

SCIENTIFIC REPORTS



OPEN

Grading facial expression is a sensitive means to detect grimace differences in orofacial pain in a rat model

Megan M. Sperry¹, Ya-Hsin Yu², Rachel L. Welch¹, Eric J. Granquist³ & Beth A. Winkelstein^{1,4}

Although pre-clinical models of pain are useful for defining relationships between biological mechanisms and pain, common methods testing peripheral sensitivity do not translate to the human pain experience. Facial grimace scales evaluate affective pain levels in rodent models by capturing and scoring spontaneous facial expression. But, the Rat Grimace Scale (RGS) has not assessed the common disorder of temporomandibular joint (TMJ) pain. A rat model of TMJ pain induced by jaw loading (1 hr/day for 7 days) was used to investigate the time course of RGS scores and compare them between different loading magnitudes with distinct peripheral sensitivity profiles (0N—no sensitivity, 2N—acute sensitivity, 3.5N—persistent sensitivity). In the 3.5N group, RGS is elevated over baseline during the loading period and one day after loading and is correlated with peripheral sensitivity ($\rho = -0.48$, $p = 0.002$). However, RGS is not elevated later when that group exhibits peripheral sensitivity and moderate TMJ condylar cartilage degeneration. Acutely, RGS is elevated in the 3.5N loading group over the other loading groups ($p < 0.001$). These findings suggest that RGS is an effective tool for detecting spontaneous TMJ pain and that spontaneous pain is detectable in rats that develop persistent TMJ sensitivity, but not in rats with acute resolving sensitivity.

Pre-clinical rodent studies commonly assess pain by measuring hypersensitivity in the relevant dermatome for the pathology or etiology being studied¹. However, that reflex outcome does not translate to the human experience of pain because it fails to capture the cortical and affective components involved in the perception of pain^{1–4}. In many studies, peripheral sensitivity is measured using mechanical stimuli to illicit a withdrawal response, defining a threshold for sensitivity^{5–10}. Although this method is useful for investigating some forms of sensitivity, it assesses only the evoked component of pain³ and nociceptive pain pathways¹¹. Hypersensitivity measurements do not evaluate the presence or extent of the non-nociceptive pathways that influence the *experience* of pain, including corticolimbic interactions that produce negative affect, drive emotional learning, and reorganize circuitry¹². Further, the use of mechanical and/or thermal hypersensitivity does not readily translate into clinical outcomes because hypersensitivity is identified *only* in a subset of patients who report pain³. In contrast, spontaneous pain is identified almost universally among chronic pain patients^{3,13,14} and has been identified as the principal symptom of clinical pain^{13,15}. However, methods have recently been developed to measure spontaneous pain in pre-clinical animal models of inflammatory and neuropathic pain to improve translatability to the human experience^{16–20}.

In most patients, spontaneous pain is reported using pain scales^{21,22}, questionnaires^{23,24}, and descriptions of sensations^{15,25}. However, those methods cannot be used for non-verbal populations; instead, facial expression has been adopted as a surrogate measure of spontaneous pain. The Facial Action Coding System transferred human facial expression movement into action units²⁶ and the Neonatal Facial Coding System is now widely used in infant populations²⁷. Pre-clinical pain researchers used these approaches as models to develop objective methods to evaluate if pain is present in animals¹⁶; the mouse grimace scale (MGS) adapted that scoring to utilize five facial features as indicators to evaluate pain¹⁶ and a four-feature facial coding system was also developed for rats

¹Department of Bioengineering, University of Pennsylvania, Pennsylvania, USA. ²Department of Endodontics, School of Dental Medicine, University of Pennsylvania, Pennsylvania, USA. ³Oral & Maxillofacial Surgery, School of Medicine, University of Pennsylvania, Pennsylvania, USA. ⁴Department of Neurosurgery, School of Medicine, University of Pennsylvania, Pennsylvania, USA. Correspondence and requests for materials should be addressed to B.A.W. (email: winkelst@seas.upenn.edu)

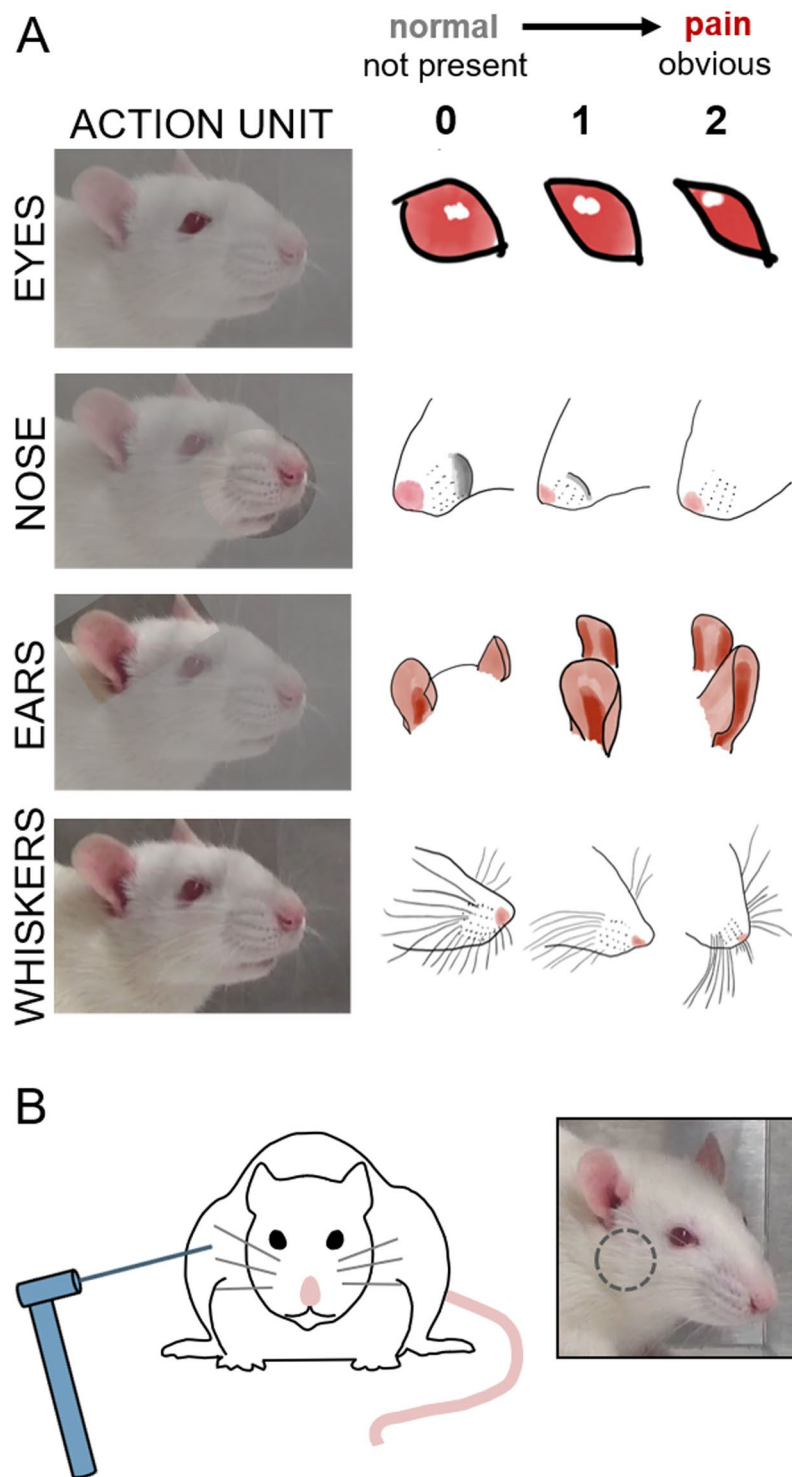


Figure 1. Sensitivity was measured by (A) RGS scoring of facial features, including the eyes, nose, ears, and whiskers, and (B) mechanical reflex testing in the TMJ region (designated by the grey circle).

(Fig. 1)²⁸. The MGS and Rat Grimace Scale (RGS) both capture spontaneous pain and have been hypothesized to represent a measure of the animal's affective response to pain^{16,28}. Although the RGS has been used to assess pain in models of intraplantar complete Freund's adjuvant (CFA) injection, intra-articular kaolin/carrageenan injection, plantar incision, laparotomy, experimental tooth movement, and acute chemotherapy-induced mucositis^{11,17–19,28,29}, its use in orofacial pain has been limited to orthodontic tooth movement¹⁷. There are no studies investigating spontaneous pain of the temporomandibular joint (TMJ) using RGS as an assessment approach.

Orofacial pain near the jaw and ear can originate from mechanical dysfunction and/or low-grade inflammation of the TMJ, which articulates the mandible to the bones of the skull³⁰. TMJ pain affects 5–12% of the

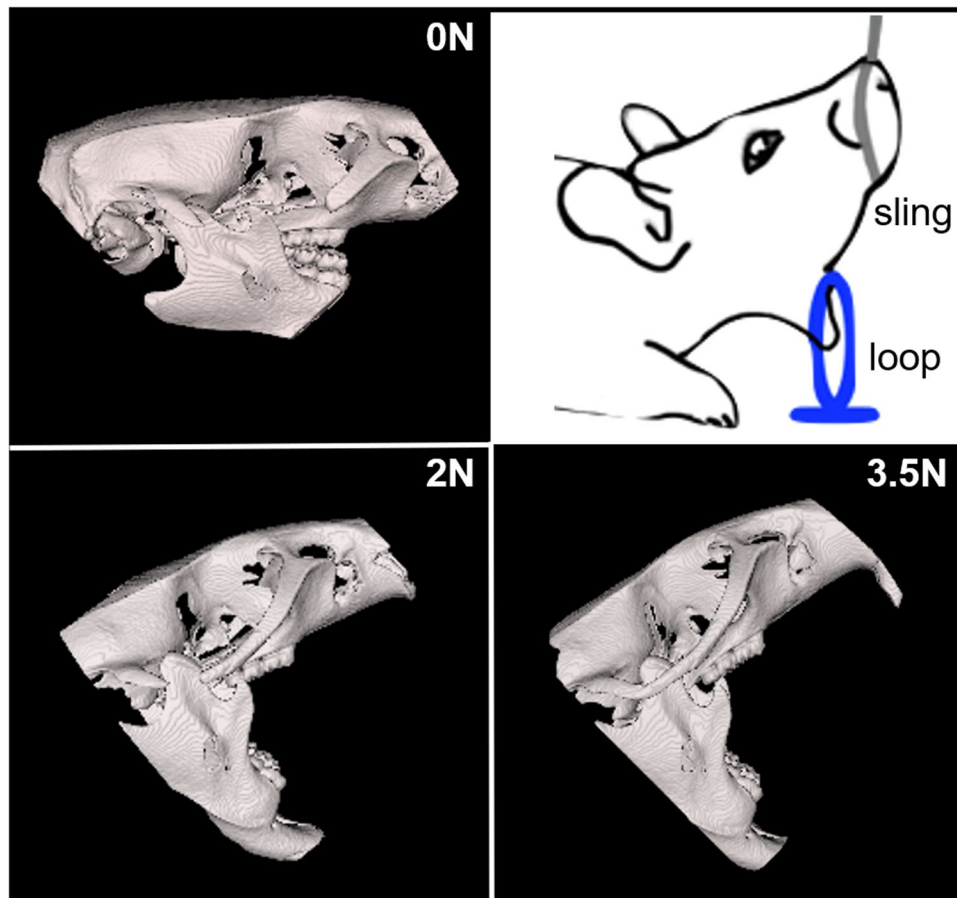


Figure 2. Three-dimensional CT reconstructions of the rat skull during the corresponding magnitudes of mechanical loading (0N, 2N, 3.5N) to the TMJ. The mandible was fixed by a nylon loop and the maxilla was held open by a sling with the applied load.

population³¹ and can be a debilitating, chronic condition associated with both increased nociceptive activity and alterations in brain circuitry^{30,32,33}. To date, studies have primarily relied on evoked measurements, such as mechanical hyperalgesia, to assess TMJ pain and the effect of potential therapeutics on TMJ sensitivity^{5,7,34,35}. Based on the anatomic location of the TMJ, RGS is expected to be a particularly robust measure for pain in this region and may also be a useful tool for studying other orofacial pain syndromes. Surrogate measurements of spontaneous pain have been used in studies of orofacial pain, such as bite force, grooming behavior, guarding, and weight loss^{1,11,35}. However, those behaviors are difficult to quantify, and present challenges in differentiating whether they indicate stress, pain, paresthesia, and/or avoidance behavior³⁶.

In this study, RGS was implemented to assess its ability to detect TMJ pain induced using repeated TMJ loading that produces moderate osteoarthritic pathology in the joint^{5,7}. Assessments were measured at baseline before any procedures, during the days of the TMJ loading period, and after the loading period, and were compared with conventional mechanical reflex testing with von Frey filaments. RGS scores were also compared across separate groups of rats receiving different degrees of TMJ loading (0N, 2N, 3.5N) (Fig. 2), known to induce different reflex responses. Exposure to repeated TMJ loading at different magnitudes produces distinct TMJ sensitivity profiles evoked in response to von Frey filament stimuli^{5,7}. Both the 3.5N and 2N groups are sensitive during and immediately after TMJ loading; however, after termination of loading, sensitivity returns to baseline levels in the 2N group whereas the 3.5N group remains sensitive for 8 days^{5,7}. The 0N control group does not exhibit any TMJ sensitivity at any time. Although orofacial sensitivity is commonly assessed using evoked measurements, the relationship between evoked and spontaneous pain is untested in TMJ pain. In addition, TMJs were harvested at early and late time points after the cessation of loading to evaluate the extent, if any, of structural changes in the cartilage of the condyle using Safranin O/Fast Green staining and Mankin scoring. Despite the widespread use of Mankin scoring to evaluate cartilage degeneration in animal models of TMJ osteoarthritis^{37–39}, the temporal relationship between TMJ pain and structural changes of the joint remains poorly understood.

Results

RGS scores increase for a painful insult but are only correlated to reflex responses acutely. The RGS score was assessed in a rat model of TMJ loading that has been previously shown to induce persistent orofacial sensitivity as measured by mechanical reflex testing^{5,7}. The RGS values reflect an overall rating based on presentation of four action units of the face (Fig. 1). In the rats undergoing 7 days of 3.5N loading (Fig. 2), the

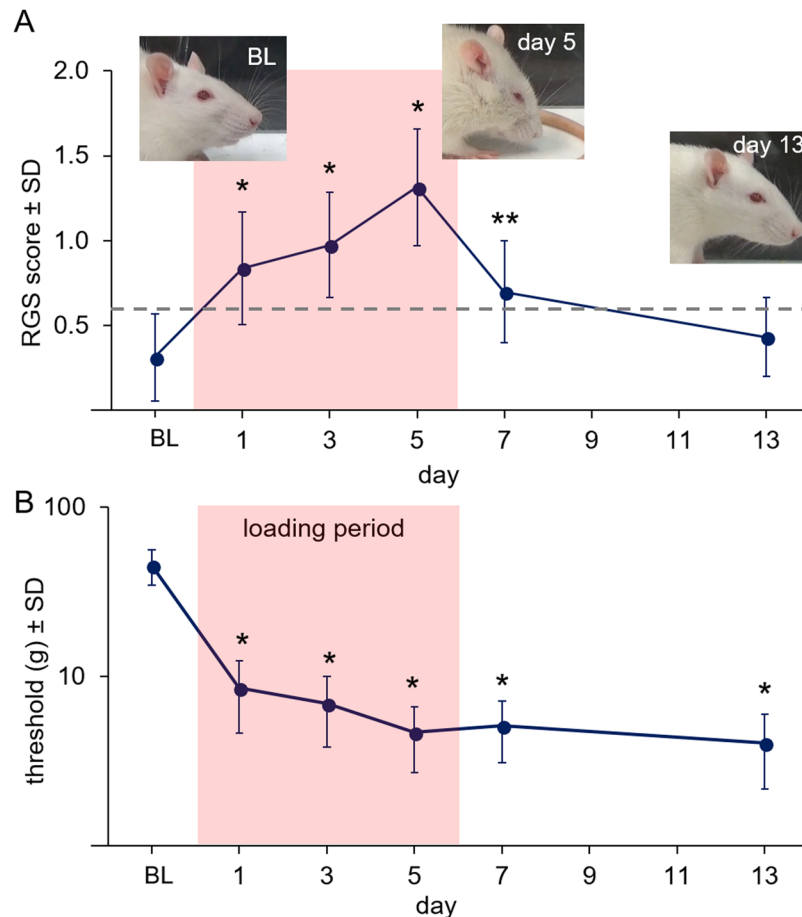


Figure 3. (A) RGS scores increase from baseline (BL) after 3.5N loading on days 1, 3, and 5 (* $p < 0.001$) and 1 day after the cessation of loading (day 7) (** $p = 0.019$). However, RGS scores return to baseline levels by day 13 ($p = 0.820$). The grey line indicates the previously reported analgesic intervention score for the RGS⁴⁶. (B) Head withdrawal thresholds (g) decrease from BL levels both during and after loading on all days (* $p < 0.01$). The loading period is labeled in (B) and shown in the pink shaded region of both panels.

RGS scores 3 hours after exposure were elevated ($p < 0.001$) immediately from baseline (0.31 ± 0.25) on day 1 (0.84 ± 0.33), day 3 (0.97 ± 0.31), and day 5 (1.31 ± 0.35) (Fig. 3A). On day 7, which was 1 day after the last day of loading, RGS scores remained elevated ($p = 0.019$) over baseline levels (0.70 ± 0.30). However, at day 13, RGS scores returned to baseline levels (0.43 ± 0.23) and were not significantly different ($p = 0.82$) from those responses.

In contrast, head withdrawal thresholds were significantly different from baseline on all days they were assessed ($p < 0.001$). Thresholds were lower ($p < 0.01$) than baseline values (45 ± 11 g) during the loading period on day 1 (9 ± 4 g), day 3 (7 ± 3 g), and day 5 (5 ± 2 g), and remained significantly decreased from baseline withdrawal thresholds ($p < 0.01$) at day 7 (5 ± 2 g) and day 13 (4 ± 2 g) (Fig. 3B). Head withdrawal thresholds and RGS scores were not correlated ($\rho = -0.28$, $p = 0.09$) when all time points were included. However, withdrawal threshold and RGS were moderately and significantly correlated ($\rho = -0.48$, $p = 0.002$) when including times during loading (days 1, 3, and 5) and at the early time after its cessation (day 7).

RGS scoring acutely detects differences that are not differentiated by reflex responses until later. We then measured RGS scores using different TMJ loading severities (3.5N, 2N, and 0N mouth-loading; Fig. 2) that have been shown to exhibit different TMJ sensitivity profiles after the cessation of loading^{5,7}. Overall, RGS scores from the 3.5N loading group were significantly higher ($p < 0.001$) than each of the 0N and 2N groups across all days tested (Fig. 4). The RGS scores for the 0N and 2N groups were not different from each other ($p = 0.999$) on any day and on average increased 0.39 ± 0.32 from baseline on each of the test days. Peak differences in RGS from baseline were observed on day 5 in each of the groups: 3.5N (1.12 ± 0.37), 2N (0.66 ± 0.36), and 0N (0.64 ± 0.25) (Fig. 4). However, RGS scores at that time were not significantly different from the scores recorded at days 1 and 3. The most robust differences between groups were observed at day 7, with RGS difference scores for the 3.5N loading group (0.87 ± 0.35) elevated over those for each of the 2N (0.21 ± 0.19 ; $p = 0.0003$) and 0N (0.25 ± 0.28 , $p = 0.006$) groups. RGS scores were significantly decreased from day 5 to day 7 only in the 2N group ($p = 0.049$).

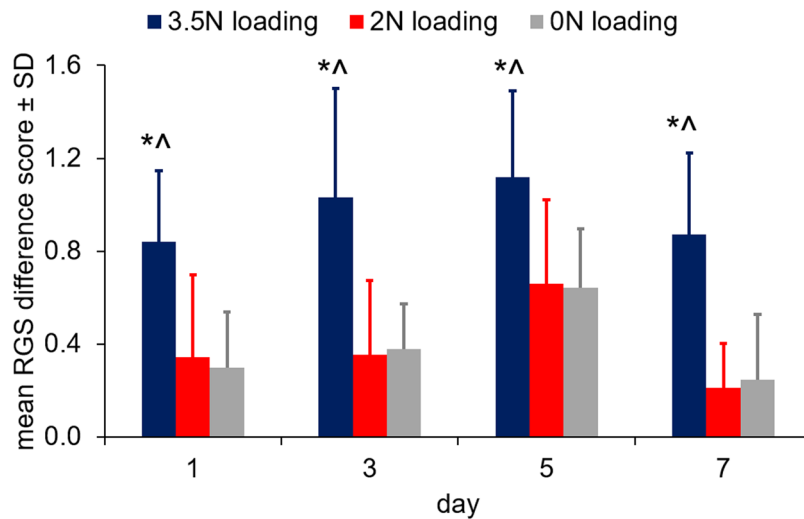


Figure 4. RGS scores are greater in the 3.5N loading group than either the 2N (* $p < 0.05$) or 0N ($\wedge p < 0.03$) group during the loading period (days 1, 3, 5) and immediately after the cessation of loading (day 7). The scores for the 2N and 0N loading groups are not different from each other at any day.

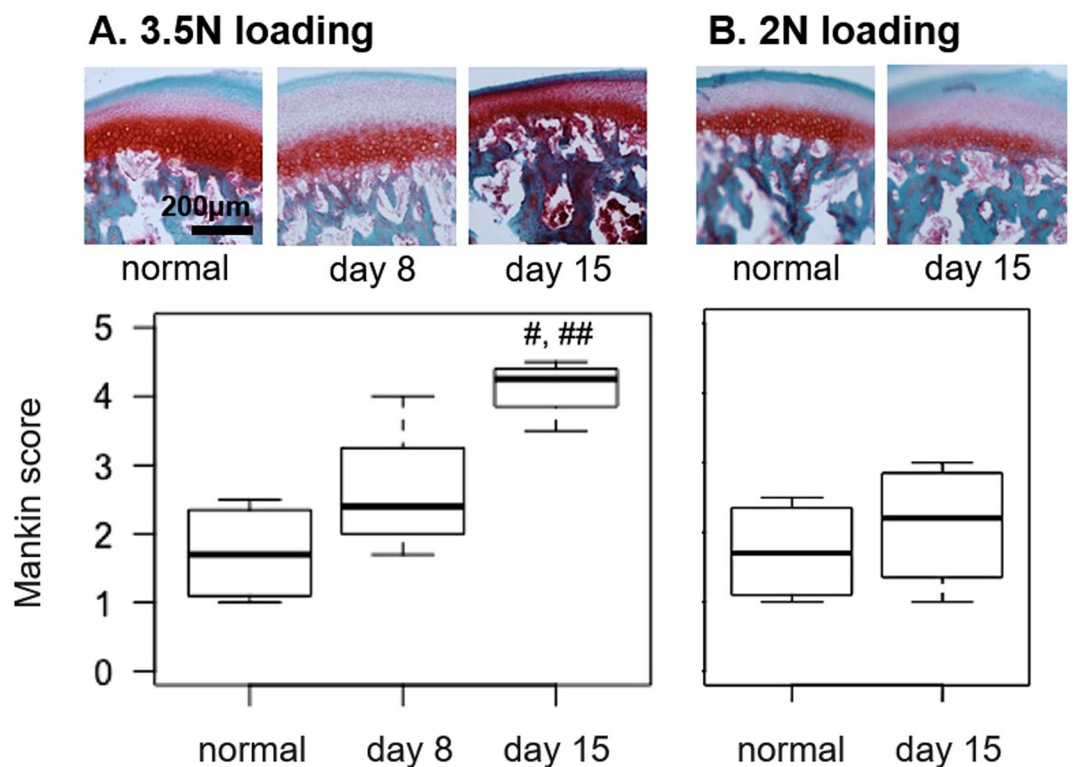


Figure 5. (A) Mankin scores of the TMJ condyle are unchanged at day 8 after 3.5N loading but increase at day 15 over normal (* $p = 0.004$) and day 8 levels (** $p = 0.047$). (B) Mankin scores are unchanged from normal levels at day 15 after 2N loading.

TMJ cartilage is altered after loading that acutely elevates RGS scores. TMJ condyle cartilage was evaluated using Safranin O/Fast Green staining and the Mankin scoring system to determine if RGS scores parallel changes in joint structure that are associated with cartilage osteoarthritis^{37,40}. Mankin scores for the 3.5N loaded TMJs were greater at day 15 (4.10 ± 0.37) than scores for normal tissue (1.71 ± 0.64 ; $p = 0.004$) and those at day 8 (2.60 ± 0.86 ; $p = 0.047$) after 3.5N loading (Fig. 5). Mankin scores for the 2N loaded TMJs at day 15 (2.1 ± 0.92) were not different ($p = 0.548$) from the scores for normal TMJs (Fig. 5), indicating no structural changes in that model.

Discussion

This study used the RGS to assess spontaneous pain in a rat model of TMJ pain and compared RGS between groups with different TMJ sensitivity profiles to evaluate its ability to differentiate pain states (Figs 3 and 4). Although facial grimace scales have been used to characterize visceral, inflammatory, and neuropathic pain in rodent models of skin and joint inflammation, cervical radiculopathy, peripheral neuropathy, bladder infection, and abdominal constriction^{16,18,19,41}, this is the first report of facial grimace for TMJ pain and in the context of repeated joint over-loading, which is a major cause of joint degeneration and pain^{42–45}. Since RGS is elevated over baseline during the TMJ loading period and one day after its cessation (Fig. 3), affective components of pain appear to be detectable both acutely after a single joint loading and also throughout repeated loading sessions. These same time points also exhibit sensitivity detected by reflex testing (Fig. 3), indicating that both nociceptive and spontaneous pain are present within this acute time frame. RGS scores during loading are elevated over the analgesic intervention score (RGS = 0.67), which has been previously determined as the quantitative standard separating painful from non-painful states⁴⁶. However, unlike reflex testing which detects TMJ sensitivity one week after cessation of loading (day 13)⁷, RGS is *not* elevated (Fig. 3). This difference between outcomes suggests that RGS may be sensitive enough to serve as an early detector of persistent TMJ pain and may even be predictive of long-term differences in TMJ sensitivity (Fig. 3).

The difference between RGS scores and reflex thresholds at day 13 is consistent with the finding that RGS scores return to baseline within 48 hours of neuropathic injury¹⁹ and acute inflammation¹⁸. Certain inflammatory conditions, like those simulated by a CFA injection into the masseter muscle, elevate RGS scores for up to 3 days after that initiating event⁴¹. Although RGS most sensitively detects pain at acute time points^{16,18,19,28}, the disappearance of facial grimace is not necessarily indicative of the resolution of spontaneous pain²⁸. Natural adaptations have led prey species, like rodents, to inhibit facial grimace as soon as possible so that they do not become the target of predators⁴⁷. By hiding painful expressions, spontaneous TMJ pain may persist for longer than is detectable by RGS. In fact, this hypothesis has been suggested previously for visceral, inflammatory, and neuropathic conditions^{19,28}. Other tests of spontaneous behaviors, like conditioned place preference and activity monitoring, also fail to capture robust pain behaviors after acute time points⁴⁸. Given these limitations, long-term monitoring of spontaneous pain remains challenging for chronic conditions.

Although the sensitivity of RGS is limited to acute times^{16,19,28}, RGS is differentially modulated in the current study between groups, and is only elevated in the group that exhibits persistent TMJ sensitivity (3.5N loading) after the insult is removed (Fig. 4). The fact that this response is only observed in the group with persistent pain^{5,7} suggests a relationship between early facial grimace and long-lasting TMJ sensitivity. Inflammatory and incisional pain models demonstrate similar early facial grimace and hypersensitivity¹⁸; however, little attention has been paid to the connection between heightened RGS and long-lasting peripheral sensitivity¹⁹. Conversely, for the 2N loading group in which TMJ sensitivity resolves after the loading period^{5,7}, facial grimace is not different from responses in the group receiving only anesthesia (0N loading) and exhibiting no TMJ sensitivity at any time point^{5,7}. Together, the differential TMJ pain models used here represent a range of pain states: combined peripheral nociception and affective pain (3.5N loading), peripheral nociception that resolves (2N loading), and no pain (0N loading). The combined use of facial grimace and peripheral reflex testing in these models, as well as in incisional, inflammatory, and neuropathic pain models, emphasizes that when spontaneous pain is detectable, hypersensitivity is also present; however, hypersensitivity can be present in the absence of measurable spontaneous pain^{18,19}. Although the RGS scores for the groups without reflex responses (2N and 0N) are not different (Fig. 4), this study only evaluated RGS scores at two time points after a stimulus: 3 hours and 24 hours (on day 7) after loading and anesthesia. It is possible that these selected time points may not capture the peak differences between groups, due to either anesthesia effects¹⁹ or a delayed onset of peak spontaneous pain¹⁶. Since differences between injured and anesthesia-only groups have been reported to peak within 3 to 9 hours after anesthesia^{18,19,28}, more pronounced differences between such groups may be evident at later times when anesthesia effects have further subsided.

In addition to evaluating limited acute time points, this study also restricted evaluation to a single long-term time point (day 13). Because TMJ structure is *slowly* altered due to repeated jaw loading, with moderate thinning and cellular disorganization of the TMJ cartilage apparent by day 15 (Fig. 5), it is possible that RGS scores may increase again at later times when cartilage damage worsens. In fact, destabilization of knee ligaments in rats, which leads to synovial inflammation and cartilage destruction, has been shown to increase RGS scores 14 days after the initial surgery⁴⁹, suggesting that spontaneous pain in osteoarthritis conditions may have a slower onset. Detection of delayed and episodic pain is particularly relevant to patients with painful TMJ disorders, who often experience cycles of pain ranging from a dull, constant ache to severe pain brought on by movement of the mandible³⁰. As such, measuring both evoked and spontaneous pain over longer periods may provide clinically relevant information for accurately modeling both the nociceptive and affective components of TMJ pain and enable testing new treatments through these episodes.

This study is also limited by assessing only a *single* sex and strain of rat. TMJ pain was investigated in female rats because patients who seek care for TMJ pain are disproportionately female, with a female-to-male ratio between 3:1 and 9:1³⁰. However, no sex differences have been identified using RGS measurements collected after intraplantar inflammatory injection or laparotomy²⁸, suggesting that facial grimace may not be inherently different between sexes. It is possible that TMJ loading and/or pain presents with different effects in male and female rats since there are known sex differences in the mechanisms of immune-neuronal interaction⁵⁰. Therefore, measuring RGS in male rats using this model would be helpful in determining if the associated affective pain responses are consistent across sexes.

Although osteoarthritic changes in the TMJ have been associated with increased orofacial sensitivity and reduced bite force^{35,51}, this is the first study to investigate the relationship between joint structure and affective pain. Degeneration of the TMJ cartilage is only detectable in those rats with acute spontaneous pain and

persistent evoked sensitivity (Figs 3–5), which suggests that early affective pain may be associated with, or even predict, structural changes in the joint that are associated with its overloading. However, the fact that structural changes develop *after* spontaneous pain appears to resolve suggests that early spontaneous pain may be driven by factors other than the structural modifications. In fact, the acute pain may be due to the increased intraarticular inflammatory factors that have been reported in the 3.5N loading group⁷. Direct induction of intraarticular inflammation, as by TNF α or CFA injection in the TMJ, has also been reported to acutely increase neuronal activity in the trigeminal subnucleus caudalis and to induce joint tissue degeneration^{52,53}. However, inhibiting inflammatory factors in the TMJ is needed to further test the association between affective pain and acute TMJ inflammation after repeated joint loading.

Overall, this study not only supports the use of, but highlights the need for using, the RGS to monitor spontaneous pain in mechanical joint loading models of TMJ pain. Further, heightened RGS scores after TMJ loading appear to portend persistent TMJ peripheral sensitivity. As such, RGS may be a helpful modality for forecasting extended periods of peripheral sensitivity and for identifying clinical pain states, in TMJ and possibly other models of clinical pain. Although the non-invasive, unprovoked collection of RGS data is simple to execute, its scoring can be a labor-intensive process²⁸. Recent advances in real-time RGS²⁰ and convolutional neural networks for automated facial grimace scoring⁵⁴ provide methods to streamline its use in both clinical and pre-clinical research applications. Therefore, more detailed investigation of the relationships between early facial grimace and persistent peripheral sensitivity would be valuable for real-time decision making in many environments. Despite strong evidence of on-going evoked sensitivity after the resolution of facial grimace across multiple models of clinical pain^{18,19,41}, it is not yet known if acutely heightened RGS is predictive of persistently evoked sensitivity for other types of pain, particularly in models with tunable reflex responses. Although additional studies are needed to investigate the predictive value of acute RGS more broadly, this study demonstrates that spontaneous pain is sensitive enough for early and reliable detection of conditions that will develop persistent TMJ sensitivity, but not those with acute resolving sensitivity.

Methods

Animals. All studies used adult female Holtzman rats (HsdHot:Holtzman Sprague Dawley), weighing 268 ± 21 g at the start of the study (obtained from Envigo (Indianapolis, IN)). Rats were housed in groups of 2–3 in standard polycarbonate caging (AnCare; Bellmore, NY), with 0.25-inch corncob bedding (Bed-o-Cobs; The Andersons Lab Bedding Products; Maumee, OH) and ad libitum access to food (LabDiet 5001; LabDiet; St Louis, MO) and water (acidified to pH = 3). Rats were housed in an Association for Assessment and Accreditation of Laboratory Animal Care accredited vivarium under a 12:12 hour light:dark cycle in a temperature-controlled environment in accordance with recommendations set forth in The Guide for Care and Use of Laboratory Animals (8th edition)⁵⁵. All procedures were approved by the University of Pennsylvania Institutional Animal Care and Use Committee and adhered to the guidelines for research and ethical issues of the International Association for the Study of Pain⁵⁶.

The RGS scores were first measured in rats ($n = 8$) undergoing the TMJ loading that is known to induce sustained sensitivity in the TMJ region as assessed by mechanical reflex testing⁷. Facial grimace was assessed in those rats undergoing daily mouth-opening by 3.5N to determine if and when it is present after TMJ loading, as well as to directly compare RGS score with the time course of peripheral sensitivity. Based on those findings, the RGS was then compared between groups ($n = 8$ /group) with different TMJ loading severities (0N, 2N, and 3.5N), which are known to have different peripheral sensitivity profiles⁵⁷.

Rat Grimace Scale evaluation & scoring. Facial grimace temporal patterns were evaluated at baseline (before procedures) and then daily using digital video recordings and RGS scores. When measurement corresponded to days of loading or other procedures requiring anesthesia, digital video recordings were acquired at 3 hours after exposure to isoflurane in order to allow the rats recovery time after anesthesia. All recording and monitoring procedures were performed in a quiet environment, and personnel remained out of visual contact with the rats for the duration of each session. Rats were placed singly in a $23 \times 10 \times 10$ cm³ transparent Plexiglas chamber with a removable stainless-steel top. A digital video camera (Sony HDR-CX380/B High Definition Handycam) was placed in front of the wider side of the box. Rats were videotaped for 30 minutes and videos were saved as mp4 files. In the subset of rats known to have sustained TMJ sensitivity (3.5N loading), digital videos were acquired at baseline, every other day during the loading period (days 1, 3, and 5), and after loading on days 7 and 13 to assess long-term responses. To compare between groups with different sensitivity profiles (3.5N, 2N, and 0N), digital videos were acquired at baseline, during loading (days 1, 3, and 5), and after loading on day 7.

Ten images were captured from each 30-minute video session at 3-minute intervals as portable network graphic (PNG) files. In order to be included in grading and analysis, each image was required to have a clear view of the four action units (eyes, nose/cheek, ears, and whiskers) (Fig. 1) and to not be taken during grooming, sleeping, or active sniffing activity^{16,19,28}. In the event that an image did not meet the requirements for extraction at the 3-minute interval, the video was advanced to the next immediate time point when an image could be used. The image-capture operator was blinded to the group, time point and rat identification. All images were randomized prior to scoring.

The scorer was given training using the standard scoring method¹⁹, before scoring any of the images from this study. Each image was scored for the action units: orbital tightening, nose/cheek flattening, ear curling, and whisker bunching (Fig. 1)^{19,28}. For each image, the scorer assigned an intensity value of 0 (absent), 1 (moderately appearing), or 2 (obviously present) for each of the four action units. If the action units of the image could not be scored by the rater, the value would be assigned as “not scored” and not included in the average value for that image. The RGS score for each image was taken as the average of the action unit scores. The mean RGS score at

an individual time point for each rat was the average of the RGS scores across the 10-images acquired. All images included in this study were scored by a single observer who was blinded to the procedures.

Mechanical reflex testing. Mechanical reflex testing was performed at baseline before any procedures and every other day during and after loading (days 1, 3, 5, and 7). In a subset of rats, reflex testing was also performed at day 13 to assess later effects at times after the loading protocol and acute TMJ sensitivity. Reflex testing was performed at 8:00 AM each morning, before any other procedures. Thresholds for eliciting a head-withdrawal was evaluated using a series of von Frey filaments with increasing strengths from 0.6 g to 60 g (Stoelting; Wood Dale, IL)^{5,7}. Three rounds of reflex testing were performed on the bilateral TMJs, with a 10-minute rest period between each round; each von Frey filament was applied five times to each side. If three of the five stimulations elicited a head withdrawal or immediate pawing of the stimulated areas, a positive response was recorded. The lowest-strength filament evoking a response was taken as the sensitivity threshold only if the next higher filament also elicited a response. If a rat was unresponsive to all filaments, the maximum filament strength (60 g) was recorded as the threshold.

Procedures for TMJ loading. TMJ sensitivity was induced by mechanical loading of the jaw by mouth-opening^{5,7}. Briefly, rats were anesthetized with isoflurane inhalation anesthesia (4–5% for induction; 2–3% for maintenance) mixed with oxygen during loading. Rats were placed in a ventilated acrylic chamber in the prone position. The mandible was held stationary with a nylon loop and the maxilla was opened by a sling attached to 0N, 2N, or 3.5N load for 1 hour/day (Fig. 2), applied daily for 7 continuous days (days 0 to 6). Rats were monitored during recovery from anesthesia. Weight and activity were monitored daily; no decreases in weight (>10% body weight) or activity were observed during the course of the study.

Safranin O/Fast Green staining & Mankin scoring of the TMJ. Cartilage structure was assessed in naïve TMJs, 2N-loaded TMJs at day 15, and 3.5N-loaded TMJs at days 8 and 15 (n = 4/group). Rats were anesthetized with sodium pentobarbital (65 mg/kg), perfused with phosphate buffer saline (PBS), and fixed by 4% paraformaldehyde perfusion. TMJs were harvested, stored in 30% sucrose in PBS at 4 °C, and later decalcified using 0.25 M ethylenediaminetetraacetic acid for 3 weeks at 4 °C. Samples were embedded in Tissue-Tek OCT Compound (Saukura Finetek; Torrance, CA), sagittally sectioned (18 µm thickness), and thaw-mounted onto slides (3 sections/slide) to assess TMJ structure. TMJs were stained with Safranin O/Fast Green (Sigma Aldrich; St. Louis, MO) to evaluate morphological changes in the articular cartilage early after loading (day 8) and at a later time point (day 15). Stained TMJs were imaged at 10X magnification using the EVOS FL Auto Imaging System (Thermo Fisher Scientific, Waltham, MA). Cartilage degradation was measured by two blinded observers using the modified Mankin scale^{37,40}. The Mankin scale scores cellular and background staining, chondrocyte arrangement, and structural condition of the cartilage, with a score of 0 for normal cartilage and 10 for maximally degenerate cartilage^{37,40}.

Data analyses & statistics. All data are expressed as mean ± standard deviation. The mean and standard deviation of RGS scores were calculated across all rats in each group for each time point. A repeated measures ANOVA tested for differences between days in RGS measurements for the 3.5N group.

The average withdrawal threshold for each rat was calculated from the three thresholds recorded for the left and right TMJ regions (6 measurements total) at each time. The mean and standard deviation of reflex thresholds were calculated across all rats in the 3.5N loading group for each day. Due to a non-normal distribution of the mechanical reflex data for the 3.5N loading group, the Kruskal-Wallis rank sum test was used to compare head withdrawal threshold between days in that group and the threshold at a given day was specifically compared to baseline with the pairwise Wilcoxon rank sum test, corrected for multiple comparisons. For that same group, the Spearman correlation coefficient was calculated between RGS and head withdrawal thresholds to quantify the relationship between those two measurements.

When comparing RGS values between loading groups, an RGS difference score was determined by subtracting the baseline scores from the scores measured at a given time point. Difference scores were compared between groups using a repeated measures ANOVA with day and group as factors. A Tukey post-hoc test compared groups on specific days. All statistical tests were performed in R software (version 3.2.3, The R Foundation for Statistical Computing) and results were considered statistically significant at $p < 0.05$.

A one-way ANOVA compared Mankin scores for the 3.5N loading group across time points and relative to normal tissue. A t-test compared normal Mankin scores to the 2N loading group at day 15.

Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

- Mogil, J. S. & Crager, S. E. What should we be measuring in behavioral studies of chronic pain in animals? *Pain* **112**, 12–15 (2004).
- Loeser, J. D. & Melzack, R. Pain: An overview. *Lancet* **353**, 1607–1609 (1999).
- Mogil, J. S. Animal models of pain: Progress and challenges. *Nat. Rev. Neurosci.* **10**, 283–294 (2009).
- Zaki, J., Wager, T. D., Singer, T., Keyser, C. & Gazzola, V. The anatomy of suffering: understanding the relationship between nociceptive and empathic pain. *Trends Cogn. Sci.* **20**, 249–259 (2016).
- Nicoll, S. B., Hee, C. K., Davis, M. B. & Winkelstein, B. A. A rat model of temporomandibular joint pain with histopathologic modifications. *J. Orofac. Pain* **24**, 298–304 (2010).
- Miller, R. E. *et al.* CCR2 chemokine receptor signaling mediates pain in experimental osteoarthritis. *Pnas* **109**, 2–7 (2012).
- Kartha, S., Zhou, T., Granquist, E. J. & Winkelstein, B. A. Development of a rat model of mechanically induced tunable pain and associated temporomandibular joint responses. *J. Oral Maxillofac. Surg.* **74**, 54.e1–54.e10 (2016).

8. Liu, X.-J. *et al.* TLR signaling adaptor protein MyD88 in primary sensory neurons contributes to persistent inflammatory and neuropathic pain and neuroinflammation. *Sci. Rep.* **6**, 28188 (2016).
9. Ita, M., Crosby, N., Bulka, B. & Winkelstein, B. Painful cervical facet joint injury is accompanied by changes in the number of excitatory and inhibitory synapses in the superficial dorsal horn that differentially relate to local tissue injury severity. *Spine (Phila. Pa. 1976)* (2016).
10. Kartha, S. *et al.* Superoxide dismutase-loaded porous polymersomes as highly efficient antioxidants for treating neuropathic pain. *Adv. Healthc. Mater.* **6**, 1–6 (2017).
11. Whittaker, A. L. & Howarth, G. S. Use of spontaneous behaviour measures to assess pain in laboratory rats and mice: How are we progressing. *Appl. Anim. Behav. Sci.* **151**, 1–12 (2014).
12. Vachon-Pressseau, E. *et al.* The emotional brain as a predictor and amplifier of chronic pain. *J. Dent Res* **95**, 605–612 (2016).
13. Dworkin, R. H. An overview of neuropathic pain: Syndromes, symptoms, signs, and several mechanisms. *Clin. J. Pain* **18**, 343–349 (2002).
14. Foss, J. M., Apkarian, A. & Chialvo, D. Dynamics of pain: Fractal dimension of temporal variability of spontaneous pain differentiates between pain states. *J. Neurophysiol.* **95**, 730–736 (2005).
15. Backonja, M. M. & Stacey, B. Neuropathic pain symptoms relative to overall pain rating. *J. Pain* **5**, 491–497 (2004).
16. Langford, D. J. *et al.* Coding of facial expressions of pain in the laboratory mouse. *Nat Methods* **7**, 447–449 (2010).
17. Liao, L. *et al.* Evaluation of pain in rats through facial expression following experimental tooth movement. *Eur. J. Oral Sci.* **122**, 121–124 (2014).
18. De Rantere, D., Schuster, C. J., Reimer, J. N. & Pang, D. S. J. The relationship between the Rat Grimace Scale and mechanical hypersensitivity testing in three experimental pain models. *Eur. J. Pain (United Kingdom)* **20**, 417–426 (2016).
19. Philips, B., Weisshaar, C. & Winkelstein, B. The rat grimace scale can evaluate spontaneous neuropathic pain in a model of cervical radiculopathy. *Comp. Med.* **67**, 34–42 (2016).
20. Leung, V., Zhang, E. & Pang, D. S. J. Real-Time application of the Rat Grimace Scale as a welfare refinement in laboratory rats. *Sci. Rep.* **6**, 1–12 (2016).
21. Williamson, A. & Hoggart, B. Pain: a review of three commonly used rating scales. *J. Clin. Nurs.* **14**, 798–804 (2005).
22. Neogi, T. The epidemiology and impact of pain in osteoarthritis. *Osteoarthr. Cartil.* **21**, 1145–1153 (2014).
23. Plesh, O., Curtis, D., Levine, J. & McCall, W. D. Amitriptyline treatment of chronic pain in patients with temporomandibular disorders. *J. Oral Rehabil.* **27**, 834–841 (2000).
24. Côté, P., Cassidy, J. D., Carroll, L. J. & Kristman, V. The annual incidence and course of neck pain in the general population: A population-based cohort study. *Pain* **112**, 267–273 (2004).
25. Thakur, M., Dickenson, A. H. & Baron, R. Osteoarthritis pain: nociceptive or neuropathic? *Nat. Rev. Rheumatol.* **10**, 374–80 (2014).
26. Ekman, P. & Friesen, W. The repertoire of nonverbal behavior: categories, origins, usage, and coding. *Semiotica* **1**, 49–98 (1969).
27. Grunau, R. V. E. & Craig, K. D. Pain expression in neonates: facial action and cry. *Pain* **28**, 395–410 (1987).
28. Sotocinal, S. G. *et al.* The Rat Grimace Scale: A partially automated method for quantifying pain in the laboratory rat via facial expressions. *Mol. Pain* **7**, 55 (2011).
29. Whittaker, A. L., Leach, M. C., Preston, F. L., Lymn, K. A. & Howarth, G. S. Effects of acute chemotherapy-induced mucositis on spontaneous behaviour and the grimace scale in laboratory rats. *Lab. Anim.* **50**, 108–118 (2016).
30. Scrivani, S. J., Keith, D. A. & Kaban, L. B. Temporomandibular disorders. *N Engl J Med* **359**, 2693–2705 (2008).
31. NIDCR. Facial Pain. Available at, <http://www.nidcr.nih.gov/datastatistics/finddatabytopic/facialpain/> (2014).
32. Wilcox, S. L. *et al.* Anatomical changes within the medullary dorsal horn in chronic temporomandibular disorder pain. *Neuroimage* **117**, 258–266 (2015).
33. Youssef, A. M. *et al.* Differential brain activity in subjects with painful trigeminal neuropathy and painful temporomandibular disorder. *Pain* **155**, 467–475 (2014).
34. Villa, G. *et al.* Temporomandibular joint inflammation activates glial and immune cells in both the trigeminal ganglia and in the spinal trigeminal nucleus. *Mol. Pain* **6**, 89 (2010).
35. Chen, Y. *et al.* Temporomandibular joint pain: A critical role for Trpv4 in the trigeminal ganglion. *Pain* **154**, 1295–1304 (2013).
36. Krzyzanowska, A. & Avendano, C. Behavioral testing in rodent models of orofacial neuropathic and inflammatory pain. *Brain Behav.* **2**, 678–697 (2012).
37. Shen, P. *et al.* Injecting vascular endothelial growth factor into the temporomandibular joint induces osteoarthritis in mice. *Sci. Rep.* **5**, 16244 (2015).
38. Mino-oka, A., Izawa, T., Shinohara, T., Mori, H. & Yasue, A. Roles of hypoxia inducible factor-1 α in the temporomandibular joint. *Arch. Oral Biol.* **73**, 274–281 (2017).
39. Rogers, A. W., Cisewski, S. E. & Kern, C. B. The Zonal Architecture of the Mandibular Condyle Requires ADAMTS5. *J. Dent. Res.* **1–8**, <https://doi.org/10.1177/0022034518777751> (2018).
40. Xu, L. *et al.* Early-onset osteoarthritis of mouse temporomandibular joint induced by partial disectomy. *Osteoarthr. Cartil.* **17**, 903–908 (2009).
41. Asgar, J. *et al.* The role of TRPA1 in muscle pain and mechanical hypersensitivity under inflammatory conditions in rats. *Neuroscience* **310**, 206–215 (2015).
42. Tanaka, E., Detamore, M. S. & Mercuri, L. G. Degenerative disorders of the temporomandibular joint: etiology, diagnosis, and treatment. *J. Dent. Res.* **87**, 296–307 (2008).
43. Kuroda, S. *et al.* Biomechanical and biochemical characteristics of the mandibular condylar cartilage. *Osteoarthr. Cartil.* **17**, 1408–1415 (2009).
44. Ita, M. E., Zhang, S., Holsgrove, T. P., Kartha, S. & Winkelstein, B. A. The Physiological Basis of Cervical Facet-Mediated Persistent Pain: Basic Science and Clinical Challenges. *J. Orthop. Sport. Phys. Ther.* **47**, 450–461 (2017).
45. Sperry, M. M. *et al.* The interface of mechanics and nociception in joint pathophysiology: insights from the facet and temporomandibular joints. *J. Biomech. Eng.* **139**, 021003 (2017).
46. Oliver, V. *et al.* Psychometric assessment of the rat grimace scale and development of an analgesic intervention score. *PLoS One* **9** (2014).
47. Stasiak, K. L., Maul, D. O. N. & French, E. Species-Specific Assessment of Pain in Laboratory Physiology of Pain. *Comp. Med.* **42**, 13–20 (2003).
48. Klinck, M. P. *et al.* Translational pain assessment: Could natural animal models be the missing link? *Pain* **158**, 1633–1646 (2017).
49. Iqbal, S. M. *et al.* Lubricin/Proteoglycan 4 binds to and regulates the activity of Toll-Like Receptors *In Vitro*. *Sci. Rep.* **6**, 1–12 (2016).
50. Sorge, R. E. *et al.* Different immune cells mediate mechanical pain hypersensitivity in male and female mice. *Nat. Neurosci.* **18**, 1–5 (2015).
51. Wang, X. D., Kou, X. X., Mao, J. J., Gan, Y. H. & Zhou, Y. H. Sustained Inflammation Induces Degeneration of the Temporomandibular Joint. *J. Dent. Res.* **91**, 499–505 (2012).
52. Tashiro, A., Okamoto, K. & Bereiter, D. A. Chronic inflammation and estradiol interact through MAPK activation to affect TMJ nociceptive processing by trigeminal caudalis neurons. *Neuroscience* **164**, 1813–1820 (2009).
53. Yang, H. *et al.* TNF accelerates death of mandibular condyle chondrocytes in rats with biomechanical stimulation-induced temporomandibular joint disease. *PLoS One* **10**, 1–18 (2015).

54. Tuttle, A. H. *et al.* A deep neural network to assess spontaneous pain from mouse facial expressions. *Mol. Pain* **14**, 174480691876365 (2018).
55. Institute for Laboratory Animal Research. *Guide for the care and use of laboratory animals: 8th Ed. Guide for the Care and Use of Laboratory Animals*, <https://doi.org/10.2307/1525495>, <https://grants.nih.gov/grants/olaw/guide-for-the-care-and-use-of-laboratory-animals.pdf> (2011).
56. Zimmermann, M. Ethical guidelines for investigations of experimental pain in conscious animals. *Pain* **16**, 109–110 (1983).

Acknowledgements

The authors thank Dr. Blythe Phillips and Christine Weisshaar for their consultation on RGS protocols and scoring procedures, Harrison Troché for cropping and organizing the images for RGS scoring, Jessie Frank and Vanessa Moody for Mankin scoring, and Sonia Kartha for guidance and help with some of the TMJ loading and tissue harvest protocols. We also thank the Small Animal Imaging Facility at the University of Pennsylvania for assistance with CT image acquisition and the Cell and Developmental Biology Microscopy Core at the University of Pennsylvania for assistance with microscopy. This project was funded by the Catherine D. Sharpe Foundation, the Oral and Maxillofacial Surgery Foundation, the Oral and Maxillofacial Surgery Schoenleber Research Fund from the University of Pennsylvania School of Dental Medicine, and fellowship training from the NIH/NIAAMS (T32-AR007132).

Author Contributions

M.M.S., E.J.G. and B.A.W. planned the study design. M.M.S. and Y.Y. acquired RGS videos and performed the studies. M.M.S. performed reflex testing. R.L.W. scored the RGS images. M.M.S. compiled data and performed statistical tests. M.M.S. and Y.Y. prepared the figures. M.M.S., E.J.G. and B.A.W. drafted the manuscript and all authors reviewed the manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018