Classification of Generic Manufacturers and Competition in the Pharmaceutical Industry

Mark Benmuvhar
14 November, 2022

Presented in Partial Fulfilment for the Masters of Science in Data Science and Analytical Storytelling

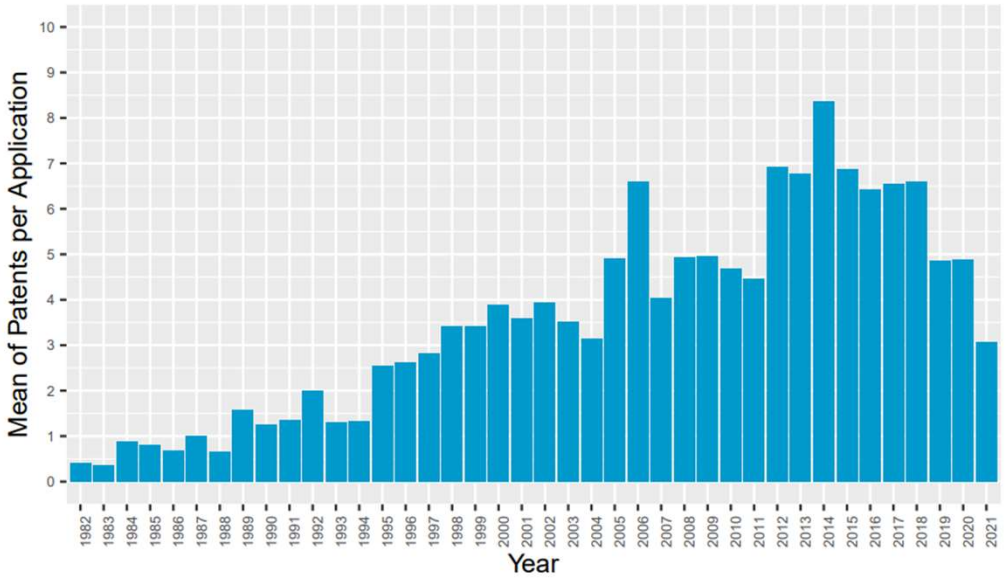
Truman State University

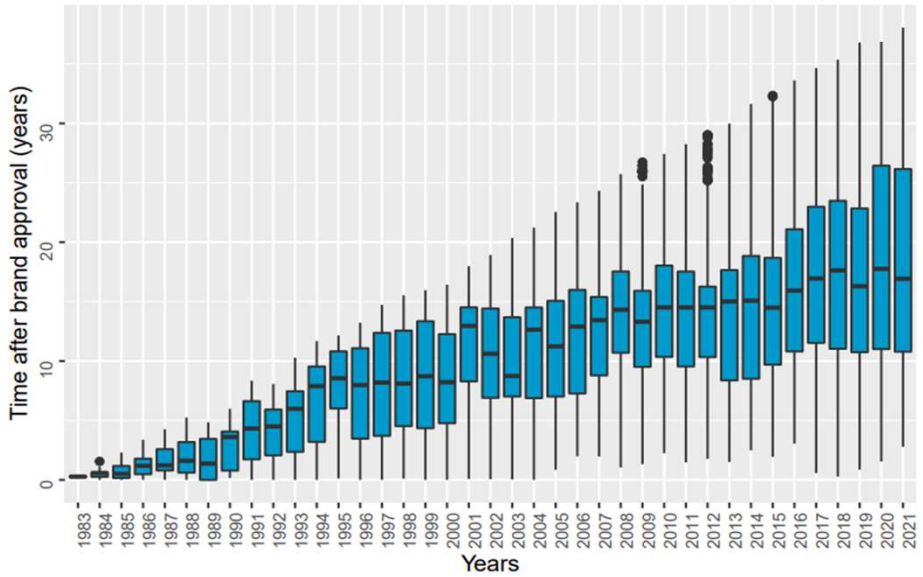
Background

The United States Pharmaceutical Industry provided almost 194 billion daily doses of pharmaceuticals in 2021, representing around $776 billion in wholesale costs and $80 billion in patient out-of-pocket costs (IQVIA, 2022). Pharmaceuticals enter commercial trade after the Food and Drug Administration (FDA) issues a marketing authorization based on either a new drug application (NDA) for a novel ("branded") drug product or an abbreviated new drug application (ANDA) for a generic drug product. An NDA documents extensive studies of safety and efficacy for the drug molecule and the drug product and can represent over a billion dollars of development costs and over a decade of research expenditures (Sun et al, 2022).

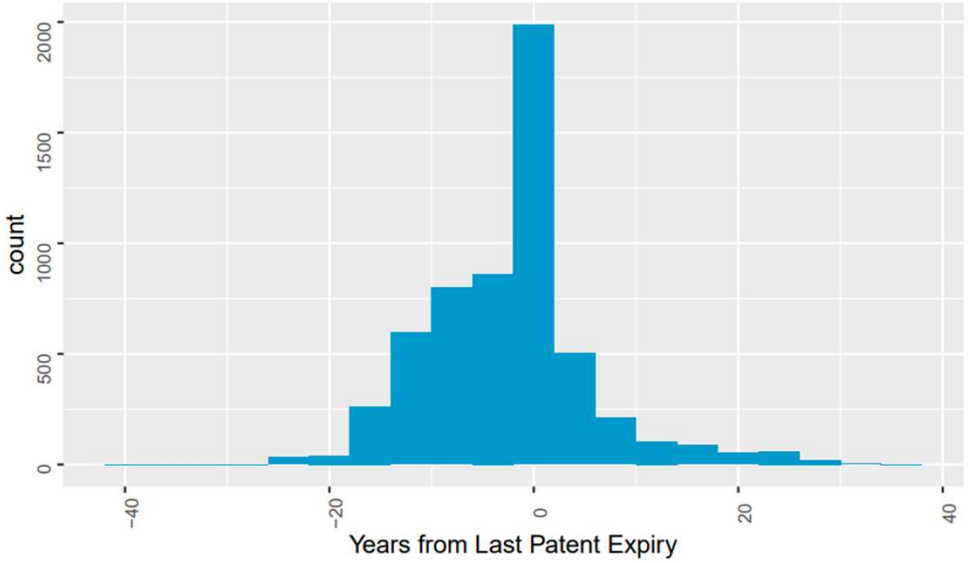
In contrast, a generic sponsor of an ANDA does not need to prove safety and efficacy through clinical trials. Instead, they demonstrate "therapeutic equivalency" to the branded product through a small clinical study in volunteers showing that there are no differences in the rate or amount of drug absorbance into the body. Further, the products must be pharmaceutically equivalent by having the same active pharmaceutical ingredient, same dosage form, same strengths, and same route of administration. (Center for Drug Evaluation and Research (CDER), 2017). Since clinical effectiveness is a function of drug concentration, demonstration of similar drug concentrations makes it unnecessary for generic products to duplicate extensive clinical performance trials required for branded product approval (Midha & McKay, 2009). As a result, two to five years of generic drug development might only cost around $2.5 to 5 million dollars, or around 1/400th the cost of the branded pharmaceuticals.

The 1984 Drug Price and Competition Act balanced the intellectual property rights of innovative brand pharmaceuticals with a regulatory framework for generic product development. Novel products became eligible for additional exclusivity periods and extensions to the twenty years of patent protection based on regulatory review time. In exchange, the brand companies could not pursue litigation against a generic pharmaceutical company for patent infringement until after an ANDA was submitted. Generic drug companies gained the right to pursue necessary clinical and development studies without risk of patent infringement litigation.



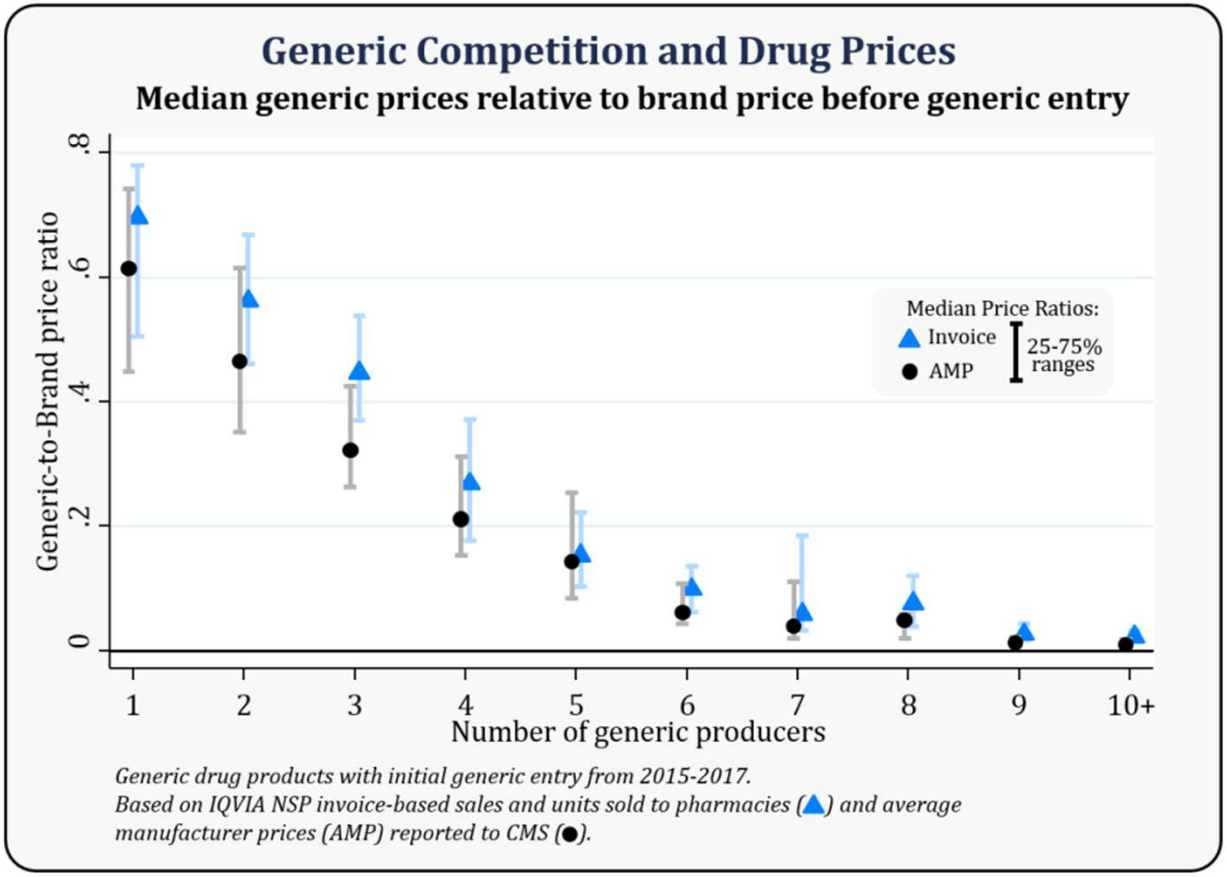


(upper left): The average number of patents for each approved new drug product has steadily increased since the Drug Price and Competition Act was passed in 1984. Because patent expiry dates may be staggered due to regulatory agreements, the increased number of patents provide opportunity for brands to extend the period without generic competition. *(upper right):* In parallel, the median time between generic and brand approval has also increased. This may reflect both time elapsed while waiting for patents to expire and generic companies finding profitability in older drug products. *(lower right):* While many generics are approved in the year following patent expiry, a large majority are approved in the years preceding expiry.



Generic competition erodes drug price. The first generic product to market is sold at a significant discount. After 6 or more competitors, the price is as little as 95% of wholesale cost (Conrad and Lutter, 2019).

Many companies who are approved to make a drug no longer market the product. Declining drug price and low contribution to company revenue were correlated with drug product shortages between 2013 and 2017 (FDA, 2019).



Conrad, R. and Lutter, R. (2019, December). Generic Competition and Drug Prices: New Evidence Linking Greater Generic Competition and Lower Generic Drug Prices. U.S. Department of Health and Human Services, Food and Drug Administration. https://www.fda.gov/media/133509/download. Accessed 11/7/2022.

Analysis Overview

Objective

- Do patents, dosage forms, ingredients, and packaging provide a basis for predicting generic competition?
- Can generic competition be predicted relative to patent expiration?

Data Sources

Approved Drug Products with Therapeutic Equivalence Evaluations. 25th to 42nd Editions.

Each year, the FDA publishes the commonly referenced "Orange Book", an official record of drug products approved for marketing in the United States. Further, it specifies the relationship between the original drug products and equivalent generics. Claimed patent protections and regulatory exclusivity in effect for each drug product application are also listed. While an electronic version is available, details of patents and bioequivalence categories are retired as patents expire or products are discontinued. Therefore, historical PDF versions were required.

Structured Product Labels

Each approved drug product carries an exhaustive labeling of ingredients, packaging, dosing instructions, clinical studies, side effects and other information. The FDA requires all labels to follow a standardized XML format, which facilitates consistent extraction of data fields. The National Library of Medicine maintains the comprehensive repository of labels for all currently marketed products.

Data Processing

For product application and patent records, custom modules were written in Python to parse and clean data following PDF conversion using modules from pypdf2 and tabula. For product labels, custom modules written in Python processed the data after using element.etree to parse the XML structures.

Data Analysis

Data sets were formed comprising of features including counts of patents, the offset between brand approval and patent expiry, dosage forms, ingredients, and packaging components. Products were filtered to include brand products with patent expiry dates between 1990 and 2010, which ensured time for generic products to reach market. Each brand product was categorized as to whether any generic product was approved before the first or last patent expiry.

Training sets were formed from 70% of the samples, and minority classes were up-sampled to 50% Data were scaled between 0 and 1 based on minimum and maximum values.

All analysis was performed in R implementing packages available from CRAN, including xgboost for boosted gradient analysis, e1071 for support vector machine analysis, and rpart for decision tree analysis.

Boosted Gradient | Baseline Features| Competition before Last Patent Expiry | Multiple Administration Routes

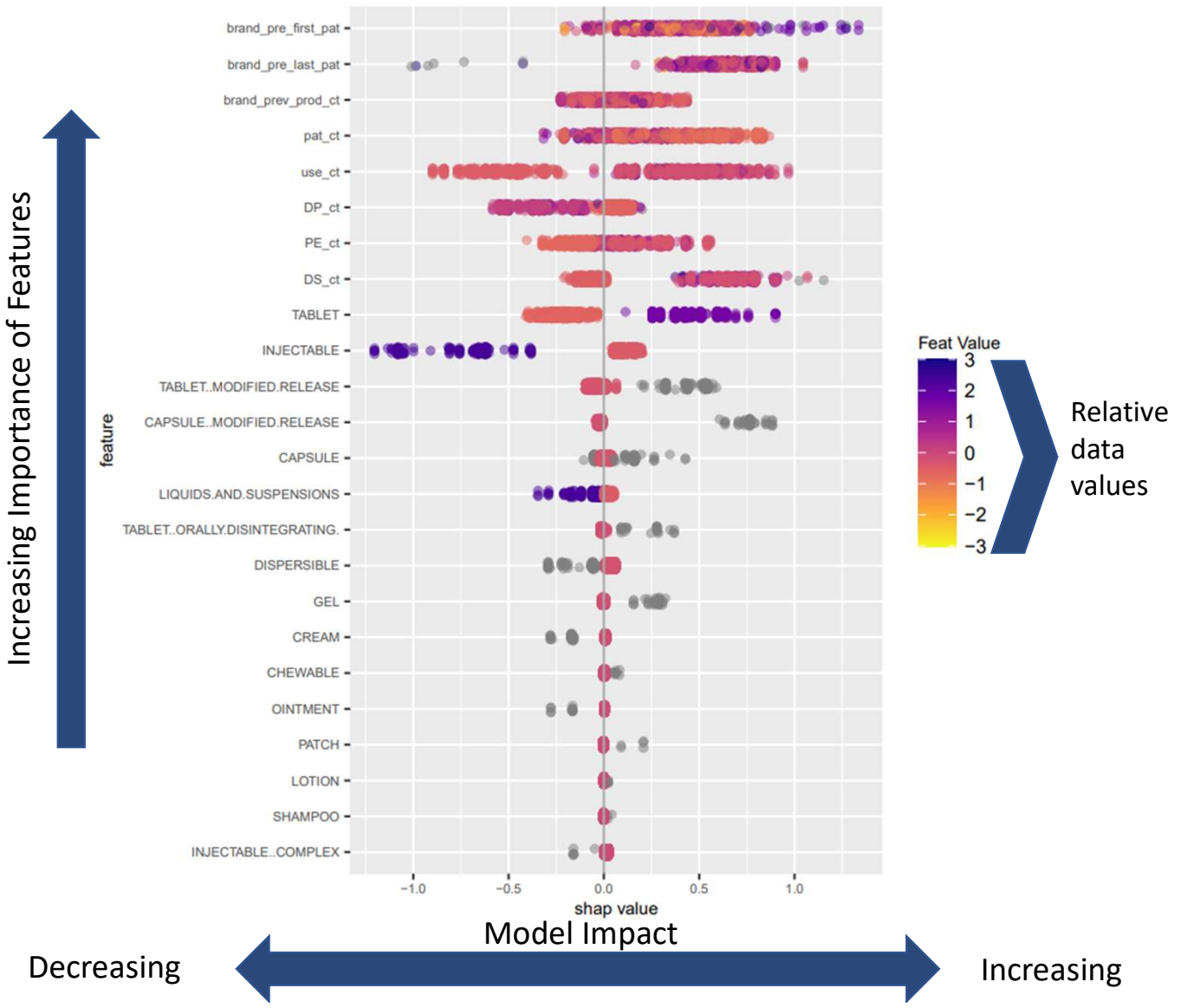
Model evaluation

- Overall model accuracy is statistically significant (greater than no-information rate (NIR)
- Incremental improvement in accuracy with inclusion of ingredients
- Combination of ingredients and packaging features did not produce significant accuracy

Key Features:

- Long approval time before first patent expiry predicts presence of generic competition
- Tablet dosage form predicts presence of generic competition
- Injectable and liquid dosage forms predicts absence of generic competition

	Added Feature Set	Acc. Train	Acc. Test	No Information Rate	Pval
Base Patent and Approval Date Features	--	0.967	0.686	0.536	**< 0.001**
	Ingredients	0.998	0.749	0.635	**< 0.001**
	Packaging	1.000	0.713	0.627	**< 0.001**
	Ingredients and Packaging	0.997	0.670	0.635	0.172

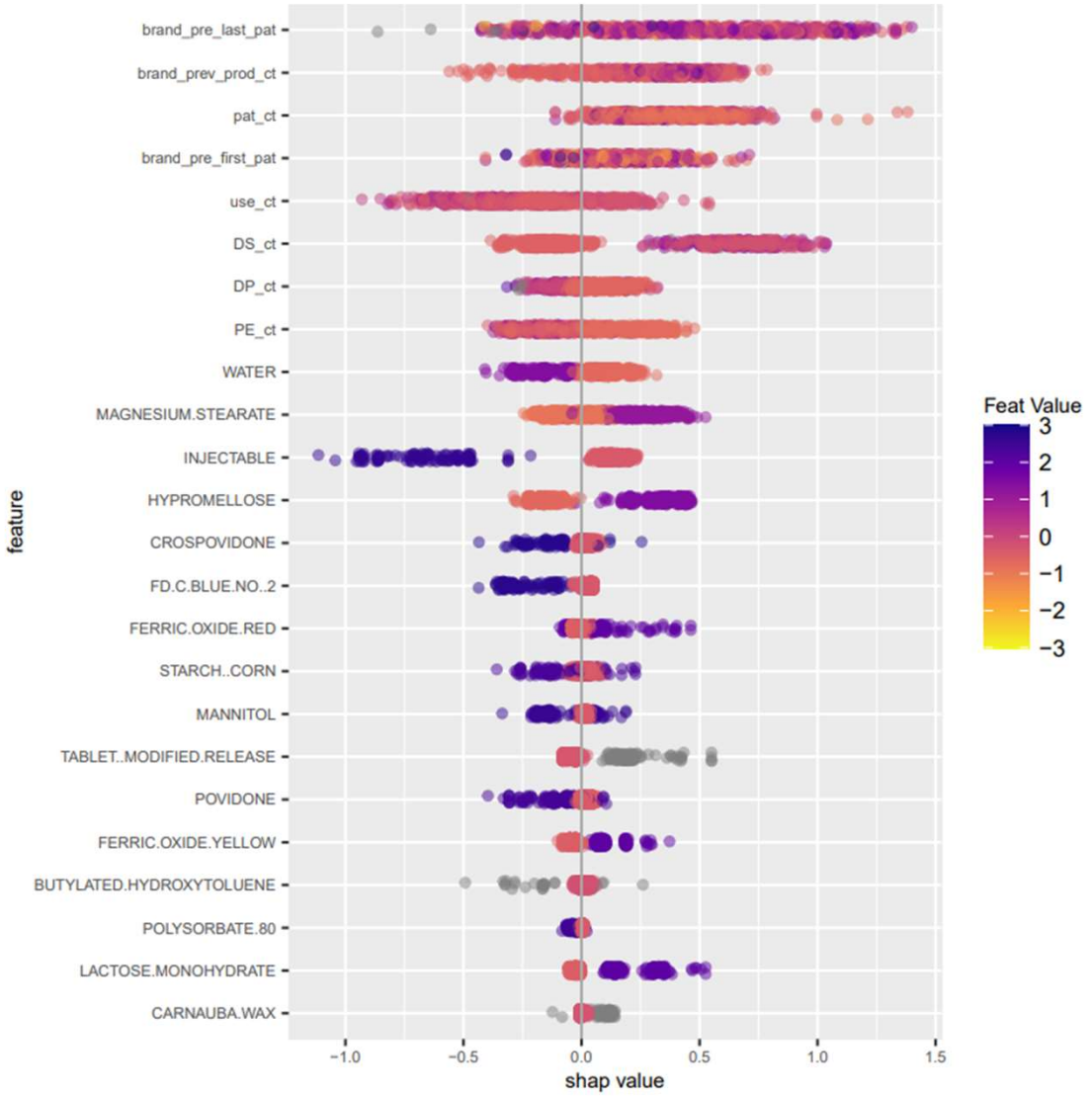


Boosted Gradient | Baseline Features with Ingredients |Competition before Last Patent Expiry | Multiple Administration Routes

Key Features

- Low previous product count a predictor of absence of competition
- Presence of water predicts absence of competition. This aligns with injectables and liquid observation
- Common tablet ingredients (crospovidone, starch, and mannitol) predict absence of competition
- Other common tablet ingredients (magnesium stearate, hypromellose, lactose monohydrate) predict competition

Model Outcomes	Added Feature Set	Accuracy		No Information Rate	Pval
		Train	Test		
Base Patent and Approval Date Features	--	0.967	0.686	0.536	**< 0.001**
	Ingredients	0.998	0.749	0.635	**< 0.001**
	Packaging	1.000	0.713	0.627	**< 0.001**
	Ingredients and Packaging	0.997	0.670	0.635	0.172



Boosted Gradient | Baseline Features with Ingredients |Competition before First Patent Expiry | Oral Administration Routes

Model evaluation

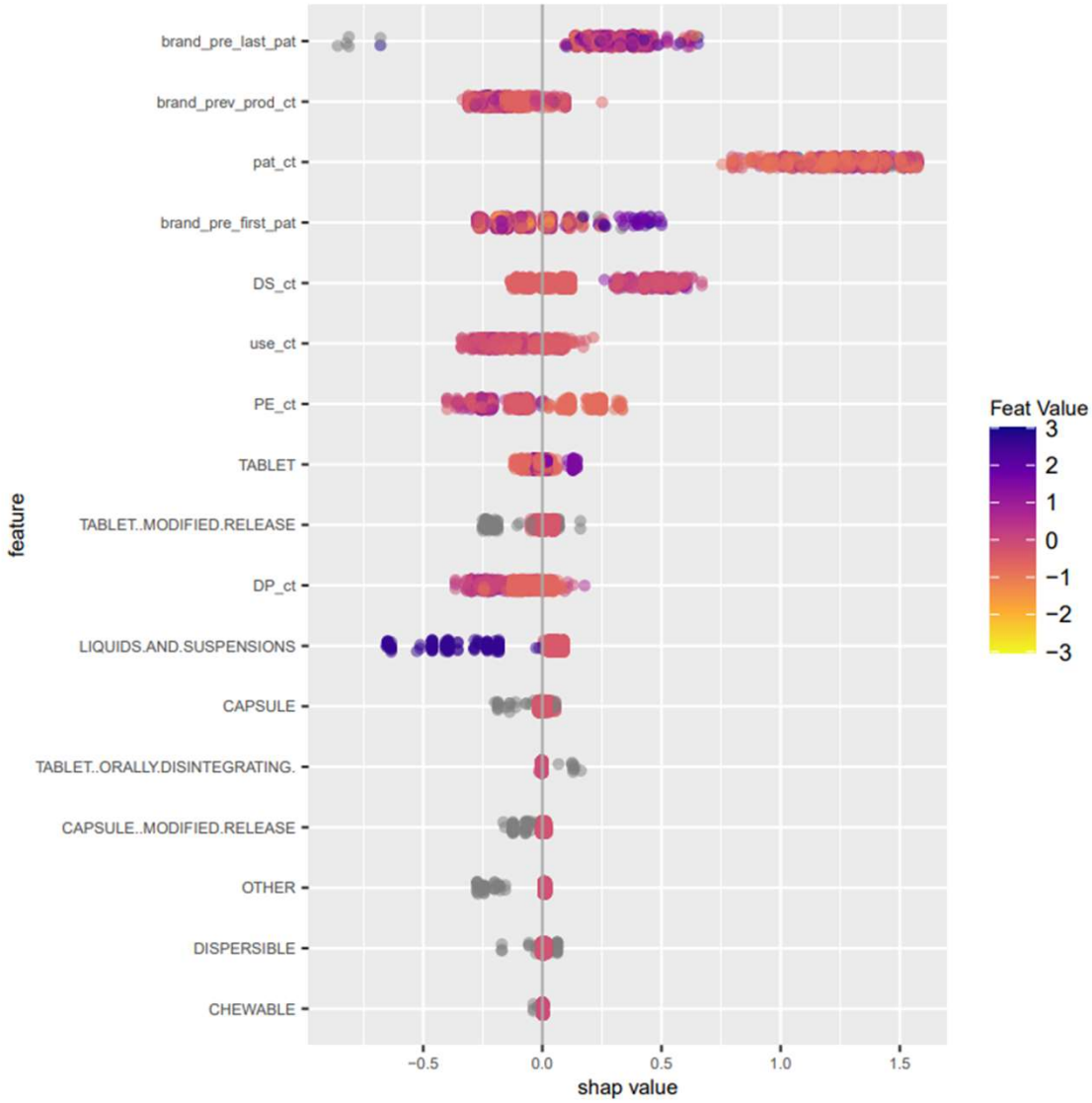
- Overall model accuracy is statistically significant (greater than no-information rate (NIR)
- Incremental improvement in accuracy with inclusion of ingredients
- Only boosted gradient model had accuracy for oral administration products with competition before the first patent expiry

Key Features

Increased time before first patent expiry indicative of competition

Liquids and suspensions predictive of absence of competition.

Features	Prediction	Acc. Train	Acc. Test	NIR	Pval
Base and Ingredients	pre-first	0.986	**0.805**	0.725	0.015



Outcomes

Conclusion

- Justified the importance of predicting competition in the generic pharmaceutical industry
- Successfully extracted and compiled data from a multi-volume annual report in PDF format
- Predicted competition before first and last patent expiry using data derived from the FDA Orange Book

Follow-up Studies

- Evaluate hyperparameter tuning
- Incorporate product sales or market size data

Questions? Contact Mark Benmuvhar:

e-mail mrb6844@truman.edu

Bibliography

Beall, R., Darrow, J., Kesselheim, A. (2018, June). A method for approximating future entry of generic drugs. Value in Health. Volume 21, Issue 12, P1382-1389

Branco, P., Torgo, L., & Ribeiro, R. P. (2016). A survey of predictive modeling on imbalanced domains. ACM Computing Surveys (CSUR), 49(2), 1-50.

Center for Drug Evaluation and Research (CDER). (2017, August). What is the approval process for generic drugs. U.S. Department of Health and Human Services, Food and Drug Administration. https://www.fda.gov/drugs/generic-drugs/what-approval-process-generic-drugs

Center for Drug Evaluation and Research. (1998, December). Guidance for industry: Variations in a drug product that may be included in a single ANDA. U.S. Department of Health and Human Services, Food and Drug Administration. https://www.fda.gov/files/drugs/published/Variations-in-Drug-Products-that-May-Be-Included-in-a-Single-ANDA.pdf

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). New York, NY, USA: ACM. https://doi.org/10.1145/2939672.2939785

Congressional Budget Office (2021, April). Research and development in the pharmaceutical industry. https://www.cbo.gov/publication/57025

Conrad, R. and Lutter, R. Generic Competition and Drug Prices: New Evidence Linking Greater Generic Competition and Lower Generic Drug Prices. U.S. Department of Health and Human Services, Food and Drug Administration. https://www.fda.gov/media/133509/download. Accessed 11/7/2022.

Deloitte Centre for Health Solutions. (2021, April). Nurturing growth: Measuring the return from pharmaceutical innovation 2021. https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/life-sciences-health-care/Measuring-the-return-of-pharmaceutical-innovation-2021-Deloitte.pdf

Dima, A. Lukens, S., Hodkiewicz, M. Sexton, T. Brundage, M. (2021). Adapting natural language processing for technical text. Applied AI Letters. 2021;2:e33.

Drug Database (N.D.). FDA orange book patents. https://drugdatabase.info/fda-orange-book-patents/. Accessed 10/15/2022.

Hastie, T., Tibshirani, R., Friedman, J. (2017). The Elements of Statistical Learning. 2nd Edition. Springer. 245, 312-313.

Hornecker, J. Generic drugs: History, approval process, and current challenges. U.S. Pharm. 2009; 34(6) (Generic Drug Review Supplement) 26-30.

IQVIA Institute for Human Data Science (2022, April). The use of medicines in the U.S. 2022: Usage and spending trends and outlook to 2026. https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/the-use-of-medicines-in-the-us-2022/iqvia-institute-the-use-of-medicines-in-the-us-2022.pdf

Kusynová, Z., Pauletti, G. M., van den Ham, H. A., Leufkens, H., & Mantel-Teeuwisse, A. K. (2022). Unmet Medical Need as a Driver for Pharmaceutical Sciences - A Survey Among Scientists. Journal of pharmaceutical sciences, 111(5), 1318–1324. https://doi.org/10.1016/j.xphs.2021.10.002

Ledley, F. D., McCoy, S. S., Vaughan, G., & Cleary, E. G. (2020). Profitability of large pharmaceutical companies compared with other large public companies. JAMA, 323(9), 834–843. https://doi.org/10.1001/jama.2020.0442

Bibliography

Lichtenberg F. R. (2019). How many life-years have new drugs saved? A three-way fixed-effects analysis of 66 diseases in 27 countries, 2000-2013. International health, 11(5), 403–416. https://doi.org/10.1093/inthealth/ihz003

Lundberg, S. and Lee, S. A unified approach to interpreting model predictions. CoRR, 1705.07874 (2017). http://arxiv.org/abs/1705.07874

Lötsch J, Malkusch S, Ultsch A (2021) Optimal distribution-preserving downsampling of large biomedical data sets (opdisDownsampling). PLoS ONE 16(8): e0255838.

Midha, K. K., & McKay, G. (2009). Bioequivalence; its history, practice, and future. The AAPS journal, 11(4), 664–670. https://doi.org/10.1208/s12248-009-9142-z

National Library of Medicine. (2022, August). SPL resources. DailyMed. https://dailymed.nlm.nih.gov/dailymed/spl-resources-all-drug-labels.cfm:

Olson, L. and Wendling, B. (2013 April). The effect of generic drug competition on generic drug prices during the Hatch-Waxman 180-day exclusivity period. Federal Trade Commission. https://www.ftc.gov/reports/estimating-effect-entry-generic-drug-prices-using-hatch-waxman-exclusivity

Reddy,M.,Ganesh,G.,Babu,B.,Jagadeesan,R. & MR,P.(2022).Risk Assessment of Failures in Generic Drug Development and Approval Procedure under Competitive Generic Drug Therapy and Patent Challenge Exclusivities Provided by the United States Food and Drug Administration. Acta Marisiensis - Seria Medica,68(1) 28-34. https://doi.org/10.2478/amma-2022-0004

Reiffen, D & Ward, M. (2005). Generic Drug Industry Dynamics. The Review of Economics and Statistics. 87. 37-49. 10.2139/ssrn.390102.

S.1538 - 98th Congress (1983-1984): An act to amend the Federal Food, Drug, and Cosmetic Act to revise the procedures for new drug applications, to amend title 35, United States Code, to authorize the extension of the patents for certain regulated products, and for other purposes. (Drug price competition and patent term restoration act). (1984, September 24). http://www.congress.gov/

Sarica, S., & Luo, J. (2021). Stopwords in technical language processing. PloS one, 16(8), e0254937. https://doi.org/10.1371/journal.pone.0254937

Sertkaya, A, Lord, A., Berger, C. (2021, December). Cost of generic drug development and approval. Eastern Research Group. Office of the Assistant Secretary for Planning and Evaluation. U.S. Department of Health and Human Services. https://aspe.hhs.gov/reports/cost-generic-drugs.

Sun, D. Gao, W. Hu, Hongxiang. Zhou, Simon. (July 2022). Why 90% of clinical drug development fails and how to improve it? Acta Pharmaceutica Sinica B, Volume 12, Issue 7, 2022, Pages 3049-3062,https://doi.org/10.1016/j.apsb.2022.02.002.

Tamimi N., Ellis, P. (2009 October). Drug development: from concept to marketing! Nephron Clin Pract 2009;113:c125–c131 DOI: 10.1159/000232592

United States Food and Drug Administration. (2005 – 2022). Approved drug products with therapeutic equivalence evaluations. 25th to 42nd Editions. Retrieved from https://www.thefdalawblog.com/orange-book-archives/. Accessed 9/2/2022.

United States Food and Drug Administration. (2019, Revised 2020, February) Drug shortages: root causes and potential solutions. U.S. Department of Health and Human Services. https://www.fda.gov/media/131130/download. Accessed 11/12/2022