



Letter frequency analysis of proprietary prescription drug names in the United States: Minding the Zs and Qs

Ron Carico Jr ^{a,*}, Keaton Kaplan ^a, Kyler Gator Hazelett ^a, Megan Dillon ^b, Kelly Melvin ^a

^a Marshall Health, 1600 Medical Center Drive, Huntington, WV 25701, United States of America

^b Marshall University School of Pharmacy, 1538 Charleston Ave, Huntington, WV 25701, United States of America

ARTICLE INFO

Keywords:

Prescription drug names

Names

Letter frequency

ABSTRACT

Background: Proprietary or brand names of prescription drugs are generally with letters that are unusual in common English. There is little academic research exploring if this perception is true, despite the fact that manufacturers pay millions of dollars to research and develop drug names that conform to regulatory standards while remaining marketable.

Objectives: To assess the extent to which letters used in prescription drug names may deviate from common English and to test if prescription drug names show measurable trends in letter frequency over time.

Methods: The names of all prescription drugs approved in the United States between 1985 and 2020 were downloaded. Duplicates were removed and products without a proprietary name were excluded. Letter frequency analyses were then conducted on all letters in these names as an aggregate and year-over-year. Letter frequencies were compared to a validated academic reference, a corpus derived from all Google Books data, and the scoring system from the board game Scrabble.

Results: Regardless of the comparator, prescription drug names use letters that are not common in typical English. Letters A (11.96% of all observed letters), V (3.08%), X (2.31%), and Z (1.91%) are all overrepresented in prescription drug names, while E (10.23%), H (0.90%), T (6.30%), and S (4.21%) are underrepresented. The letters C and N are becoming less common over time (frequency decrease of 0.10 percentage points and 0.12 percentage points per year, respectively), while V, Y, and Z are becoming more common (frequency increases of 0.61 to 0.86 percentage points per year).

Conclusions: Proprietary prescription names use letters that are unlike words used in everyday American English, and there are measureable trends in letter selection. It remains to be seen how drug manufacturers will cope with an increasingly-narrow naming space as more products continue to be approved over time.

1. Introduction

Proprietary names of prescription drug products—also known as trade names or brand names—have been the subject of commentary from mass media news organizations and other layperson-oriented outlets for years. These reports describe proprietary drug names as unusual or humorous, and often hold that medication names are strange and becoming stranger over time.^{1–4} Academic commenters have stated that manufacturers have an “insatiable proclivity” for the letters X and Z in journals such as *New England Journal of Medicine*.⁵ However, proprietary drug names are not arbitrary: they are the result of a process that must take into account regulatory requirements and clinical considerations,⁶ as well as marketing concerns applicable to the United States of America (U.S.) and internationally. These external factors sometimes work at cross-purposes, which may contribute to identifiable trends in proprietary drug names.

With respect to regulation, all newly-approved prescription drug products in the United States are given at least two names: a generic name—which uniquely refers to the product's active ingredient(s)—and a proprietary name. The proprietary name is usually held as a trademark by the first manufacturer of the new product, but both names are subject to regulatory approval. First, proprietary names must be unique, and will be screened for similarity to other proprietary or nonproprietary names.⁶ There are four main sources of confusion for medication names, each of which must be captured by this screening step: “different drugs with similar names, formulations with the same brand name containing different drugs, the same drug marketed in formulations with different names, and abbreviated drug names.”⁷ For a clinically-relevant example of the first type of confusion, the antidepressant Brintellix® had its name changed after being confused with the antiplatelet Brilinta®.⁸ To illustrate the scope and gravity of the problem, the Institute for Safe Medication Practices offers a list of confused drug names

* Corresponding author at: 1115 20th Street, Huntington, WV 25703, United States of America.
E-mail address: caricor@marshall.edu (R. Carico).

to create “safeguards to reduce the risk of errors and minimize harm” amongst dispensing clinicians.⁹ Additionally, proposed proprietary names are not likely to meet regulatory approval if they contain letters or numbers, include references to product-specific attributes (e.g., “Nametabs” for a tablet formulation), or are not pronounceable.⁶ Close scrutiny will be applied to proprietary names that allude to the product's intended use or manufacturer, or names that may overstate the product's effectiveness (e.g., by incorporating a variation of the word “cure” when the product only mitigates disease).⁶ Proprietary names that violate these guidelines are unlikely to be approved in the U.S.

The constraints placed upon these companies by regulatory requirements are compounded by marketing considerations. Prescription drug developers are often private corporations that are interested in using the proprietary name of a product to maximize profit through branding and/or advertising. For instance, manufacturers would likely prefer to use the same proprietary name in all markets around the world, as doing so will help with brand recognition and minimize the amount of retooling needed for branding and labeling across diverse countries. This means that the ideal proprietary name would conform with regulatory requirements in the U.S. and all international markets. This also limits which characters can be used in a proprietary name: some languages use diacritics or accent marks as pronunciation guides, while these characters are practically forbidden by U.S. regulatory guidance.⁶ Similarly, some letters that are commonly used in English may be all-but-absent in other languages, and some characters may be pronounced differently in different parts of the world (such as the letter “j” in English and Spanish).

Moreover, the ideal proprietary name would be memorable and pleasant, while adhering to regulatory limits. In the U.S., this may mean that a proprietary drug name should be fewer than three syllables with a sound that evokes positive connotations or sentiment.¹⁰ Drug manufacturers often hire marketing firms to develop proprietary names for their products. As many as five such names may need to be submitted for review, and each name may cost more than \$1 million to develop, research, and market-test, according to an interview with an executive from a company that may have been involved in development of 75% of proprietary names approved in the U.S.¹⁰

Despite the high amount of attention from the lay press, the potential safety implications of easily-confused medication names, and the obvious monetary interest in proprietary drug product names, relatively little administrative pharmacy research exists that explores trends in proprietary drug names. The purpose of this study is to quantify the extent the letters used in proprietary drug names may deviate from common English and to test if prescription drug names show measurable trends in letter frequency over time.

2. Methods

The compilation of approvals of new molecular entity drug and new biologic agents was downloaded from the U.S. Food and Drug Administration (FDA) Center for Drug Evaluation and Research.¹¹ This list contained all new molecular entities and new combination prescription products approved between 1985 and 2020; it did not include labels with new indications for previous approvals, new variations of previously approved active ingredients (e.g., new salt forms), or new dosage forms of previously approved products (e.g., a liquid formulation of a medication previously available as a tablet).¹² Non-alphabetic characters, including spaces, were removed from proprietary names in the dataset. Products were excluded from analysis if they were listed as “marketed without a proprietary name.” Duplicated proprietary names were excluded after the chronologically first mention. Proprietary names that included references to dosage forms (e.g., “[Medication Name] Depot”) were retained as-is for analysis, as were products whose proprietary names seemed to be identical to the name of the active ingredient. The FDA's dataset included details on the medication type (drug versus biologic) and administrative classifications applied to each new drug approval. These administrative classifications included drug versus biologic agent status, priority review versus standard review, orphan versus non-orphan drug, and whether the approval was eligible for so-called “accelerated approval,” “breakthrough medication,”

fast track approval,” or “qualified infectious disease” status. Additionally, the dataset included details on whether the medication led to the issuance of a priority review voucher or whether a priority review voucher was redeemed in the approval of the medication. These variables were summarized to better characterize the medications approved by the FDA since 1985.

The primary analysis centered on the letters in each medication's proprietary medication name. Each letter in the product proprietary names was assigned a frequency score using three widely-available rankings of letter frequency in common English. The principle analysis was performed with a list of letter frequencies compiled by Mayzner and Tresselt from a corpus of 20,000 English words randomly selected from fiction and nonfiction sources in the 1960s.¹³ A secondary analysis was constructed using a non-peer reviewed update to the Mayzner-Tresselt frequency list that was compiled from the corpus of Google Books data by Norvig in 2012.¹⁴ This reference was used to face-validate the output of the older Mayzner-Tresselt listing against a more modern corpus. Finally, frequency scores were also assigned using the point values from the commercially-available board game Scrabble, where less-frequently used letters are assigned higher point values.¹⁵ This reference was used for its familiarity to professional and lay readership in the U.S. For example, the letter E occurred in 13.3% of all letter observations in the Mayzner-Tresselt rankings and 12.5% of all observations of the Norvig rankings. Meanwhile, the Scrabble point value for the letter E is 1. Conversely, the letter Z occurred in 0.06% of all observations in Mayzner-Tresselt rankings and 0.09% of all observations in Norvig rankings; this letter is assigned a point value of 10 in Scrabble scoring. A full list of Mayzner-Tresselt, Norvig, and Scrabble frequency scoring used for all letters is available in the online supplement.

The proprietary names from the dataset were then assigned frequency scores for each scale by adding Mayzner-Tresselt letter frequencies, Norvig letter frequencies, or Scrabble scores for each letter in each name. As an example calculation, adding the Mayzner-Tresselt frequencies for the word “paper” would result in a total of 30.37 percentage points (1.53 percentage points for each P, 8.10 points for the A, 13.32 points for the E, and 5.89 points for the R). To correct for longer drug names, these total frequency scores were divided by the length of their respective proprietary names to create a letter frequency-per-letter score, which should be a reasonable representation of how unusual each proprietary name may seem to a person fluent in English. For example, dividing the total Mayzner-Tresselt score for the word “paper” (30.37 points) by its length (5 letters) gives a value of 6.07 percentage points-per-letter, making the average letter frequency score for this word comparable to that of the letter S for the Mayzner-Tresselt corpus. These letter frequency-per-letter scores were plotted against the year of U.S. FDA approval. After inspecting visually for linearity and testing for potential heteroscedasticity using White's test, univariate linear regressions were performed to assess how letter rarity in U.S. prescription drug products may be changing over time, using frequency-per-letter as the dependent variable and year as the independent variable.

Frequencies for each individual letter were also calculated across all years and for each year in the dataset. A linear regression model was constructed for each letter over time, with the year of approval as the independent variable and the percentage frequency of that letter in that year acting as the dependent variable. The output of each of these models was also inspected visually for linearity before White's test was used to assess for heteroscedasticity. The rate of change in letter frequency for each year was calculated, with its accompanying *p*-value and correlation coefficient, R^2 . As a safeguard against false positives due to multiple comparisons, *a priori* limitations were set on interpretation of these values. Results were only regarded as statistically and practically significant if White's test resulted in a *p*-value of greater than 0.05, the *p*-value of the rate of change in letter frequency was less than 0.05, and the correlation coefficient was greater than 0.30. As a post hoc test of the assumptions of linear regression, the individual letter models' residual plots were also inspected visually to assess the assumption of independence of residuals and Q-Q plots were inspected visually to assess the assumption of normality in residuals.

Letter rarity scoring was calculated using Microsoft Excel 365. Tests of heteroscedasticity and linear modeling were conducted in SPSS 64-bit version 26.0.0.0. This study was conducted using only publicly available data. Thus, the study was deemed exempt from review by the Institutional Review Board of Marshall University.

3. Results

The FDA dataset contained 1149 entries. Nine of these were explicitly stated to be marketed without a proprietary name, and 3 were found to be duplicates of items listed previously in the dataset. This left 1137 medication names for analysis. A breakdown of the types of approvals listed in the dataset, including types of reviews (New Drug Applications versus Biologics License Applications), priority status of reviews, and orphan drug status of reviews is provided in Table 1. The number of approvals per year from 1985 to 2020 is plotted in Fig. 1. While there is a general trend toward more approvals per year over time, the R^2 value of 0.1525 suggests that only 15% of the variation in number of approvals is correlated with the year of approval. The average length of proprietary names was 7.41 characters, with a standard deviation of 2.17 characters.

The scatter plots and regressions of letter frequency scores against year of approval is shown in Fig. 2. Mayzner-Tresselt letter frequency per letter decreases over time in a homoscedastic manner (White's test p -value: 0.443), with each letter becoming, on average, 0.01% rarer each year according to the slope of the regression equation. The R^2 value of this model suggests that 38.5% of the variation in letter rarity per letter is accounted for by

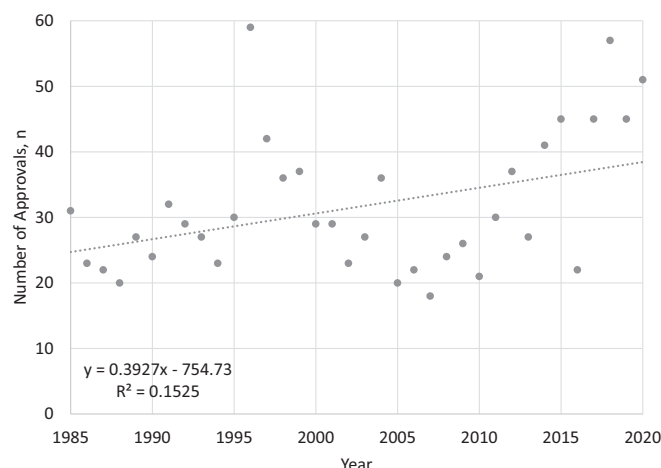


Fig. 1. Number of newly approved proprietary products by year, 1985 to 2020.

variation in year. Similar results were obtained with Norvig letter frequency per letter (visible in online supplement), which homoscedastically (White's test p -value 0.728) decreased by 0.02% each year, with variation in year of observation accounting for 46.5% of the variation in this measure. Finally, Scrabble scores per letter homoscedastically (White's test p -value 0.123) increased by 0.0096 points per letter per year, and the year of observation accounted for 41.9% of the variation in this measure.

The frequency occurrence of each letter, the observed percentage frequency of each letter and the expected Mayzner-Tresselt frequencies are shown in Fig. 3. Table 2 also summarizes White's test results for each letter's linear modeling of frequency over time, the rate of change in frequency of each letter per year, the p -value for the rate of change per year, and the correlation coefficient for the model. The letter A was the most frequently observed letter in the dataset, accounting for approximately 12% of all letters observed; this made it almost 4 percentage points more frequent than expected based on Mayzner-Tresselt ratings. Other letters that were more frequently observed than expected included V (3.08% of observed letters, approximately 2 percentage points more common than expected), X (2.31% of all observed letters, approximately 2 percentage points more common than expected), and Z (1.91% of all observed letters, approximately 1.8 percentage points more common than expected). Letters that were less common than expected included H, which comprised 0.90% of all letters in the dataset, making it some 6.8 percentage points less common than would have been predicted by the Mayzner-Tresselt frequencies. Other less common than expected letters included E (approximately 3 percentage points less common than expected), T (approximately 3 percentage points less common than expected), and S (approximately 2 percentage points less common than expected).

When constructing linear regression models for how individual letter frequencies have changed over time, White's test suggested that significant heteroscedasticity was present for the letters H and J, thus implying that a linear regression model may be misspecified for these letters. The post hoc visual inspection of residual plots and Q-Q plots suggested that linear regression may not appropriately model the frequency-over-time changes for the letters J, Q, and W; these letters were absent from the dataset (i.e., had a frequency of 0%) in several years. Statistically significant changes over time were present at the 0.05 level for B, C, I, K, N, O, Q, V, Y, and Z. However, only C, N, V, Y, and Z met the prespecified critical value of 0.30 for the correlation coefficient. Relatedly, these letters all had a p -value of less than 0.001 on the statistical significance testing of their slopes. The letters C and N are becoming less common over time, according to the model: the frequency for the letter C is decreasing by 0.100 percentage points per year and the frequency of N is decreasing by 0.119 percentage points. By contrast, V, Y, and Z are becoming more common: V is increasing by 0.064 percentage points per year, Y is increasing by 0.086 points, and Z is increasing by 0.061 points.

Table 1

Summary of proprietary prescription products newly approved by the United States Food and Drug Administration, 1985–2020.

Total n (%)	1137 (100)
Type of approval	
New drug application	959 (84.3)
Biologic license application	178 (15.7)
Review type	
Priority	569 (50.0)
Standard	568 (50.0)
Orphan drug status	
Orphan	342 (30.1)
Not orphan	795 (69.9)
Accelerated approval status	
Accelerated	112 (9.9)
Not accelerated	843 (74.1)
Not applicable ^a	182 (16.0)
Breakthrough medication status	
Breakthrough	93 (8.2)
Not breakthrough	263 (23.1)
Not applicable ^a	781 (68.7)
Fast track approval status	
Fast track	223 (19.6)
Not fast track	540 (47.5)
Not applicable ^a	374 (32.9)
Qualified infectious disease medication	
Qualified infectious disease	16 (1.4)
Not qualified infectious disease	253 (22.3)
Not applicable ^a	868 (76.3)
Priority review voucher issued	
Voucher issued	34 (3.0)
Voucher not issued	444 (39.0)
Not applicable ^a	659 (58.0)
Priority review voucher redeemed	
Voucher redeemed	7 (0.6)
Voucher not redeemed	471 (41.4)
Not applicable ^a	659 (58.0)

^a "Not Applicable" refers to products that were submitted or approved before the relevant program began.

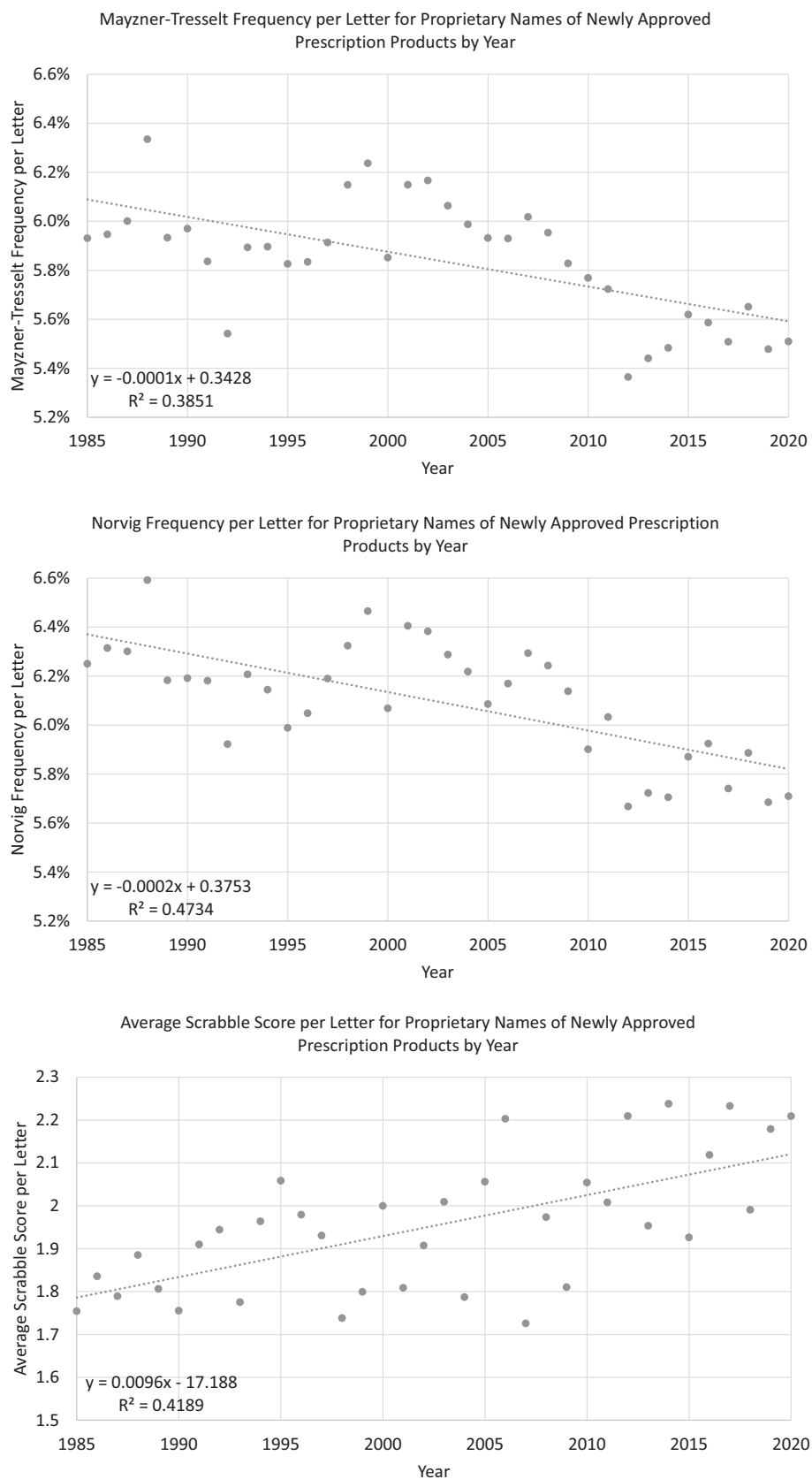


Fig. 2. Average letter frequency scores for proprietary names of newly approved prescription products by year.

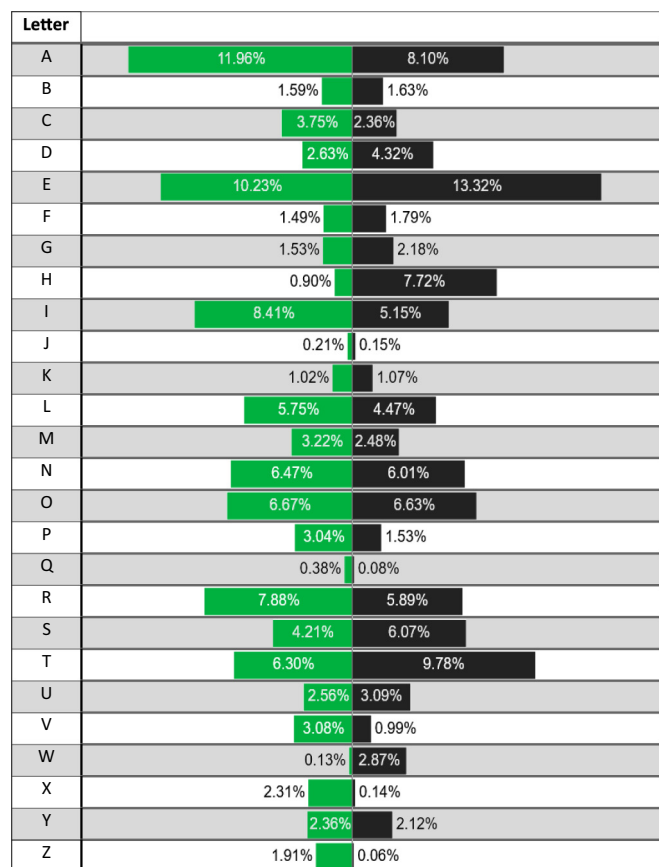


Fig. 3. Letter frequencies in prescription proprietary names (green) versus Mayzner-Tresselt Corpus (black).

4. Discussion

This analysis of proprietary prescription medication names in the U.S. supports the assertion that prescription drug names are comprised of letters that are uncommon in everyday English, and that the divergence with conversational English is becoming more pronounced over time. This finding is consistent when assessing letter frequencies using a peer-reviewed corpus from 1965, a corpus derived from Google Books data in 2012, or the scoring-based system used in the commercial board game Scrabble. Across all time points, letters such as X, Z, and V are overrepresented, while letters such as E, H, and W are underrepresented. This trend has been consistent since 1985, which was the first year in the dataset. If temporal trends continue for another 35 years, the Mayzner-Tresselt average letter-frequency-per-letter will drop from 5.5% in 2020 to 5.15% in 2055. This would be equivalent to a 2.8 percentage point drop in frequency per 8 letters, which would be roughly equivalent to substituting an “m” in an eight-letter word with “x” or “z.” A substitution such as this would change Dormalin (approved in 1985) to Dorzalin, or change Imcivree (approved in 2020) to Ixcivree.

These findings, and their implications for proprietary naming of future medications, have ramifications that extend beyond novelty academic interest. Prescription drug developers face competing interests in selecting a name for their new products. From a marketing perspective, manufacturers will want medication names that are memorable and distinct, and with positive connotations.^{1,4,10} However, FDA best practices and procedures on proprietary medication name development constrain manufacturers' selections by screening for names that may be confused for other product names, names that contain non-alphabetic characters, or for names that make clinically inappropriate implications.⁶ Thus, from a safety perspective, manufacturers should avoid medication names that might look or sound like existing medication names. The FDA has made an algorithm-based

Table 2

Changes in letter frequencies over time.

Letter	White test result	Slope (95% CI)	p-value	R ²
A	0.330	0.069 (−0.016, 0.154)	0.109	0.074
B	0.725	0.045 (0.018, 0.072)	0.002	0.227
C	0.470	−0.100 (−0.141, −0.059)	<0.001	0.420
D	0.195	−0.022 (−0.059, 0.014)	0.225	0.043
E	0.215	−0.045 (−0.104, 0.014)	0.128	0.067
F	0.398	−0.017 (−0.045, 0.012)	0.234	0.041
G	0.656	0.010 (−0.020, 0.040)	0.512	0.013
H	0.006	<i>0.010 (−0.054, −0.013)</i>	<i>0.002</i>	<i>0.246</i>
I	0.462	0.062 (0.015, 0.108)	0.012	0.173
J	0.044	<i>0.018 (0.007, 0.029)</i>	<i>0.002</i>	<i>0.243</i>
K	0.650	0.042 (0.017, 0.067)	0.002	0.255
L	0.164	0.013 (−0.029, 0.055)	0.539	0.011
M	0.593	−0.030 (−0.069, 0.010)	0.135	0.064
N	0.364	−0.119 (−0.170, −0.068)	<0.001	0.397
O	0.499	−0.107 (−0.177, −0.037)	0.004	0.221
P	0.213	−0.017 (−0.046, 0.012)	0.244	0.040
Q	0.118	0.021 (0.006, 0.035)	0.008	0.191
R	0.324	−0.022 (−0.074, 0.030)	0.401	0.021
S	0.233	0.037 (−0.008, 0.083)	0.103	0.076
T	0.523	−0.006 (−0.062, 0.049)	0.819	0.002
U	0.113	0.010 (−0.023, 0.043)	0.558	0.010
V	0.451	0.064 (0.033, 0.096)	<0.001	0.340
W	0.976	0.004 (−0.007, 0.008)	0.949	0.000
X	0.310	−0.018 (−0.051, 0.014)	0.261	0.037
Y	0.140	0.086 (0.054, 0.118)	<0.001	0.471
Z	0.542	0.061 (0.032, 0.090)	<0.001	0.348

Bold: $p < 0.05$

Italics: statistically significant result of White's test suggests linear regression may not be appropriate

phonographic and orthographic assessment tool available to the public¹⁶; manufacturers can use this tool to assess if their proposed product name may be similar to other product names in a way that may compromise patient safety or intellectual property rights. As the pool of approved proprietary names continues to grow, it is reasonable to assume that selecting a sufficiently distinctive proprietary name will become exponentially more difficult.

Further, optimizing the product name for simplicity places constraints on name length and optimizing for marketability places constraints on letter and phoneme selection. Marketing and memorability, which intersect with temporal trends in public taste and perception, may push manufacturers toward using letters that are less common. This seems to be reflected in the trend of manufacturers to increase use of the letters V, Y, and Z, apparently at the cost of letters such as C and N. However, this process is likely to be saturable: once enough names have been approved using these letters, manufacturers will need to seek out novel combinations to use to avoid potential look-alike/sound-alike issues. This process may even push the public's perception of certain letters as “trendy” or “interesting.” It may be reasonable for manufacturers to begin to explore heretofore-overlooked letters, such as H and W, despite issues that these letters may create in the U.S. and abroad. For instance, the letter H is often aspirated or silent in languages like Spanish or French; medications that use these letters prominently in their names may need to be renamed for other markets. Alternatively, underused letters may create potentially-undesired connotations. For example, care may need to be taken to avoid creating letters with the word W that do not imply uncertainty by looking or sounding like questions.

This study is subject to important limitations. First, it uses only U.S. FDA data since 1985, and is silent on trends in proprietary drug naming before this period. Second, this study only analyzes individual letters; analysis of letter combinations (e.g., two-letter “bigrams”) is beyond the present scope, but may be an interesting target for further research. Third, this study made no attempt to analyze nonproprietary generic medication names. These names are given in accordance with international guidelines, which often standardized by medication type and are less concerned with memorability or marketability. Further, these names may have fixed groups of letters that may skew temporal trends. For instance, approval of beta

blocker products in consecutive years may lead to a spike in the frequency of the “-olo” letter cluster. Future research in this area could address these limitations and build on this work by assessing if newer medication names are less likely to be confused for other medication names, borrowing analytic methods from linguistics to test letter combinations, analyzing the cost of developing medication names, or assessing how nonproprietary names may have changed over time. Additionally, construction of non-linear hurdle models may be needed to accurately model frequencies of letters that are frequently absent from proprietary medication names in given years, such as J, Q, and W.

5. Conclusion

This analysis supports the assumption that proprietary prescription names are unlike words used in everyday American English, and that there are measureable trends in letter selection in these names over time. These trends may be the result of sociocultural and regulatory interplay. It remains to be seen how drug manufacturers will cope with the increasingly-narrow space made available to them for drug names as more products continue to be approved over time.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.rcsop.2022.100146>.

References

- Hirst EJ. Why drug names are so increasingly weird. *Chic Tribune* 2015.Mar 20. Accessed Apr 19 2022: <https://www.chicagotribune.com/business/ct-confusing-drug-names-0323-biz-20150320-story.html>.
- Tirrell M, Chemi E, Fahey M. What's in a name? For drugs, a lot of Zs and Xs. *CNBC* 2015. Apr 4. Updated Jul 24 2019. Accessed Apr 19 2022: <https://www.cnbc.com/2015/04/09/whats-in-a-name-for-drugs-a-lot-of-zs-and-xs.html>.
- Schultz D. With a Name Like Xalkori... *Slate*. Aug 22. Accessed Apr 19 2022: <https://slate.com/technology/2013/08/drug-name-confusion-fda-regulations-and-pharma-create-bizarre-new-names.html> 2013.
- Nordrum A. Why do prescription drugs have such weird names? Blame branding consultants and the FDA. *Int Bus Times* 2015.Jun 24. Accessed Apr 19 2022: <https://www.ibtimes.com/why-do-prescription-drugs-have-such-weird-names-blame-branding-consultants-fda-1981819>.
- Garnick MB. Possible confusion in names of new treatments for prostate cancer. *N Engl J Med* 2013;368(2):194. <https://doi.org/10.1056/NEJMc1213236>.Jan 10.
- U.S. Food and Drug Administration. Best Practices in Developing Proprietary Names for Human Prescription Drug Products; Guidance for Industry. Dec 9 2020. Accessed Apr 19 2022: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/best-practices-developing-proprietary-names-human-prescription-drug-products-guidance-industry>.
- Aronson JK. Medication errors resulting from the confusion of drug names. *Expert Opin Drug Saf* 2005;3:167–172. <https://doi.org/10.1517/eods.3.3.167.31069.3> March.
- Brintellix (vortioxetine). Renamed Trintellix (vortioxetine) in U.S. to Avoid Name Confusion. *Takeda*. May 1. Accessed Apr 19 2022: <https://www.takeda.com/newsroom/newsreleases/2016/brintellix-vortioxetine-renamed-trintellix-vortioxetine-in-u.s.-to-avoid-name-confusion/> 2016.
- Institute for Safe Medication Practices. ISMP's List of Confused Drug Names. Institute for Safe Medication Practices. Accessed Apr 19 2022. <https://www.ismp.org/tools/confuseddrugnames.pdf>
- Dorschner J. What's in a name? Big business, for a company marketing drugs. *Stan Times* 2004.Apr 27. L2. Accessed Apr 19 2022: <https://www.southcoasttoday.com/article/20040427/news/304279949>.
- U.S. Food and Drug Administration. Compilation of CDER New Molecular Entity (NME) Drug and New Biologic Approvals. May 3. Accessed Apr 19 2022: <https://www.fda.gov/drugs/drug-approvals-and-databases/compilation-cder-new-molecular-entity-nme-drug-and-new-biologic-approvals> 2021.
- U.S. Food and Drug Administration. Manual of Policies and Procedures 5018.2: NDA Classification Codes. Nov 4. Accessed Apr 19 2022: <https://www.fda.gov/media/94381/download> 2015.
- Mayzner MS, Tresselt ME. Tables of single-letter and digram frequency counts for various word-length and letter-position combinations. *Psychon Monogr Suppl* 1965;1(2):13–32.
- Norvig P. English Letter Frequency Counts: Mayzner Revisited. Jan 4. Accessed Apr 19 2022: <http://norvig.com/mayzner.html> 2013.
- Scrabble FAQ. Hasbro Gaming. Accessed Apr 19 2022: <https://scrabble.hasbro.com/en-us/faq>.
- U.S. Food and Drug Administration. Phonetic and Orthographic Computer Analysis (POCA) Public Version 1.0. Accessed Apr 19 2022: https://poca-public.fda.gov/name_search.