDwR, and 56% of joints with DDw/oR. Finally, the cross-sectional MRI-based study by Bertram et al. (2001) found TMJ DJD in 25% of joints with no DD, 20% of joints with DDwR, and 72% of joints with DDw/oR. Thus, a normal disc position or DDwR may coexist with DJD.

The current study shows that, in the presence of joint-specific DJD, 155 joints had baseline diagnoses of normal/indeterminate or DDwR, and 137 (88.4%) of these joints were still diagnosed as normal/indeterminate or DDwR an average of 7.9

y later (Fig. 3). Given this study's large sample size and the use of MRI and CT, these findings confirm the previous reports that DDw/oR is not an essential precursor for DJD (Bertram et al. 2001; Emshoff et al. 2001; Alkhader et al. 2010). This conclusion is strengthened by our findings that baseline TMJ soft tissue diagnoses have a significant association with TMJ hard tissue diagnostic changes, while baseline hard tissue diagnoses do not predict changes in TMJ soft tissue diagnoses.

Diagnostic Validity

This study's findings depend on the reliability and validity of the radiologists' image interpretation. Their interrater reliability was good to excellent. In addition, having 2 radiologists render consensus diagnoses, rather than depending on a single radiologist's interpretation, theoretically enhances the validity of the image interpretation. Bias in radiologic interpretation was also reduced by not allowing radiologists to simultaneously view a subject's baseline and follow-up images. Side-by-side assessments are performed in clinical practice but susceptible to detection bias (Kallman et al. 1989). Finally, the credibility of these results is supported because they follow biologically plausible patterns: minimal progression from normal soft and hard tissue diagnoses, the most progression and reversal from midstage DDwR and grade 1 DJD, and the least reversal from end-stage DDw/oR and grade 2 DJD. Baseline imaging techniques were upgraded at follow-up, but as explained earlier (see TMJ Imaging Acquisition section), this change likely had minimal effect on the baseline/follow-up comparisons.

Generalizability of Results

We previously reported that the baseline subject population was representative of the general population (Schiffman et al. 2010). In the present study, we reported differences between the sample included in the TMJ Impact Project and the subjects not included. There is no basis for suggesting that our main clinical findings would not hold for a population-based study assessing the course of TMJ DD and DJD.

Clinical Relevance

We have shown with relative certainty that the model for temporomandibular disorder progression (no DD/DJD → DDwR → DDw/oR → DJD) is only 1 possible outcome and that primary DJD may precede DD. Other outcomes include reversal over time to less advanced diagnoses, especially with grade 1 DJD. The most important observation is that, among the majority of subjects, soft and hard tissue diagnoses were stable. These findings indicate that TMJ DD and DJD are neither typically reversible nor progressive but rather that the TMJ is stable and adaptive, with the potential for improvement in status and with healing occurring over time. Currently, we cannot predict which individuals will progress, and no treatments, including TMJ surgery, can predictably prevent progression of

either soft or hard tissue disorders (Schiffman et al. 2014). Therefore, current clinical practice should focus on symptom management rather than structural intervention for the TMJ.

Conclusion

The classic model of progression is uncommon, as no change was the most typical finding and diagnostic reversal occurred. Longitudinally, diagnostic reversal and progression in hard and soft tissue were statistically supported after accounting for radiologic diagnostic disagreement. Finally, baseline joint soft tissue diagnoses have a highly significant association with joint hard tissue diagnostic changes, while baseline hard tissue diagnoses do not predict changes in joint soft tissue diagnoses.

Author Contributions

Acknowledgments

References

Ahmad M, Schiffman EL. 2016. Temporomandibular joint disorders and orofacial pain. Dent Clin North Am. 60(1):105–124.

Ahmad M, Hollender L, Anderson Q, Kartha K, Ohrbach R, Truelove EL, John MT, Schiffman EL. 2009. Research diagnostic criteria for temporomandibular disorders (RDC/TMD): development of image analysis criteria and examiner reliability for image analysis. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 107(6):844–860.

Alkhader M, Kuribayashi A, Ohbayashi N, Nakamura S, Kurabayashi T. 2010. Usefulness of cone beam computed tomography in temporomandibular joints with soft tissue pathology. Dentomaxillofac Radiol. 39(6):343–348.

Bertram S, Rudisch A, Innerhofer K, Pumpel E, Grubwieser G, Emshoff R. 2001. Diagnosing TMJ internal derangement and osteoarthritis with magnetic resonance imaging. J Am Dent Assoc. 132(6):753–761.

Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, Haythornthwaite JA, Jensen MP, Kerns RD, Ader DN, et al. 2008. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain. 9(2):105–121.

Emshoff R, Rudisch A, Innerhofer K, Bosch R, Bertram S. 2001. Temporomandibular joint internal derangement type III: relationship to magnetic resonance imaging findings of internal derangement and osteoarthrosis. An intraindividual approach. Int J Oral Maxillofac Surg. 30(5):390–396.

Hui SL, Walter SD. 1980. Estimating the error rates of diagnostic tests. Biometrics. 36(1):167–171.

Kallman DA, Wigley FM, Scott WW Jr, Hochberg MC, Tobin JD. 1989. New radiographic grading scales for osteoarthritis of the hand: reliability for determining prevalence and progression. Arthritis Rheum. 32(12):1584–1591.

Kurita H, Uehara S, Yokochi M, Nakatsuka A, Kobayashi H, Kurashina K. 2006. A long-term follow-up study of radiographically evident degenerative changes in the temporomandibular joint with different conditions of disk displacement. Int J Oral Maxillofac Surg. 35(1):49–54.

Ohrbach R, Larsson P, List T. 2008. The jaw functional limitation scale: development, reliability, and validity of 8-item and 20-item versions. J Orofacial Pain. 22(3):219–230.

Plummer M. 2013. JAGS user manual [computer program]. Version 3.4.0 [accessed 2016 October 24]. http://www.stats.ox.ac.uk/~nicholls/MScMC MC14/jags_user_manual.pdf.

Plummer M, Stukalov A, Denwood M. 2015. rjags: bayesian graphical models using MCMC [computer program]. Version 3-15 [accessed 2016 October 25]. http://CRAN.R-project.org/package=rjags.

R Development Core Team. 2015. R: a language and environment for statistical computing [computer program]. Version 3.2.2 [accessed 2016 October 24]. https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf.

Rasmussen OC. 1981. Description of population and progress of symptoms in a longitudinal study of temporomandibular arthropathy. Scand J Dent Res. 89(2):196–203.

Sale H, Bryndahl F, Isberg A. 2013. Temporomandibular joints in asymptomatic and symptomatic nonpatient volunteers: a prospective 15-year follow-up clinical and MR imaging study. Radiology. 267(1):183–194.

Schiffman EL, Truelove EL, Ohrbach R, Anderson GC, John MT, List T, Look JO. 2010. The research diagnostic criteria for temporomandibular disorders: I. Overview and methodology for assessment of validity. J Orofac Pain. 24(1):7–24.

Schiffman EL, Velly AM, Look JO, Hodges JS, Swift JQ, Decker KL, Anderson QN, Templeton RB, Lenton PA, Kang W, et al. 2014. Effects of four treatment strategies for temporomandibular joint closed lock. Int J Oral Maxillofac Surg. 43(2):217–226.

von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. 2014. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Int J Surg. 12(12):1495–1499.

VonKorff M, Ormel J, Keefe FJ, Dworkin SF. 1992. Grading the severity of chronic pain. Pain. 50(2):133–149.

Wilkes CH. 1989. Internal derangements of the temporomandibular joint: pathological variations. Arch Otolaryngol Head Neck Surg. 115(4):469–477.

Zain-Alabdeen EH, Alsadhan RI. 2012. A comparative study of accuracy of detection of surface osseous changes in the temporomandibular joint using multidetector CT and cone beam CT. Dentomaxillofac Radiol. 41(3):185–191.