Culturomic Analysis of Trends in Otolaryngology
Aaron D. Tward, M.D., Ph.D., and Daniel J. Lee M.D., F.A.C.S.
Department of Otolaryngology, Massachusetts Eye & Ear Infirmary, Boston, MA
Department of Otology and Laryngology, Harvard Medical School, Boston, MA

ABSTRACT

Objective: 1. Quantitatively describe historical trends in diagnoses and treatments in otolaryngology as measured by frequency of word use. 2. Quantitatively describe historical trends in terminology use in otolaryngology.

Method: A variety of search terms of n-grams (words or short phrases) specific to the practice of otolaryngology was used to search the English Language Corpus of Google Books. Trends in n-gram frequency across years were correlated to known historical milestones in otolaryngology.

Results: A variety of search terms was used the results of which quantitatively document historical trends in otolaryngology. For example, the frequency of “stapes” in the corpus increased sharply around 1957, correlated with increases in the frequency of the terms “otosclerosis” and “stapedectomy”, whereas “incus” showed no similar increase in frequency. These frequencies correlate well with the popularization of stapedectomy for otosclerosis. Many additional analyses are presented with similarly strong correlations. In addition, trends of word use in otolaryngology were quantified by the same technique. Conclusion: Culturomic searches accurately demonstrate historical trends in otolaryngology, permitting both quantitative analysis of known trends and revelation of previously unknown trends in diagnosis, treatment, and terminology.

INTRODUCTION/ METHODS

The study of historical trends is typically based upon the directed analysis of a small number of relevant written works, which are then used to draw inferences about the concept being studied. Such an approach, although powerful, introduces a number of intrinsic biases, and does not provide quantitative information about the concept being studied (1). As part of the Google’s book scanning project, 11% of all books ever written have been scanned, and a digitized and indexed corpus of words and phrases (n-grams) representative of 4% of all books ever published was generated. The result is a searchable database that permits the frequency of each n-gram relative to all n-grams in English to be assessed from the years of 1500-2000 (1). The frequency of n-gram use in books provides a direct reflection of trends in terminology and an indirect reflection of the use of diagnoses and treatments within a medical field. We have selected candidate terms on the basis of both previously known milestones in otolaryngology and from terminology in common use and used the n-gram viewer at googlelabs to search the English language corpus (2). All graphs plot n-gram frequency as a fraction of all n-grams (ordinate) for a given year (abscissa).

RESULTS

Figure 2: Diseases of the mucous membranes. As described by Dr. Lighthill in 1865: “Catarrh is a disease of the mucous membrane of the nasal passages and those cavities of the head communicating with them (magnified as it appears in its first stages, it is apt to progress to become instrumental in causing the tear or impairment of smell, taste, hearing and sight, and of creating serious constitutional derangements, not infrequently terminating in consumption.” (3)

Figure 3: Reconstructive options in Head and Neck Surgery

Figure 4: Otosclerosis and the popularization of stapedectomy

Figure 5: The genesis of laryngoscopy

Figure 6: Subspecialties in Otolaryngology: Head and Neck Surgery

Figure 1: Trends in popular terminology for hearing loss

DISCUSSION/CONCLUSIONS

The use of quantitative methods to track historical trends based upon word frequencies is viable, even in a field as specialized as Otolaryngology. Future work aimed at correlating n-gram use frequencies with well documented epidemiological data may further validate this approach as a rapid, robust method to track diagnoses and treatments. This approach may also be expanded beyond this pilot study to compare trends in different languages and quantify cultural differences in interest and application of the ideas being studied.

REFERENCES