Original articles

Readability assessment of psychiatry journals

Skye P. Barbic
Faculty of Medicine, Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada; Social Aetiology of Mental Illness (SAMI) Canadian Institute of Health Research (CIHR) Initiative, Centre for Addiction and Mental Health, Toronto, Canada; sbarbic@mail.ubc.ca
Karen Roberts
Post Doctoral Fellow; SAMI CIHR Strategic Training Initiative, Centre for Addiction and Mental Health, Toronto, Canada; roberts.karen.elizabeth@gmail.com
Zachary Durisko
Post Doctoral Fellow. SAMI CIHR Strategic Training Initiative, Centre for Addiction and Mental Health, Toronto, Canada; zachary.durisko@gmail.com
Cheolsoon Lee
Department of Psychiatry, Gyeongsang National University Hospital, School of Medicine, Gyeongsang National University, Jinju, Republic of Korea; Social Aetiology of Mental Illness (SAMI) Canadian Institute of Health Research (CIHR) Initiative, Centre for Addiction and Mental Health, Toronto, Canada; cheolsoon.lee@camh.ca
Rosemary Yachouh
York University, Faculty of Health, Toronto, ON, Canada; rosieyachouh@gmail.com
Jason Guriel
Evidence Exchange Network, Toronto, ON, Canada; Jason.guriel@camh.ca
Kwame McKenzie
Professor of Psychiatry University of Toronto, Medical Director Centre for Addiction and Mental Health, CEO Wellesley Institute, Toronto, Canada; Kwame.mckenzie@camh.ca

Abstract: ***Introduction*:** Peer-reviewed journals in psychiatry are the primary source of new information for researchers and clinicians. New evidence is emerging faster than ever. To maintain the highest standards of practice, clinicians and researchers need to make sense of the latest research. To facilitate this, papers need to be clear, concise, and highly readable. Clear communication is especially important in fields like psychiatry, which brings together many different research approaches. The aim of the study was to assess the readability of the most prominent journals in psychiatry.

***Method*:** We tested the readability over time of articles from eight of the most widely cited psychiatric journals. We sampled 504 articles from different issues and years of publication (2002-2013) and looked at their Abstracts, Introductions, and Discussion sections using five validated readability metrics. We also compared the readability of psychiatry journals to general medicine journals for the year 2013.

***Results*:** The readability level of all psychiatry journals was "very difficult" across time. Psychiatry journals were harder to read than general medical journals ($P<0.001$). We found that the strongest predictors of poor readability in psychiatry journals were high impact factor, article type (review), and high number of words per sentence. The Discussion was the most readable section of psychiatry papers, followed by the Abstract, and Introduction.

Conclusion: Psychiatry articles require a high level of effort because they are difficult to read. Authors and editors should strive to make articles as readable as possible. This may increase the uptake of evidence and improve practice.

Keywords: Readability, Flesh-Kincaid, Psychiatry, Evidence-based Medicine, Communication

Introduction

Psychiatry is a multi-disciplinary medical specialty. People need to work together and communicate to study, diagnose, and treat mental disorders. Safe, high-quality patient care relies on the use of new knowledge from various areas of research. For instance, to understand how best to treat a mental illness, one may need to bring together biological theory, molecular genomics, and psychosocial research. Given this diversity of approaches, harnessing and communicating specialized knowledge can be challenging. Clear communication is critical to providing the best possible mental health care.

Peer-reviewed publications are the primary access point for new medical knowledge.[1] To maximize impact, researchers submit their findings to journals that are widely read by clinicians. In psychiatry, clinicians report that their main source of information comes from journals articles.[1] But with 500,000 new articles added to PubMed each year, navigating psychiatry literature can be overwhelming.[2-4] The degree to which psychiatrists keep up with new literature is unknown.

A key aspect of information uptake is the "ease with which written materials are read" or *readability*.[5,6] Readability is directly correlated with a person's ability to understand content.[5-7] Currently, no official guidelines govern how readable peer-reviewed articles should be. Several studies have analyzed the content of online patient information in psychiatry and other medical fields.[5,8,9] However, we are unaware of any study that has assessed readability of peer-reviewed research articles.

This study asked: *how readable are research articles published in psychiatry journals?* We hypothesized three

things. First, the readability of psychiatry journals would be similar to that of popular general medicine journals. Second, based on the growth rate of scientific publication,[10] we hypothesized that peer-reviewed articles in psychiatry have become more readable over time. Third, considering that the Abstract is the most readily available part of a research paper and is often available for free online,[11] we hypothesized that the Abstract would be the easiest section of a paper to read. A secondary objective of this study was to estimate the extent to which journal and research article characteristics predict readability.

Methods

Article Selection

We used *Web of Science* to identify the eight psychiatry journals with the highest five-year impact factors. We developed a standardized operating procedure for article selection (see Appendix A).

From each journal, we identified the number of issues per year and randomly selected three issues per year (2002, 2004, 2006, 2008, 2010, 2012, and 2013). At the time of study conception, in order to assess trends in readability across time, we decided to evaluate every second year from 2002-2012. Once data collection for this time period was completed, all academic journal entries for the year 2013 had been published. As a result, to be current in our search we decided to also include issues from the year 2013.

From each issue, we identified the number of original research articles published and randomly selected three papers for a total of 63 from each journal. For both randomization procedures, we used an online calculator (www.random.org). For comparison, we repeated the same process, picking one paper each from every second issue of three commonly read general medicine journals for the year 2013 (24-26 total general medicine papers per journal).

Measurement

Five raters tested readability using a standardized operating procedure (see Appendix A). In order to achieve an inter-rater reliability score of greater than 90% and have minimal discrepancies between raters, all raters piloted the procedure three times with the same five randomly selected articles. Next, we assigned each rater a unique list of articles to evaluate. We removed extraneous text (such as subtitles, references, parenthetical results and statistics, and any material in parentheses with fewer than three words, which typically explains abbreviations).

Next, we assessed the readability of the Abstract, Introduction and Discussion sections on their own. Then, we assessed the sections together, for an overall score. We copied the sections of text into a free online calculator (http://www.readability-score.com/) to generate five readability test scores: (1) *Flesch Reading Ease Score* (FRES)[12], (2) *Flesch-Kincaid Grade Level* (FKGL)[13], (3) *Gunning-Fog Score* (GFOG)[14], (4) *Coleman-Liau Index* (CLI)[15], and (5) the *Simple Measure of Gobbledygook* (SMOG)[16]. The program also compiled an average grade level score based on the latter four tests. This program has been used in previous tests of readability of health literature.[17, 19] Appendix B provides a detailed description of each test. In brief, tests differ in calculation and interpretation. We thought it best to include several of these metrics instead of selecting one potentially biased measure. Each measure of readability focuses on and weights different components of writing differently (see algorithms in Appendix B). However, for the purpose of simplicity in this paper, we interpret findings from the FRES only. A higher score on the FRES indicates that material is easier to read. Scores ranging from 60-70 are *standard difficulty* (8th-9th grade); 50-60 *fairly difficult* (10th-12th grade); 30-50 *difficult* (13th-16th grade); and less than 30 are *very difficult* (college graduate). Readability scores are standardized to correspond to grade levels in the United States. We also recorded the word count and average number of words per sentence for each section. Additionally, we recorded the number of times each article had been cited (as of the week of June 15, 2014) using a Google Scholar™ search.

Sample Size calculation

To estimate our required sample size, we reviewed 72 research papers from the year 2012 (4 raters, 18 papers each). Based on this, we performed *a priori* sample size calculations using *G*Power v3.0.10*. We assumed a medium effect size of 0.40, alpha=0.05, power=0.80, and a correlation of 0.50 between seven time points. This yielded a required sample of 376 papers.

Statistical Analysis

We calculated the median word count per article (with interquartile range [IQR] and 25-75 percentiles) for all journals. We measured the association of readability scores between journals and impact factor with a Spearman rank-order correlation. We conducted a one-way analysis of variance (ANOVA) to examine differences in journal, year of publication, and their interaction. For significant main effects, we used Tukey's HSD *post hoc* individual pairwise comparisons between journals and time points. We compared readability scores between psychiatry and general medicine journals in the year 2013 using two-tailed independent t-tests.

We used multiple linear regression models to evaluate the effect of impact factor, publication year, study type (*original research* or *review*), words per sentence, rater, and citation count on FRES readability scores. We included predictor variables at the P=0.15 level in unadjusted analyses and used P=0.10 for retention. We confirmed findings with stepwise forward analysis. We considered models with P-values of less than 0.05 to be significant. For all analyses, we inspected residual and diagnostic plots to identify any outliers, major deviations from normality, linearity, or multi-collinearity. We conducted all analyses with *SPSS v20.0*.

Results

As shown in Table 1, the eight psychiatric journals with the highest impact factors were: *The American Academy of Child and Adolescent Psychiatry*, *The American Journal of Psychiatry*, *Archives of General Psychiatry* (now known

as *JAMA Psychiatry*), *The British Journal of Psychiatry*, *Biological Psychiatry*, *Neuropsychopharmachology*, *Schizophrenia Bulletin*, and *Molecular Psychiatry*. Seven journals are based in the United States. One journal specializes in child and adolescent psychiatry. None of the journals are currently open-access. Of note, despite being one of the most popular psychiatric journals, we excluded *World Psychiatry* from our study because the types of articles published in this journal differed from the other journals (eg, commentaries, perspectives, special reports).

Table 1. Selection of journals, issues and 5 year impact factor

Journal	Issues/ year	No of issues selected/ year	5-year Impact Factor
Psychiatry Journals			
Archives of General Psychiatry[a]	12	3	14.5
American Journal of Psychiatry	12	3	14.4
Molecular Psychiatry	12	3	14.0
Biological Psychiatry	12	3	9.8
Schizophrenia Bulletin	4-6	3	8.9
Neuropsychopharmacology	12	3	7.8
Journal of the American Academy of Child and Adolescent Psychiatry	12	3	7.2
British Journal of Psychiatry	12	3	7.1
General Medicine Journals[b]			
New England Journal of Medicine	52	26	51.7
The Lancet	52	26	36.4
Journal of the American Medical Association	48	24	29.3

[a] Currently called JAMA Psychiatry

[b] Readability only assessed for the year 2013

As shown in Table 1, we included a total of 504 peer-reviewed psychiatry articles (63 per journal) for this study. Of these, 473 were original research (94%), and 31 (6%) were reviews. We also assessed a total of 76 general medicine articles from one year (2013) from the *New England Journal of Medicine (NEJM), the Lancet,* and *Journal of the American Medical Association (JAMA).*

How readable are psychiatry journals?

We found that 460 studies (92%) had a total FRES reading score of less than 30 (corresponding to *very difficult*). Overall, the median FRES score was 17.7 (10.6-24.8). Table 2 summarizes the median readability scores. All other readability scores were highly correlated with the FRES (r_s>0.8, *P*<0.01).

How readable are psychiatry journals compared to general medicine journals?

General medical articles were significantly more readable than psychiatry articles (t_{75}= 4.8, *P*<0.001). The median FRES of the combined Abstract/Introduction/Discussion (with interquartile range) for general medicine articles in 2013 was 24.9 (19.2-30.6). We found that 49 studies (68%) had a total FRES reading score of less than 30 (corresponding to *very difficult*).

How does readability differ between psychiatry journals?

Based on median readability scores, the *Journal of the American Academy of Child and Adolescent Psychiatry* was the most readable, followed by *Molecular Psychiatry* and the *British Journal of Psychiatry*. *Schizophrenia Bulletin* and the *Archives of General Psychiatry* were the least readable. As shown in Table 3, we observed statistically significant differences in the total mean readability scores of several psychiatry journals. Specifically, we found the *Journal of the American Academy of Child and Adolescent Psychiatry* to be more readable than the *American Journal of Psychiatry, Schizophrenia Bulletin, Archives of General Psychiatry,* and *Biological Psychiatry* ($F_{7, 496}$ = 6.3, *P*<0.001).

How does readability change over time?

The trajectory of readability of each psychiatry journal differed across time. Figure 1 shows that the trajectory of the median FRES score was quite variable for four journals (*American Academy of Child Psychiatry, American Journal of Psychiatry, Schizophrenia Bulletin, and Neuropsychopharmachology*), with annual median scores increasing and decreasing frequently over the time period. The remaining journals showed more stable, although lower, readability scores across time.

What are the most readable sections of psychiatry journals?

Among psychiatry journals, the readability of each article section was significantly different (*P*< 0.001). Discussion sections were the most readable (median FRES 20.9, IQR 12.1), followed by the Abstract (median FRES = 17.5, IQR 14.3), and the Introduction (median FRES = 14.1, IQR 14.3). We found that 417 (83%) Abstracts, 460 (93%) Introductions, and 433 (86%) Discussions met the criteria for "*very difficult*" (FRES < 30). There was a main effect of journal and time, with the *Journal American Academy of Child and Adolescent Psychiatry* having a greater FRES score for all sections compared to articles from the other journals (see pairwise comparisons in Table 3). We found similar trends with the other readability tests.

For general medicine journals, the most readable journal section was the Abstract (median FRES=30.9, IQR 11.2), followed by the Discussion (median FRES=23.7, IQR 16.7), and the Introduction (median FRES=15.7, IQR 7.1). We found that 57% of psychiatry Abstracts were *extremely difficult* to read (<20 FRES). In contrast, only 11% of general medicine Abstracts met this criterion, with median FRES scores ranging from 28-47.

What predicts readability?

Year of publication and impact factor both significantly affected readability (*F*=22.05, R=0.35, R^2 =0.12, *P*<0.001; Table 4). As impact factor increased, readability decreased

(ß = -0.38), suggesting readability is lower for psychiatry journals with higher impact factors. We also found that for every incremental increase in publication year (every two years 2002-2012, 2013), readability decreased (ß = -0.24), suggesting readability is declining. We also found that readability of review articles was lower than original research articles (ß = -1.75). As shown in Table 4, we did not find evidence to support a significant relationship between citation rate and readability (ß = 0.004, $P = 0.32$).

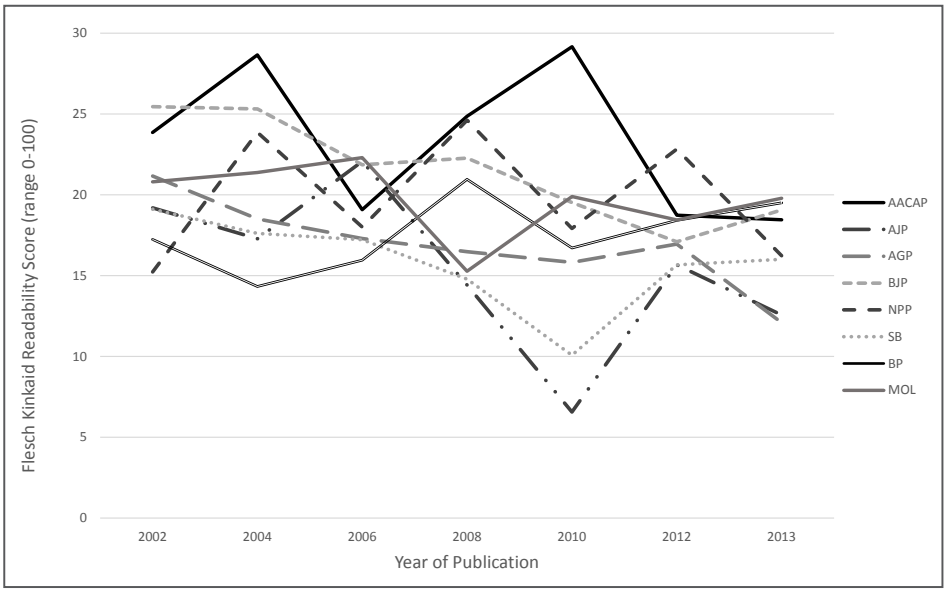
Discussion

Most psychiatry articles (92%) scored as *very difficult* to read, compared to general medicine journals (68%). We observed this trend in all psychiatry journals across time. On average, we found that the higher a psychiatry journal's impact factor, the more difficult the text was to read. The Introduction and Abstract were the most difficult sections. This may reflect the authors' attempt to cover as much material as possible in short sections within word limit requirements. Interestingly, psychiatry journals are more difficult to read than general medical journals. This is especially true for Abstracts. Our study showed that 57% of psychiatry Abstracts were *extremely difficult* to read. In contrast, only 11% of general medicine Abstracts met this criterion.

The journal of the *American Academy of Child and Adolescent Psychiatry* was the most readable psychiatry journal. However, even it was more difficult than the average general medicine article. Our results show that psychiatry journals are (not surprisingly) more difficult to read than online patient information.[5,8,20] For example, one study found 165 online webpages from ophthalmology as *difficult*[5]. Our study, on the other hand, found 92% of articles to be *very difficult* (see Table 2). Given that the target audience tends to be highly educated, this may seem appropriate. However, the increasing demands on psychiatrists and health services make it hard to keep up with new knowledge. Future research is needed to understand the extent to which poor readability may delay or prevent new knowledge from getting into practice.

Our group also evaluated the association between an article's readability on its citation rate. We found that reading level alone did not predict how often the article was cited, even when we took into account publication year and journal (see Table 4). Other studies have shown similar results when assessing the popularity of online health materials.[20] This suggests that other factors such as article topic, and visual and layout features may influence how often an article is read and cited. Our study looked only at the text of commonly read sections of journal articles. Future efforts should be made to understand how other materials, such as figures, tables, images, and color, impact the citation rate.

Evidence-based psychiatry integrates the best research with clinical expertise and patient values.[21] In psychiatry, new evidence emerges daily. This evidence can inform healthcare decisions that have an impact on the lives of patients and their families. For clinicians, keeping abreast of new evidence and incorporating it into their daily practice is time consuming. It may even be impossible given the 13 million general medicine references and over 4800 biomedical journals in Medline alone.[22] Despite the exponential increase in the number of articles being published yearly, our study highlights that the readability of articles in psychiatry remains consistently difficult across time (see Figure 1). Clinicians are busy, and making the uptake of new information as easy and efficient as possible will likely improve healthcare.[4,21,22] Abstracts and systematic reviews can synthesize and summarize studies, and increase efficient access to the evidence. However, our results suggest that these reviews may be at least as difficult to read as primary research, and that Abstracts are often the most difficult to read passages.



AACAP= American Academy of Child and Adolescent Psychiatry; AJP= American Journal of Psychiatry; AGP= Archives of General Psychiatry (now called JAMA psychiatry); BJP=British Journal of Psychiatry; BP= Biological Psychiatry; NPP= NeuroPsycho Pharmacology; SB= Schizophrenia Bulletin; MOL=Molecular Psychiatry.

Figure 1. Trajectories of Flesch readability ease of journals from years 2002-2013

Readability is also important for researchers. Clear literature reviews, research questions, methods, results, and explanations of findings are paramount for understanding, but also for informing future research and advancing theory. In today's knowledge landscape, mental health research crosses many different disciplines. Using common language (ie, not jargon) that is easy to read and understand will help research transcend the barriers between disciplines. This will advance knowledge, drive innovation, and improve the standard of care for people with mental health problems.

Making psychiatry research more readable may also expand its audience. Patients, family members, and an increasingly scientifically-literate global public would be able to read such work. The Internet offers increased access to information, but only if the material can be read and understood by users.[23] With the growing movement toward Open Access publishing, primary research articles are increasingly a part of freely available public health information. Editors and authors may consider this wide range of target audiences and how this information can be used by the general public, other research disciplines, clinicians, and policymakers.

Our results suggest several ways to make peer-reviewed psychiatry journals more readable. First, readability can be improved by keeping sentences short and to the point. Ninety-five percent of people will understand a sentence of eight words having only read it once. Only 4% will understand a sentence of 27 words after a single reading.[24] Second, the Abstract is the most-read section of a paper. It is read by laypeople, busy clinicians, individuals from other disciplines, and those without paid access to the full article.[20] We recommend that Abstracts include all of the main points of the article, with minimal jargon, short sentences, and a clear take-home message. Most Abstracts are fewer than 300 words. This encourages concision but sometimes sacrifices clarity. Testing the readability of an Abstract can take less than 30 seconds using a free online tool. Third, we encourage authors to adopt an active rather than a passive voice. In other words, they should say, "We analyzed data" rather than "Data were analyzed" to increase readability.[13] Fourth, authors should consider that their work will increasingly be accessed online or on a hand-held device. Reading speed and attention decrease online, so we recommend the use of simple language and precise sentence structure.[20] We hope these recommendations will improve the spread of information and enhance patient care.

Using these recommendations, the present paper is an example appropriate for an American equivalent of grade 12 reading level (FRES=36.3).

Despite accessing 504 articles, we recognize that this is still a relatively small representation of available information in the pool of peer-reviewed psychiatry articles. Future assessment should include a more elaborate evaluation of language, research study design, research topic content, context, open access, and the impact of time on readability. As well, we also recognize that we only compared general medicine articles from one year. Future work should include multiple years for more robust comparisons to be made.

Conclusions

Readability formulas have been around for over 70 years. In the past, the sheer time it took to use the formulas posed a challenge. However, several free online readability tools are now available. We found that on average, articles published in peer-reviewed psychiatry journals are *very difficult* to read. This trend has been stable for each journal across at least the last twelve years (2002-2013). Given the large amount of evidence clinicians and researchers need to sift through each day, we suggest that articles need to be more readable. Clear writing can increase the potential of a published article to guide meaningful change in clinical practice and research. Psychiatry journals may benefit from adopting readability guidelines or minimums to ensure that the work they publish can be read, understood, and implemented.

Table 2. Median reading level (25-75 percentile) and word count of commonly read journals in psychiatry using common readability formulas on the Abstract, Introduction, and Discussion of randomly selected original research articles from 2002-2013 (n=504 total psychiatry articles; N=78 general medicine articles)

Journal	FRES[a]	FKGL	GFOG	SMOG	Coleman Liau	Average of all measures	Word count	Words/ Sentence
AACAP								
Abstract	23.6 (16,34)	14.8 (13,16)	17.7 (15,20)	12.9 (12,14)	17.0 (15,19)	15.4 (13,17)	230 (194,266)	18.6 (15,22)
Introduction	17.8 (13,27)	17.0 (15,18)	20.3 (18,22)	14.7 (14,16)	17.0 (16,18)	17.3 (16,18)	634 (482,792)	24.5 (21,27)
Discussion	22.2 (14,30)	16.1 (15,17)	19.2 (17,21)	14.0 (13,15)	16.5 (16,18)	16.6 (15,17)	1194 (843,1470)	23.9 (21,26)
All	22.6 (17,29)	15.8 (15,17)	19.2 (18,20)	13.9 (13,15)	16.7 (16,18)	16.5 (15,17)	2143 (1797,2384)	22.1 (20,25)
AJP								
Abstract	13.7 (7,22)	16.9 (16,19)	20.8 (19,23)	15.0 (14,17)	18.5 (16,20)	17.9 (17,19)	248 (230,270)	23.7 (21,26)
Introduction	12.2 (4,22)	17.9 (17,19)	21.4 (17,19)	15.4 (15,17)	18.0 (16,19)	18.3 (17,20)	466 (405,647)	25.8 (23,28)
Discussion	18.2 (11,24)	17.2(16,18)	21.2 (20,24)	15.2 (14,16)	17.2 (16,18)	17.6 (17,19)	1266 (1020,1420)	26.0 (24,28)
All	17.2 (9,23)	17.4 (16,18)	21.1 (20,22)	15.3 (14,16)	17.7 (16,19)	17.6 (17,19)	1993 (1746,2217)	25.3 (24,28)
AGP								
Abstract	15.9 (10,23)	16.3 (15,19)	19.8 (18,20)	14.4 (13,16)	18.3 (18,20)	17.0 (16,19)	276 (251,292)	21.1 (19,24)
Introduction	11.1 (4,18)	18.3 (17,20)	22.3 (21,24)	16.1 (15,17)	18.2 (17,19)	18.8 (18,20)	542 (394,688)	27.0 (24,30)
Discussion	18.5 (13,23)	17.6 (16,18)	21.2 (20,22)	15.2 (14,16)	17.2 (16,18)	18.1 (17,19)	1258 (1042,1470)	27.2 (25,29)
All	17.0 (11,22)	17.0 (16,18)	21.1 (2022)	15.2 (15,16)	17.6 (17,18)	18.0 (17,19)	2057 (1813,2354)	26.5 (24,28)
BJP								
Abstract	20.0 (11,33)	15.6 (14,17)	18.9 (17,21)	13.5 (13,15)	17.6 (16,20)	16.0 (15,18)	161 (141,172)	18.8 (17,21)
Introduction	16.0 (12,24)	17.0 (16,19)	21.0 (20,23)	15.2 (14,16)	17.6 (16,19)	17.9 (17,19)	195 (178,359)	24.8 (23,28)
Discussion	23.6 (17,28)	16.5 (15,18)	20.2 (19,21)	14.4 (14,15)	16.7 (16,17)	16.8 (16,18)	1110 (796,1366)	26.0 (24,28)
All	23.5 (16,26)	16.2 (15,17)	20.1 (19,21)	14.3 (14,15)	16.9 (16,19)	16.8 (16,18)	1534 (1100,1879)	24.5 (23,27)

NPP								
Abstract	17.6 (11,23)	16.9 (15,18)	21.0 (18,22)	15.2 (13,16)	17.2 (16,19)	17.3 (16,19)	241 (219,258)	24.0 (21,27)
Introduction	14.7 (11,22)	18.1 (17,20)	21.7 (20,23)	15.9 (15,17)	15.8 (15,18)	18.2 (17,19)	605 (527,764)	28.1 (26,32)
Discussion	22.7 (16,26)	17.1 (16,18)	20.7 (20,22)	15.0 (14,16)	15.8 (14,17)	17.4 (16,18)	1485 (1152,1828)	27.5 (25,31)
All	20.0 (15,25)	17.2 (16,18)	21.0 (20,22)	15.2 (14,16)	16.1 (15,17)	17.4 (16,18)	2346 (1943,2782)	27.3 (25,30)
SB								
Abstract	13.2 (2,24)	17.1 (15,19)	21.0 (19,23)	15.1 (14,16)	18.6 (17,21)	18.4 (16,20)	212 (172,243)	238 (20,28)
Introduction	12.0 (6,20)	18.6 (17,20)	22.0 (20,23)	16.0 (15,17)	18.4 (17,19)	19.1 (17,20)	609 (406,850)	26.2 (24,30)
Discussion	19.7 (13,26)	17.0 (16,18)	20.4 (19,22)	14.7 (14,16)	17.3 (16,18)	14.7 (14,16)	1123 (759,1530)	25.5 (23,29)
All	16.7 (9,24)	17.4 (16,19)	21.0 (19,22)	15.1 (14,16)	17.8 (17,19)	18.0 (17,19)	2143 (1565,2431)	25.0 (23,28)
BP								
Abstract	20.0 (13-28)	15.0 (13-17)	18.8 (17-21)	13.1 (11-15)	18.4 (17-20)	15.8 (14-18)	231 (195-247)	19 (12-23)
Introduction	11.6 (5-18)	17.8 (17-20)	22.0 (21-23)	15.7 (15-17)	18.4 (18-19)	18.6 (17-20)	517 (414-635)	25.4 (22-29)
Discussion	18.2 (15-23)	17.1 (16-18)	20.8 (20-22)	14.9 (14-16)	17.5 (17-19)	17.7 (16-19)	1184 (899-1413)	25.5 (23-29)
All	17.6 (14-21)	16.6 (16-18)	20.3 (20-22)	16.6 (14-16)	17.1 (16-18)	17.2 (17-18)	1991 (1734-2305)	23.8 (21-27)
MP								
Abstract	16.6 (8-26)	16.7 (15-19)	20.2 (18-22)	14.7 (13-16)	17.3 (16-19)	16.8 (15-19)	216 (184-241)	23.6 (19-27)
Introduction	14.8 (8-20)	17.4 (17-18)	21.2 (20-23)	15.3 (15-16)	17.4 (16-19)	17.8 (17-19)	479 (369-623)	25.1 (22-28)
Discussion	22.5 (14-28)	16.8 (16-18)	20.6 (19-22)	14.7 (14-16)	15.9 (15-17)	17.0 (16-18)	989 (775-1484)	26.1 (23-29)
All	21.1 (13-26)	16.8 (16-18)	20.2 (19-22)	14.6 (14-16)	16.4 (16-17)	16.8 (16-18)	1761 (1432-2221)	25.5 (23-27)
NEJM*								
Abstract	28.1 (23-26)	15.6 (13-16)	19.3 (16-21)	13.7 (12-15)	15.1 (14-16)	16.0 (13-17)	263 (247-296)	24.0 (18-26)
Introduction	17.7 (4-28)	17.8 (16-20)	21.5 (19-23)	15.4 (14-16)	17.4 (15-18)	17.8 (16-20)	276 (240-331)	27.8 (24-31)
Discussion	23.6 (17-30)	16.6 (15-18)	20.3 (18-21)	14.9 (13-15)	16.0 (15-17)	17.1 (15-18)	888 (764-1088)	25.7 (23-29)
All	24.1 (19-29)	16.0 (15-17)	19.7 (18-21)	14.4 (13-15)	16.0 (15-17)	16.7 (16-17)	1440 (1267-1595)	24.8 (22-27)
Lancet*								
Abstract	30.6 (24-34)	14.4 (14-15)	17.5 (17-19)	12.9 (12-14)	15.4 (14-17)	14.6 (14-16)	322 (250-362)	21.4 (19-24)
Introduction	17.8 (11-25)	17.3 (15-18)	20.1 (19-21)	14.9 (14-16)	17.8 (17-18)	17.6 (16-17)	259 (202-323)	25.8 (23-27)
Discussion	24.3 (21-29)	15.9 (15-17)	19.2 (19-21)	13.9 (14-15)	13.9 (13-15)	16.6 (16-17)	981 (819-1202)	24.7 (24-27)
All	23.0 (21-30)	15.7 (15-17)	19.1 (18-20)	13.9 (13-15)	13.9 (13-14)	16.3 (15-17)	1578 (1362-1713)	23.8 (23-26)
JAMA*								
Abstract	36.6 (29-41)	10.9 (9-13)	13.1 (12-15)	9.7 (8-11)	12.9 (11-15)	11.3 (8-13)	386 (342-429)	11.5 (10-13)
Introduction	11.3 (9-13)	17.9 (16-19)	20.5 (19-22)	14.8 (14-19)	17.2 (16-19)	17.9 (16-19)	286 (268-316)	23.0 (20-29)
Discussion	16.3 (15-18)	16.3 (15-18)	19.9 (18-21)	14.1 (13-15)	17.0 (16-18)	16.8 (15-18)	1070 (887-1336)	23.0 (20-26)
All	28.1 (18-33)	14.9 (13-16)	16.9 (16-19)	12.5 (12-14)	16.0 (15-17)	14.8 (14-16)	1766 (1546-1989)	19.0 (17-20)

[a] All scores from readability tests were highly correlated with the FRES (r_s >0.8, P <0.01).

*= statistics for the year 2013 only. AACAP= American Academy of Child and Adolescent Psychiatry; AJP= American Journal of Psychiatry; AGP= Archives of General Psychiatry; BJP= British Journal of Psychiatry; BP= Biological Psychiatry; NPP= Neuro Psycho Pharmacology; SB= Schizophrenia Bulletin; MP= Molecular Psychiatry NEJM= New England Journal of MEdicine; JAMA= Journal of the American Medical Association

Table 3. Pairwise comparisons of journals and publication year according to the readability scores

Outcome	Journal F(7, 503) P	Pairwise Comparisons**	Year F(6, 503) = P	Pairwise Comparisons**	Journal x Year F(42, 503) P
FRES – Abstract	5.66, <0.001	AACAP > AJP, SB, AGP, NPP, MP; BJP > SB; BP > SB	1.64 0.14	-	1.53 0.02
FRES – Introduction	4.14*, 0.002	AACAP > SB, AGP, BP	2.81* 0.001	2004 > 2013	0.95 0.57
FRES – Discussion	5.41*, <0.001	AACAP > AJP, SB, AGP, BP	3.54* <0.001	2004 > 2013	1.39 0.06
FRES – All	6.68* <0.001	AACAP > AJP, SB, AGP, BP	3.47* <0.001	2004 > 2010	1.30 0.10

* P < .05; ** Bonferonni correction at p=0.001

FRES= Flesch Kinkaid Ease of Reading Score; AVG= Average readability grade score. This score is calculated as a mean of the Flesch Kinkaid grade level, Gunning-Fog Score (GFOG)[14], (4) Coleman-Liau Index (CLI)[15], and (5) the Simple Measure of Gobbledygook; AACAP= American Academy of Child and Adolescent Psychiatry; AJP= American Journal of Psychiatry; AGP= Archives of General Psychiatry; BJP=British Journal of Psychiatry; BP= Biological Psychiatry; NPP= Neuro Psycho Pharmacology; SB= Schizophrenia Bulletin; MP=Molecular Psychiatry.

Table 4: Linear Regression Model for Flesch Readability Ease Score (FRES) of Psychiatry Journals (R=0.35, R2 =0.12; N = 504)

Characteristic	Beta	SE.	P-value
Year of Publication[a]	-0.24	0.12	<0.01
Impact Factor	-0.38	0.13	<0.01
Type of Article[b]	-1.75	1.64	0.28
Number of times cited	0.004	0.004	0.32
Rater[c]	0.66	0.27	0.16

SE – standard error[a] *Years included 2002, 2004, 2006, 2008, 2010, 2012, & 2013*

[b] *Original research article referent group (versus reviews)*

[c] *5 raters*

Acknowledgements

Financial support was provided by the Canadian Institutes of Health Research (CIHR) Social Aetiology of Mental Illness (SAMI) Training Program at the Centre for Addiction and Mental Health (CAMH), Toronto, Ontario, Canada. We would like to thank the Evidence Exchange Network at the CAMH for supporting this project and improving the readability of the paper. We also thank Emma Ware, PhD, Laurie Manwell, PhD, and David Barbic, MD for feedback and edits on the project and manuscript. Finally, we would like to thank David Child for his technical assistance throughout the data collection process.

References

1 Jones T, Hanney S, Buxton M, Burns T. What British psychiatrists read: questionnaire survey of journal usage among clinicians? *The British Journal of Psychiatry.* 2004;185(251-257).

2 Haynes RB. Where's the meat in clinical journals? *American College of Physicians Journal Club.* 1993;119:A23-24.

3 Balas E, Boren S. Managing Clinical Knowledge for Health Care Improvement. In: van Bemmel JH, McCray AT, eds. *Yearbook of Medical Informatics*. Stuttgart:: Schattauer Verlagsgesellschaft 2000:65-70.

4 Morris ZL, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. *Journal of the Royal Society of Medicine.* 2011;104 (12):510-520.

5 Edmunds MR, Barry RJ, Denniston AK. Readability Assessment of Online Ophthalmic Patient Information. *JAMA Opthamology.* 2013;131(12):1610-1616.

6 Tulbert BH, Snyder CW, Bodell RT. Readability of patient-orientated dermatology resources. *Journal of CLinical Aesthetics and Dermatology.* 2011;4(3):27-33.

7 Baker GC, Newton DE, Bergstresser PR. Increased readability improves the comprehension of written information for patients with skin disease. *Journal of the American Academy of Dermatology.* 1988;19(6):1135-1141.

8 Edmunds MR, Denniston AK, Boelaert K, Franklyn JA, Durrania OMT. Patient Information in Graves' Disease and Thyroid-Associated Ophthalmopathy: Readability Assessment of Online Resources *Thyroid.* 2014;24(1):67-72.

9 Berland GK, Elliott MN, Morales LS, et al. Health Information on the InternetAccessibility, Quality, and Readability in English and Spanish. *JAMA.* 2001;285(20):2612-2621.

10 Larsen PO, von Ins M. The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Scientometrics.* 2010;84(3):575-603.

11 Durbin CG. How to read a scientific paper. *Respiratory Care.* 2009;54(10):1366-1371.

12 Flesch R. A new readability yardstick. *Journal of Applied Psychology.* 1948;32(3):221-233.

13 Kinkaid JP, Fishburn RP, Rodgers RL, Chissom BS. *Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel.* Springfield, IL: Naval Technical Training Command; 1975.

14 Gunning R. *The technique of clear writing: revised edition.* New York, NY: McGraw-Hill Book Company; 1969.

15 Coleman M, Liau TL. A computer readability formula designed for machine scoring. *Journal of Applied Psychology.* 1975;60:283-284.

16 McLaughlin GH. SMOG Grading-a New Readability Formula. *Journal of Reading.* 1969;12(8):639-646.

17 Tirlapur SA, Leiu C, Khan KS. Quality of information on the internet related to bladder pain syndrome: a systematic review of the evidence. *International Urogynecology Journal* 2013;24(8):1257-1262.

18 Hariprasad T, Trivedi A, Hannan MF. Readability and comprehensibility of over-the-counter medication labels. *Renal Failure.* 2014;36(3):473-477.

19 Brosnan S, Barron E, Sahm LJ. Health literacy and the clozapine patient. *Perspectives in Public Health.* 2012;132(1):39-42.

20 Lagassé LP, Rimal RN, Smith KC, Storey JD, Rhoades E. Correction: How Accessible Was Information about H1N1 Flu? Literacy Assessments CDC Guidance Documents for Different Audiences. *PLoS ONE.* 2011;6(11):e23583.

21 Guyatt G, ed *Users' Guide to the Medical Literature.* 2nd ed. New York: McGraw-Hill Professional; 2008.

22 Gopalakrishnan D, Ganeshkumar P. Systematic Reviews and Meta-analysis: Understanding the Best Evidence in Primary Healthcare. *Journal of Family Medicne and Primary Care.* 2013;2(1):9-14

23 McLellan F. Like hunger, like thirst": patients, journals, and the Internet. *LANCET.* 1998:SII39-SII43.

24 Isbell P. *A Guide to Letter Writing.* England: Industrial Society 1979.

Appendix A: Standardized Operating Procedure for the Readability Study

The following Standardized Operating Procedure has been developed for the Readability Study. It has two steps: (1) Getting journal articles, and (2) Testing Readability.

STEP 1-Getting Journal Articles:

1. Select journal you will review.
2. Go to X folder and note the issues of each journal to be uploaded.
3. For each issue that has been randomly selected, note how many original research articles there are. For example, in issue 2 of Schizophrenia Bulletin, there are 14 articles.
4. Next, go to website www.random.org
5. On the right hand side there is a "true random number generator". There are two boxes called "min" and "max". For "min", enter the number 1. For "max", enter the number of articles for the issue you are looking at. For example, for issue 2 of Schizophrenia Bulletin, Rater would put "14" for max.
6. Click the "generate" button. This will produce a number. This is the first article of that issue that you need to upload. For example, in issue 2 of Schizophrenia Bulletin, the first random number Rater 1 generated is "7". So, she will upload the 7th article from issue 2.
7. Generate two more numbers (cannot be the same number).
8. Record these numbers on the MASTER TABLE and then upload the corresponding articles from that issue. For example: Rater randomly generated numbers 7, 5, and 2.
9. Upload the Original Research Articles.
10. Save PDFs of articles on X-drive folder of corresponding year and journal you are reviewing.
11. Upload references to Endnote Web Software. http://www.myendnoteweb.com/EndNoteWeb.html

STEP 2-Testing Readability

1. InX-drive, open up excel **DATA extraction sheet**
2. Open up your first PDF article to be tested.
3. There is one data extraction file for each rater with their name on it. Each Excel file has two sheets. Sheet 1=data, sheet 2=article details (specifications of the article "SPEC")
4. Record the following on Sheet 2 called "SPECS":
    a. Journal
    b. Year
    c. Issue
    d. Article #
    e. Title of paper
    f. Author
    g. Type of study design
5. Highlight the text from the Abstract.
6. Copy and Paste this text into square on http://www.readability-score.com/
7. Remove references from text. This includes all authors and year. For example, if there is a reference (Barbic, 2002), remove this. Only remove references in parentheses.
8. Remove all headings from text (ie Introduction).
9. Remove anything in brackets that is not a sentence (ie p=0.005, <, >, n=67, F=7.86, etc…)
10. Remove any text that is in a box.
11. Remove all figures, tables, figure headings, table headings, etc…
12. Look away from screen, go for a walk, have a coffee… for at least 30 seconds
13. Look at text again and make sure all references and headings have been removed.
14. Record readability statistics from right hand side of screen including:
    a. Flesch-Kincaid Reading Ease
    b. Flesch-Kincaid Grade Level
    c. Gunning-Fog Score
    d. Coleman Liau
    e. SMOG Index
    f. AVERAGE grade level
    g. Word count
    h. Words/sentence
15. Record these values in **your NEW** data extraction sheet-saved in X-drive.
16. Copy and paste the clean text into Microsoft Word for later use.
17. Repeat steps 5-16 for the INTRODUCTION and the DISCUSSION
18. Take all text from Word and re-paste ALL text back into readability testing box.
19. Repeat steps 14 and 15 for ALL text.
20. Repeat steps 1-19 for each article of each issue selected.

Appendix B: Selected readability formulas used for study

Formula	Equation [a]
Flesch Reading Ease	Reading ease score=206.835-(1.015 x average no. of words per sentence)
Flesch-Kinkaid Grade level	Reading grade level= (0.39 x average no. of words per sentence) + (11.8 x average no. of syllables per word) – 15.59
Gunning Fog	Reading grade level= 0.4 (average no. of words per sentence + no. words with 3 or more syllables x [100/no.of words])
SMOG	Reading grade level= 3 +√[no. of words with three or more syllables] x [30/no. of sentences]
Coleman-Liau	Reading grade level= 0.0588 (average number of letters/100 words) – 0.296 (Average number of sentences/100 words) – 15.8
Average	Average of scores resulting from equations above (except FRE)

[a] Equations are adjusted for nonstandard sample size