Readability and Suitability of Over-the-Counter Medication Labels

Young R. Lee*, Christie L. Baker, Gilbert G. Aguirre
Department of Pharmacy Practice, Texas Tech University Health Sciences Center School of Pharmacy, USA

*Corresponding author: Young R. Lee, Pharm.D., BCPS, BCCCP, Associate Professor, Department of Pharmacy Practice, Texas Tech University Health Sciences Center School of Pharmacy, 1718 Pine Street Abilene, TX, USA 79601, Tel: 1-325-696-0449; Fax: 1-325-676-3824; E-mail: young.lee@ttuhsc.edu

Introduction

Over-the-counter (OTC) medications are easily accessible and are used widely in various populations. Because drug pharmacodynamics differs between geriatric and pediatric patients, counseling these populations could positively impact patient outcomes. Older adults in the United States are the largest consumers of OTC medications and often self-administer them without their healthcare provider’s knowledge [1]. One study found that 57% of geriatrics taking an OTC medication chronically required intervention from a pharmacist due to unsafe self-administration [2]. The pediatric population is not immune to adverse events involving OTC medications either. Between 2004 and 2005, there were an estimated 1,500 children under the age of two were treated in the emergency department related to use of OTC cough and cold medications [3]. Analgesics, a commonly utilized category of OTC medications, have also been associated with more adverse events due to inappropriate administration. Wilcox et al. found that 54% of surveyed nonsteroidal anti-inflammatory drug (NSAID) users were not aware of their side effects and 26% used more than the recommended dose written on the label [4]. Because OTC medications are often purchased without consulting a healthcare provider, the Drug Facts label in the packaging must be easily understandable to allow patients to safely administer their therapy.

The average reading level in the United States has been estimated to be 8th grade; The Food & Drug Administration (FDA) has recommended that nonprescription labels be written at a 4th or 5th grade level, not to exceed an 8th grade level [5]. These recommendations do not translate into regulations that mandate OTC drug labels. A previous study on the readability of OTC drug labels found that none of the evaluated drug labels met this recommendation [6]. In fact, the study concluded that the average grade level needed to comprehend OTC drug labels. 

Abstract

Objective: The objective of the study was to evaluate the appropriateness of the readability and suitability of non-prescription drug labels for the average patient in the United States. Methods: 149 drug labels were evaluated. Readability was assessed using 5 readability formulas with and without the ‘Inactive Ingredients’ list for comparison. Suitability was assessed using the Suitability Assessment of Materials (SAM) score for the hardest and easiest drug labels in each drug class as determined by Flesch-Kincaid Grade Level. Results: In Readability, the overall average Flesch-Kincaid Grade Level for the labels including the ‘Inactive Ingredients’ list was 9.5. The overall average Flesch-Kincaid Grade Level for the labels excluding the ‘Inactive Ingredients’ list was 6.6. In Suitability, all labels assessed were scored as “adequate” or higher (40% or higher per the SAM scoring tool). Conclusion: The readability and suitability scores of the evaluated OTC medication labels excluding ‘Inactive Ingredients’ reveal a reading grade level and comprehensibility appropriate for the average patient to easily read and understand. Practice Implications: This study, along with evaluating how much readability of a medication label affects the incidence of adverse drug reactions, would likely provide more insight into the importance of easily readable OTC labels.

Keywords: over-the-counter, OTC, readability, suitability, SAM, drug labeling, comprehension, nonprescription drugs, health literacy
labels was about twice the recommended grade level at 16 ± 5 (using the Flesch-Kincaid Grade Level score) [6]. Labels with adequate readability and suitability should correlate with patient safety.

Understanding the importance for a patient to not only read but to comprehend medication labels, this study evaluated about 150 OTC medications for readability and suitability. The suitability assessment of materials (SAM) instrument is used to assess the ease of understanding, motivation, and learning stimulation. The objective of the study was to evaluate the appropriateness of the readability and suitability of non-prescription drug labels for the average patient in the United States.

Methods

OTC labels

Commonly used OTC medications were selected from seven different classes (149 labels). In order to appropriately identify and select “common” therapies, all medications were selected from the aisles of a local pharmacy chain in Abilene, Texas. A pharmacy’s purchase inventory based on demand, thus the displayed medications are the most likely to be seen and purchased by patients. The drugs were categorized based on the section of the pharmacy where it was sold. Labels for each medication were obtained from the National Institutes of Health’s Daily Med Database [7]. The text was first evaluated including the entire Drug Facts Label; the second evaluation excluded the Inactive Ingredients. Both series of computations were each assessed by two investigators.

Readability scores

The Flesch Reading Ease score, Flesch-Kincaid Grade Level score, Gunning Frequency of Gobbledygook (FOG) Index, Coleman-Liau Index, and Simplified Measure of Gobbledygook (SMOG) Index were used to measure readability of 149 over-the-counter drug labels. For consistency, these calculations were completed by the program readable.io. The Flesch Reading Ease score is a widely used and reliable readability test and generates a score of 0 to 100, where a higher score indicates that text is easier to read [8]. A conversion table is necessary to read the score as a grade level, but the Flesch-Kincaid Grade Level Score is a modified Flesch Reading Ease formula that generates the grade level education a person needs to understand the material [8]. The Flesch Reading Ease score has a correlation of -0.88 and the Flesch-Kincaid Grade Level score has a correlation of 0.91 in normed reading tests with grade level understanding [8]. Both have been used extensively in reporting readability of healthcare-related text [6, 9-14]. The Gunning Fog Index is another readability test that is useful in unassisted reading, such as healthcare-related text. It also has a correlation of 0.91 in normed reading tests with grade level understanding and tends to report a higher grade level score than the Flesch-Kincaid method, but lower scores than the SMOG Index [8]. Drug label texts were also scored using the SMOG Index, another method that reports a grade level for comprehension, tends to report higher scores than the other methods and has a correlation of 0.88 with grade level understanding [7]. The SMOG Index is widely used to measure readability of healthcare-related text and has been recommended by the US National Institute of Health for this purpose [12-15]. The Coleman-Liau Index was used to score labels as well, which has also been used in healthcare readability scoring [13, 16]. This instrument was developed later (1975) than the other readability instruments, which is likely why it has not been used as extensively as other readability formulas [17]. It is distinct in that it evaluates length of words and sentences, rather than number of syllables [17].

Suitability

The Suitability Assessment of Materials (SAM) score of the most and least readable drugs (selected using Flesch-Kincaid Grade Level) within each class was calculated as a percentage score (14 drug labels), then interpreted as “not suitable” (0-39%), adequate (40-69%), and superior (≥70%). The drugs scoring the lowest per category on the Flesch-Kincaid Grade Level assessment are defined as most readable drugs. Likewise, the drugs scoring the most per category are defined as the least readable. The labels selected were from the labels evaluated without the Inactive Ingredient list. SAM scores the suitability and comprehensibility of text by rating content, literacy demand, graphics, layout, type, learning stimulation and motivation, and cultural appropriateness [18]. Content is determined by evaluating purpose, content about behaviors, scope, and whether a summary is included. Literacy demand is determined by reading grade level, active voice, common vocabulary, context and learning aids such as appropriate headers. The graphics category is determined by purposeful graphics, type and relevance of the graphic, explained tables and captioned graphics. Layout and typography is determined by how easy the text is to follow, appropriate typography and subheadings. Learning stimulation and motivation is determined by interactions utilized such as prompting questions, behaviors modeled and self-efficacy. Lastly, cultural appropriateness is determined by logic/ language/ experience and cultural image/ examples. The SAM score is useful for analyzing more than grade level readability, such as the layout of the OTC medication label to evaluate how comprehensible the text is. It has previously been used for scoring suitability of healthcare-related educational text [19, 20]. The most and least readable labels were assessed manually by two investigators.

Verification of scores

The readability scores were verified by calculating them twice, deriving mean scores between two raters and rounded grade level scores up to the nearest whole number. SAM scores were verified by scoring labels twice and deriving a mean percentage score between two raters.
### Results

#### Readability including the inactive ingredient list

The mean scores for all labels assessed per readability formula were 48.7, 9.5, 11.1, 10.8, and 10.2 (Flesch Reading Ease, Flesch-Kincaid Grade level, Coleman-Liau Index, SMOG Index, and Gunning Fog Index, respectively). Mean scores were calculated within each medication class. All medication classes were found to be written above the FDA-recommended maximum 8th grade reading level in all categories using the Flesch-Kincaid Grade Level formula: pain medications (n=15), sleep aids (n=14), infant and children medications (n=22), allergy medications (n=36), cough and flu (n=10), and gastrointestinal medications (n=44). The cold and flu medications were determined to be the hardest categories to read via the Flesch-Kincaid Grade Level formula (10+5 + 1.5). Refer to Table 1 for additional formula scores by class. The mean readability score among NSAIDs (n=24) assessed was slightly lower than the rest of the drug labels per the Flesch-Kincaid Grade Level formula (9.66, 9.70 respectively).

### Table 1. Mean and standard deviation readability scores by medication class

Readability excluding the inactive ingredient list

The mean scores for all labels assessed per readability formula were 59.7, 6.6, 9.1, 8.8, and 7.5 (Flesch Reading Ease, Flesch-Kincaid Grade level, Coleman-Liau Index, SMOG Index, and Gunning Fog Index, respectively). Mean scores were calculated within each medication class. All medications were found to be written below the FDA-recommended maximum 8th grade reading level in all categories using the Flesch-Kincaid Grade Level formula: pain medications (n=15), sleep aids (n=14), infant and children medications (n=22), allergy medications (n=36), cough medications (n=8), cold and flu (n=10), and gastrointestinal medications (n=44). The sleep aid medications were determined to be the hardest categories to read via the Flesch-Kincaid Grade Level formula (7.0 + 0.4). Refer to Table 1 for additional formula scores by class. The mean readability score among NSAIDs (n=24) assessed was slightly lower than the rest of the drug labels per the Flesch-Kincaid Grade Level formula (6.0, 6.6 respectively).

Suitability

Suitability Assessment of Materials was evaluated for 14 labels, selected by the highest and lowest Flesch Reading Ease score of each medication class determined without the Inactive Ingredient list. All labels were scored as “adequate” or higher (Table 2). Aleve® of the Pain medication class and the Generic Dayquil of the Cold/Flu class both scored higher than 70%, deeming them “superior.”

Table 2. Suitability of selected drug labels

<table>
<thead>
<tr>
<th>Label</th>
<th>Mean percentage score</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aleve®</td>
<td>74%</td>
<td>Superior</td>
</tr>
<tr>
<td>Percogesic®</td>
<td>55%</td>
<td>Adequate</td>
</tr>
<tr>
<td>Generic Doxylamine</td>
<td>59%</td>
<td>Adequate</td>
</tr>
<tr>
<td>Generic Caffeine</td>
<td>53%</td>
<td>Adequate</td>
</tr>
<tr>
<td>Children's Nasacort®</td>
<td>55%</td>
<td>Adequate</td>
</tr>
<tr>
<td>Generic Liquid Glycerin Suppositories</td>
<td>69%</td>
<td>Adequate</td>
</tr>
<tr>
<td>Generic Nasal Triamcinolone</td>
<td>65%</td>
<td>Adequate</td>
</tr>
<tr>
<td>Zyrtec D®</td>
<td>54%</td>
<td>Adequate</td>
</tr>
<tr>
<td>Delsym®</td>
<td>54%</td>
<td>Adequate</td>
</tr>
<tr>
<td>Robitussin 12 Hour Cough Relief®</td>
<td>67%</td>
<td>Adequate</td>
</tr>
<tr>
<td>Generic Dayquil</td>
<td>76%</td>
<td>Superior</td>
</tr>
<tr>
<td>Cordicin Maximum Strength Flu HBP®</td>
<td>57%</td>
<td>Adequate</td>
</tr>
<tr>
<td>Immodium AD®</td>
<td>56%</td>
<td>Adequate</td>
</tr>
<tr>
<td>Bonine®</td>
<td>60%</td>
<td>Adequate</td>
</tr>
</tbody>
</table>
Discussion

Readability scores were inconsistent to those of previous studies within the last 5 years with all assessed drug labels scoring at or below the maximum reading grade level recommended by the FDA when the Inactive Ingredient list was excluded. Additionally, this study does not align with a previous study’s finding that NSAID labels were particularly more difficult to read [6]. Overall, this study refutes the results of a recent study by Trivedi et al. When designing this study, it was determined that the inclusion of the inactive ingredient list from the Drug Facts Label greatly increases the difficulty of the readability. Because a drug’s safety and administration is not contingent on the patient’s understanding of the inactive ingredient list, this portion of the label was excluded. The exclusion of the inactive ingredients and the addition of 3 more readability scores and a suitability score were all beneficial changes to the methods of the aforementioned study. Evaluating the labels with more readability scores was important to determining the true and more standardized readability of each label. Additionally, the suitability was an extra benefit to the study. Readability is evaluated by formulas but suitability allows analysis of the labels in a more subjective way. It is one thing to count the letters and complexity of labels, it is another thing to examine how easy it is to read based on aesthetic. Suitability scores support the overall conclusion made by the readability studies. By changing the methods to exclude the inactive ingredients, the readability of NSAID labels was positively affected as well. This conclusion suggests that the reported adverse reactions and hospitalizations to OTC medications are not directly caused by incomprehensible medication labels by the average population. A study by Sansgiry et al. in 1997 also evaluated the readability of OTC labels, but used the Label Readability Guidelines recommended by the Nonprescription Drug Manufacturers Association (NDMA) [21]. This study also found use of all uppercase font and small font size can reduce readability, along with lack of boldface and use of hyphenation.

Regarding suitability, all labels were rated adequate or higher. These results indicate that much of OTC labels’ suitability could be improved in terms of more purposeful graphics, use of examples, cultural appropriateness, etc. Specifically, scores for the Layout section of the score had the highest average score indicating that (summary not applicable for all labels), indicating that the labels have appropriate layout factors, typography is practical and appropriate subheadings were used. Graphic, Learning Stimulation/Motivation and Cultural Appropriateness scored the worst across the categories.

The Flesch Reading Ease, Flesch-Kincaid Grade level, Gunning Fog, and SMOG index have all been used in previous studies that evaluate readability of OTC medications, as well as in evaluating health-care text [6, 22]. Unique to previous OTC label readability studies, the Coleman-Liau Index was used in this study. This method has been used in the evaluation of healthcare patient education material [23-25]. The five aforementioned instruments were utilized to account for the different formulas each platform used to calculate scores to better compare findings with previous studies.

The strengths of this study include an increased number of labels evaluated compared to previous studies, the inclusion of the SAM instrument and altering the methods by excluding the inactive ingredient list. From what has been researched, the SAM instrument has not been used to assess suitability of over-the-counter medication labels. While it provides a unique perspective towards comprehensibility of OTC labels, it is mostly useful as a predictor of suitability in this scenario. In addition, readability and suitability instruments can provide an estimation for reading ease. However, label comprehension studies using active participants would provide data that is more accurate. To make a more robust assessment of the SAM score, it would be ideal to find evaluators with various health literacy. These evaluations could be performed to further examine the readability of OTC labels and confirm the results of studies using readability instruments. It should also be mentioned that even though the drug label is short, the SAM score effectively evaluates suitability. These medication labels are often the only information a patient received about the OTC medications and should therefore be able to convey the already readable text in the best way possible. Without a suitable label, the information is obscured and the patient may not be able to adequately assess the information to the detriment of safety. Analyzing the incidence of adverse events in relation to readability of the medication label is also an important future direction in that it may spur future OTC labels to be made even more readable.

A 2014 study conducted in South Africa examined caregivers’ ability to administer pain killers to children based on information provided [26]. Researchers assessed paracetamol labels, inserts, and patient information leaflets by conducting face-to-face surveys with sixty caregivers and six pharmacists [26]. Regarding the label specifically, 10% of caregivers did not understand one or more scientific terms and 12% answered scientific terms incorrectly [26]. These findings support our findings that removing scientific terms such as inactive ingredients might help improve readability of labels. In addition, similar to results of our study, 54% of participants of the South African study indicated that font size was too small. However, one of the major drawbacks to this study is that only one OTC product was evaluated. Our current study evaluated several OTC products across different classes. We can expand upon the current study in future studies by including patient evaluations of different health literacy, cultural, and socioeconomic backgrounds. In addition, we would further be able to evaluate how inactive ingredients may negatively impact OTC label scores.

A 2018 review on the effectiveness of nonprescription drug labels found comprehension of warnings or specific
statements to be high [27]. For example, it was found that comprehension of dosage and duration of therapy for omeprazole was 95% and 91% respectively. Comprehension of safety warnings were similarly high for lovastatin at approximately 90%. These studies coincide with our results, upon removing inactive ingredients, the readability of OTC labels increased dramatically. Previous studies also only examined one OTC product at a time. The review also identified health-literacy and visual impairment to be independently associated with improper dosing. In order to properly evaluate OTC labels, examining the comprehension of essential aspects a label need guide future studies while assessing multiple drug products. This may help guide how drug labels are scored in the future with proper insights on which areas of a label should be highlighted with regards to language and format.

Conclusion

The readability scores of 149 OTC medication labels excluding Inactive Ingredients reveal a reading grade level appropriate for the average patient to easily read and understand. SAM scores of 14 OTC labels are adequate but show room for improvement to increase comprehensibility of the label. Label comprehension studies would likely provide more insight into real-world readability of medication labels. Prospective studies like this, along with evaluating how much readability of a medication label affects the incidence of adverse drug reactions, would likely provide more insight into the importance of easily readable OTC labels.

Abbreviations

OTC: over-the-counter; NSAID: Nonsteroidal Anti-Inflammatory Drug; FDA: Food & Drug Administration; SAM: Suitability Assessment of Materials; FOG: Frequency of Gobbledygook; SMOG: Simplified Measure of Gobbledygook; NDMA: Nonprescription Drug Manufacturers Association

Conflict of Interest

Author Young Lee, author Christie Baker and author Gilbert Aguirre declare that they have no conflict of interest.

Funding

No funding was received for this research.

Declarations of Interest

None.

Acknowledgements

We thank Nathan Baker, Pharm.D, for his assistance in data collection and study design. Additionally, we are grateful to Adam Hilgemeier, Pharm.D, candidate for assisting in data collection.

References


