Static versus interactive online resources about dementia: A comparison of readability scores

Julie M. Robillard PhD\textsuperscript{a,}* 
Arlo B. Sporn\textsuperscript{a}

\textsuperscript{a}Division of Neurology, Department of Medicine, University of British Columbia, Vancouver, Canada; *Corresponding author: jrobilla@mail.ubc.ca


Static and social online health resources are widely and easily accessed, but not necessarily easily understood. Recent guidelines suggest that health-related text should be written at a grade level between 5 and 6. This guideline is particularly critical when considering persons with dementia; as the symptoms of dementia progress, patients experience losses in their reading comprehension abilities. The aim of this study was to evaluate the readability of online information about dementia returned through keyword searches on search engines (N=200 webpages), as well as contained within discussion forums about dementia (n=1068 discussion threads). Grade scores were analyzed and compared using the following formulas: Flesch-Kinkaid, SMOG and Gunning FOG. For webpages, the mean grade level of text as calculated by the Flesch-Kincaid, SMOG, and Gunning FOG readability formulas were 11.8, 13.9, and 13.2 respectively; for discussion threads, the mean scores were 7.66, 9.88, and 9.75, respectively. Information about treatment was the most difficult to read. Results show that information on discussion forums is significantly easier to read than is information found on static websites, and that overall, online information about dementia is written at a reading level that is too challenging for the general public. These findings have important implications for health information-seeking, and lowering readability should be a priority for providers of both static and interactive online information about dementia.

Keywords: dementia, Alzheimer, readability, eHealth, discussion forums

INTRODUCTION

Online health information

In the United States of America (USA), the vast majority of the adult population uses the Internet: a report by the Pew Internet Project shows that as of 2015, 84 per cent of adults in the USA are Internet users\textsuperscript{1}. Older adults are also spending time online: of adults aged 65 and older, 58 per cent use the Internet\textsuperscript{1}. The Internet is frequently used as a source of health information in both the USA, where 72 per cent of internet users say that they have looked online for health information\textsuperscript{2}, and in the United Kingdom, where nearly half of the individuals who searched the Internet within the previous three months used it to search for medical or health-related information\textsuperscript{3}.

Many studies have looked into the quality of various types of online health information, yielding mixed results\textsuperscript{4-5}. Different aspects of the quality of information that have been considered therein have included factors such as content features (e.g., authoritativeness, clarity, breadth of information covered, currency of information), ethical features (e.g., disclosure of authors, conflicts of interest), and accessibility features (e.g., usability, readability)\textsuperscript{4-5}. As the online health environment is rapidly evolving, new interactive options are becoming available, allowing users to participate in a dynamic fashion, contributing as well as receiving content on platforms such as social media and online discussion forums\textsuperscript{6-7}. However, research into the quality of information on these platforms is in early stages\textsuperscript{8-9}, despite evidence suggesting they are widely used to disseminate health information that is relevant to the older population, such as that pertaining to dementia\textsuperscript{10}. As online health information plays an increasingly important role in health decision-making, either by the affected individual or their caregiver, it is critical to ensure it can be understood by its target audience.

Dementia

Dementia is described as “a clinical syndrome caused by neurodegeneration and characterized by inexorably progressive deterioration in cognitive ability”\textsuperscript{6-11}. With increasing prevalence and high costs of care, dementias have become a matter of major public health concern\textsuperscript{12}, in addition to being one of the most feared health issues by older adults\textsuperscript{13}. Worldwide, it is estimated that...
46 million people were living with dementia in 2015, resulting in a worldwide cost of 818 billion United States (US) dollars, with a projected cost of over one trillion dollars by 2018.

There are currently no effective treatments that halt or reverse the progression of most dementias, including Alzheimer disease (AD), the most common cause of dementia. The treatments currently available for AD merely target particular cognitive symptoms of dementia, and while the effectiveness of these drugs is statistically significant, it is also clinically marginal. Research into future treatment for dementias is ongoing, but without an effective treatment, there has been a shift in attention towards the importance of its prevention; recent studies have brought to light the possibility that the onset of dementia symptoms can be delayed by targeting risk factors. In fact, many interventions looking into factors such as exercise and nutrition are currently being investigated. However, the success of these preventative and postponing measures rests on the accessibility of the information about them. Therefore, whether about treatment or prevention of dementias, it is critical for the ageing population to be able to access and understand relevant information.

Health literacy and readability
The potential usefulness of high-quality online health information is squandered if its audience cannot understand it. Health literacy is defined as “the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions” and plays a key role in one’s ability to understand health information. Low health literacy is linked to poor health outcomes as well as earlier mortality in older adults. In a European survey nearly half of respondents showed limited health literacy, while in the US, only 12 per cent of adults had proficient health literacy in 2003, with only 3 per cent of adults aged 65 years and older showing proficient health literacy.

Reading ability impacts a person’s capacity to understand health information and as such is a key component of health literacy. The American Medical Association (AMA) recommends that health information should be written at a US fifth or sixth grade level for the general population, or between a US third and fifth grade level for information targeted at groups with a higher risk for limited literacy. Readability is defined as “a measure of the ease with which a passage of text can be read,” and the readability of a sample of text can be determined by the use of readability formulas. Several formulas have been developed for this purpose, such as the Flesch Reading Ease, Flesch-Kincaid, Gunning FOG, and SMOG readability formulas (Figure 1).

Studies that have looked into the readability levels of online health information pertaining to medical conditions have reliably shown that the information is written at a reading level above the AMA’s recommended range. However, little attention has been paid to the readability of online information about AD and other dementias, despite the large number of online resources available for both people with dementia and their families and caregivers. Also unknown is how the readability of online health information on static webpages compares to that of text exchanged in online discussion forums, which provide a more interactive information-seeking experience. To fill these gaps, the aim of this study was to determine the readability of online information written about AD and other dementias, written on (1) static pages returned from online searches for ‘demen-
Readability of online dementia information

tia’ and ‘Alzheimer’ keywords, and (2) the discussion forums that exist on those returned websites.

**METHODS**

**Sampling: search-returns**

Keyword searches were conducted on the five most popular online search engines in English (Google, Bing, Yahoo, Ask, AOL) using keywords ‘dementia’ OR ‘Alzheimer’, and collecting the Uniform Resource Locators (URLs, ex. http://www.alz.org/) returned on the first three pages of results for each search. Inclusion criteria for each URL returned were that (1) webpage hosts text-based, educational information specifically about AD or other dementias; (2) the information provided is intended for a public audience; (3) content is in English language; and (4) text is 100 words or longer. Following the initial data collection, for each URL, menu items or headers that were relevant to AD or other dementias (e.g., ‘Alzheimer’s facts’, ‘Warning signs’) were collected and categorized (e.g., ‘Overview’, ‘Symptoms / Stages / Signs’). We conducted a frequency analysis of the thematic content of menu items and headers and identified and retained information within the most common categories for further analysis: ‘Overview’ (present on 79 per cent of included URLs), ‘Symptoms / Stages / Signs’ (64 per cent), and ‘Treatment/Medication’ (55 per cent). Each page within these categories was then further mined for links to information under the same three categories and additional pages that were (1) hosted on the same base domain and (2) accessible within three clicks of the original URL were collected. The unit of analysis for the search-returns portion of this study consisted in “blocks of text”, defined as a single topic under a single URL. Henceforth, these blocks of text will be referred to as “search-returned” blocks of text. For a detailed schematic of the search-returned data collection process (Figure 2).

Each search-returned block of text was subject to minor edits to ensure consistency across the sample. The following edits were made: (1) headers that were not fully-formed sentences were removed; (2) periods were added at the end of headers that were fully formed sentences but did not end in a period; (3) lists of bullet points were removed from the text along with the preceding sentence if over half of the bullet points were shorter than five words long; (4) periods were added to the ends of bullet points that were missing them; (5) numbers were removed from numbered lists of bullet points; (6) references and advertising texts were removed; and (7) non-ASCII symbols were removed (Table 1).

![Figure 2. Sampling methodology for search-returned and discussion forums data](image-url)
Table 1. Examples of edits made to blocks of text

<table>
<thead>
<tr>
<th>#</th>
<th>Rule</th>
<th>Text before edits</th>
<th>Text after edits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Headers that were not fully-formed sentences were removed.</td>
<td>Early signs and symptoms.</td>
<td>The first symptoms of Alzheimer’s vary from person to person. Memory problems are typically one of the first signs of cognitive impairment related to Alzheimer’s disease.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The first symptoms of Alzheimer’s vary from person to person. Memory problems are typically one of the first signs of cognitive impairment related to Alzheimer’s disease.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Headers that were fully formed sentenced bud did not end in a period had a period added at the end.</td>
<td>Consult your health-care provider.</td>
<td>Consult your health-care provider.</td>
</tr>
<tr>
<td>3</td>
<td>Lists of bulled points were removed along with their preceding sentence if fewer than 50 per cent of the bullet points were shorter than five words in length.</td>
<td>However, medication, environmental influences and some medical conditions also can cause symptoms or make them worse. In early stages, people may experience behavior and personality changes such as: Irritability Anxiety Depression</td>
<td>However, medication, environmental influences and some medical conditions also can cause symptoms or make them worse.</td>
</tr>
<tr>
<td>4</td>
<td>Periods were added to the ends of bullet points that were missing them.</td>
<td>Common symptoms associated with dementia are: Trouble completing everyday tasks like cooking or cleaning. Inability to find things that have been misplaced Decreased ability to focus and pay attention. [...]</td>
<td>Common symptoms associated with dementia are: Trouble completing everyday tasks like cooking or cleaning. Inability to find things that have been misplaced. Decreased ability to focus and pay attention. [...]</td>
</tr>
<tr>
<td>5</td>
<td>Numbers were removed from lists of numbered bullet points.</td>
<td>1. Memory loss sufficient to disrupt daily life – such as forgetting recently learned information, important dates or events, asking for the same information over and over, relying more and more on memory aids or family members.</td>
<td>Memory loss sufficient to disrupt daily life – such as forgetting recently learned information, important dates or events, asking for the same information over and over, relying more and more on memory aids or family members.</td>
</tr>
</tbody>
</table>

**Sampling: discussion forums**

Each URL from our original sample was mined for links to interactive discussion forums. For each of the four forums identified (alzconnected.org, forum.alzheimers.org.uk, lbda.org), we identified sub-forums related to the three thematic categories of interest (‘Overview’, ‘Symptoms/Stages/Signs’, or ‘Treatment/Medication’). On each of these relevant sub-forums, we collected threads based on these inclusion criteria: (1) thread is listed on the sub-forum’s first three pages and (2) thread contains two or more responses. For each thread, the initial post and all of its responses were analyzed together in one text file and consisted of a unit of analysis. We removed text such as user information, and time of posting for each text submission, leaving only the content of the response for readability calculation. **Figure 2** provides a schematic of the data collection process for discussion forums.

**Readability calculation**

Search-returned blocks of text and discussion threads were analyzed for their readability scores using the online readability calculator tool available at http://www.online-utility.org/ as has been used in previous readability studies24,25. We determined readability using multiple readability calculation formulas, as has been recommended in order to ensure the validity of the readability scores26. Flesch-Kincaid (F-K), SMOG, and Gunning FOG readability scores, which output the reading grade level required for adequate comprehension of the text, were recorded for each block of text, along with the number of words therein. Additionally, the Flesch-Reading-Ease (FRE) readability score, which yields a quantitative result between 100 (very easy to read) and 0 (unreadable) was recorded for each sample of text26.

**Statistical Analysis**

Descriptive statistics such as mean, median, and standard deviation were calculated for each sample. We used the Shapiro-Wilk test to determine whether or not the readability levels of our sample texts were normally distributed, finding that while those for search-returned blocks of text were normally distributed, those within the discussion forums were not. Therefore, in testing for differences in readability levels between different
subsets of the data, we used two-sample t-tests for comparisons within the search Returned data, and Mann-Whitney U tests for comparisons within the discussion forum data or between the search-returned data and the discussion forum data.

**Results**

**Sample**
The online search for keywords ‘Alzheimer’ and ‘dementia’ returned 355 URLs. After removing duplicates and search returns that did not meet inclusion criteria, 95 unique URLs remained over 61 unique websites. A total of 742 menu items or headers that were relevant to AD or dementia were collected and categorized as follows: ‘Overview’ (158), ‘Signs/Symptoms/Stages’ (96), ‘Treatment/Medications’ (76), ‘Living with’ (60), ‘Diagnosis/Tests’ (59), ‘Causes’ (56), ‘Risk factors/Statistics’ (43), ‘Types’ (42), ‘Prevention’ (35), ‘Care’ (26), and ‘Help’ (18). Of these categories, only headers under ‘Overview’, ‘Signs/Symptoms/Stages’, and ‘Treatment/Medication’ were present in over 50 per cent of the 95 unique URLs collected (Figure 2). Therefore, the 89 URLs that contained at least one header under one of those three categories were retained for further analysis.

Further mining for links under those three categories within each remaining URL resulted in a total of 200 unique URLs. After collecting text from each of those URLs, our final sample consisted of 218 blocks of text, distributed across ‘Overview’ (62), ‘Signs/Symptoms/Stages’ (72), and ‘Treatment/Medication’ (84). Of these, 102 were found on AD or dementia advocacy websites, and 102 were found on health or medical information sites. The remaining blocks of text were found on general reference websites (6), or news websites (8). On the four discussion forums surveyed, 14 relevant sub-forums matched the three categories of interest and were retained for further analysis. From the 14 sub-forums, a total of 1068 discussion threads were collected and analyzed (Figure 2).

**Readability scores**
The mean F-K, SMOG, and Gunning FOG readability grade scores for the entire sample of search-returned blocks of text were 11.86 ± 2.08, 13.27 ± 1.69, and 14.05 ± 2.36, respectively (Table 2).

In further analyzing our search-returned data, we found that between blocks of text under ‘Overview’ and ‘Signs/Symptoms/Stages’ there were no significant differences in their average readability grade scores for the FK (t(132)=0.85, p=0.40), SMOG (t(132)=1.45, p=0.15), and FRE (t(132)=0.07, p=0.95) formulas, while Gunning FOG readability grade scores were significantly higher (i.e. more difficult) in ‘Overview’ texts than in ‘Signs/Symptoms/Stages’ texts (t(132)=2.11, p=0.04). Across all four readability formulas (FK, SMOG, Gunning FOG, FRE), scores for texts under ‘Treatment/Medication’ were significantly more difficult to read than those under either ‘Overview’ (p<0.05 for all pairs, t-values 3.36, 4.06, 3.72, 3.12, respectively) or ‘Signs/Symptoms/Stages’ (p<0.05 for all pairs, t-values 4.49, 5.62, 6.17, 3.17, respectively). We found no differences in the readability scores of texts according to the types of websites (e.g., government, news, advocacy) on which they were hosted.

The median F-K, SMOG, and Gunning FOG readability grade scores of discussion threads on online discussion forums were 7.40, 9.75, and 9.48 (Table 2). The median FRE score was 70.00, corresponding to a ‘standard’ reading difficulty level. Of the discussion threads analyzed, 413 (39 per cent) scored at or below the AMA’s recommended readability range for at least one of the readability formulas that output a grade level score (F-K, SMOG, and Gunning FOG). Comparisons between readability scores in threads in sub-forums under different categories showed that for all readability formulas (FK, SMOG, Gunning FOG, FRE), texts under ‘Treatment/Medication’ sub-forums were significantly harder to read than those under either ‘Overview’ (p<0.05 for all pairs, U-values 23272, 18801, 21982, 20634, respectively) or ‘Signs/Symptoms/Stages’ (p<0.05 for all pairs, U-values 48553, 40949, 46194, 38743, respectively) sub-forums, and ‘Overview’ threads were significantly harder to read than

**Table 2. Overall readability scores for search-returned text and discussion forum threads**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Search-returned text</th>
<th>Discussion forum threads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flesch-Kincaid (Grade level)</td>
<td>11.86</td>
<td>7.40</td>
</tr>
<tr>
<td>SMOG (Grade level)</td>
<td>13.27</td>
<td>9.75</td>
</tr>
<tr>
<td>Gunning Fog (Grade level)</td>
<td>14.05</td>
<td>9.48</td>
</tr>
<tr>
<td>Flesch-Kincaid Reading Ease (FRE) Score</td>
<td>38.33</td>
<td>70.00</td>
</tr>
<tr>
<td>FRE Score interpretation</td>
<td>“Difficult”</td>
<td>“Standard”</td>
</tr>
</tbody>
</table>

2018 Vol. 17, No 1
Readability of online dementia information

The goal of this study was to determine the readability levels of online AD and dementia information. We found that (1) the readability grade scores (FK, SMOG, and Gunning FOG) of both search-returned texts and discussion forum threads were significantly higher than the AMA’s recommended readability grade level (5-6); (2) the readability levels of search-returned blocks of text pertaining to AD were significantly more difficult than were levels. Interestingly, our results also showed that for only the Gunning FOG readability formula, within the search-returned section of this study, there was a significant difference between the readability levels between ‘Overview’ and ‘Signs/Symptoms/Stages’ text blocks, while no difference was found between those two categories for any of the other readability formulas. Though we would have expected all four readability formulas to show consistent results with each other, it is possible that differences in the parameters that each readability formula uses (such as an average number of words per sentence, or an average number of syllables per word) could account for this difference.

The misunderstanding or misinterpretation of online health information by Internet users who are ill-equipped to adequately understand content on easily accessible websites may have negative con-
sequences when this information is used in health decision-making. It has been shown that people with lower levels of health literacy have poorer health outcomes than do those with higher health literacy\(^{18}\), and in the case of users with dementia, the cognitive impairment that accompanies the progression of their illness\(^{29}\) may make it even harder to understand the online health information they need.

Our work adds to the growing number of studies evaluating the readability levels of online health information. Many other groups have similarly found that texts related to a large variety of health conditions (e.g., Uterine Artery Embolization, cancers) returned by online searches on popular search engines such as Google are written at readability levels that are more difficult than they should be, as recommended by the AMA\(^ {20,29}\). With regards to the type of information, Walsh and Volsko measured the readability levels of general online health information and found that information about treatments was particularly hard to read in comparison with most other categories of information, with the exception of information about diagnosis and screening, which was the most difficult to read\(^ {22}\). This discrepancy with our finding that ‘Treatment/Medication’ information has the most difficult readability levels could be explained by the variable nature of diagnostic procedures and treatment across health conditions. For both diagnostic procedures and treatments, descriptions may require the use of complex scientific names (e.g., imaging modalities, biomarkers for diagnosis, pharmaceutical names for treatment). In our sample, difficult readability levels observed in treatment-related content was at least partially driven by the polysyllabic words for currently approved AD drugs such as acetylcholinesterase and N-methyl-D-Aspartate (NMDA) antagonists. This may in part account for findings from a recent Australian study, which looked into the readability levels of Australian online health information pertaining to 12 commonly searched conditions and found that while the information about all 12 conditions was written above the recommended readability levels, information about dementia was the hardest to read of all 1225.

A unique aspect of our study is the analysis of data from discussion forums: we found that the readability levels of threads within AD- and dementia-centered discussion forums are lower than those of search-returned texts. This bears further investigation, given the emergence and increasing popularity of social media and other interactive online platforms as a source for and place to discuss health information\(^ {7,10}\). Perhaps more important are the implications of this finding. Previous research has shown that the thematic content and quality of health information shared on interactive platforms varies greatly\(^ {4,10}\). For example, a 2016 study by our group found that low quality online resources about the prevention of AD are more likely to endorse products or services\(^ {4}\). If these online resources are more accessible to the general population due to lower readability levels, they may become a preferred source of information despite containing information of lower quality than expert-curated websites. Future research should investigate the impact of readability on the selection of online resources.

We appreciate the limitations of our current study. We only looked at four online discussion forums in our sample as we were interested in those contained within our original sample of URLs. Despite capturing a wide scope of data through a large number of threads (N=1068) in our sample, those four discussion forums may not represent the entirety of the online interactive discussion about AD and dementia, and therefore, a larger sample size may have resulted in a different readability outcome. Additionally, in selecting the texts to be analyzed, though the three categories we considered (‘Overview’, ‘Signs/Symptoms/Stages’, and ‘Treatment/Medication’) reflected the most common categories of information found on the webpages returned in our initial search, it is possible that other categories of information would have yielded different readability levels. Other inherent limitations to our study are products of our use of readability formulas. For one, these formulas only take linguistic characteristics of the text into account, and therefore cannot consider other factors that may increase or decrease readability levels, such as webpage formatting and layout, complementary images or charts, and videos. However, these specific formulas were chosen in order to allow for a quantified comparison with the AMA’s suggested readability level for health information. Also, the use of complex, polysyllabic medical terminology inherent to the discussion of AD and other dementias naturally makes for difficult readability levels, since each of the readability formulas use the average number of syllables per word in the text in their formulas.

In summary, we have shown that the readability levels of online AD and dementia information are more difficult than the AMA recommends for health information in general. Accordingly, we propose three recommendations to hosts of online information about AD or other dementias, according to the results of our study: (1) the readability levels of hosted information should be monitored, and lowered if possible, (2) specific attention should be paid to readability levels when discussing treatments, and (3) prior to publishing new information about AD or dementia, readability calculation should be a priority. As
readability calculators are readily available for free, either online or through word processing software, health care professionals can easily assess this metric prior to making recommendations about online resources to their patient communities. Compliance with these recommendations will allow for greater accessibility to critical online information about all aspects of the global epidemic of dementia and will promote informed health-decision making from expert online sources in the growing population of computer-literate older adults worldwide.

Acknowledgments
Funding for this study was provided by the Canadian Consortium on Neurodegeneration in Aging, the Canadian Institutes for Health Research, the Canadian Foundation for Innovation and the Vancouver Coastal Health Research Institute. The authors gratefully acknowledge Dr. Jesse Hoey, PhD, Associate Professor of Computer Science at the University of Waterloo, for providing insightful comments on the manuscript.

References
Readability of online dementia information


