**Chapter Five – Discrimination in Treatment of Minority Groups**

**Abstract**

**This report shows results from the *use of algorithms in Excel* that provide descriptions of health disparities. A significant advantage is that the analyst does not need background knowledge of the subject. Instead, the world’s author-specialists serve as mentor-guides via their ideas. As such, students can learn new topics rapidly and begin to develop new descriptions and knowledge-generating strategies.**

**Introduction**

**In a sense, critical thinking can be transformed from quoting past authors to one of considering the world’s view of the ideas composing a topic. Translating those observations/opinions into quantitative evidence is an example of data-driven critical thinking.**

**This report described the steps involved in comparing disease and treatment in a composite group labeled as *Disparities* with a specific subgroup labeled as *LGBT* (Lesbians, Gays, Bisexuals, and Transgenders). The comparison is an example of the many possible when the total population is divided into subgroups. The LGBT group is defined by sexual orientation rather than demographic or socioeconomic characteristics.**

**Historically, linguists employed frequency of use of a key term as a measure in assessing importance in communication. Those frequently used informative terms (i.e., nouns, adjectives, or gerunds) were considered of greater importance in determining meaning. The premise was that subject experts would tend to employ favored terms in describing their subject.**

**As with key terms and their interpretation, frequency of occurrence of ideas was considered to be an indicator of importance. In addition to presenting simple pairs of terms, authors may tend to use a specific pair as a bridge to more complex ideas. This is done by including additional informative terms within the same sentence. In considering health disparities, the additional terms might be descriptors of sexual minorities. As such, terms such as bisexual, gay, lesbian, and transgender could be included with terms such as health disparities and specific treatments. These complex combinations would express thoughts that were both specific and relevant.**

**Exhibit 1 shows examples of sentences and corresponding ideas. The document is identified by the PubMed assigned identification number. Software-identified informative terms are highlighted in red. Missed terms are in blue. The idea set consists of the terms representing the idea, the time period entered into PubMed, the assigned identification number and the number of the sentence containing the idea.**

**Exhibit 1. Sentences and Idea Sets Illustrating Determination of Capture Accuracy.**

***For women, having difficulty taking medications openly at home was associated with a substantial decrease in the probability of being on HAART* (highly active antiretroviral therapy) *in the adjusted model (0.59, 95% CI 0.47-0.70 vs. 0.78, 95% CI 0.74-0.83), whereas no significant differences were observed for heterosexual or gay/bisexual men.* *(16536681 – 2006)* (Sayles 2006). The ideas extracted by the text analysis software included:**

***Primary Related Time Period Doc ID Sentence***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Bisexual** | **Gay** | **2005-2009** | **16536681** | **7** |
| **Bisexual** | **heterosexual** | **2005-2009** | **16536681** | **7** |
| **Bisexual** | **Sexual** | **2005-2009** | **16536681** | **7** |
| **Bisexual** | **Women** | **2005-2009** | **16536681** | **7** |

***We also investigated the possible role of HIV infection among gay men and higher rates of psychological distress among lesbians, gay men, and bisexually and homosexually experienced heterosexual individuals in generating these health disparities. (17463371 – 2007)* (Cochrane 2007). Ideas extracted included:**

***Primary Related Time Period Doc ID Sentence***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Bisexual** | **Disparities** | **2005-2009** | **17463371** | **2** |
| **Bisexual** | **Distress** | **2005-2009** | **17463371** | **2** |
| **Bisexual** | **Gay** | **2005-2009** | **17463371** | **2** |
| **Bisexual** | **Health** | **2005-2009** | **17463371** | **2** |
| **Bisexual** | **heterosexual** | **2005-2009** | **17463371** | **2** |
| **Bisexual** | **Hiv** | **2005-2009** | **17463371** | **2** |
| **Bisexual** | **Infection** | **2005-2009** | **17463371** | **2** |
| **Bisexual** | **Lesbian** | **2005-2009** | **17463371** | **2** |
| **Bisexual** | **men** | **2005-2009** | **17463371** | **2** |
| **Bisexual** | **psychological** | **2005-2009** | **17463371** | **2** |
| **Bisexual** | **Sexual** | **2005-2009** | **17463371** | **2** |

**The advantages associated with displaying ideas as pairs of informative terms include:**

1. **The identification of specific ideas in building idea structures.**
2. **The ability to combine the pairs into more complex arrangements.**

**These advantages suggest that the hypothesis of interest could be – *Can idea frequency and/or type be measures of medical discrimination in dealing with patients from sexual orientation minorities?***

**Procedures in Text Analysis**

**The scholarly literature dealing with health disparities were retrieved from PubMed for the period 1990 through July 2015. There were 11,553 documents and contained 2,279,032 ideas.**

**The analytic software performs the following functions:**

1. **Separates the text into individual sentences using punctuation to recognize endings.**
2. **Identifies informative terms (nouns, adjectives, or gerunds) using characteristic endings and contextual relationships. The latter captures terms that authors have linked with recognized informative terms. Those terms may have endings different from those usually used in describing the grammatical words of interest. Examples of these contextually captured terms are: health, disparities, gay, lesbian, transgender. The term – bisexual – would be captured using the adjective ending.**
3. **Combine informative terms in pairs within each sentence.**
4. **Generate idea records consisting of the pair of terms and bibliographic data indicating the involved document, date of publication and the location of the sentence within the document.**
5. **Store these idea records in excel files for subsequent analytic use.**

**The idea analysis software’s performance across topics yielded a median of 85% (66% - 99%) capture of vocabulary used by the authors. In the disparity subject, the capture of informative terms was 99%. Capture of ideas exceeded 95% across topics and over 99% in analysis of disparity ideas.**

**The hypotheses that could be generated include:**

1. **Ideas can act as measures of different subgroups in the population.**
2. **Differences in use of ideas may be definable characteristics of individual subgroups.**
3. **Idea frequency is a viable attribute useful in comparing different subgroups.**
4. **Idea frequency offers a way to translate observations made by specialists into quantitative evidence.**

**Different Ideas – Different Sub-Groups**

**The process applied to health disparities required managing of approximately 2.2 million ideas. Preliminary organizations suggested the possibility that discrimination of medical care in terms of access and treatment might be different in different subgroups of the population. Specific groups included – age, ethnicity, gender, race, socioeconomic groups and minorities identified by sexual orientation.**

**Traditional methods of analysis of literature have been an integral part of knowledge generation and utilization. When performed using established manual methods, the results could be considered as summary opinions rather than evidence. This leads to a paradox between evidence based on observation/opinion vs. that resulting from formalized study.**

**In a situation where *description is the standard*, putting the disparate observations/opinions together could be a way to build the required evidence. However, in doing so, what is needed is an effective measure of the phenomenon. That requirement is satisfied by the idea – a combination of informative terms linked by the author-specialist and presented within the domain of the sentence. Ideas are independent building blocks reflecting specific aspects of the conceptual structure under consideration. The question considered in this report was – *could ideas from the world’s authors be translated into quantitative evidence supporting a new description of the topic as well as hypotheses suitable for further testing?***

**Summary**

**The Contextual Analysis approach is a way to organize millions of ideas presented by world’s specialists. The result is a transparent and formalized path to critical thinking. By employing algorithms intended to organize and clarify the individual data points (observations and or opinions), students could duplicate the creative paths developed by specialists. Accelerating management of large volumes of information is a necessity in a world of search engines delivering millions of documents and websites. In a similar fashion, computer-supported algorithms make the processing feasible.**

**The idea analysis approach also has the advantage of being quantitative providing evidence that is more objective. If so, the data suggest that medical treatment discrimination is present, as has been claimed by numerous authors through time.**