Developing a Visual Analytics Tool to Explore the Readability Levels of Health-related Documents

Himaja Chintalapalli[1], Anunita Nattam[1], Hexuan Liu, PhD[2], Danny T.Y. Wu, PhD, MSI[1]

[1]College of Medicine; [2]College of Education, Criminal Justice, and Human Services, University of Cincinnati, Cincinnati, OH

Introduction

Readability assessment is crucial for ensuring that written materials are comprehensible and accessible to diverse audiences, especially in the field of healthcare since it empowers patients to make informed health decisions.[1–3] Traditional readability measures, such as the Flesch-Kincaid Grade Level (FKGL), Gunning Fog Index (GFI), and Simple Measure of Gobbledygook (SMOG), have been widely adopted as valuable tools for evaluating the complexity of written materials.[4–6] These measures offer a standardized approach to assessing text difficulty. To better support effective communication, more advanced tools are needed to evaluate readability with greater consistency and accuracy. In this study, we aimed to develop a visual analytics tool that can assess readability, offering a more standardized analysis to answer common research questions.

Methods

The study began with synthesizing readability analysis methods and visualizations by conducting a small-scale review. The review aimed to establish a foundation by identifying common analysis and visualization types, which were summarized in Table 1. Specifically, the top 30 articles from PubMed (keywords: "readability assessment" and "health information") [7–14] were analyzed and included if it was accessible, not a review, and assessed the readability of health-related documents. A total of 6 papers [1–6] published from our previous work plus an additional 8 from the PubMed search were reviewed to extract the types of analysis. This review revealed 11 analysis types in three broad groups: Data Visualization (Bar Chart, Density Plot, Boxplot, Flowchart, and Pie Chart), Data Summary and Measures (Grouping, Frequency Distributions, Readability Measures), and Statistical Testing (Parametric Testing, Non-Parametric Tests, and Post-Hoc Tests). **(Table 1).** No additional analysis types were identified after the 10th paper in the PubMed search.

The included papers were synthesized to generate a set of common research questions. This was done by starting with the research questions of our latest publication (Nattam and Vithala et al.)[2] and adding the research questions of the remaining papers one at a time. This process was done by the first author and reviewed by the second author. Once the common research questions were synthesized, the visual analytics tool was designed accordingly to provide an interactive dashboard to answer the research questions. In other words, **Table 1** served as the basis for the design of the visual analytics tool, which considered the analysis groups with a goal to cover as many analysis types as possible.

Results

The synthesis generated four common research questions (RQs): 1) What is the distribution between the source and type of the health-related document? 2) Do readability measures generate similar scores in terms of grade level? 3) Is the overall grade level equal or below 8th grade as recommended? 4) Are the readability levels different among the sub-groups of documents? This resulted in 4 corresponding interactive visualizations in our visual analytics tool: 1) show the distribution of sources and topics using a pie chart (**Figure 1**) to answer the 1st RQ, 2) examine if the readability measures agree with each other using a side-by-side bar chart and Fleiss' Kappa (**Figure 2**) for the 2nd RQ, 3) test if the overall grade level is 8th grade or below using a bar chart and a non-parametric test when the distribution is not normal (**Figure 3**) for the 3rd RQ, 4) compare the readability levels among the groups using pair-wise test and Bonferroni correction and show the distribution in a density plot or a box plot (**Figure 4**) for the 4th RQ. Of note, all figures were generated based on the dataset of Nattam and Vithala et al.[2] as an example.

In terms of data pipeline, the data input of the tool includes a csv file with the first column being the document identifier, followed by a set of columns for pre-defined categories (e.g., source, content type) and another set of columns for readability scores (e.g., Flesch Kincaid Grade Level, Gunning Fog Index, SMOG Index). The data input also includes a dictionary file describing the column definitions.

Discussion

We summarized the analysis types and designed an R-shiny application to provide a standardized and complete solution for readability assessment. The app offers a systematic process to explore the readability scores and conduct hypothesis testing to answer common research questions in a readability assessment study. This process is independent of the readability measures used and the domain of the research questions. We are in the process of

implementing this application and aim to demonstrate the system in the VAHC workshop to collect additional feedback. Our future work involves formative evaluation (e.g., usability testing) when the app is fully developed and implementation of the app on a cloud-based platform for dissemination.

Table 1. This table shows the types of analysis found in each paper. The types of analyses were categorized into three sub-categories: Data Visualizations, Data Organization and Evaluation, and Statistical Testing.

	Types of Analysis Present in Articles										
	Data Visualization					Data Summary and Measures			Statistical Testing		
Types of Analysis/Articles	Pie Chart	Flowchart	Bar Chart	Density Plot	Box Chart	Grouping	Frequency Distributions	Readability Measures	Non-parametric Tests	Post-Hoc Tests	Parametric Testing
(Wu et al., 2016) PMID: 26269536	0	0	0	1	1	1	1	1	1	1	1
(Su et al., 2019) PMID: 31259026	0	0	0	0	1	1	1	1	1	1	1
(Karthik et al., 2022) PMID: 33832394	0	0	1	0	0	1	1	1	1	0	1
(Su et al., 2022) PMID: 35308941	0	0	1	0	0	1	1	1	1	0	1
(Wu et al., 2013) PMID: 23920636	0	0	0	0	1	1	1	1	0	1	1
(Nattam et al., 2023) PMID: 37647115	0	1	0	0	0	1	1	1	1	1	1
(Okuhara et al., 2021) PMID: 34682926	0	1	0	0	0	1	1	1	0	0	1
(Seth et al., 2016) PMID: 26818318	1	0	1	0	0	1	1	1	0	0	1
(Alkhuwaiter et. Al, 2024) PMID: 38764566	0	0	0	0	0	1	1	1	0	1	1
(Aboalshamat, 2024) PMID: 39196637	0	1	0	0	0	1	1	1	1	0	0
(Abu-Heija, 2019) PMID: 31106084	0	0	0	0	1	1	1	1	0	0	0
(Correa, 2020) PMID: 32096225	1	0	0	0	0	1	1	1	1	0	1
(Powell, 2021) PMID: 38993224	0	0	1	0	0	1	0	1	1	1	1
(Michel, 2022) PMID: 35819724	0	0	0	0	0	1	0	1	1	1	1

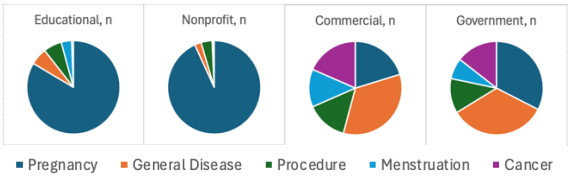

Figure 1. This pie chart illustrates the distribution between four different groups. It showcases the "Online OB/GYN PEMs sources based on topics". The data was sourced from Nattam and Vithala et al.[2]

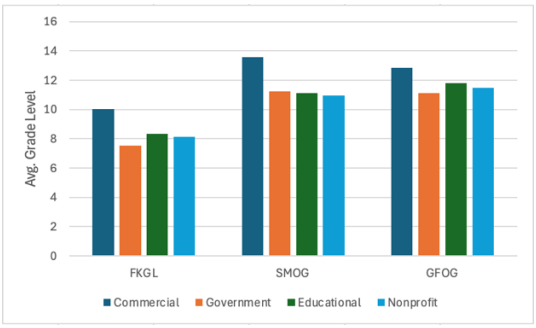

Figure 2. This figure displays the readability (grade) level by each source and readability measure. A dropdown menu allows users to select which category to use. Fleiss' kappa coefficient is calculated to show the agreement between multiple measures on the same set of ítems. The data was sourced from Nattam and Vithala et al.[2]

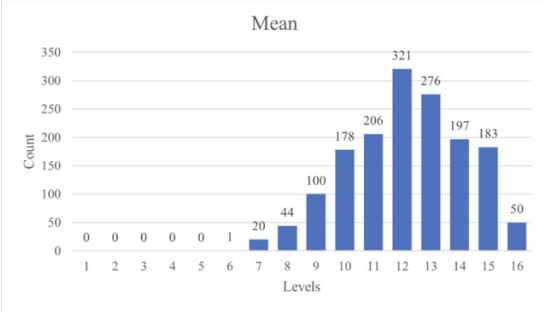

Figure 3. This bar chart compares the distribution of a readability measure against an 8th grade threshold. The user can choose which measure they want to use. A non-parametric test will determine if the overall grade level is significantly different from an 8th grade level. The data was sourced from Nattam and Vithala et al.[2]

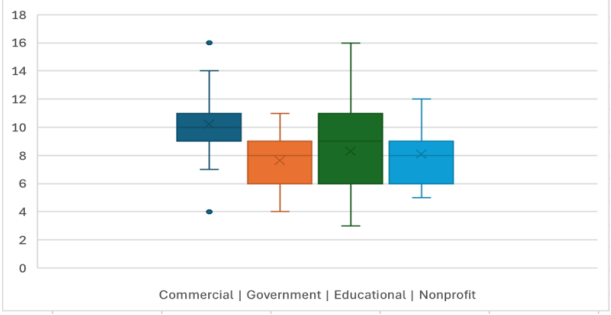

Figure 4. This figure displays the distribution of readability measures for three different groups. A density plot or box plot is used to visualize the distributions. Users can select the specific groups and measures to compare from two different dropdowns. The data was sourced from Nattam and Vithala et al.[2]

Reference

1. Wu DTY, Hanauer DA, Mei Q, et al. Assessing the readability of ClinicalTrials.gov. *J Am Med Inform Assoc JAMIA*. 2016;23(2):269-275. doi:10.1093/jamia/ocv062
2. Nattam A, Vithala T, Wu TC, et al. Assessing the Readability of Online Patient Education Materials in Obstetrics and Gynecology Using Traditional Measures: Comparative Analysis and Limitations. *J Med Internet Res*. 2023;25:e46346. doi:10.2196/46346
3. Su WC, Dufendach K, Wu DTY. Assessing the Readability of Freely Available ICU Notes. *AMIA Jt Summits Transl Sci Proc AMIA Jt Summits Transl Sci*. 2019;2019:696-703.
4. Karthik N, Barekatain K, Vu H, Wu DTY, Ehrlich JR. A READABILITY COMPARISON OF ONLINE SPANISH AND ENGLISH PATIENT EDUCATION MATERIALS ABOUT VISION HEALTH. *Ophthalmic Epidemiol*. 2022;29(2):182-188. doi:10.1080/09286586.2021.1910316
5. Wu DTY, Hanauer DA, Mei Q, et al. Applying multiple methods to assess the readability of a large corpus of medical documents. *Stud Health Technol Inform*. 2013;192:647-651.
6. Su WC, Mehta KY, Gill K, Yeh P, Chih MY, Wu DTY. Assessing the Readability of App Descriptions and Investigating its Role in the Choice of mHealth Apps: Retrospective and Prospective Analyses. *AMIA Annu Symp Proc AMIA Symp*. 2021;2021:1139-1148.
7. Okuhara T, Okada H, Goto E, Kiuchi T. Readability Assessment of HPV Vaccination and Cervical Cancer Information: A Systematic Scoping Review. *Healthcare*. 2021;9(10):1246. doi:10.3390/healthcare9101246
8. Seth AK, Vargas CR, Chuang DJ, Lee BT. Readability Assessment of Patient Information about Lymphedema and Its Treatment: *Plast Reconstr Surg*. 2016;137(2):287e-295e. doi:10.1097/01.prs.0000475747.95096.ab
9. Alkhuwaiter SS, Alssafi F, Alshunaiber R, Ahmed YB, Albdah B. Quality and Readability Assessment of Web-based Arabic Health Information on Early Childhood Caries. *Saudi J Med Med Sci*. 2024;12(2):188-193. doi:10.4103/sjmms.sjmms_443_23
10. Aboalshamat K. Assessment of the Quality and Readability of Web-Based Arabic Health Information on Halitosis: Infodemiological Study. *J Med Internet Res*. 2024;26:e54072. doi:10.2196/54072
11. Abu-Heija AA, Shatta M, Ajam M, Abu-Heija U, Imran N, Levine D. Quantitative Readability Assessment of the Internal Medicine Online Patient Information on Annals.org. *Cureus*. 2019;11(3):e4184. doi:10.7759/cureus.4184
12. Correa DJ, Milano L, Kwon CS, et al. Quantitative readability analysis of websites providing information on traumatic brain injury and epilepsy: A need for clear communication. *Epilepsia*. 2020;61(3):528-538. doi:10.1111/epi.16446
13. Michel C, Dijanic C, Abdelmalek G, et al. Readability assessment of patient educational materials for pediatric spinal deformity from top academic orthopedic institutions. *Spine Deform*. 2022;10(6):1315-1321. doi:10.1007/s43390-022-00545-1
14. Powell LE, Cisu TI, Klausner AP. Bladder Cancer Health Literacy: Assessing Readability of Online Patient Education Materials. *Bladder Cancer Amst Neth*. 2021;7(1):91-98. doi:10.3233/BLC-200387