Methods

Comments were collected from videos submitted by users of the YouTube video streaming service between [start date] and [end date]. The videos were affiliated with wearing one of the following including braces, Invisalign, or direct to consumer aligners. Using [insert the methods used for sentiment analysis], sentiment analysis was performed to add scores to the comments. Each comment was given a positive score and a negative score according to the sentiment analysis. Positive scores ranged from 0 to 5, and indicated favorable experiences with the specific wearables. Negative comments ranged from -5 to 0, and indicated unfavorable experiences.

Comments were categorized into positive, negative, or neutral classes. To do so, positive and negative scores were summed together. Comments whose scores summed to greater than 0 were denoted as a positive class, while those that summed to less than 0 were denoted as a negative class. In the case where scores summed to 0, comments were denoted as neutral. That is, there was not a clear indication from the sentiment analysis whether the user that provided said comment had a favorable or unfavorable experience with the specific wearable.

Wordclouds were generated for each combination of wearable by positive and negative comments for the top 200 most frequent words using the *wordcloud* package (version 2.6). Rudimentary text analysis was performed to clean the comments. Punctuations, numbers, words of length 1, and stop words were all removed. Additionally, specific words including "get", "getting", "got", "just", "can", "now", "im", "will", "braces", "invisalign", "direct", and "teeth" were removed prior to creating the wordcloud.

We further reported frequencies and percentages of comments in each stratum. A Chi-square test of independence was performed to assess the association between wearable type and class of comment. The neutral class of comments were removed from statistical analysis. All statistical analyses was performed in R (version 3.6.1).

Results

A total of 8438 comments were collected from videos submitted on YouTube between [start date] and [end date]. The majority of comments were in response to braces-related videos with 5766 (65.3%), while Invisalign and direct to consumer had 2418 (28.7%) and 254 (3.0%) comments each. After categorized the comments into classes, we found that there were 2762 positive (32.7%) and 2387 negative (28.3%) comments. The remaining 3289 (39.0%) comments were neutral and were removed from statistical analyses. Positive and negative comments resulted in an overall average score of 1.49 (sd=0.64) and -1.75 (sd=0.79), respectively.

The cross-tabulation of wearable type and comment class is provided in Table 1. There were 1838 (53.1%) comments that were categorized as braces-related and positive. The most frequent terms included “like”, “love”, “excited”, and “beautiful” (Figure 1). There were 833 (54.8%) and 91 (53.8%) comments that were categorized as Invisalign-related and positive, and direct to consumer-related and positive respectively. The most frequent terms were found to be similar to that of the braces-related and positive (Supplementary Figures 1 (A-B)). Additionally, there were 1621 (46.9%) comments that were categorized as braces related and negative. The most frequent terms included words like “hurt”, “bad”, “pain”, and “hate” (Figure 2). There were 688 (45.2%) and 78 (54.8%) comments that were categorized as Invisalign-related and positive, and direct to consumer-related and positive, respectively. As before, the most frequent terms were found to be similar to that of the braces-related and negative (Supplementary Figures 2 (A-B)). Results from the Chi-square test of independence indicated that there was not a significant association between wearables and positive or negative comments (p= 0.5679).