

Interactive Visualization Tool for Drugs and Therapeutic Agents in COVID-19 Open Research Dataset (CORD-19) Instruction

Simon Munyan, Peggy Li, Matthew Rever, Anna Hiszpanski and T. Yong Han
Lawrence Livermore National Laboratory

Introduction

The COVID-19 Open Research Dataset (CORD-19) Visualization in Tableau allows users to explore the field of research on COVID-19 treatment research. Through a dashboard view of information on chemicals and efficacies related to treatments, users can formulate their own hypotheses and form conclusions based on the data. This document serves as an overview of the various controls, panels, and functions within the CORD-19 visualization to enable users to intuitively navigate and utilize this tool.

Chemicals were extracted using **ChemDataExtractor** (M.C. Swain & J.M. "ChemDataExtractor: A Toolkit for Automated Extraction of Chemical Information from the Scientific Literature", *J. Chem. Inf. Model.* 2016, 56 (10), pp 1894–1904.) and other methods described in A.M. Hiszpanski, B. Gallagher, K. Chellappan, P. Li, S. Liu, H. Kim, J. Han, B. Kailkhura, D.J. Buttler, T.Y.-J. Han. "Nanomaterial Synthesis Insights from Machine Learning of Scientific Articles by Extracting, Structuring, and Visualizing Knowledge." *J. Chem. Inf. Model.* 2020, 60 (6), pp. 2876–2887.

Controls

The controls found in the upper left corner of the dashboard allow users to filter by two Anatomical Therapeutic Chemical (ATC) levels and by efficacy, as well as reset all filters. The ATC classification system categorizes chemicals based on five levels. On the dashboard, users can filter chemicals based on the first two ATC levels. ATC Level 1 classifies chemicals based on their anatomical/pharmacological group¹. ATC Level 2 classifies chemicals based on their pharmacological/therapeutic subgroup¹. Efficacy describes what purpose a chemical serves in the body. Every chemical has a set of efficacies associated with it. For example, the common pain-reliever ibuprofen has the ATC code M01AE01. The first ATC level of ibuprofen is M, which means the chemical works on the musculo-skeletal system. The second ATC level of ibuprofen is M01, meaning it is an anti-inflammatory and anti-rheumatic product. Ibuprofen has several efficacies, the most well-known being anti-inflammatory, anti-pyretic (i.e. fever reducer), and analgesic (i.e. pain reliever). The user can reset the ATC levels and any applied filter with the Reset Filters button. Note that this button does not reset the Selected Efficacy.

Panels

The dashboard is composed of six panels which serve to present different views of the data. These panels are affected by the controls and user inputs, such as clicking on an efficacy or

chemical in one of the panels. The relationships between the controls and each panel as well as the relationships between panels will be discussed.

Top Efficacies

The “Top Efficacies” panel in the top left is a bar plot which shows the number of papers published about each efficacy. The controls for ATC Level 1 and Level 2 affect this panel by including only the efficacies in the “Top Efficacies” panel which correspond to chemicals that fall under the ATC levels. Additionally, the user can select one of the efficacies displayed on the panel, which will set the Selected Efficacy control to the one selected on the panel. This action applies a filter of the selected efficacy to the other panels, causing the titles of these panels to include the selected efficacy.

Efficacies co-occurring with [Selected Efficacy]

The “Efficacies co-occurring with [Selected Efficacy]” panel in the bottom left is a bar plot which shows the efficacies that most frequently occur with the selected efficacy. The panel can also be used as a filter if the user selects an efficacy in the panel. This filter will affect the other panels by filtering the chemicals based on the Selected Efficacy and the efficacy chosen on the panel. The filter can be removed either by clicking in empty space in the panel or by clicking on the Reset Filters button.

Efficacies sharing chemicals with [Selected Efficacy]

The “Efficacies sharing chemicals with [Selected Efficacy]” panel in the bottom center is a bar plot which shows the efficacies that most frequently share chemicals with the selected efficacy. The panel can also be used as a filter in the same way as the “Efficacies co-occurring with [Selected Efficacy]” panel. It is important to note that if efficacies are chosen in both panels, then the chemicals will be filtered by the Selected Efficacy and the two efficacies chosen in the panels. This may cause few to no chemicals to appear in the “Top Chemicals for [Selected Efficacy]” panel, since it is being filtered by three efficacies. These filters can be removed either by clicking in empty space in the panel or by clicking on the Reset Filters button.

Top Chemicals for [Selected Efficacy]

The “Top Chemicals for [Selected Efficacy]” panel is a bar plot which shows the most frequently occurring chemicals that fall under the Selected Efficacy as well as any efficacies chosen in the two previously discussed panels. This panel can be used as a filter if the user selects a chemical in the panel. This filter affects the “Chemicals Timeline” and “Monthly Change in Paper Count” panels by filtering out all chemicals except the selected chemical. These filters can be removed either by clicking in empty space in the panel or by clicking on the Reset Filters button.

Chemicals Timeline

The “Chemicals Timeline” panel is a scatter plot which shows what proportion of the papers published in a given year are about a certain chemical. This panel is filtered by the Selected Efficacy as well as any efficacies or chemicals chosen by the user in the three previously discussed panels. The “Chemicals Timeline” panel can also act as a filter by the user selecting either a chemical or year. Hovering over a point in the scatter plot causes a tooltip to appear for that point. The user can select the chemical in the tooltip to filter by that chemical. Additionally, the user can select “Go to ChEMBL” to open a webpage in the ChEMBL database for that chemical. Selecting a year, either in the tooltip or on the x-axis, filters the “Monthly Change in Paper Count” panel by that year. These filters can be removed either by clicking in empty space in the panel or by clicking on the Reset Filters button.

Monthly Change in Paper Count

The “Monthly Change in Paper Count” panel is a heat map which shows the percent change in the number of papers published about a given chemical each month. The percent change is calculated based on the change in the number of papers from the previous month for which data is available. Hovering over a color bar shows the tooltip which contains the year, month, chemical, and percent change in papers.

(1) WHO | 2. Anatomical Therapeutic Chemical (ATC) Classification. WHO 2017.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344