Quantifying Patient Sentiment of Treatments for Lateral Elbow Tendinopathy Using Machine Learning

McGinley, BM.[1]; Whittingslow, DC[1]; Ghareeb, PA [1,2] ; Lamplot, JL [1,2]; Bowers, RL [1,2]; Daly, CA [1,2]; Wagner, ER [1,2]; Gottschalk, MB [1,2]
1. Emory University School of Medicine 2. Emory Orthopaedics & Spine Center

emoryhealthcare.org/ortho

Objectives:

- In the management of Lateral Elbow Tendinopathy (LET), no best practices exist following failure of conservative management. To better understand patient perceptions amidst clinical ambiguity, we performed a sentiment analysis of tweets regarding various treatments of LET.

Methods:

- A computer program was written to search Twitter and extract data on all posts with specific key words. The data collected from each tweet included the username, body of the tweet, and number of likes, comments, and retweets. A proof-of-concept test was performed before analyzing tweets involving LET with conservative management, platelet rich plasma (PRP), steroids, or surgery. Advertisements were filtered and removed.

- Each tweet was input into multiple natural language processing (NLP) programs. The performance of each NLP program was assessed using the aforementioned proof-of-concept dataset to determine which NLP would work best for our LET dataset. We then analyzed the sentiment of the tweet dataset related to LET. Based on their key words, tweets were divided into four groups: conservative management, PRP, surgery, or steroids. Mean sentiment scores were calculated and statistically compared for each group.

- Kolmogorov-Smirnov-Lilliefors test of normality was performed. Mean sentiment scores for tweets containing each key word were compared using the nonparametric Kruskal-Wallis one way test of variance with Tukey's Post Hoc Analysis.
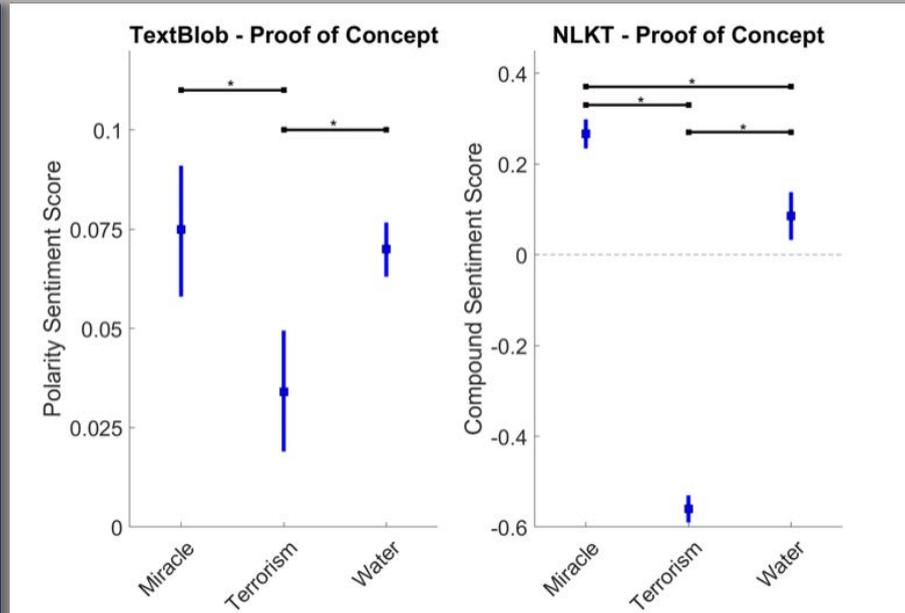

Fig 1. Proof of Concept Results Comparing TextBlob to NLTK. The NLKT algorithm found significance between each of the three control groups, whereas the Textblob algorithm could not differentiate the sentiment of miracle (positive) and water (neutral), therefore, NLKT was determined to be more suited for our later LET treatment comparison (Blue bars = confidence intervals. * indicates statistical significance (p<0.05)).

Fig 3. Number or tweets, Likes, Retweets, and Comments for each group of tweets analyzed.
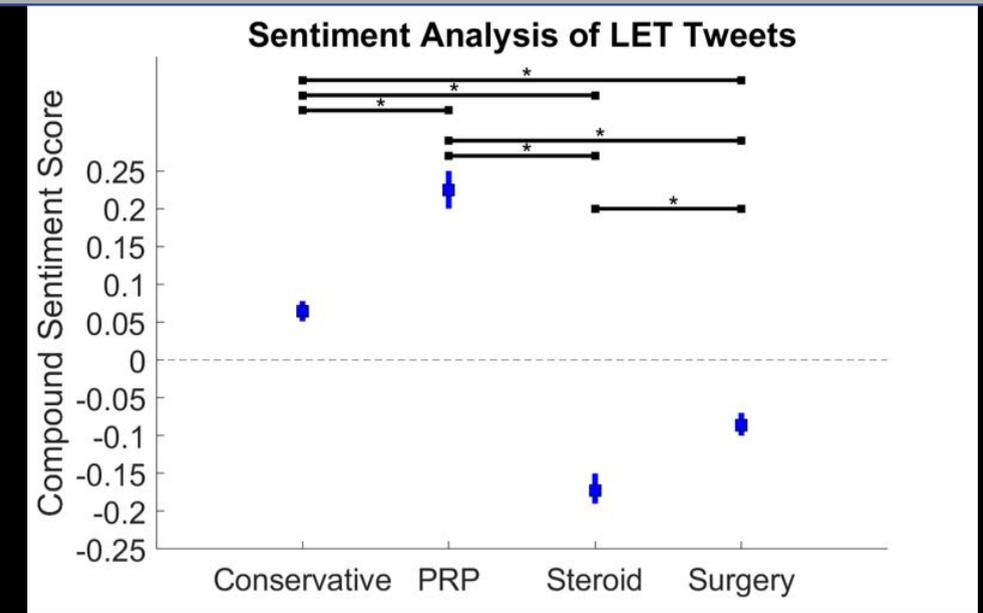

Fig 2. Sentiment Comparison using NLTK algorithm on LET Tweets. Average compound scores for each group is depicted. More negative scores indicate a more negative sentiment with a score of 0 indicating neutrality. (Blue bars = confidence intervals. * indicates statistical significance (p<0.05))
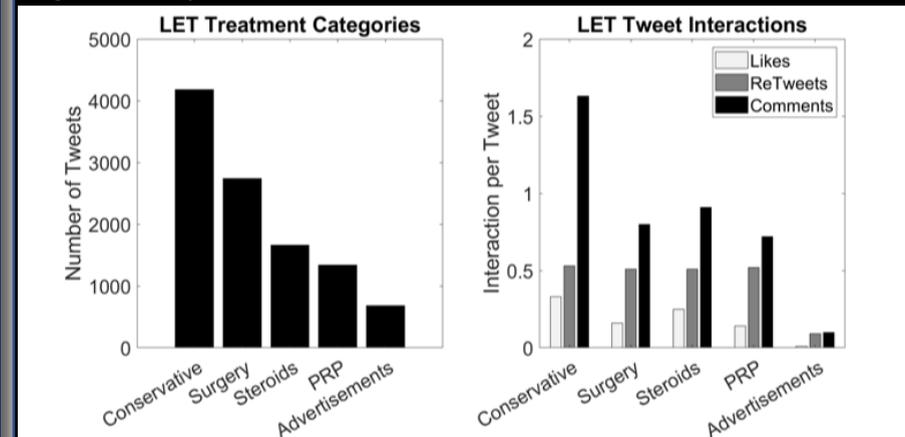
Results:

One NLP program calculated significantly different sentiment scores for groups of negative, neutral, and positive tweets (Fig 1). Using the same NLP program, A total of 10,602 tweets related to LET were analyzed. Mean sentiment scores were significantly different (p<0.05) between each group, descending from PRP (0.22) to conservative (0.06) to surgery (-0.09) to steroids (-0.17) (Fig 2.). The conservative management group had the most likes, retweets, and comments per tweet (Fig 3).

Conclusions:

For the management of LET, social media users portray PRP most positively, followed by conservative management, surgery, and steroids in descending order. Further, PRP and conservative therapy have positive sentiment on average, while surgery and steroids were negatively portrayed by twitter users.