

# Web scraping Project

## **Scraping the Personal Experiences of Antihypertensive Drug Side Effects On WebMD Database**

22th, April, 2019

**SANGWOO, LEE**

# Introduction



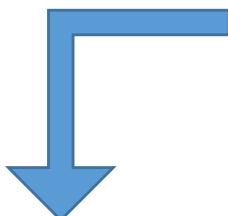
You need to  
take medicine!



Google



151/99 mmHg



# Google

Search Google or type a URL

Antihypertensive drug side effects



# Introduction

All Images News Videos Shopping More Settings Tools

About 6,810,000 results (0.77 seconds)

Some common side effects of high blood pressure medicines include:

- Cough.
- Diarrhea or constipation.
- Dizziness or lightheadedness.
- Erection problems.
- Feeling nervous.
- Feeling tired, weak, drowsy, or a lack of energy.
- Headache.
- Nausea or vomiting.

[More items...](#) • May 21, 2017

High blood pressure medicines: MedlinePlus Medical Encyclopedia

<https://medlineplus.gov/ency/article/007484.htm>

[About this result](#) [Feedback](#)

People also ask

What is a common side effect of antihypertensive medications?
What are the 4 worst blood pressure drugs?
What is the most popular medication for high blood pressure?
Can blood pressure medication cause shortness of breath?

[Feedback](#)

High blood pressure medicines: MedlinePlus Medical Encyclopedia

<https://medlineplus.gov> › Medical Encyclopedia

May 21, 2017 – Some common side effects of high blood pressure medicines include: Cough. Diarrhea or constipation. Dizziness or lightheadedness. Erection problems. Feeling nervous. Feeling tired, weak, drowsy, or a lack of energy. Headache. Nausea or vomiting.

High Blood Pressure Medication Side Effects - WebMD

<https://www.webmd.com> › Hypertension › Guide

Apr 17, 2017 – Any medication can cause side effects, and high blood pressure (HBP) medications are no exception. However, many people do not have side ...

High Blood Pressure Drug List: Side Effects, Uses, Types & Names

[https://www.medicinenet.com/high\\_blood\\_pressure\\_medicine/article.htm](https://www.medicinenet.com/high_blood_pressure_medicine/article.htm)

There are a number of types and classes of drugs available for the management and treatment of high blood pressure (hypertension). ... This information will provide a list of high blood pressure medication generic and brand names, their uses (treatment, management, or prevention of ...

How to choose a high blood ... · Angiotensin receptor ... · Beta-blockers · Diuretics

High Blood Pressure Medications List and Side Effects - RxList

[https://www.rxlist.com/high\\_blood\\_pressure\\_medications/drugs-condition.htm](https://www.rxlist.com/high_blood_pressure_medications/drugs-condition.htm)

Read about high blood pressure medication side effects. ... African-American patients respond to some antihypertensive medications better than others.

Diuretics · Beta-blockers · Calcium channel blockers · Peripheral adrenergic ...

A Profile of Adverse Effects of Antihypertensive Medicines in a Tertiary

...

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC557994/>

by AO Olowofela - 2017 - Cited by 8 - Related articles

A Profile of Adverse Effects of Antihypertensive Medicines in a Tertiary Care Clinic ... Keywords:



# Introduction

- Objective

Scrape webMD website for review opinions,  
with focus on quantitative and qualitative analysis of side effects  
( <https://www.webmd.com> )

# Data Extraction(1/2)

- Used pkg: scrapy
- Analyzed dataset definition

<https://www.webmd.com>

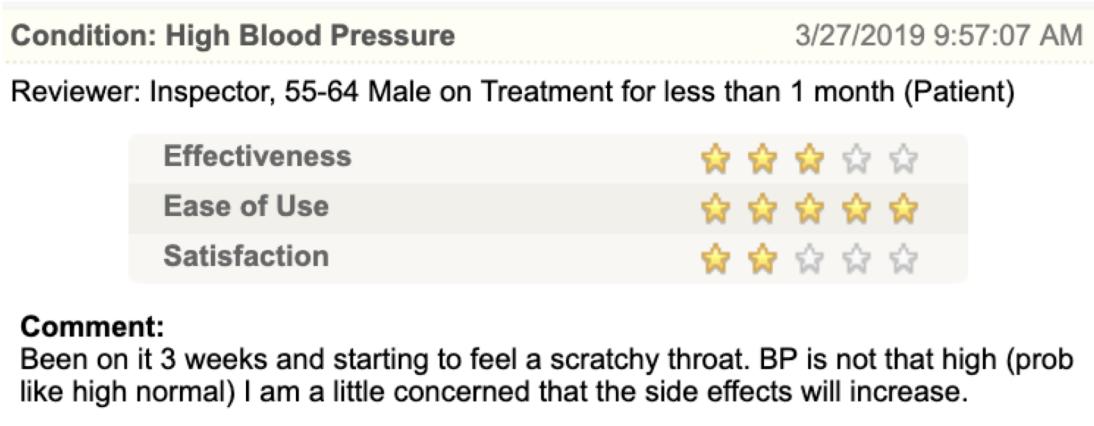
- ➔ ‘drugs and supplements’ ➔ ‘conditions’
- ➔ enter ‘hypertension’ or ‘high blood pressure’
- ➔ scraped from 9 most reviewed drugs

The screenshot shows the WebMD Drugs & Medications A-Z page. On the left, there are three vertical navigation bars: 'DRUGS AND MEDICATIONS CENTER' with links to Find a Drug, My Medicine, Pill Identifier, Interaction Checker, Find a Vitamin, and Find a Pharmacy; 'VITAMINS & SUPPLEMENTS CENTER' with links to Find a Vitamin or Supplement, Vitamin Lifestyle Guide, and Assess Your Vitamin Needs; and 'FIRST AID RESOURCES'. The main content area has a teal header 'Drugs & Medications A-Z'. Below it, a navigation bar has 'Drugs' (unselected) and 'Conditions' (selected). A search bar contains the text 'hypertension'. Below the search bar, a message states: 'The display and use of drug information on this site is subject to express terms of use. By continuing to view the drug information, you agree to abide by such terms of use.' A note below says: 'Considering taking medication to treat undefined? Below is a list of common medications used to treat or reduce the symptoms of undefined. Follow the links to read common uses, side effects, dosage details and read user reviews for the drugs listed below.' A bold heading '479 medications found for undefined' is displayed, followed by a sub-heading 'Sorted by User Reviews'. A table lists three medications: 'A drug', 'B drug', and 'C drug', each with 'On Label' indication, 'RX' type, and a link to '4259 Reviews'.

Drug Name	Indication ⓘ	Type ⓘ	User Reviews
A drug	On Label	RX	<a href="#">4259 Reviews</a>
B drug	On Label	RX	<a href="#">4259 Reviews</a>
C drug	On Label	RX	<a href="#">1269 Reviews</a>

# Data Extraction(2/2)

- Example of review opinion



- Scrapped dataset example

index	screenname	age	gender	whois	drugname	condition	TimeOnMed	easeofuse	effectiveness	satisfaction	comment
0	Inspector	55-64	Male	Patient	Drug A	High Blood Pressure	less than 1 month	5	3	2	been on it 3 weeks and starting to feel a scratchy throat bp is not that high prob like high normal i am a little concerned that the side effects will increase
1	ileewms	65-74	Female	Patient	Drug B	High Blood Pressure	2 to less than 5 years	5	4	4	have been taking for almost 2 years seems to help lower bp have noticed hair loss occasional dizziness when tilting my head to look upward hoping for something better with no side effects
2		45-54	Female	Patient	Drug C	High Blood Pressure	1 to less than 2 years	5	5	5	i ve been taking lisinopril for over a year with great success my blood pressure was consistently 160 90 or higher was reluctant to take meds within a week of being on lisinopril my blood pressure dropped to 120 75 and is still that today or lower i have no side effects what so ever truly a miracle drug for me

# Data Preprocessing on 'comments' (1/2)

- Filtering on 'comment' column
  - Used stop words of nltk and own definition
- Lemmatization on 'comment' column
  - Used nltk
- A table of side effects vs. synonyms created based on 'comment' column
  - Synonyms need to be found since many reviewers use diverse expressions.
  - There are many typos.
- Side effects considered in this study: 28 categories

Cough, Pain, Allergy, Headache, Nausea & Vomit

Lethargy, Anxiety & depression, Dizzy, Blurring, Skin problem

Appetite, Diarrhea, Sleep disorder, GI symptom

Sweating, Hair loss, Breathe discomfort

Mental alteration, Abnormal sensation

Movement disorder, Constipation, Sexual dysfunction

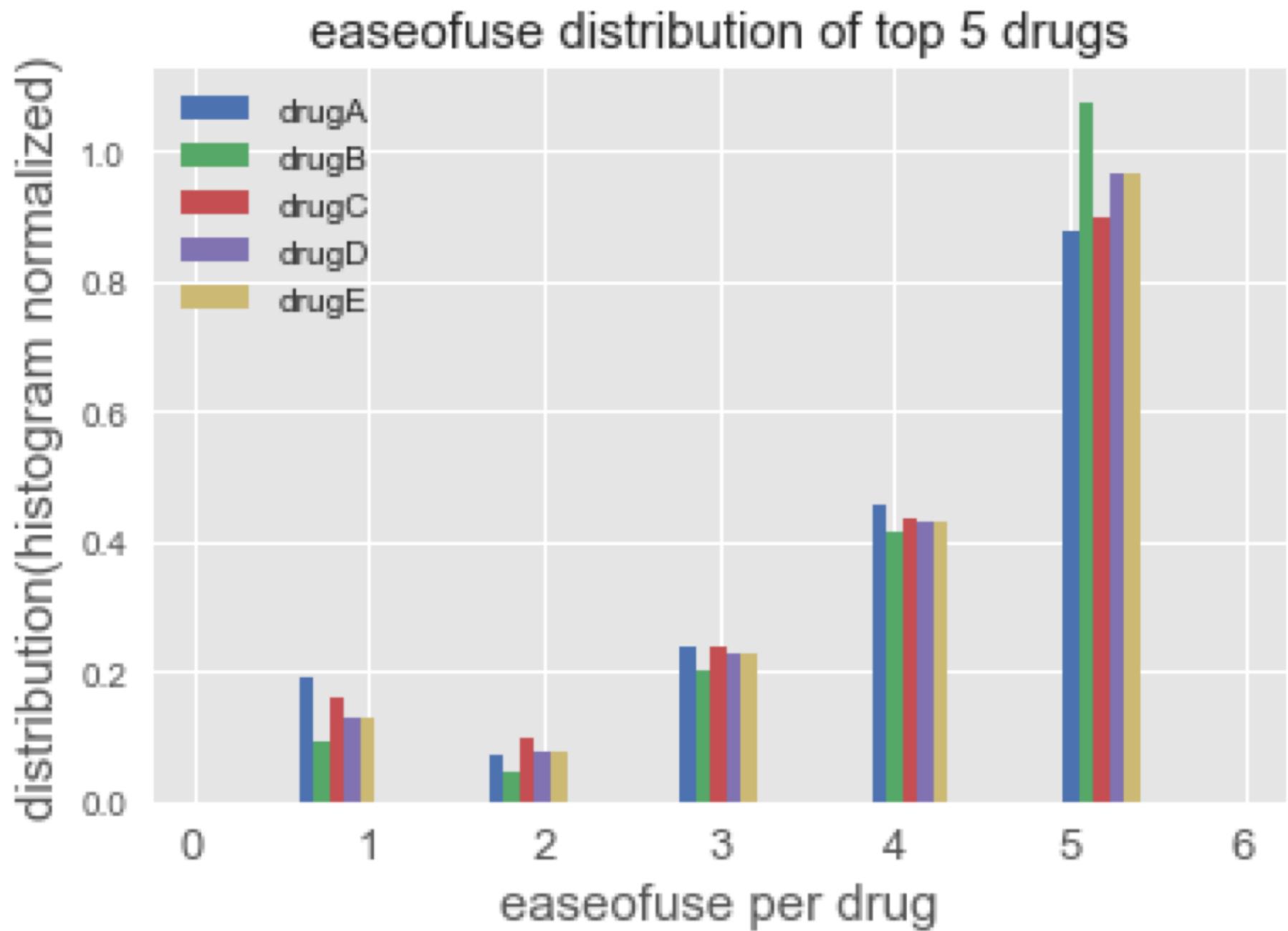
Thirsty, Dysuria, Swallow difficult, Ear problem

Voice change, Pancreatitis

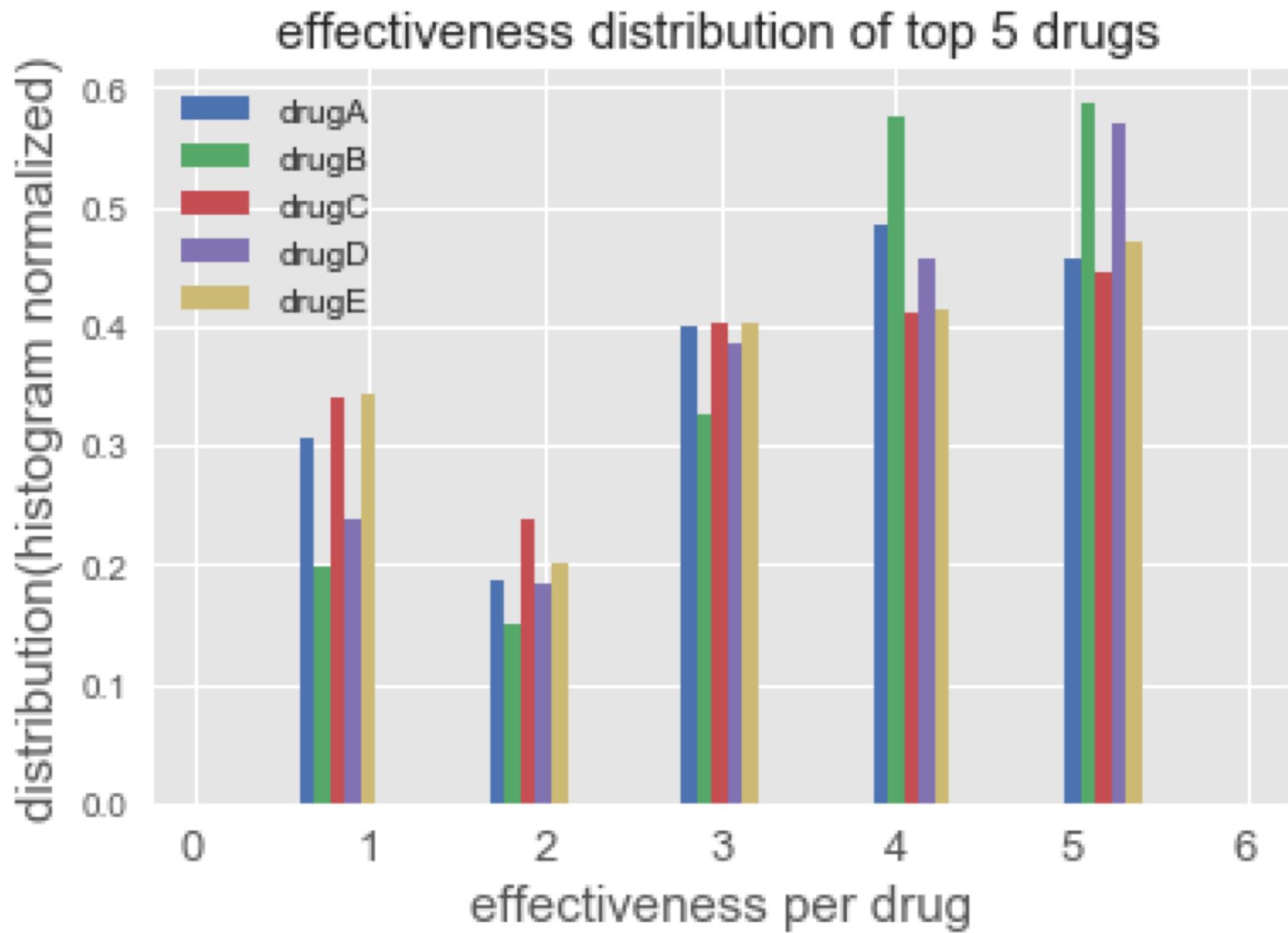
# Data Preprocessing on ‘comments’ (2/2)

- A scraped example with side effect columns added
    - 6342 reviews about 9 drugs for hypertension

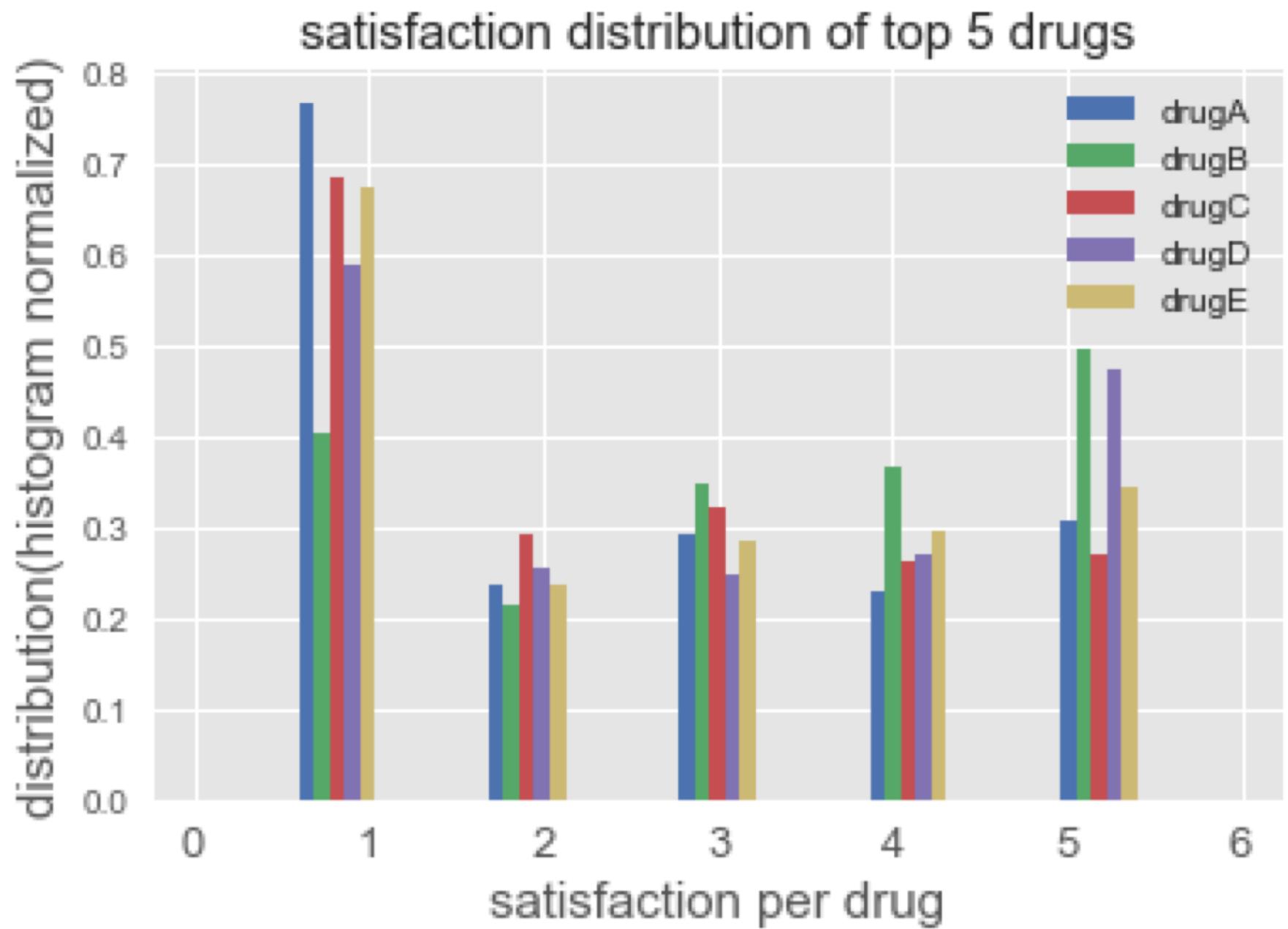
# Results<sub>(1/6)</sub>: ease of use vs drugs



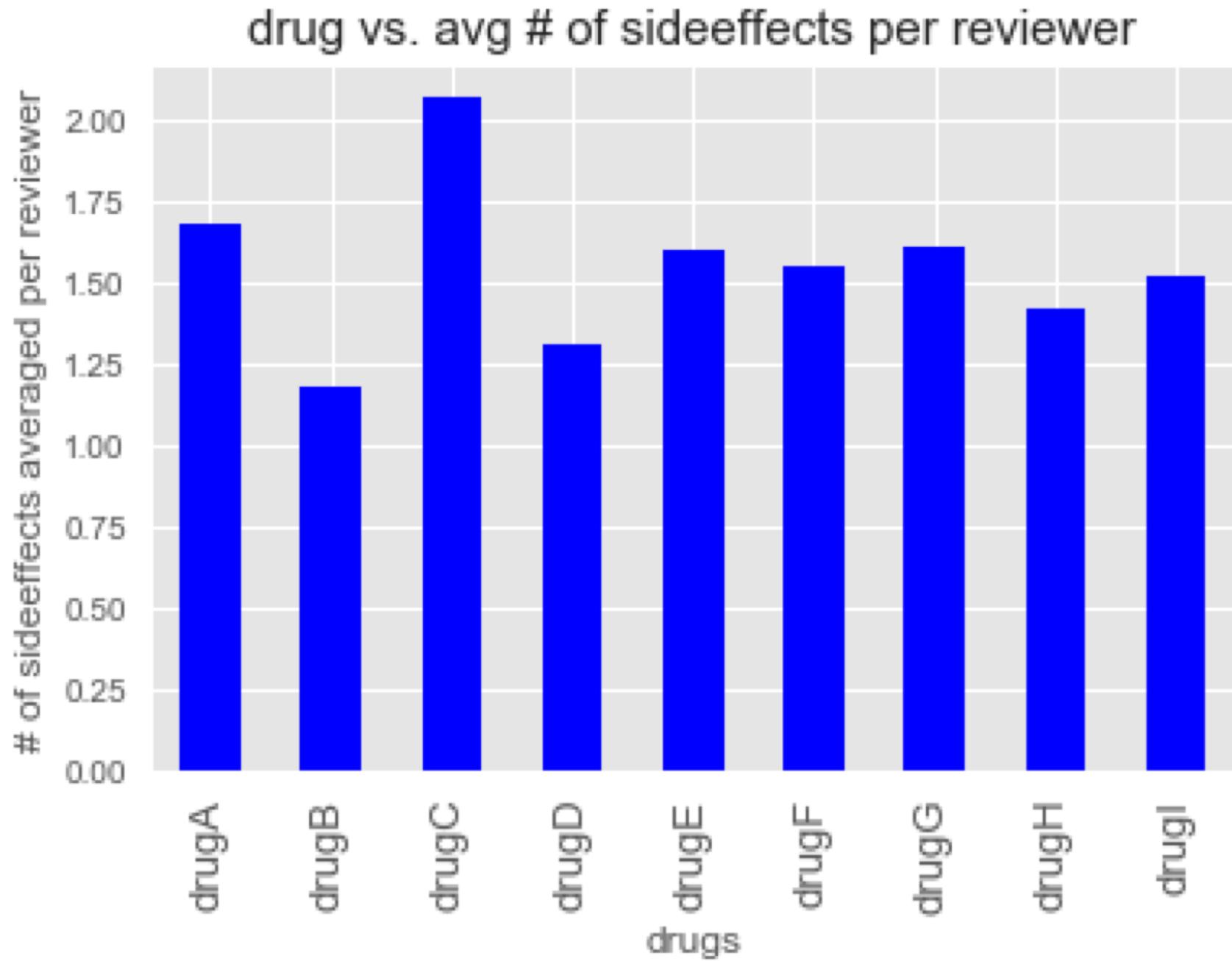
## Results<sub>(2/6)</sub>: effectiveness vs drugs



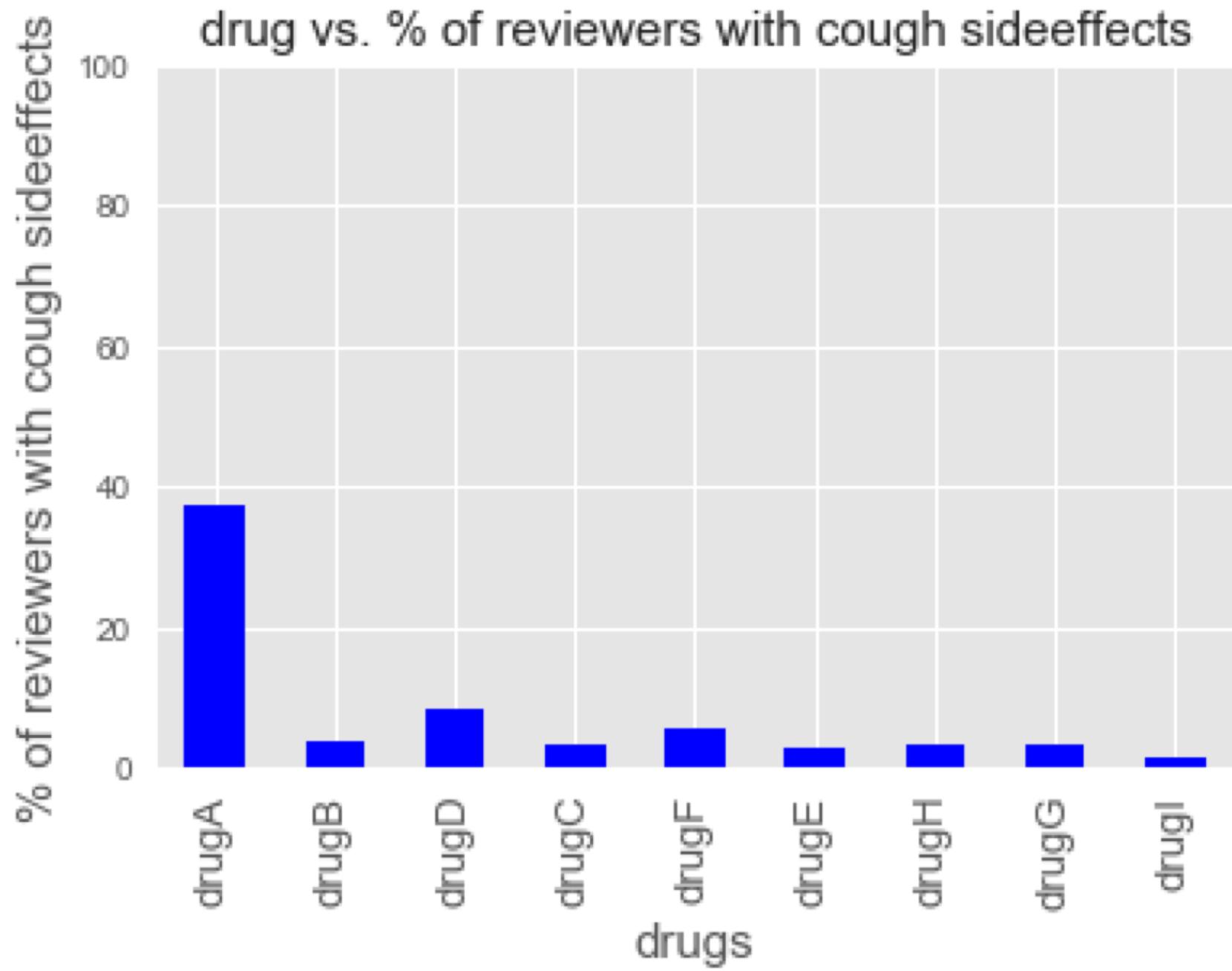
## Results<sub>(3/6)</sub>: satisfaction vs drugs



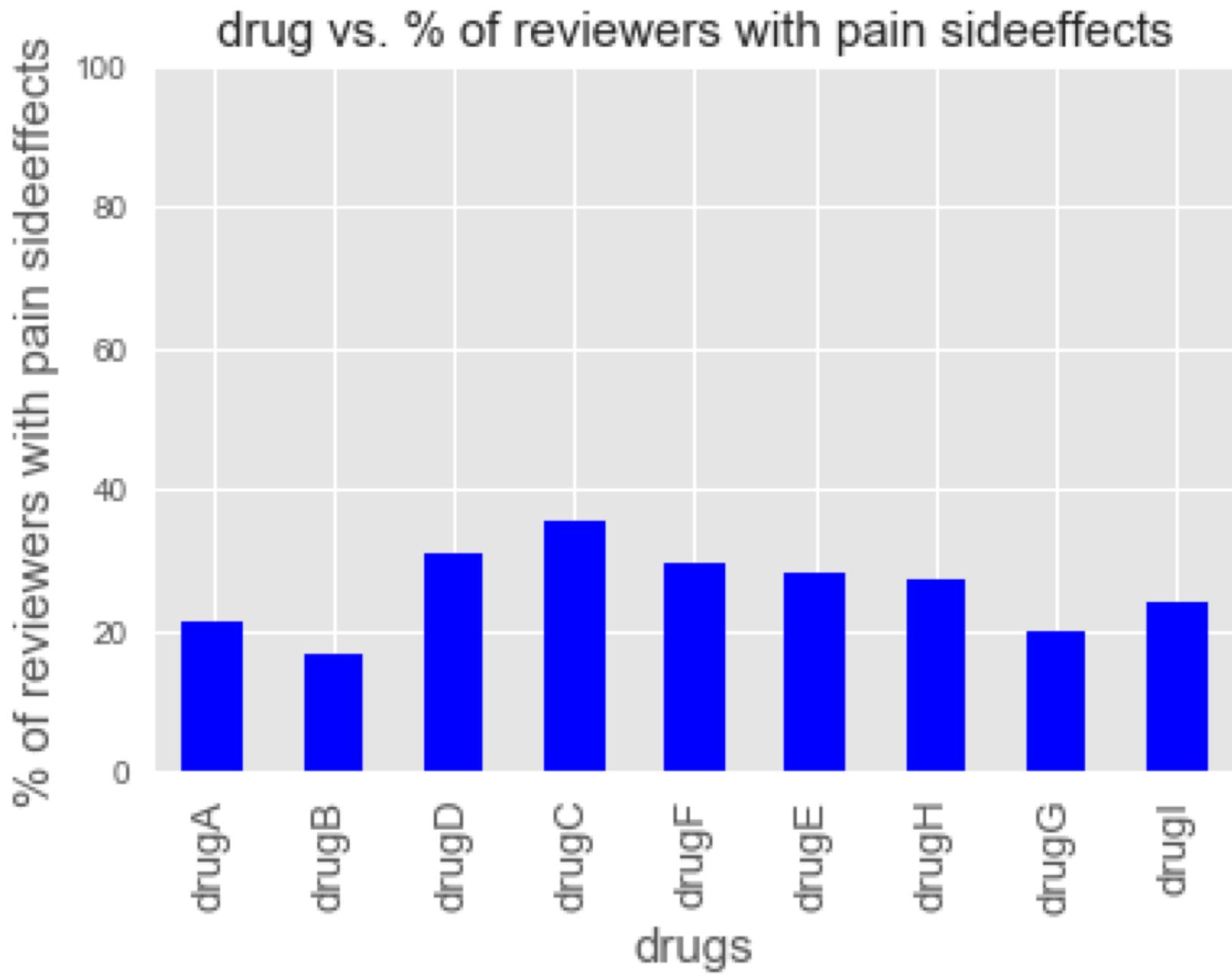
## Results<sub>(4/6)</sub>: drug vs. avg # of side effects per reviewer



# Results<sub>(5/6)</sub>: drug vs. % of reviewers with cough



# Results<sub>(6/6)</sub>: drug vs. % of reviewers with pain

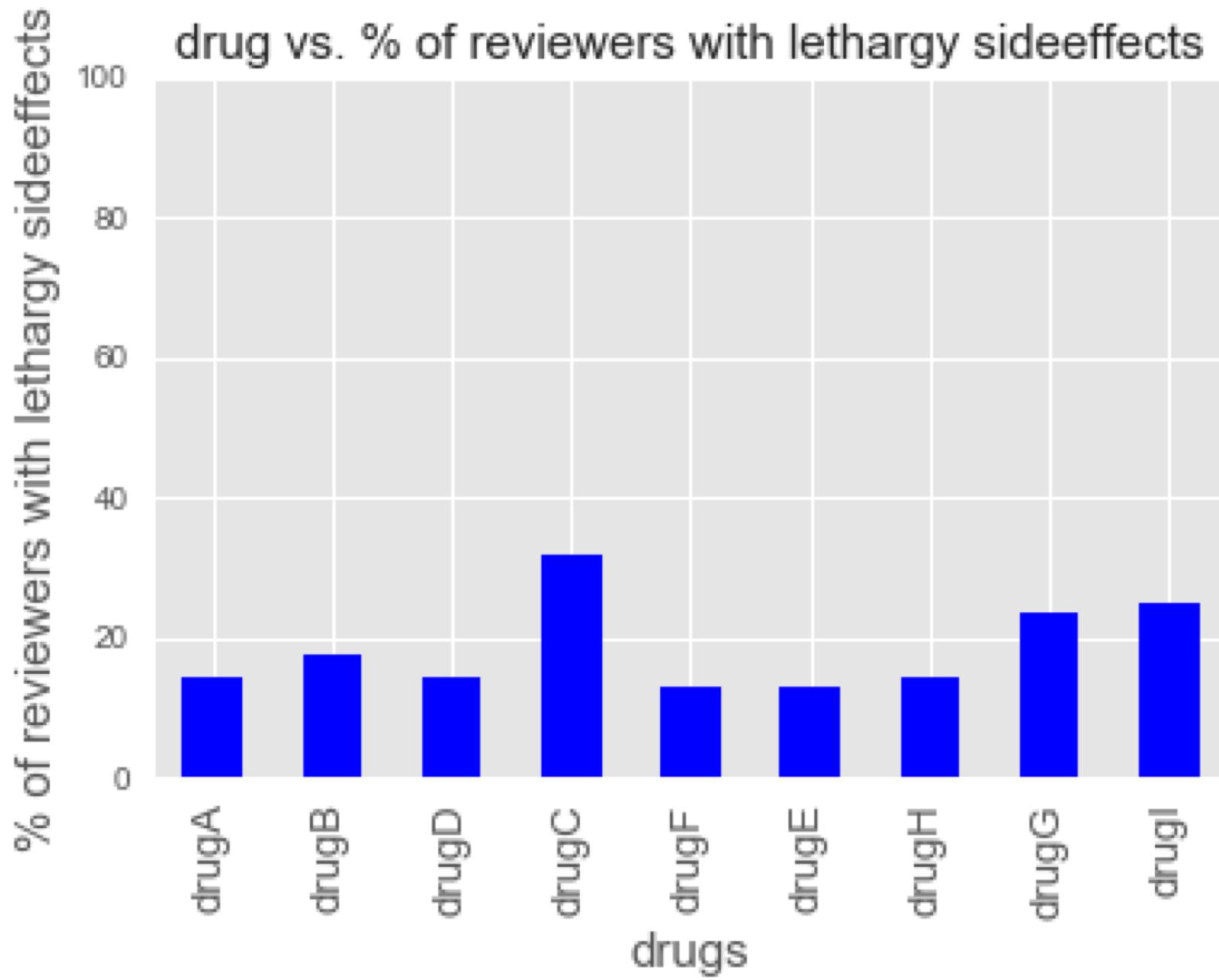


# Discussion

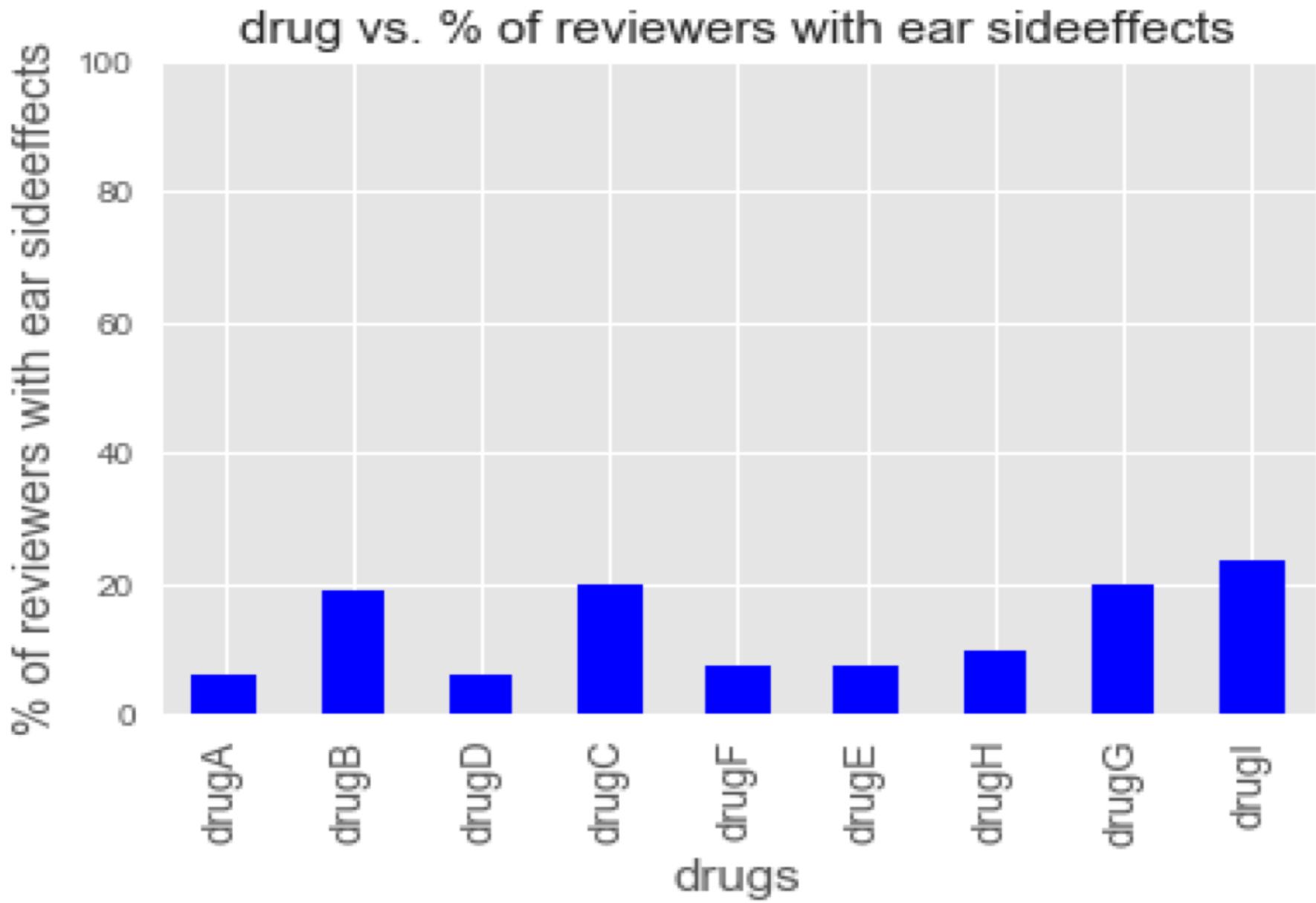
- Conclusion
  - We scraped from webMD website for review opinions and performed quantitative and qualitative analysis of side effects
- Future work
  - The table of side effects vs. synonyms need more validations.
  - Increase size of datasets
    - By choosing other drugs with other conditions
    - By scaping more number of review opinions

# Appendix

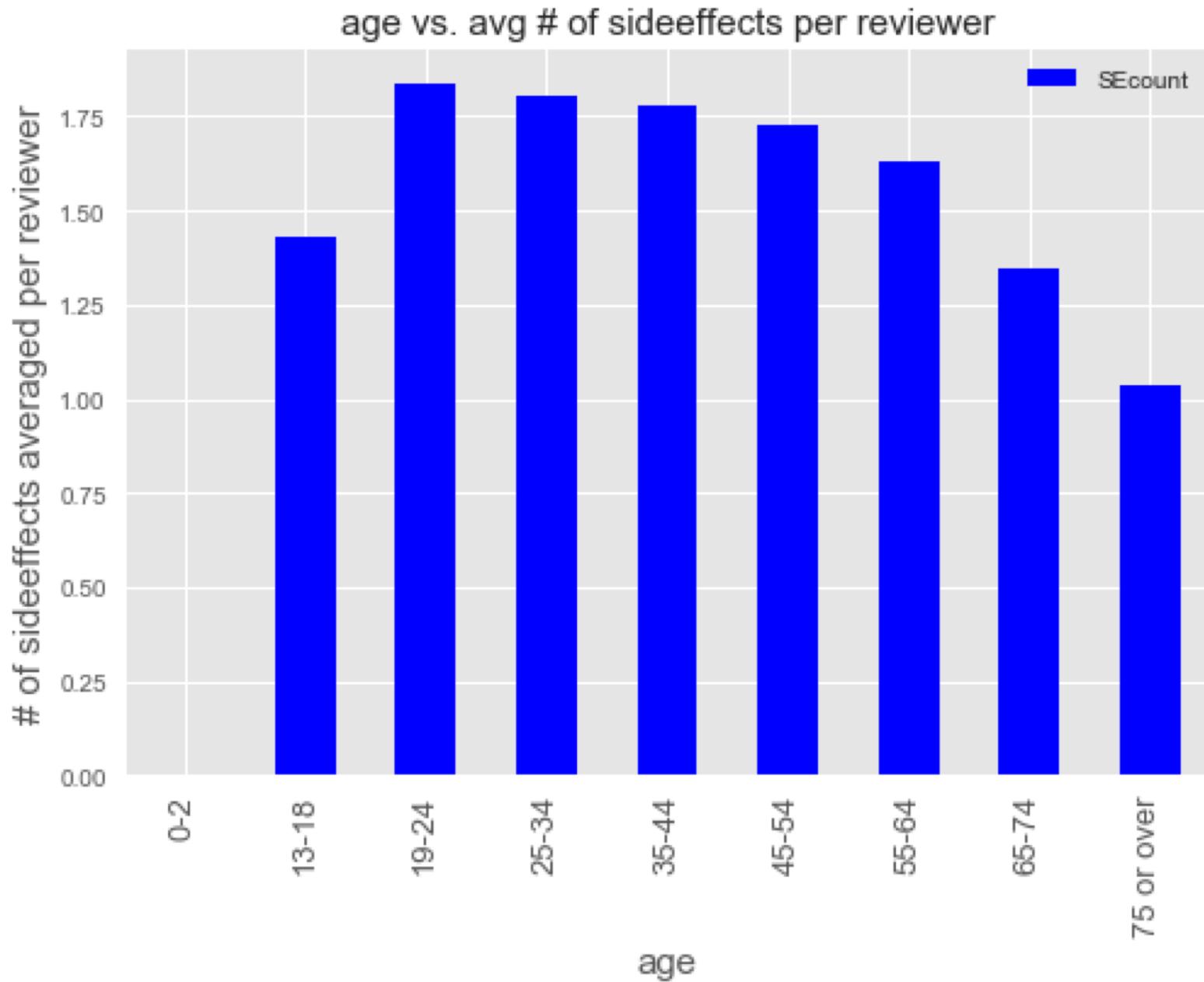
# Appendix: drug vs. % of reviewers with lethargy



# Appendix : drug vs. % of reviewers with ear problem



# Appendix: age vs. avg # of side effects per reviewer



# Appendix: gender vs. avg # of side effects per reviewer

