Andrew Ma

ISB Job interview project

January 22, 2023

Which IMID Medications are Prescribed Most Often?

Abstract

We present a survey of data on Immune mediated Inflammatory disease (IMID) medications from the Columbia Open Health Data project database. We present bar plots of some of the commonly prescribed medications to treat IMIDs and we find that TNF-alpha inhibitors are the most frequently prescribed medication for any IMID among those considered in this paper. Furthermore we find that the anchor drug methotrexate is most frequently paired with a TNF-alpha inhibitor.

Introduction

In a recent paper [1] a collection of Immune mediated Inflammatory diseases (IMIDs) were organized according to the cytokine mechanisms or cytokine hubs that are thought to cause these diseases. In [1] the authors argue that medications which inhibit targeted cytokines have an effectiveness that is comparable and oftentimes slightly better than non-targeted medications such as Tumor necrosis factor-alpha (TNF-alpha) inhibitors or Janus kinase (JAK) inhibitors with regards to patients living

with IMIDs. In certain cases it is additionally argued that targeted cytokine inhibitors benefit from needing fewer medications taken with them concurrently such as in the case of Rheumatoid arthritis.

The discussion presented in [1] is based on the results from clinical trials. It would be interesting to compare the suggested best practices from [1] with the practices used in a large population of patients in the US healthcare system. Doing survey of the medications prescribed for IMIDs is an example of Exploratory Data Analysis (EDA) and EDA a common first step when preparing any large dataset for data science or machine learning applications.

In this paper we are interested in comparing how often medications are prescribed for patients with IMIDs. Specifically we look at the counts of patients with an IMID and a particular medication prescription taken from a public database of electronic health records. We organize a small set of IMID medications into three categories: TNF-alpha inhibitors, targeted cytokine hub inhibitors, and JAK inhibitors. We then compare how often drugs from each of these categories are prescribed to patients with a specific IMID.

Additionally we compare the frequency with which an IMID medications is paired with the anchor drug methotrexate. Methotrexate is a common anchor drug often paired with other medications, e.g. a TNF-alpha inhibitor, for treating patients with IMIDs.

We make the following two hypothesis about the Columbia Open Health Data (COHD) database of patients:

1. Among patients being treated for any of following IMIDs: Rheumatoid arthritis, Crohn's Disease, Ulcerative Colitis, Axial Spondyloarthritis, or Psoriatic Arthritis, the most frequently prescribed medication will be a targeted cytokine hub inhibitor instead of TNF-alpha inhibitor or a JAK inhibitor.
2. Among patients who are prescribed the anchor drug methotrexate the least commonly paired medication i.e. medication prescribed concurrently, will be a targeted cytokine hub inhibitor or a JAK inhibitor instead of a TNF-alpha inhibitor.

The remainder of this paper is broken down into sections. Section 2 discusses the specific IMIDs and medications considered in this EDA survey. Section 3 discusses details of the COHD database used. Section 4 presents the results of the surveys and reviews how the results either support or oppose the two hypotheses stated above.

Section 2. Disease and medications considered

For hypothesis 1 we survey the following IMIDs: Rheumatoid arthritis, Crohn's Disease, Ulcerative Colitis, Axial Spondyloarthritis, or Psoriatic Arthritis. For each IMID we consider a different set of associated TNF-alpha inhibitors, targeted cytokine inhibitors, and JAK inhibitors that are listed in table 1 from [1]. We present the table again for convenience.

For hypothesis 2 we survey the COHD for methotrexate paired with the drugs listed in table 2 below:

Table 1. Clinical and Pathological Features of Immune-Mediated Inflammatory Diseases and Approved Treatments.*

Variable	Rheumatoid Arthritis	Crohn's Disease	Ulcerative Colitis	Axial Spondyloarthritis	Psoriatic Arthritis
Genetic characteristics	MHC class II (DR4) PTPN22, CTLA4	MHC class II (DRB1) Interleukin-23R, NOD2	MHC class II (DRB1) Interleukin-23R, interleukin-10R	MHC class I (B27) Interleukin-23R, ERAP1	MHC class I (C06) Interleukin-23R, A20
Drivers	Autoimmunity	Microbial dysbiosis and barrier dysfunction	Microbial dysbiosis and barrier dysfunction	Mechanical stress	Mechanical stress and metabolism
Key pathological process	Synovitis	Granuloma formation	Cryptitis and goblet-cell loss	Axial enthesitis	Enthesitis and synovitis
Cellular immune response	B cells, Tph or Tfh cells, macrophages, fibroblasts	Th1/Th17 cells, dendritic cells, macrophages	Th2/Th9/Th17 cells, neutrophils	Th17 cells, T γ/δ cells, ILC3, neutrophils	Th17 cells, T γ/δ cells, ILC3, neutrophils, fibroblasts
Key associated disease	Interstitial lung disease	Erythema nodosum	Primary sclerosing cholangitis	Anterior uveitis	Psoriasis
NSAID responsiveness	Absent	Absent	Absent	High	Moderate
Glucocorticoid responsiveness	High	High	High	Absent	Moderate
Conventional anchor drug	Methotrexate	Azathioprine	Cyclosporine	Sulfasalazine†	Methotrexate
Approved TNF-α inhibitors	Adalimumab, certolizumab, etanercept, golimumab, infliximab	Adalimumab, certolizumab (U.S.), infliximab	Adalimumab, certolizumab (U.S.), golimumab, infliximab	Adalimumab, certolizumab, etanercept, golimumab, infliximab	Adalimumab, certolizumab, etanercept, golimumab, infliximab
Approved cytokine signature drugs (targets)	Tocilizumab (interleukin-6R), sarilumab (interleukin-6R)	Ustekinumab (interleukin-12/23)	Ustekinumab (interleukin-12/23)	Secukinumab (interleukin-17A), ixekizumab (interleukin-17A)	Secukinumab (interleukin-17A), ixekizumab (interleukin-17A), ustekinumab (interleukin-12/23), guselkumab (p19, interleukin-23)
Other approved targeted therapies	Abatacept, rituximab	Vedolizumab	Vedolizumab	None	Apremilast, abatacept
Approved JAK inhibitors	Tofacitinib, baricitinib, upadacitinib, filgotinib (E.U.)	None	Tofacitinib	Upadacitinib (E.U.)	Tofacitinib, upadacitinib (E.U.)

* A20 protein is also known as tumor necrosis factor α (TNF-α)–induced protein 3. CTLA-4 denotes cytotoxic T-lymphocyte–associated protein 4, ERAP1 endoplasmic reticulum aminopeptidase 1, E.U. European Union, ILC3 innate lymphoid cells type 3, interleukin-10R interleukin 10 receptor, interleukin-23R interleukin-23 receptor, JAK Janus kinase, MHC major histocompatibility complex, NOD2 nucleotide-binding oligomerization domain–containing protein 2, NSAID nonsteroidal antiinflammatory drug, PTPN22 protein tyrosine phosphatase nonreceptor type 22, Tfh follicular helper T cell, Th helper T cell, Tph peripheral helper T cell, and U.S. United States.

† Sulfasalazine is recommended only for the treatment of peripheral spondyloarthritis.

Table 1

Methotrexate drug pairings	Drug	Drug type tag
Approved TNF-α inhibitors	Adalimumab, certolizumab, etanercept, golimumab, infliximab	tnf
Approved cytokine signature drugs (targets	Tocilizumab (interleukin-6R), sarilumab (interleukin-6R)	targeted
Approved JAK inhibitors	Tofacitinib, baricitinib, upadaci-tinib,	jak

Table 2

These drug pairings were chosen because these medications and methotrexate are all used in treatment for patients with Rheumatoid arthritis. The relationship between these drugs and their efficacy when paired with methotrexate is explained in [1].

Section 3. Database

For this paper we use data taken from the "Columbia Open Health Data" (COHD) project. The COHD is a publicly accessible dataset containing electronic health records kept at the Columbia University Iriving Medical Center [2]. These records are made accessible via a web API that allows users to retrieve patient counts of people with up to two "medical concepts". For the COHD a medical concept can be any form of medical condition, drug, or procedure.

It is highly relevant to the results of this paper to emphasize that the COHD limits information requests to at most two medical concepts. For example, it is possible retrieve data for answering questions such as "How many patients had Rheumatoid arthritis (RA) and were taking Adalimumab?" because this query contains two medical concepts: (RA) and the drug Adalimumab. However, it is not possible to make queries such as "How many patients had (RA), were taking Adalimumab, and were not taking Tocilizumab?" Because this query refers to three medical concepts: (RA), the drug Adalimumab, and the drug Tocilizumab. Furthermore, due to the way the data returns aggregate counts of patients it is not possible to make precise comparisons between queries. For example, if we have one query that tells us there were 100 patients with (RA) and taking Adalimumab meanwhile a second query tells us that there were 20 patients with Rheumatoid arthritis and taking Tocilizumab, then it is not possible to com-

pare these numbers. It could be the case that all 20 patients with (RA) and taking Tocilizumab were also taking Adalimumab concurrently or it could be the case that none of the patients with (RA) and taking Tocilizumab were taking any other medication. At most we can state that among patients with (RA) the drug Adalimumab was prescribed more often than Tocilizumab but we can't state precisely how much more often the first drug is prescribed over the second.

In this paper we query results from the "5 year hierarchical" dataset in the COHD. This dataset contains the information from 1,790,431 patients recorded between the years 2013-2017. We query the data through the COHD API interface provided through the COHD Github account [2].

Section 4. Results

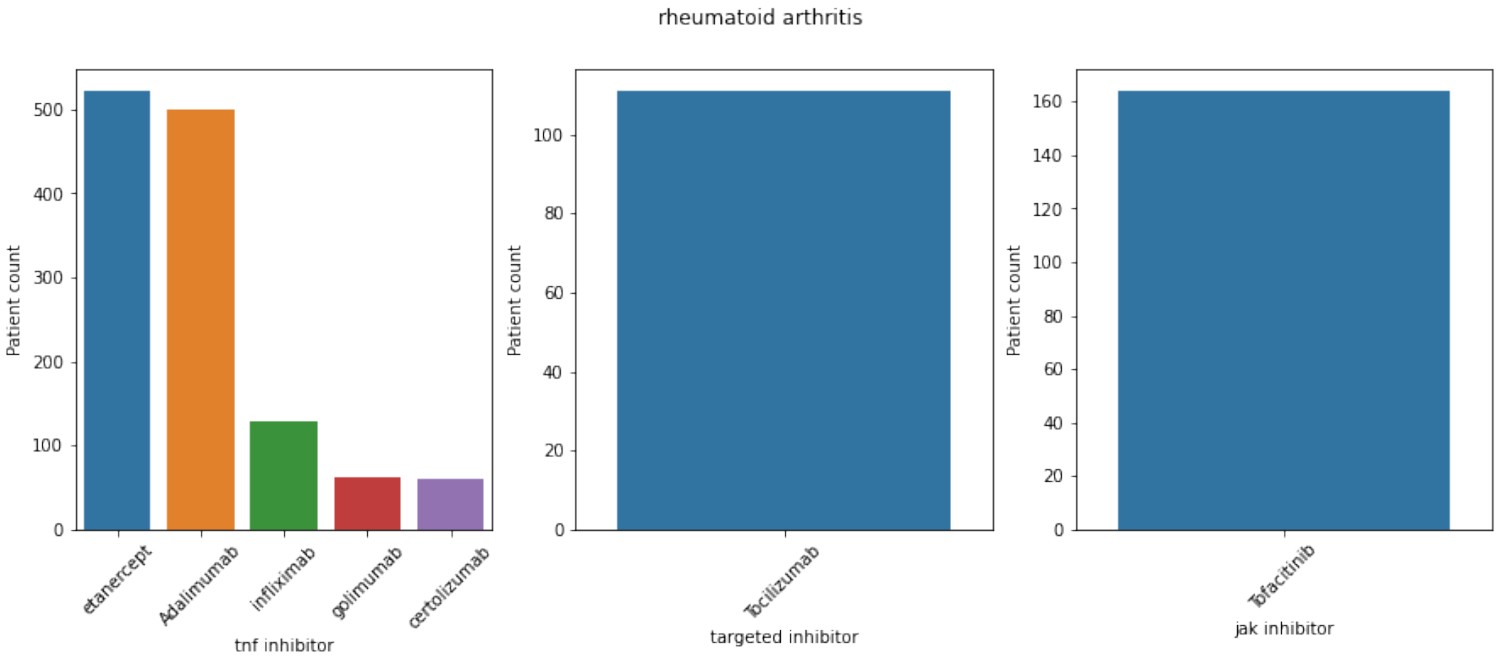
In this section we present bar plots made with the results of our surveys. We present plots of patient counts created for hypothesis 1 first and then the plots relevant to hypothesis 2 second.

Beginning with the plots for hypothesis 1, for each IMID considered there is a set of up to three plots of patient counts, one for each type of medication, TNF-alpha inhibitors, targeted cytokine hub inhibitors, and JAK inhibitors whenever applicable. If a medication is not plotted or a plot is not show then it is because the COHD did not have any data on that particular disease-drug pairing. Underneath the set of plots grouped by drug type there is a plot containing all of the drugs considered for a single IMID. The bars are organized in descending order. Details about the specific numbers

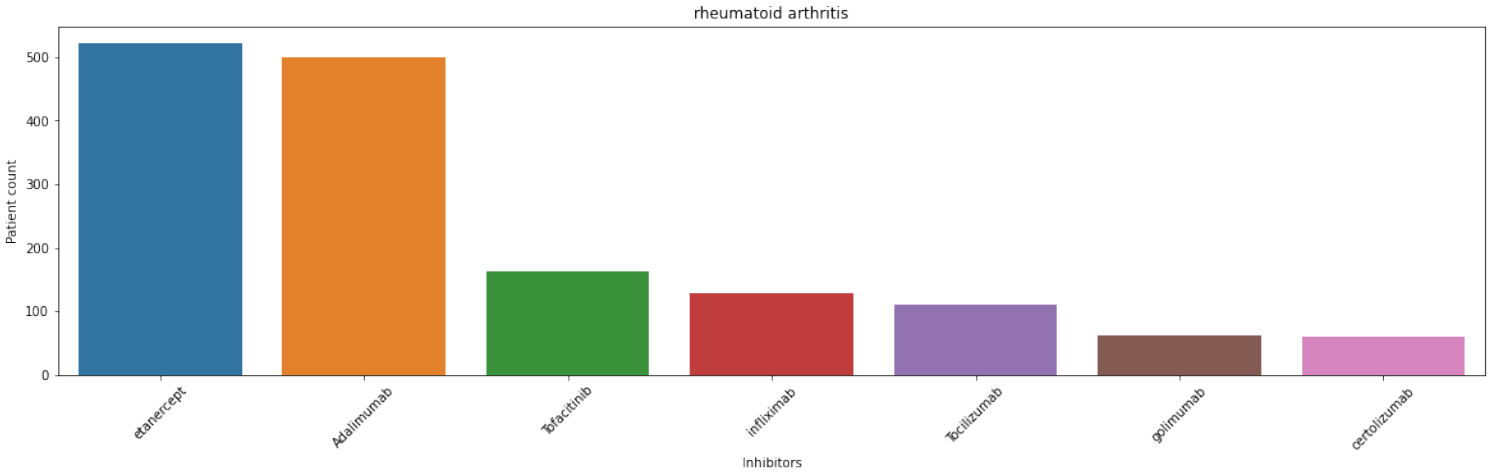
for each patient count can be found in the data frames listed in the accompanying code (add link to GitHub later).

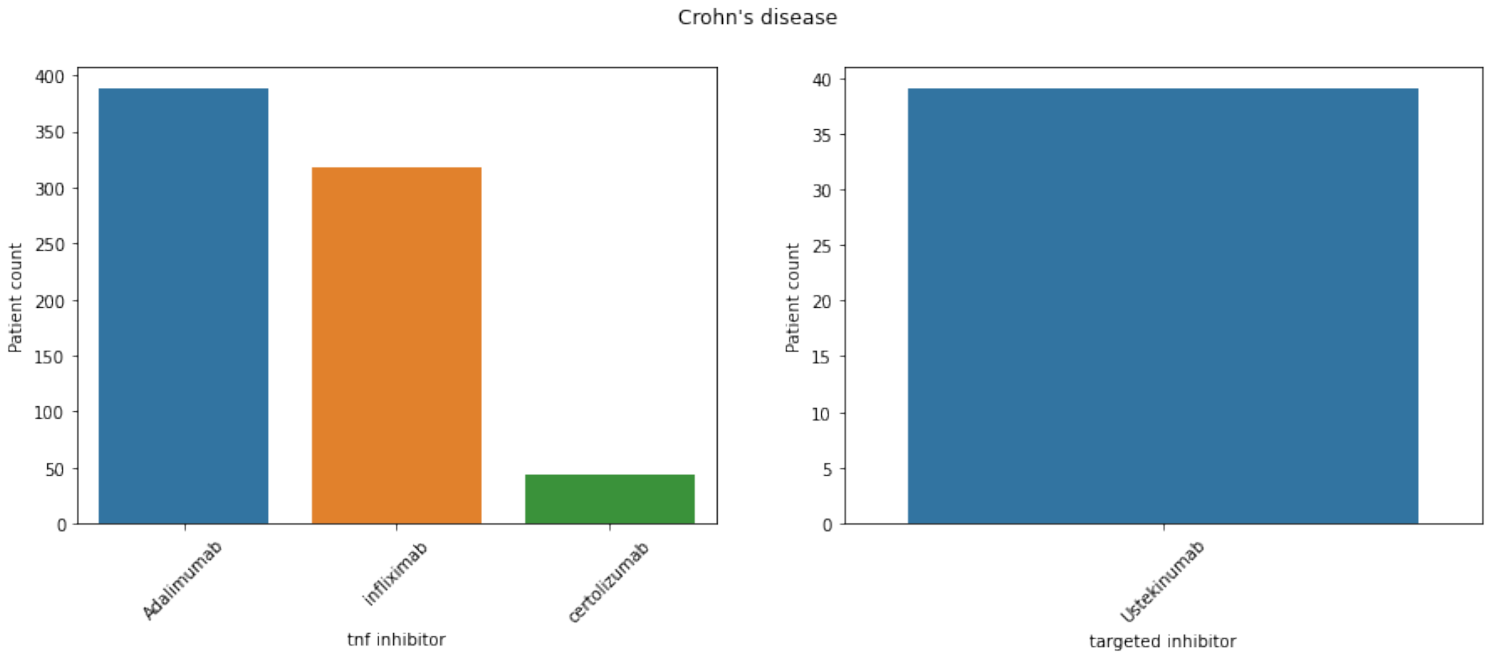
We can see that across all IMID the TNF-alpha inhibitors are more commonly prescribed. This is opposed to predication of hypothesis 1 that targeted cytokine-hub inhibitor medications would be more commonly prescribed. Surprisingly in the case of (RA), even the JAK inhibitor Tofacitinib is prescribed more often than the Toilizumab, the targeted cytokine hub inhibitor for (RA).

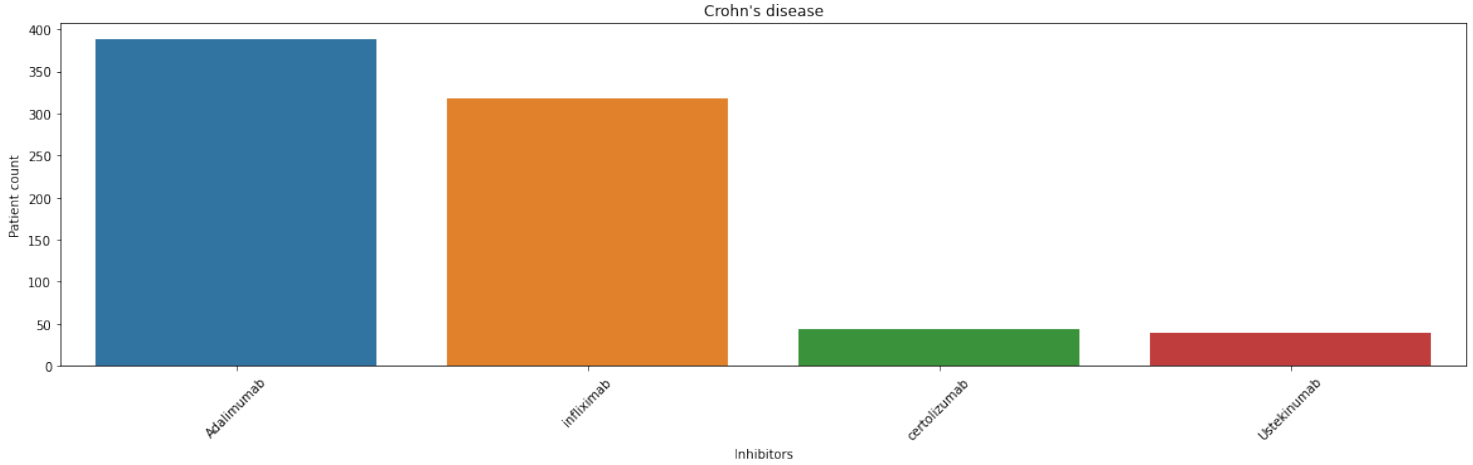
This difference could be due to TNF-alpha inhibitors and JAK inhibitors being medications that have been in use for longer time compared to the more recently used targeted cytokine hub inhibitors. In particular the paper [1] was published in 2021 but the COHD contains data on patients between 2013-2017.

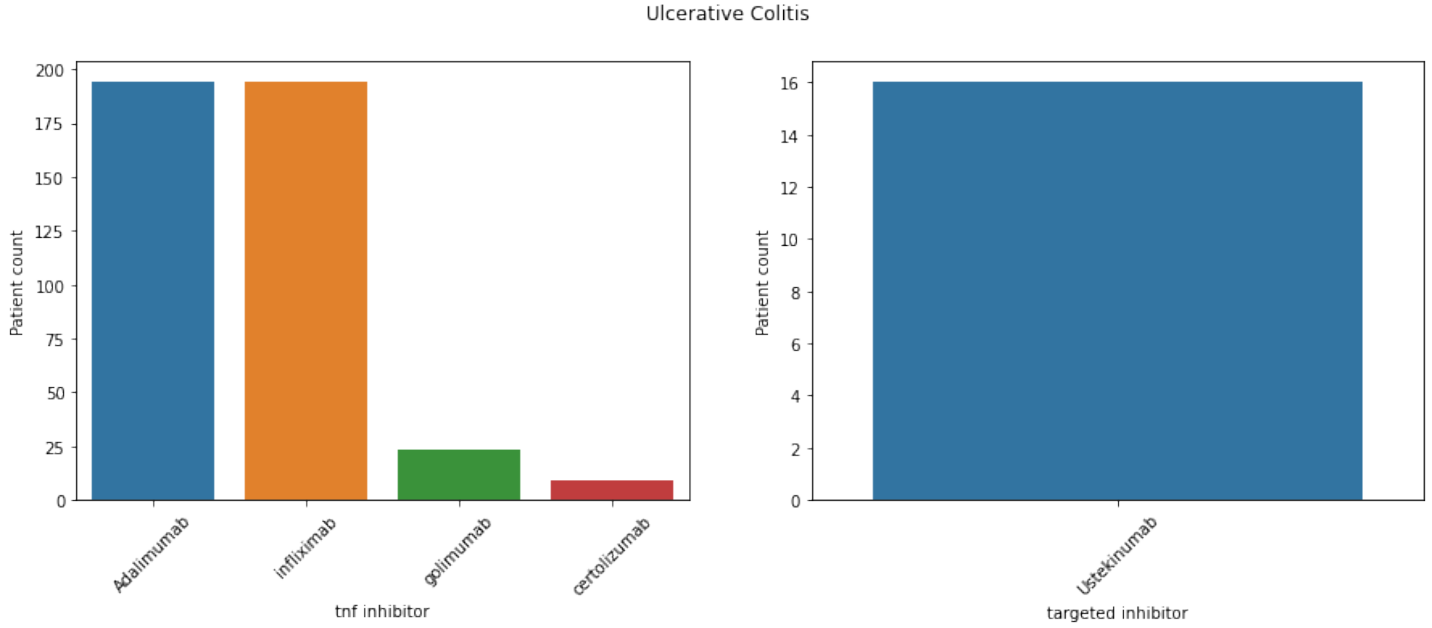


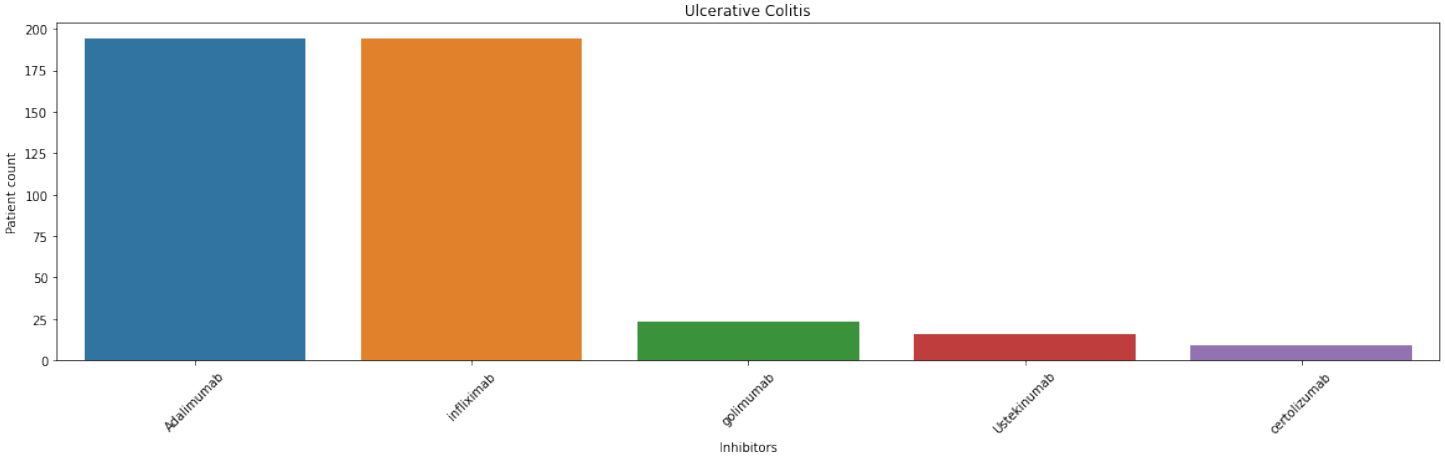
Caption

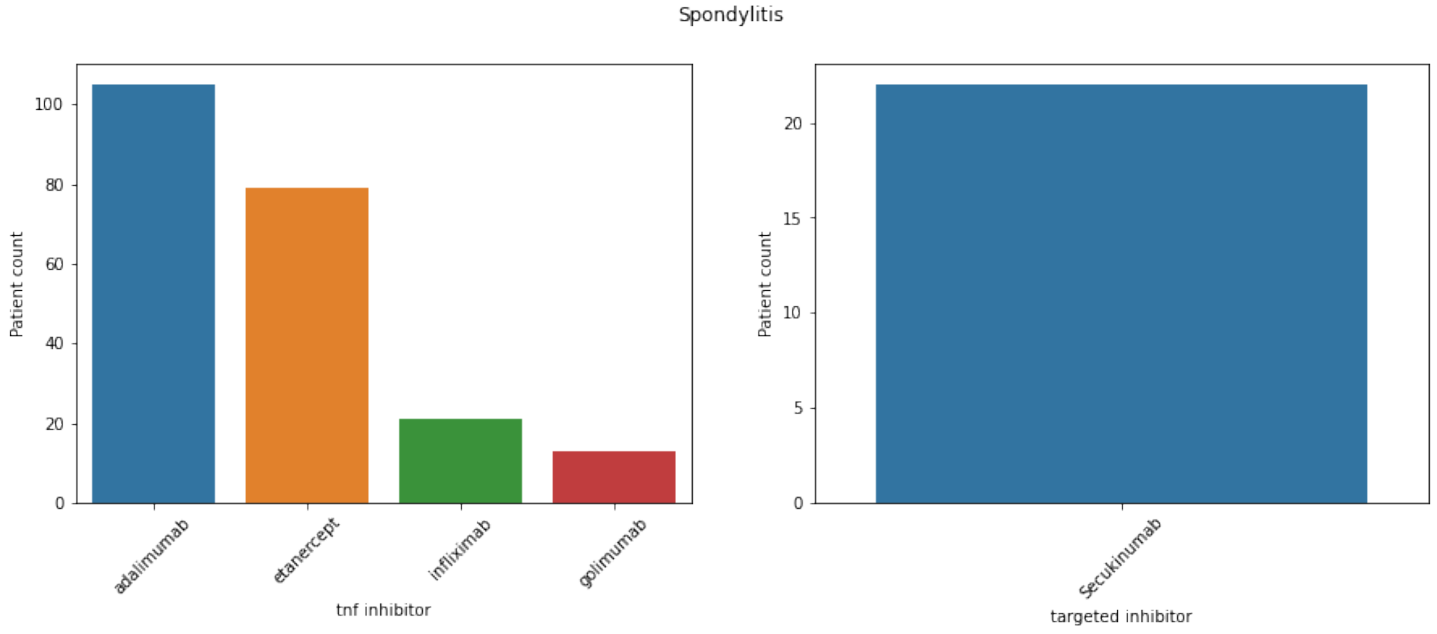

rheumatoid arthritis
Patient count
0
100
200
300
400
500
etanercept
Adalimumab
Tofacitinib
infliximab
Tocilizumab
golimumab
certolizumab
Inhibitors

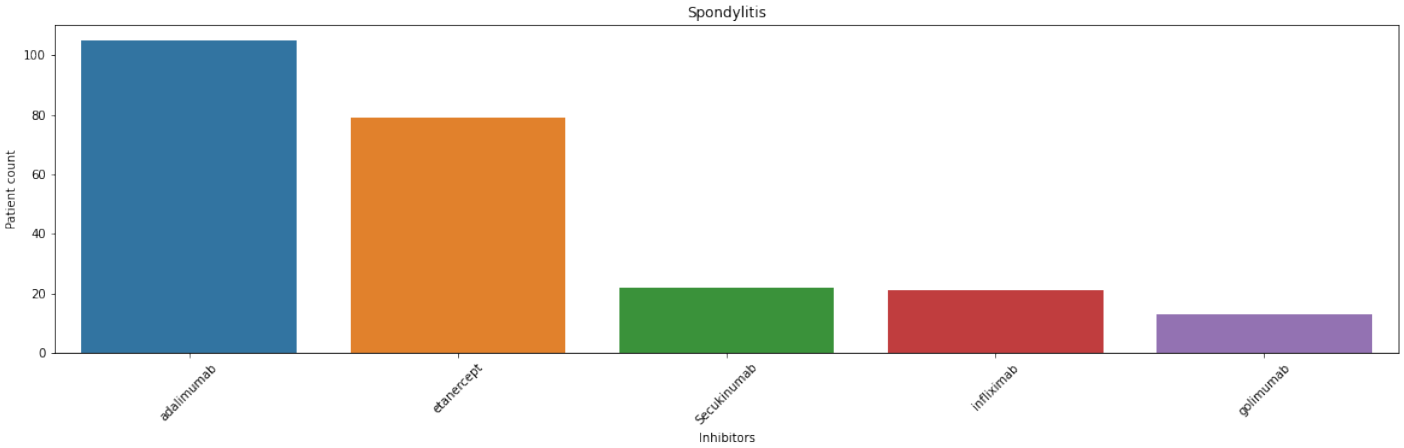

Crohn's disease
Patient count
0
50
100
150
200
250
300
350
400
Adalimumab
infliximab
certolizumab
tnf inhibitor
Patient count
0
5
10
15
20
25
30
35
40
Ustekinumab
targeted inhibitor

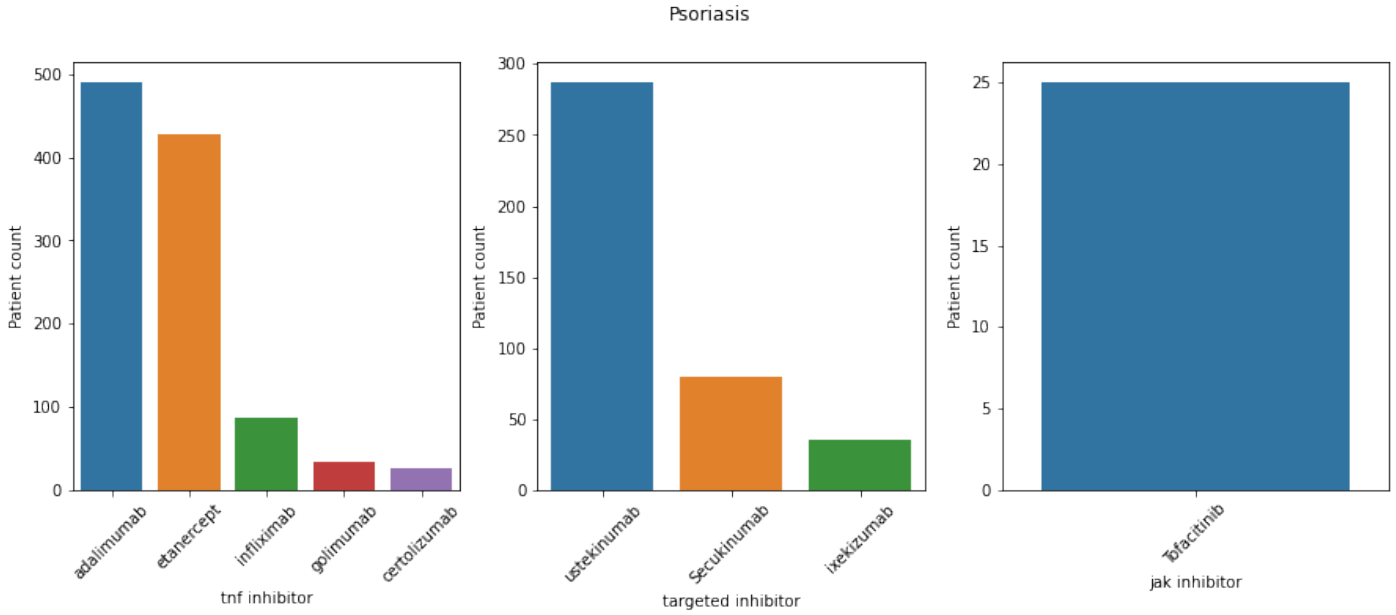

Crohn's disease
Patient count
0
50
100
150
200
250
300
350
400
Adalimumab
infliximab
certolizumab
Ustekinumab
Inhibitors

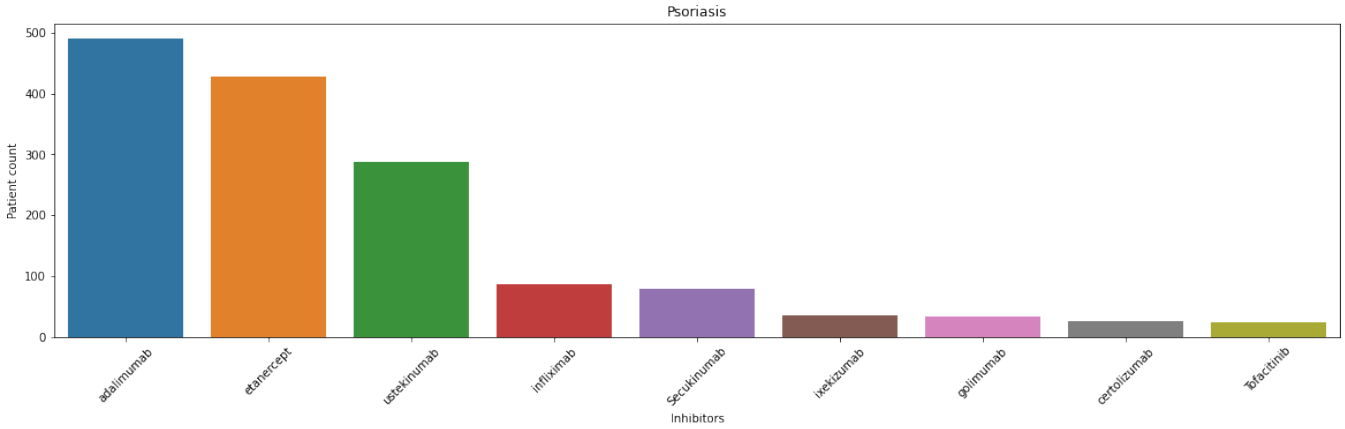

Ulcerative Colitis
Patient count
tnf inhibitor
Adalimumab
infliximab
golimumab
certolizumab
targeted inhibitor
Ustekinumab


Ulcerative Colitis
Patient count
Inhibitors
Adalimumab
infliximab
golimumab
Ustekinumab
certolizumab


Spondylitis
Patient count
tnf inhibitor
adalimumab
etanercept
infliximab
golimumab
targeted inhibitor
Secukinumab

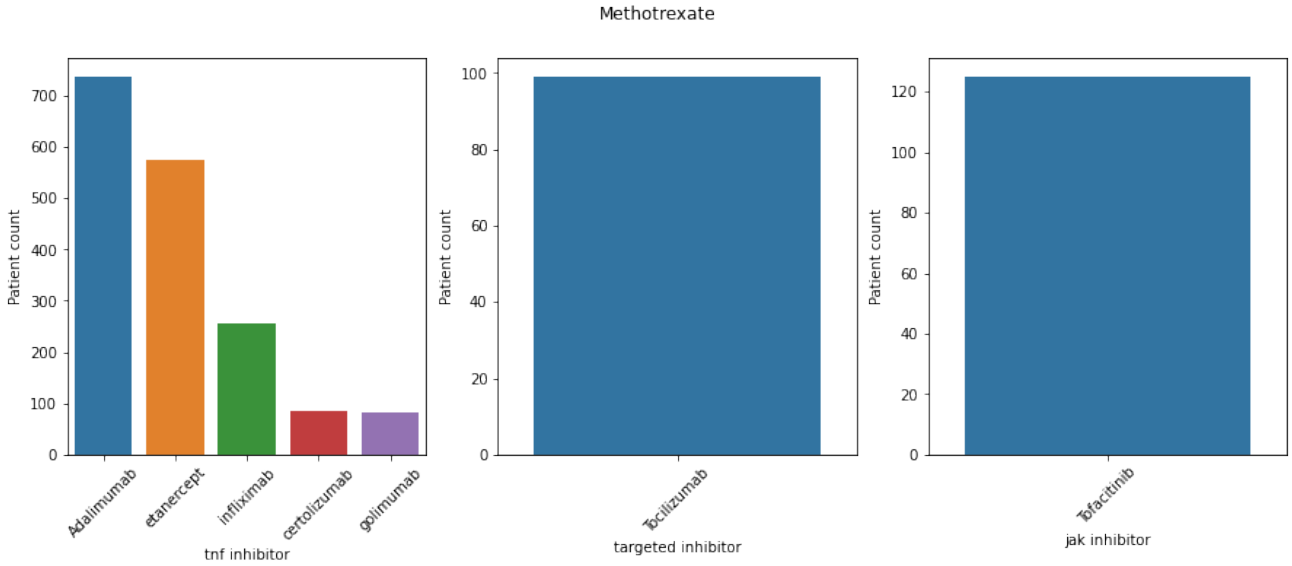

Spondylitis
Patient count
0
20
40
60
80
100
adalimumab
etanercept
Secukinumab
infliximab
golimumab
Inhibitors

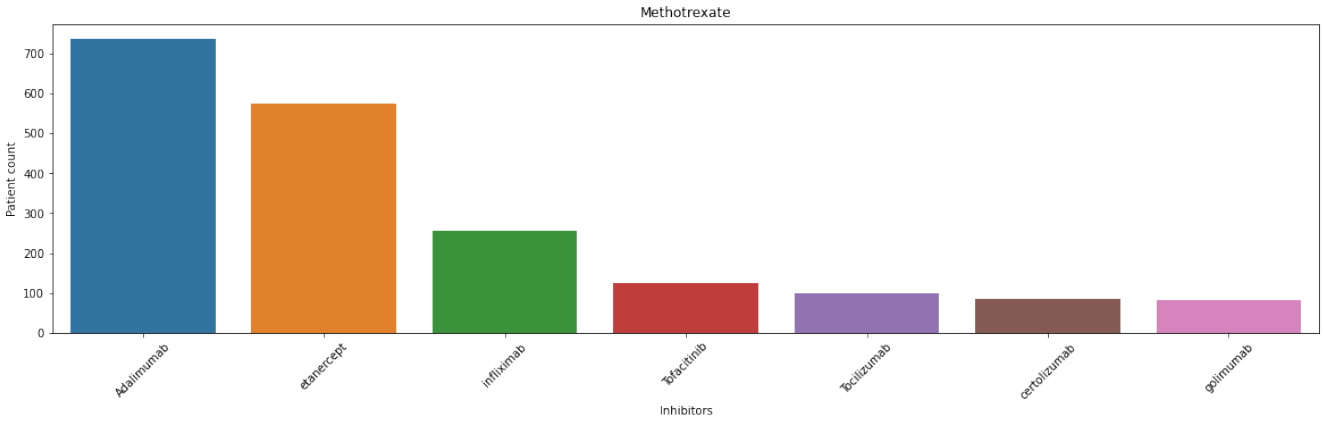

Psoriasis
Patient count
0
100
200
300
400
500
adalimumab
etanercept
infliximab
golimumab
certolizumab
tnf inhibitor
Patient count
0
50
100
150
200
250
300
ustekinumab
Secukinumab
ixekizumab
targeted inhibitor
Patient count
0
5
10
15
20
25
Tofacitinib
jak inhibitor


Psoriasis
Patient count
0
100
200
300
400
500
adalimumab
etanercept
ustekinumab
infliximab
Secukinumab
ixekizumab
golimumab
certolizumab
Tofacitinib
Inhibitors

For hypothesis 2 we present bar plots of patient counts who were prescribed both methotrexate and another IMID medication. The medications are grouped into three drug types: TNF-alpha inhibitors, targeted cytokine hub inhibitors, and JAK inhibitors. Beneath these plots there is a larger bar plot containing all of the patient counts found. Details of the specific numbers for each patient count can be found in the accompanying code (add link to GitHub later).

Based on these plots we see that the least commonly paired drug with methotrexate is the JAK inhibitor Tofacitinib. This supports our hypothesis 2. However it is important to note that TNF-alpha inhibitors were not uniformly less frequently paired with methotrexate compared to the other drug types. For example the targeted cytokine hub inhibitor ustekinumab was paired with methotrexate more often than a few TNF-alpha inhibitors however, the most frequently medications paired with methotrexate were TNF-alpha inhibitors. This seems to support an idea from [1] that certain TNF-alpha inhibitors need anchor drugs in a way that other more targeted medications do not but this data is far from conclusive.





References:

1. Reframing Immune-Mediated Inflammatory Diseases through Signature Cytokine Hubs, https://www.nejm.org/doi/full/10.1056/NEJMra1909094

2. Columbia Open Health Data, clinical concept prevalence and co-occurrence from electronic health records, https://pubmed.ncbi.nlm.nih.gov/30480666/

1. link to COHD Github / API