Big Data Analysis on Adverse Effects of Aquatic Pharmaceutical Contaminants to Human Health

Nirali P. Trivedi, Vinayak Ganesan and Phalguni Ghosh*
Department of Natural Sciences
Middlesex County College, Edison, NJ
Nirali P Trivedi: npt15@scarletmail.rutgers.edu
Vinayak Ganesan: vinayakg5142@my.middlesexcc.edu
Ghosh: PGhosh@Middlesexcc.edu

ABSTRACT
A recent publication by the USGS reported the presence of 88 pharmaceutical compounds in 308 streams across the USA. These findings point to a nation-wide environmental concern and pose a risk to human and aquatic-ecosystem's health. Hence there is a need to generate a complete picture of disease risk, by constructing a conceptual map illustrating the link between the contaminating chemicals and diseases. This study is not simple, given the complex nature of disease mechanism and the mode of action of the chemicals analyzed. Today where analytical chemistry methods are useful in detecting the presence of pharmaceuticals in water, they are insufficient to understand the direct, harmful effects of pharmaceuticals getting into humans through the ecosystem of the water bodies. Herein, we have carried out "Big Data" analysis to identify the disease-association of these 88 compounds. We have extracted data from the Comparative Toxicogenomics Database, Drugbank and used the Cytoscape software to construct the chemical-disease networks to explore the breadth of the issue. Using Cytoscape to study biological networks, begins with building a combined network for the topic of interest and mapping the curated network for analysis. Being able to visualize combined and curated data allows the researcher to understand and evaluate targeted information. In our studies the (a) 88 chemicals were linked to about 2208 data points spanning several diseases; (b) diseases were distributed in 19 classes; (c) Disease/Chemical associations were generated for the most harmful chemicals and major diseases. Hence, the disease map generated by this study can be helpful in strategizing proper healthcare efforts in the event of a contamination crisis.

KEYWORDS
big data analysis; pharmaceutical contaminants; Human Health; environment; chemical-disease network; network biology

INTRODUCTION

Water system of a country is one of the most crucial assets for a nation because of its extensive use for domestic, public as well as industrial activities. As witnessed by different research studies of the nineteenth century i.e. during the post industrialization age, the biggest setback that came along industrialization is increased contamination of wadeable streams across the United States. This trend of water contamination has not changed much with time except the shift of contamination concern from nonorganic contaminants in the past to pharmaceuticals and biochemicals in the present time. Unfortunately the impact of water contamination on human and aquatic organisms' health has only worsened with time. Past studies have shown the contaminants primarily inorganic compounds linked to multiple organ damage upon consumption and causing diseases like cancer, skin lesions and pneumonia (Tchounwou *et al*., 2012). High contamination levels of metals like Aluminum, Arsenic, Cadmium etc. in water sources was a major cause of concern as these metals were known to cause a plethora of human diseases as published in the study by National Research Council (US) Safe Drinking Water Committee. Washington (DC) (National Academies Press (US); 1982).

A recent study that gained significant attention is published by USGS scientists (Bradley, *et al*. 2020), and aimed to analyze contamination of the surface water systems across the wadeable streams of the USA. The study is crucial as it hinted towards a novel threat to human and aquatic health in the form of high contaminant levels of commonly prescribed pharmaceuticals in these water bodies. The published data of the referenced study contained information for 308 water head streams and showed the consistent presence of 88 top leading contaminants in water samples from all the sites. In this study, we tried to build on the results of the USGS study to better understand the harmful interactions of these 88 contaminants by exploring the human conditions caused by their non-therapeutic effects. The primary goal of this study is to explore, identify and visualize the harmful interactions of these commonly found contaminating pharmaceuticals on human health as it will motivate to develop new safer pharmaceutical products and design effective wastewater treatment plans to combat this sort of environmental threat. An extensive understanding of the complex diseases-chemical relationships is crucial towards addressing such environmental threats.

In this study we started with collecting relevant drug-disease interaction data available from the published studies to generate a vast database. The information collected was then made into different data networks to provide a comprehensive view and derive some important conclusion about the overall target and toxicity of these pharmaceuticals. Producing data networks is a convenient method to deal with 'Big Data' available online. The past interpretation of "Big Data" was confined to the increased volume, variety and velocity of information till the last decade but now the term is inclusive of the increasing ability to analyze and interpret these data (Hulson et al., 2019). The process requires two important tools: (a) standardized data content, format from a reliable database, and (b) software to collaboratively form networks with display, analysis features.

The major barrier for studies focused on evaluating effects of contaminants on human health is their diverse and indirect effects along with varying dosage of these contaminants in

addition to differential human susceptibility which varies on individual levels (Briggs, 2003). Although there is a vast amount of literature on the mechanisms of action of most pharmaceuticals, it is difficult to gauge the relevancy of these data. We have used the Comparative Toxicogenomics Database (CTD) (Davis, *et al*.) to extract pharmaceutical-disease data for this study. CTD publicly promotes understanding about the effects of chemicals on human health by integrating curated data reported in published literature to help explore chemical-gene and disease interaction. CTD also provides transitive inferences for each curated chemical-genes-disease connection which helps scientists to formulate novel testable hypotheses about effects of drug and chemical actions and disease etiologies (King *et al*., 2012).

Network science as used in this research study, is a good way to depict the relationships among source and target nodes and to understand the combined interactions instead of the isolated features. To better represent our curated data from CTD, we utilized Cytoscape i.e., an open source software to visually explore possible biological interconnections through its interactive visualization interface (Su et al., 2015). In our chemical-disease network, the pharmaceuticals were defined as source nodes and diseases were defined as target nodes and the relationship between them including genes, were represented as edges connecting the nodes. To enhance and specify this study we manually classified the 88 pharmaceuticals in 29 pharmaceutical categories based on the conditions they are used against by the public. We also classified the filtered 2208 linked diseases in our final datasheet into 20 categories based on their defined classes by CTD. In particular, the goal was to identify the different links of the contaminant pharmaceuticals of different categories to the potential human disease development. Our study can be helpful in reconsideration of the proper prescription and disposal regulation of these pharmaceutical contaminants in our country considering the health of its citizens as the top priority.

METHODS AND PROCEDURES

2.1 Data from Water Contamination Study

We obtained the primary data and background information on the increasing contamination level of human used pharmaceuticals in the waterbodies across the United States from the recent study published in 2020 (Bradley et al., 2020). The dataset contained 88 contaminating pharmaceuticals and their corresponding concentration in each of the four regions throughout the country: Pacific Northwest, PNSQA; California, CaSQA; Northeast, NESQA; and Southeast SESQA. A bar graph (see results and discussion) was plotted to compare the cumulative concentration in each of the contamination sites where the pharmaceuticals were detected in each of the four regions and this helped to identify the top 5 pharmaceuticals present at the highest concentrations.

2.2 Database of Chemical-Disease Associations

The publicly available Comparative Toxicogenomics Database (CTD) was used to obtain the required dataset for the pharmaceutical-disease interactions on March 01, 2020. This database contains information on how chemical exposure affects human health through complex

interactions between chemicals, genes and diseases. As of writing this paper, the CTD contained information of over 221,516 curated chemical – disease direct interactions for 13,766 unique chemicals, 51,897 genes and 7200 diseases from published peer-reviewed articles. The dataset obtained originally for this study consisted of curated pharmaceutical-disease directed interactions with 88 contaminating pharmaceuticals and these were found to be linked to 9,408 diseases entries via different gene interactions.

Figure 1 represents an overview of the extraction and manual curation of the data. The primary dataset was manually curated to remove (a) derivatives of pharmaceuticals (for e.g. entries for Amitriptyline – N oxide were removed and only Amitriptyline interactions were considered); (b) combinations of the 88 pharmaceuticals (For e.g. entries for dextromethorphan – quinidine combination were removed and only dextromethorphan entries were considered); (c) and those pharmaceuticals that did not have direct evidence in the CTD of causing harmful side-effects; i.e., no correlation with causing a disease or no role in the etiology of a disease; (d) 7 of the 88 compounds did not have sufficient data in the CTD in the context of this study. Hence, these pharmaceuticals were removed from the dataset (10-hydroxyamitriptyline, Dehydronifedipine, Desmethyldiltiazem, Guanylurea, 1,7 dimethylene, Hydroxizine, and Noverapalmil). (e) Finally, since we were interested in analyzing the extent to which pharmaceutical contaminants negatively affect human health, we removed all the therapeutic drug-disease interactions from the dataset. In summary, we extracted only the associations supported by direct evidence of marker-mechanism for causing disease. Also experimentally induced diseases were removed, since they do not represent the scenario studied here. Hence, we performed a row by row manual approach against automated curation, considering each pharmaceutical, associated disease and interaction based on their validation in previous publications. This approach led to 63 pharmaceuticals, 555 diseases and 2,208 associated diseases interactions as our final data set.

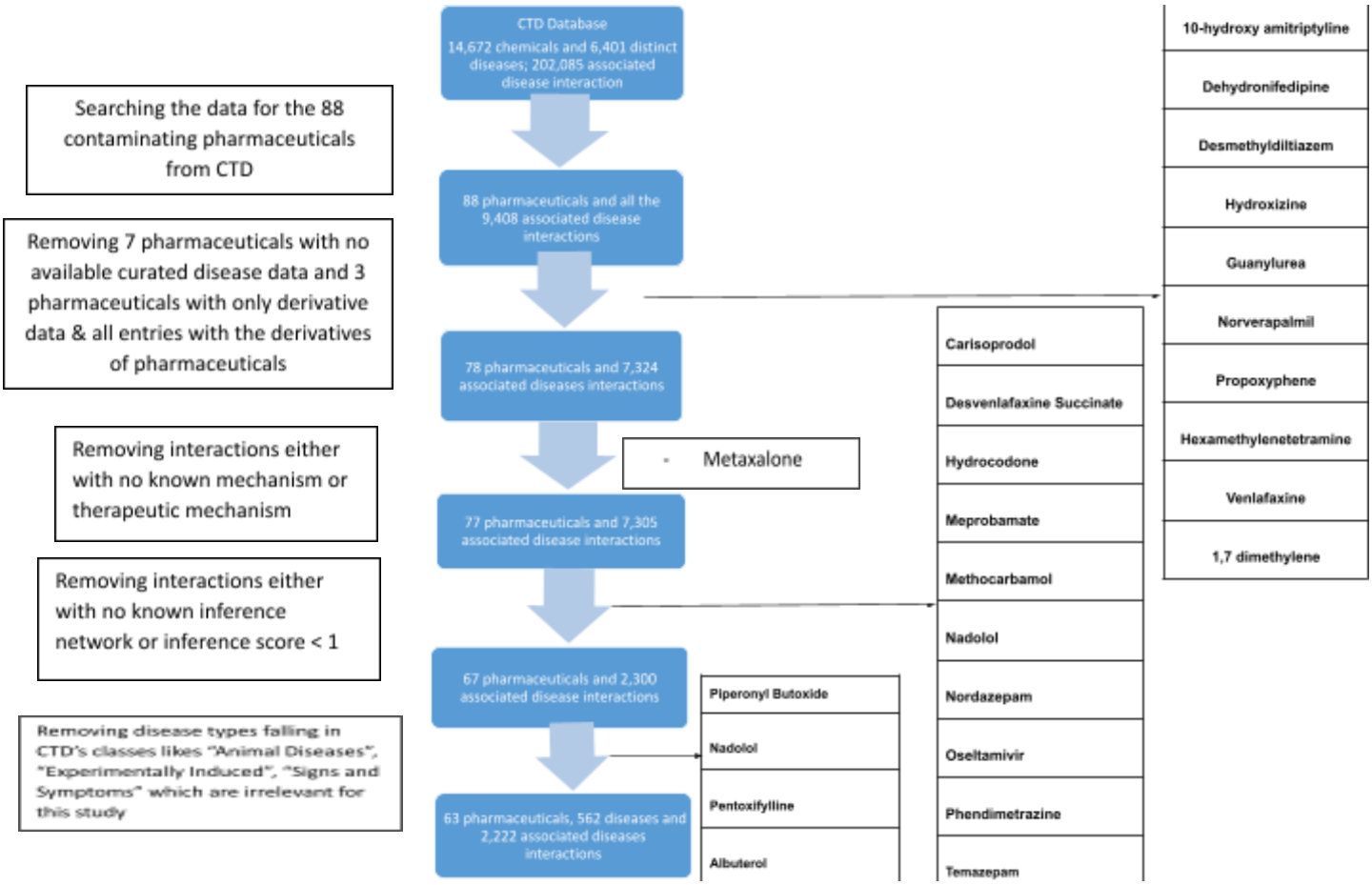


Figure 1: Data curation process. Data were collected from the CTD database as of March 1, 2020, and curated manually to identify the 63 pharmaceuticals, 555 diseases and 2208 interactions amongst them.

2.3 Classification of Pharmaceuticals

All pharmaceuticals that are present in the final database were further classified into several pharmaceutical categories. In this study the 63 final pharmaceuticals were categorized in 29 pharmaceutical classes like Antidiabetic, Analgesic, Anti-depressant, Antibiotics, ADHD related, Dermatosis, etc. The classes were assigned based on the general function of the pharmaceuticals in the medicine industry or based on the types of disease it is known to cure.

2.4 Classification of the human diseases

The 63 pharmaceuticals in our final data set were negatively correlated with 2208 diseases. Likewise, we classified the 555 diseases by the organs system of the human body they primarily affected (e.g. cardiovascular, nervous system), or the type of disease (e.g. cancer). This classification led to 19 disease categories.

2.5 Building Pharmaceutical-Disease networks

The network analyses were performed using the Cytoscape 3.8.1 software (Shannon *et. al.*). Cytoscape is a free software platform that allows the investigation and visualization of integrated diverse networks.

The final excel datasheet containing a total of 2,208 pharmaceutical-disease interactions was imported into cytoscape. Each column of the datasheet was assigned functions for building the network. The pharmaceuticals were defined as source nodes and the diseases were defined as target nodes for all the directed chemical-diseases data. Other multiple complimentary information about the nodes like pharmaceutical categories, disease-categories and inference scores were defined as source attributes and target attributes, respectively. The inference network i.e. the genes linked to each pharmaceutical - disease interaction were assigned interaction and edge attributes. One of the important edge attributes used in the study was inference score. If a larger number of genes points to a chemical-disease association, it is assigned a higher inference score in CTD (Davis et al., 2020). The supplementary information like mechanism types, reference count were excluded. All the nodes were given a node size corresponding to its degree or the number of other nodes connected to it. The link between a chemical and a disease was represented by the line or edge joining these two node. The strength of this chemical-disease interaction as given by the inference score was represented by the thickness of the edge

We have used two type of network representation throughout this study. In a pharmaceutical-based network, all the pharmaceutical nodes were color coded based on different pharmaceutical categories like analgesics, antihypertensives, CNS stimulants etc., and all diseases were represented as a single color. Conversely, in a disease-based network, the diseases were color-coded by category like, cancer, cardiovascular, digestive etc. and pharmaceuticals were represented in a single color. Similarly, different shapes in the data network can be used to make the interactions easily visible and evident as it differentiates source nodes from the target nodes. For all the data networks presented in this study elliptical nodes represent diseases and rectangular nodes represents pharmaceuticals. Inference scores and interactions involved for a specific pharmaceutical-disease pair is depicted via the directed edges. Later, in the study less complicated data networks contain different edge thickness as they represent the inference scores for each interaction. Many useful layouts were applied on the data-networks in this study to make it evenly spaced, more comprehensible, and explorative of the various characteristic each network represents.

Results and Discussion

1 Water contamination levels for all the 88 pharmaceuticals across the four regions

The study performed by USGS National Water Quality Assessment and Regional Stream Quality Assessment (NAWQA -RSQA) mentioned previously, analyzed samples from 308 wadeable streams (261 urban gradient sites and 47 low impact sites) spread across the four regions of the USA during 2014-2017. Upon analysis of all the samples collected in the study, 108 total pharmaceuticals were identified in significant quantities, among which 88 analytes were detected across all four regions (Bradley et. al., 2020). The site-specific cumulative maximum and median

concentrations of these contaminants ranged nd-36,142 ng/L and nd-8,756ng/L, respectively. Figure-2 below represents the plot comparing the contamination levels of different pharmaceuticals to analyze their influence on the total contamination. The horizontal axis lists the 88 analytes and the vertical axis displays the corresponding cumulative number of sites where each pharmaceutical was detected for the four different regions analyzed.

We specifically observed higher contamination detected in streams of eastern regions i.e. NESQA and SESQA (grey and golden bars respectively) compared to that of western regions specifically Pacific Northwest. From the 88 drugs analyzed, the top five polluters were Metformin, Caffeine, Cotinine, Lidocaine and Carbamazepine. These 5 compounds were detected in water samples of more than 40 sites in at least one of the four regions. The highest region-specific contamination of Lidocaine is observed in the north eastern region with 70 regional sites (74.5 % of NESQA sites and 22.7% of all sites) detecting the drug. Whereas, Nicotine followed by Metformin were the topmost polluters with comparatively high detection numbers across all four regions. Some important findings from the study revealed that Nicotine and its metabolite Cotinine were found in 70% and 47% of all sites, respectively. Metformin, an anti-diabetic drug for type-II diabetes was found in 68% of all the sites analyzed. Caffeine and Lidocaine were both detected in 42% of total sites. Some other contaminants with a moderate number of cumulative site detections include carbamazepine, an anti-seizure medication (41% of total sites and 57.1% of PNSQA sites); acetaminophen, an oral analgesic (26% of all sites and 67.8% of SESQA sites); fexofenadine, an anti-histamine drug (21% of all sites and 39% of SESQA sites). This data helped us to focus on the specific drugs for which we wanted to explore their drug-diseases interactions. These pharmaceuticals are given emphasis based on their high contamination outreach. The harmful effects of the top 5 polluting compounds (Metformin, Caffeine, Cotinine, Lidocaine and Nicotine) on human health has been discussed in section 5 through their data networks.

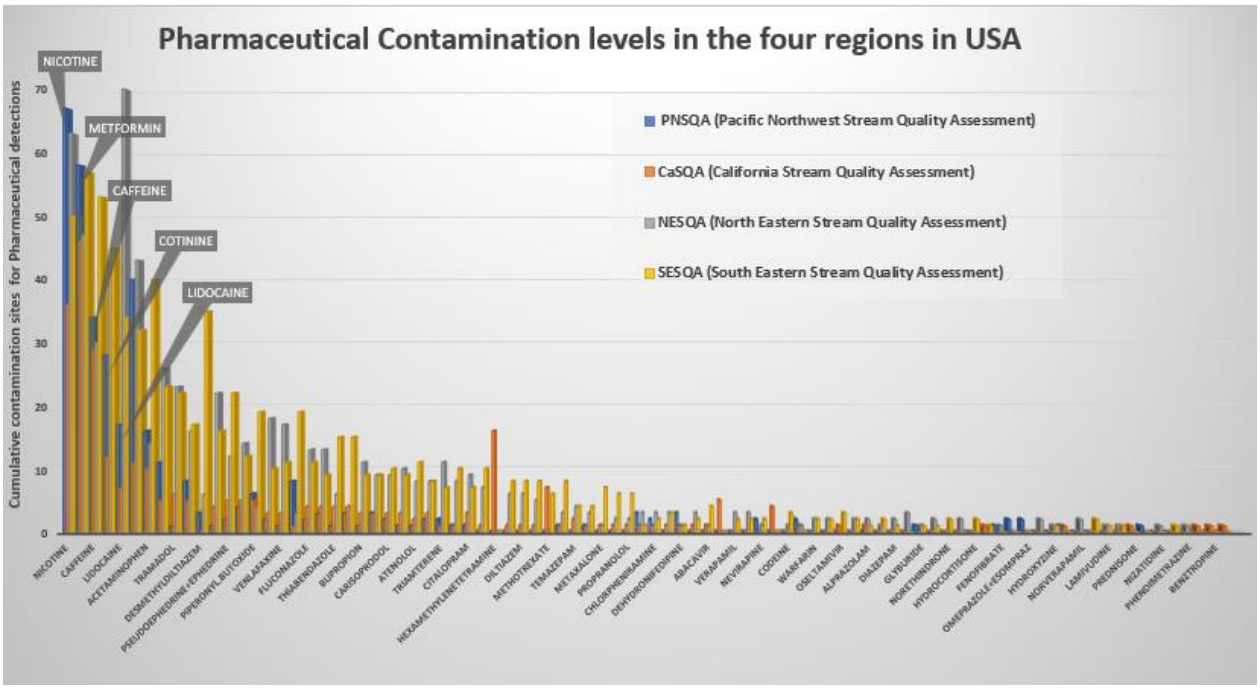


Figure 2:- Cumulative number of sample sites across USA where each pharmaceuticals were detected at least once during the 2014-2017 synoptic samplings of water from wadeable streams in the Pacific Northwest (PNSQA, blue bar), California (CaSQA, orange bar), Northeast (NESQA, grey bar) and Southeast (SESQA, golden bar) regions as part of the USGS Regional Stream Quality Assessment (RSQA), data published in the previous study earlier in 2020 (Bradley *et al*, 2020).

2. Analyzing some features of the Pharmaceuticals-Diseases Data

We studied specific features of the pharmaceuticals and associated diseases using the final dataset and their background information from CTD to categorize them and evaluate their disease interactions. The dataset was extracted to find the number of diseases linked to each pharmaceutical as shown in Figure 3a, to compare the harmfulness of the pharmaceuticals. The horizontal axis lists the pharmaceutical names and the vertical axis displays the number of diseases associated with each contaminant.

Methotrexate was associated with the highest number of diseases i.e. 172 diseases followed by Nicotine and Carbamazepine, with the number of disease associations being 148 and 132 respectively. While Benztropine and Triamterene are linked to the minimum number of diseases; 3 diseases in each case. Nicotine, besides being one of the highest contaminant drug in the waterways, is also very harmful due to its ability to cause more than 100 diseases in humans. The drug Metformin, one of the top polluters, is far less harmful than Nicotine with only 35 diseases associated with it. The detailed analysis of the individual disease networks of Methotrexate, Nicotine and Carbamazepine i.e. the 3 most harmful drugs has been discussed in section 6.

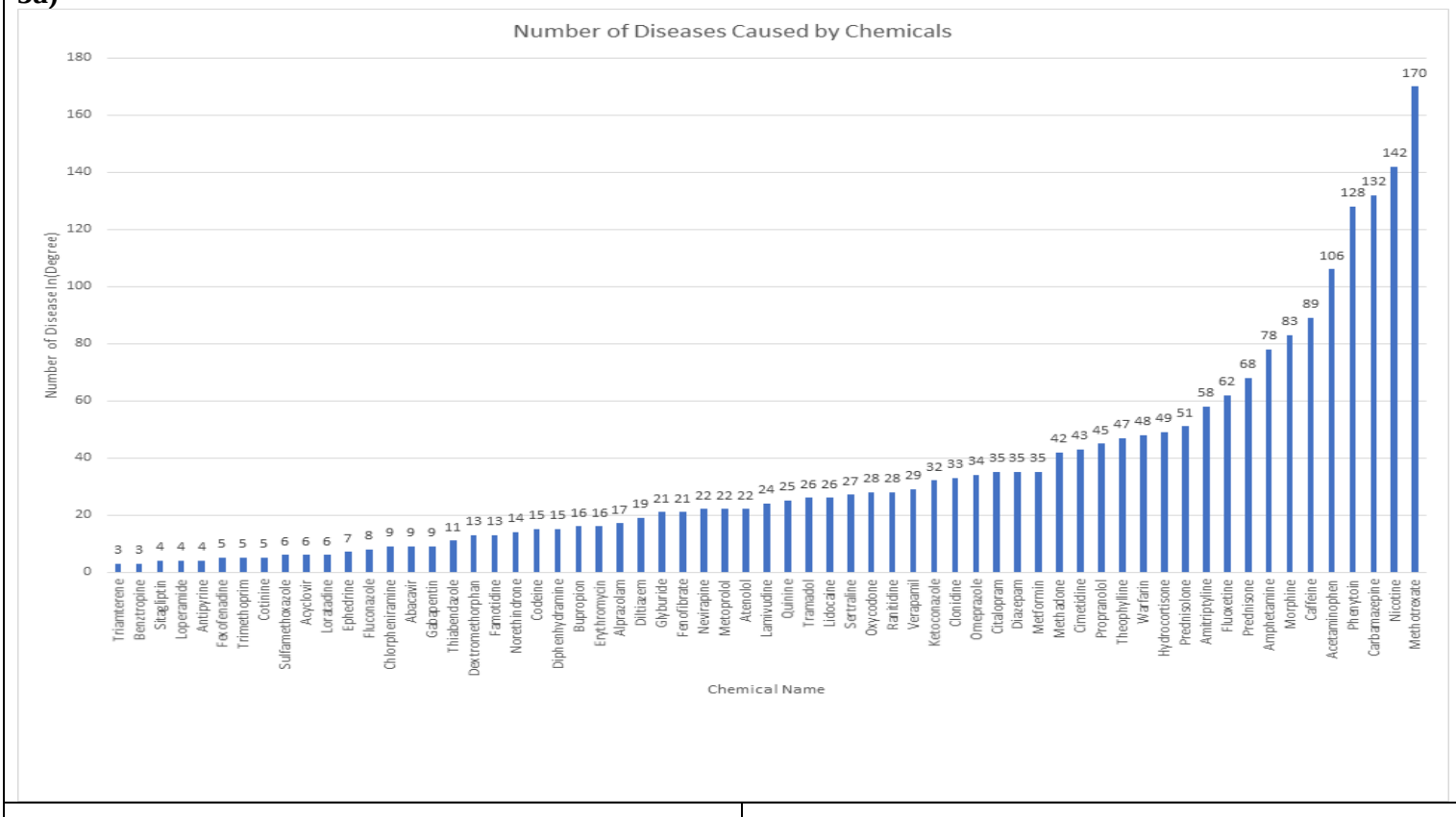
Fig. 3b shows a pie-chart distribution for the 63 pharmaceuticals on the basis of their respective function. We found the 63 compounds to be distributed amongst 29 classes of pharmaceuticals. A significant number of the contaminating drugs fell into the class of analgesic (11%), followed by anti-depressants (10%), and anti-hypertensive (8%) classes. Analgesics are the pain relieving medications that don't affect consciousness. Analgesic pathways are previously known to affect various pain processing mechanisms happening in the peripheral and central nervous system of the body. This class of drugs are regarded as one of the most valuable drugs in the pharmaceutical industry, making them available to masses over the counter which also makes it a dangerous group of medications (Creg et al., 2013, Cazacu et al., 2015). This group of drugs include : Acetaminophen, Antipyrine, Codeine, Methadone, Morphine, Oxycodone and Tramadol. Acetaminophen (commonly known as Tylenol) has broad accessibility to common people across the country leading to its contamination levels (Fig 2) along with a very high disease association (Fig 3a). Acetaminophen is one of the ten most contaminating drugs according to the USGS data of 2020.

Similar to categorizing the pharmaceuticals, we also categorized the diseases associated with the contaminants. They broadly fell into 20 categories. The pie-chart shown in Fig 3c, shows the distribution of the types of diseases. The data revealed that a significant number of diseases

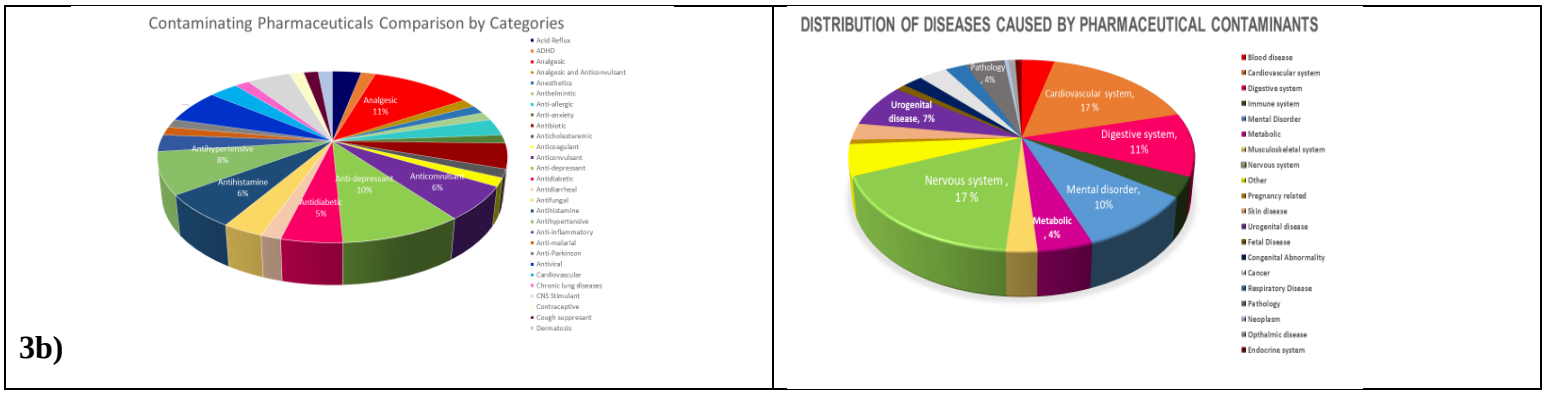
which are associated with the contaminating-pharmaceuticals belong to the cardiovascular disease and nervous system disease categories equally (17%), followed by others like, mental disorder diseases (10%), digestive system diseases (11%) and urogenital diseases (7%). According to CTD, cardiovascular diseases refers to the pathological conditions involving the cardiovascular system including heart, blood vessels and pericardium. The fact that heart disease is the leading cause of death for men, women, and people of most racial and ethnic descent in the United States might also hint at the possibilities that one of the underrated factors causing heart diseases can be the triggering interactions of the contaminating pharmaceuticals inclusive of those analyzed in this study. Several studies have reported the link between environmental pollution and cardiovascular diseases (Hatzis et al (2016), Harvard Health Letter (2011)). The common diseases that fell under this category from the dataset obtained from CTD include stroke, heart arrest, heart block, etc. which were further analyzed in more detail later in the study (Figures 7a-7e).

Based on our findings we further extended our study to further explore two disease classes i.e. a) Cardiovascular disease, one of the categories with maximum number of associated diseases and b) Cancerous diseases which often gets linked to environmental pollution (Boffetta et al. (2003), Lewandowska et al. (2019)) and in this study contain diseases with significantly high inference scores, indicating the involvement of multiple genes in the chemical-disease pathway (Fig. 8).

3a)



3c)



3b)

Figure 3a: Bar plot representing the number of diseases associated with the pharmaceuticals analysed in this study
3b: Pie-chart showing the distribution of the 63 pharmaceuticals present in our final dataset in their respective pharmaceutical categories.
3c: Pie-chart showing the distribution of the 2,208 diseases present in our final dataset in their respective disease categories

3. Building primary Pharmaceutical-Disease data networks

We used cytoscape to investigate the interactions of pharmaceuticals and their associated disease targets that could affect human health. We initially created a hair-ball like data network containing a large number of hub nodes and almost rare stand-alone pharmaceutical-disease interactions (Fig. 4a). In this network the source nodes were displayed by red squares representing the 63 pharmaceuticals and the target nodes were displayed as the blue ellipses representing the 562 unique disease of the dataset. The directed edges of the network not only connect target nodes to the source nodes, but also contain information of all the human genes involved in the pathway of interactions.

Although such hair-ball like networks contain a vast amount of information, its layout does not allow easy interpretation of the source-target relationship. One way to extract useful information from such networks is through alteration of the layout and visual attributes to bring out useful information. A second strategy is to apply a variety of analytical techniques to understand the physical characteristics of the network. Parameters like the node degree or number of connected components, average node density, average shortest-path distance, measures of centrality, are useful descriptors to quantify the observed interactions in the network. In our work we have used both strategies to infer the harmful effects of the 63 water-contaminating pharmaceuticals. We started by modifying the layout. As a first step, we color coded the pharmaceuticals (squares) and diseases (circles) by their respective classes which enabled us to generate a chemical-based network and a disease-based network.

Figure 4b. represents the pharmaceutical based network in which the pharmaceuticals are square shaped and color coded based on their respective drug category, all diseases were represented as pink circles. For all nodes, chemicals and diseases, we assigned a third attribute, size. The size of each node is directly proportional to the number of other nodes linked to it. Figure 4b, clearly shows that methotrexate, nicotine, carbamazepine, phenytoin, acetaminophen and caffeine are each linked to a large number of diseases.

The pharma-based network was further sorted and grouped by pharmaceutical class to understand how each class of pharmaceuticals cumulatively contributed to the observed diseases (Figure 4c). The data network indicated that most of the diseases are linked to analgesics (red squares) with 38 associated diseases, followed by anti-inflammatory drugs (faded blue squares) and CNS stimulants (grey squares) with 19 associated diseases each. As per this pharma-grouped network, analgesics contains the highest number of pharmaceuticals, 7 (shown as red squares), followed by the antidepressants, 6 drugs (shown as green squares), and followed by antihypertensive, 5 drugs.

Similar to the pharma-based network, we also created a disease-based network (Figure-4d). In this network all diseases were color coded by category and represented as circles, while all pharmaceuticals were represented as pink squares. All nodes were assigned a size attribute similar to the previous network. The six compounds listed before (methotrexate, nicotine, carbamazepine, phenytoin, acetaminophen and caffeine) stand out as major hubs linked to multiple diseases. Chemical, drug induced liver injury, acute kidney injury and stevens-johnson syndrome were linked to the maximum number of contaminating pharmaceuticals. This was not fairly obvious in figure 4d, since the contaminants are linked to a larger number of nodes compared to the diseases. Hence the pharmaceuticals have higher degree and therefore larger node size, compared to the diseases. In order to decipher the spectrum of the diseases, we further sorted and grouped the network by disease class to understand the prevalent disease categories (Figure 4e). As per this disease-grouped network, cardiovascular (orange dots), nervous system (green dots) and cancer (grey dots) were the leading diseases outcome from pharmaceutical pollution of the waterways. Overall our research protocol shows an elegant approach to going from the skeleton-like information in the basic pie-charts (in Figs 3b, 3c) to visually refined maps of the contaminant-disease outcome (figs 4a-4e).

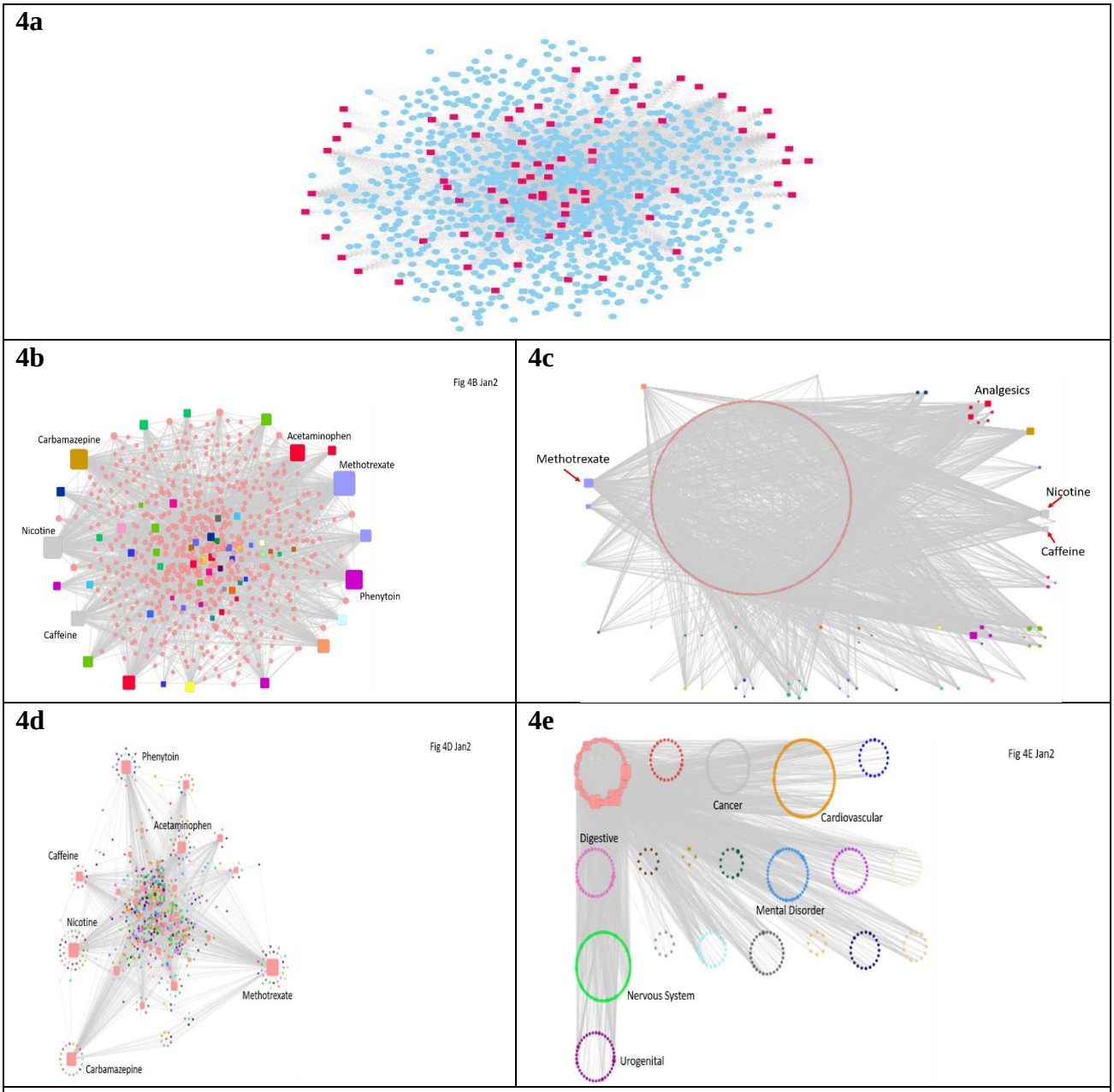


Figure 4a: Hairball network representing the interaction amongst the 63 pharmaceuticals (red squares) and 562 diseases (blue circles).
4b: Pharmaceutical based network, where the pharmaceuticals are color coded by category, and represented as squares, while all diseases were represented as pink circles. The grey lines or edges represent the chemical-disease interactions
4c: Pharmaceutical based network, sorted by pharmaceutical class
4d: Disease based network, where the diseases are color coded by category, and represented as circles, while all pharmaceuticals are represented as pink squares. The grey lines or edges represent the chemical-disease interactions
4e: Disease based network, sorted by disease category

4. Diseases caused by 5 top polluters

Following our findings in section-1, we took an in-depth look at the drug-disease networks of the top 5 contaminating pharmaceuticals i.e., Metformin, Caffeine, Cotinine, Lidocaine and Nicotine. Data networks discussed in this section are more specific and simplified then the drug-disease networks discussed earlier in the paper. This network of 5 compounds was extracted from one of the cumulative networks shown in Figure 4d. The diseases in the network are color coded based on their disease categories. This network has an additional feature, where the edges connecting the source node to target nodes are also colored according to the strength of evidence for such interactions (Figure 5). The information contained in the edges includes the inference score and the genes associated in the chemical-disease interaction. The darker the color of the line, the stronger is the evidence for such an interaction.

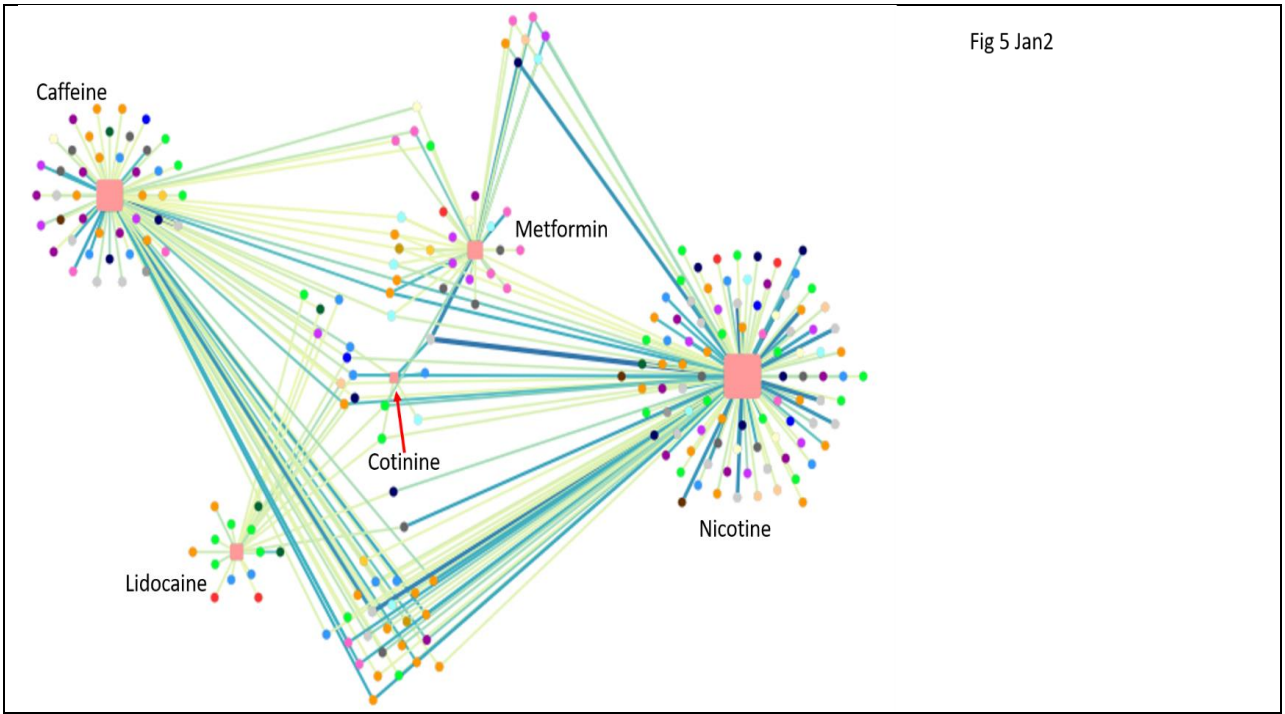


Figure 5: Subnetwork of the top 5 contaminating compounds Metformin, Caffeine, Cotinine, Lidocaine and Nicotine (pink squares) and their associated diseases (circles). Darker lines represent stronger chemical-disease association

From Fig. 5, we observed that out of the five pharmaceuticals represented, nicotine has the highest number of associated diseases followed by caffeine. Being the top most polluters,the association of these drugs with a large number of diseases (Figures 3a, 5) is very concerning. Interestingly, both drugs belong to the CNS stimulant category of drugs. On the other hand, cotinine, which is the major metabolite of nicotine, showed few disease association. One of the important information analyzed about these drugs is that all of these belong to the drug categories which are either commonly available and has a wide-spread usage or are known to be

addiction prone. Cotinine belongs to the class of antidepressant, lidocaine is an anesthetic whereas metformin is antidiabetic. Nicotine and Caffeine are the stimulators people intake by their choice and desire. Widespread use of caffeine through coffee intake and refreshment beverages probably contributes to its leading edge in water pollution. Sorting this 5-compound network by disease category shows that the combined interaction of all five drugs follows the overall trend seen in the general pharmaceutical network (Fig 4e) i.e., these 5 compounds are linked to the highest number of cardiovascular diseases, followed by the nervous system diseases and cancer.

Since nicotine stands out as one of the most polluting it is unfortunately, also associated with a large number of diseases. Hence, we have carried out an in-depth analysis of its adverse effect on human health by discussing it as a stand-alone network later in the paper.

As per our analysis, Caffeine follows nicotine in terms of the number of diseases associated with it. However, most of the disease interactions are not as harmful as those observed for nicotine. From our data network, the predominant side effects of caffeine are cancer and cardiovascular diseases but very few of these interactions are completely reliable due to their light edge color. According to the previous studies, caffeine consumption is relatively safe for healthy adults, but for some vulnerable populations, caffeine consumption could lead to impairments in cardiovascular function, sleep, and substance use (Temple, et al., 2017).

Metformin is the 4th most prescribed medicine in the US with 78,602,870 prescriptions in 2017. Our analysis shows that negative impact of metformin is minimal, cholestasis being the only predominant adverse effect. Fortunately, the diseases associated with metformin including cholestasis have comparatively low inference scores indicating that few genes are affected by the pharmaceutical. This observation is on par with the reported side effects of metformin.

Lidocaine was the 208th most prescribed drug in 2017. It is used as a local anesthetic for a large variety of surgical procedures. Moreover, it is widely available as over the counter pain relieving patches or crème, which backs up its widespread presence in the stream across the USA. As per our analysis the side effects of lidocaine is minimal, as highlighted by few diseases of lower inference score in our mapping.

Cotinine is an alkaloid found in tobacco and is also the predominant metabolite of nicotine. Cotinine was developed as an antidepressant, but was never marketed. As per our analysis, cotinine is mainly associated with lung neoplasm. Cotinine has the maximum interaction with the gene CYP2A6. This is because the protein encoded by the gene CYP2A6 is involved in the metabolism of nicotine and cotinine. Being the primary metabolite of nicotine, we would expect cotinine to be present in higher concentrations in the water bodies across the US similar to the contamination levels of nicotine.

5. Methotraxate, Nicotine, Carbamazepine are associated with highest number of diseases

Based on our observation from plot 3a, we produced individual disease networks for Methotrexate, Nicotine, and Carbamazepine (Figs 6a, 6b, 6c) i.e. the top 3 pharmaceuticals that

were associated with the maximum number of diseases. The harmfulness of the pharmaceutical in this study is determined by the number of associated non-therapeutic disease interactions. For the above mentioned drugs these have been identified as Methotrexate (170 diseases), Nicotine (142 diseases), and Carbamazepine (132 diseases).)Higher the non-therapeutic disease interactions more harmful the pharmaceutical.

Methotrexate is a chemotherapeutic agent and an immune system suppressant which is used to slow the growth of cancer cells. Methotrexate exerts it pharmaceutical action by inhibiting the enzyme DHFR **(**Dihydrofolate reductase), encoded by the *DHFR* gene, which ultimately leads to inhibition of DNA, RNA and protein synthesis. The major adverse effect of methotrexate is hepatotoxicity, also known as liver damage. As per CTD, the gene *SLC19A1* is shown to have the most interaction with methotrexate followed by *DHFR*. *SLC19A1* is a widely expressed gene that regulates the intracellular transport of folates and anti-folates drugs, like methotrexate. *SLC19A1* gene has been reported to be polymorphic in humans. Variants in the *SLC19A1* gene have been found to be correlated with various cancers and with variable response to methotrexate (Yee et al., 2011). A study with 240 ALL patients, treated with high dose of methotrexate have reported that the G80A mutation in the *SLC19A1* gene has been linked to gastrointestinal toxicity (Kishi, et al., 2007). As per our mapping studies methotrexate shows strong evidence of hepatocellular carcinoma as the predominant adverse effect. The other reported common side effect of methotrexate is linked to the nervous system (drugs.com). Headache, fatigue, drowsiness have been commonly reported side effects. These results support our mapping studies (Figure 6a) whereby we see gastrointestinal toxicity and adverse effects on the nervous system in individuals exposed to this compound.

6a Methotraxate	Methotraxate

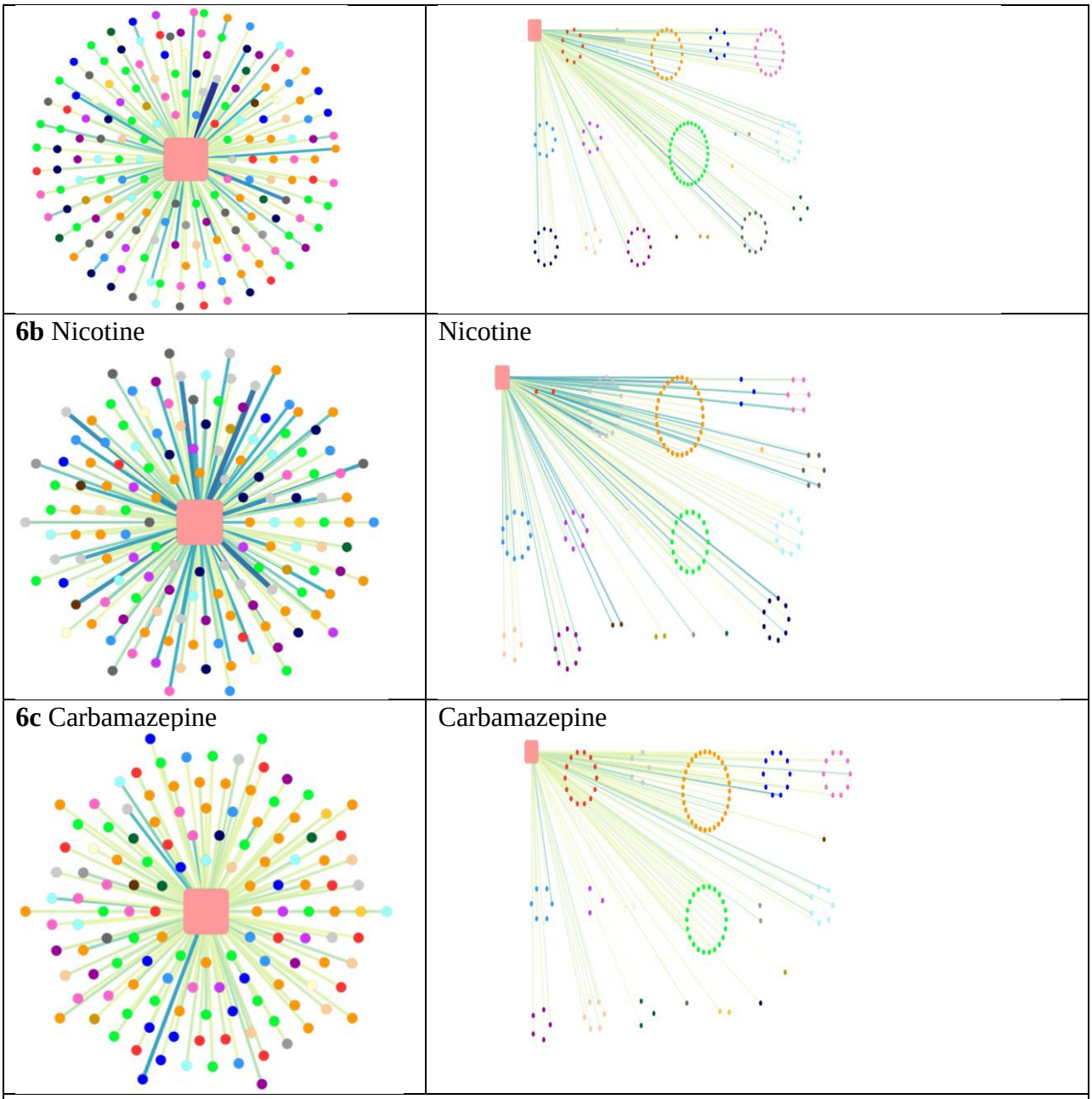


Figure 6: Methotrexate (a), Nicotine (b) and Carbamazepine (c) are each linked with the highest number of diseases. The left panels show the strength of the chemical-disease association, with stronger association being represented by darker lines. In the right panel, the diseases are sorted by category

Nicotine is a dangerous contaminant in that it is found in high concentrations in the waters and is associated with many diseases. Although nicotine is associated with lower number of disease interactions compared to methotrexate, it is more dangerous as the disease involved in the nicotine network has stronger evidence of the adverse effects (fig 6b). The results clearly show that nicotine is linked to various neoplasms (grey circles in Fig 6b) and several cardiovascular diseases. The dark, thick edge lines indicate that many genes are involved in causing those diseases and hence the presence of nicotine as a contaminating chemical is of major environmental concern. Nicotine is a highly toxic alkaloid (DrugBank). As per the Center for Disease Control and prevention, nicotine use leads to increased risk of cancer, COPD and coronary heart disease. This is evident when analyzing the top 10 genes listed by CTD as interacting with nicotine. These are *CHRNA7, CYP2A6, CHRNA4, CHRNB2, MAPK3, MAPK1, TNF, STAT3, CASP3, CHRNA5* (CTD). Of these, Nicotine acts as a receptor agonist for proteins expressed by *CHRNA7, CHRNA4, CHRNB2, CHRNA5*. However, *MAPK3, MAPK1, CASP3* play a role in apoptosis; TNF plays a role in cell death; while STAT3 is implicated in cell growth and division (proliferation), cell movement (migration), and the self-destruction of cells (apoptosis) (https://www.uniprot.org). The interaction of nicotine with the above mentioned genes lead to the observed neoplasms. The health hazards of nicotine has been widely reported in literature (Mishra *et al*. 2015)

Carbamazepine is an anticonvulsant and analgesic agent used for the management of bipolar disorder. In 2017, it was listed as the 176th most commonly prescribed medication in the United States, with more than 3.5 million prescriptions. As per our mapping analysis, shown in Fig 6c, carbamazepine is largely associated with cardiovascular and nervous system diseases. All the diseases associated with carbamazepine have low inference scores, indicating that carbamazepine affects very few genes as compared to nicotine. The greatest negative impact of this drug is linked to congenital abnormality. There has been reports of teratogenic effects of carbamazepine in children of women who were treated with this drug during pregnancy (Jones *et al*., 1989). Our analysis also shows that carbamazepine is linked to a large number of cardiovascular diseases. We will take a deeper look at the association of carbamazepine in different cardiovascular diseases in the later sections focused on cardiovascular disease sub-categories.

7a. Cardiovascular diseases : most prevalent disease category

Our preliminary analysis indicated that cardiovascular diseases are the most prevalently caused disease category linked upon combined analysis of all the drug-disease interactions considered in this study (Fiure 4e). This suggests that there is a possibility that cardiovascular problems may get initiated and/or aggravated by the consumption of the contaminated water polluted with leaching pharmaceuticals. In order to understand the explicit effect of the 63 compounds on cardiovascular diseases, we generated a sub-network of the original network focusing on cardiovascular disease interactions as shown in Fig 7a.

In this sub-network (Fig. 7a), all the pharmaceuticals are represented as squares and are color coded based on their pharmaceutical categories, as discussed before. Cardiovascular disease refers to all the pathological conditions involving the cardiovascular system including the heart;

the blood vessels; or the pericardium (CTD); and are represented as pink circles in this data network. The connecting lines represents the interactive association between the pharmaceutical and the disease as it contains the genomic information about the genes involved. Additionally, the size of each node corresponds to the number of nodes linked to it. Interestingly, six drugs: nicotine, caffeine (CNS stimulants), methotrexate (anti-cancer agent), carbamazepine, phenytoin (anticonvulsants) and amphetamine (used for ADHD), were found to be associated with majority of the cardiovascular diseases. This observation is reflected in the larger node size of these pharmaceuticals in Figure 7a. The strength of the interaction between a chemical and a disease is derived from the inference score of the interaction, which is reflected in the edge joining them. The stronger the association, the thicker and darker the line. As per CTD, the higher the inference score, the more likely the inference network has atypical connectivity.

As per our analysis, nicotine shows strong correlation for many cardiovascular diseases. There have been many reports addressing the adverse cardiovascular effects on epilepsy patients taking carbamazepine and phenytoin (Mintzer *et al*., 2009). One compound that also stands out in terms of interaction strength is hydrocortisone (shown in light blue square in figure 7a, linked by thick blue lines to diseases). Hydrocortisone linked hypertension has been linked to 42 genes, hydrocortisone linked heart failure to 30 genes, and 14 genes are known to be involved in hydrocortisone linked cardiomegaly (ctd.org). There has been reports in literature on the increased risks of steroids on the cardiovascular system (Hitti (2004), Sholter (2000)). In terms of pharmaceutical class, large number of analgesics and antidepressants are linked to cardiovascular diseases, compared to other classes. However, their impact is not as harmful as the leading offenders described above.

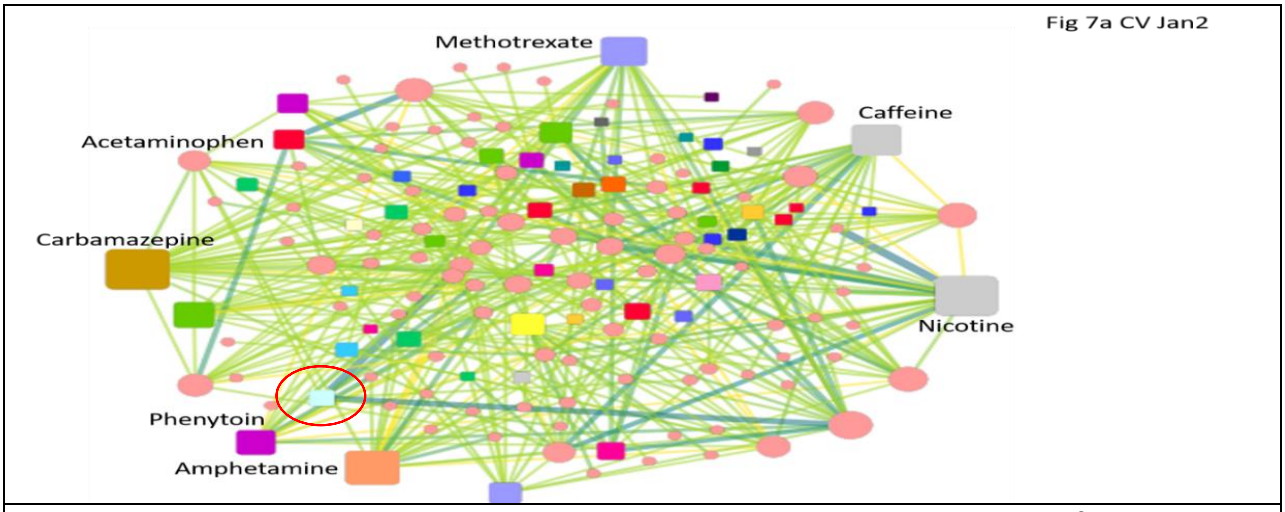


Fig 7a: Chemicals (squares) linked to cardiovascular diseases (circles). The size of each node corresponds to the number of nodes linked to it. Carbamazepine, Nicotine, Amphetamine are linked to the highest number of cardiovascular diseases

7b. Selected Cardiovascular diseases

To better analyze all the diseases under the cardiovascular disease categories and to visualize contribution of contaminating pharmaceuticals in causing specific diseases, we further isolated several sub networks. The cardiovascular pathologies we elaborated on are as follow: Bradycardia-Tachycardia (figure 7b), Hypertension-Hypotension (figure 7c), Heart block, heart

arrest, heart failure (figure 7c), and Stroke, Myocardial diseases (figure 7e). The usual suspects, nicotine, caffeine, methotrexate, carbamazepine, and phenytoin are predominantly present in all of these sub-networks. In this section, we elaborated on the key findings for each of these sub data network

Bradycardia or slow heart rate is linked to 15 pharmaceuticals whereas, tachycardia (or abnormally rapid heartbeat) is linked to 16 pharmaceuticals. Data network (figure 7b) shows that the impact of most of these compounds is mild, as represented by the thin and light-colored edges (or connecting lines). Acetaminophen stands out as having the strongest association for bradycardia and ventricular tachycardia. The cardiotoxic effect of this drug has been linked to its major metabolite norpropoxyphene (Heaney 1983). High frequency or dose of this compound poses increased risk for cardiovascular events (Chan, 2006). Nicotine is one of the pharmaceutical contaminants, which is found to be associated with both bradycardia as well as tachycardia. Fortunately, small dosage of nicotine similar to that found in contaminated water bodies is not currently known to be harmful. But severe nicotine poisoning is known to cause a biphasic response, with initial excitatory symptoms leading to many symptoms including tachycardia, followed by symptoms of bradycardia (Paik, 2018).

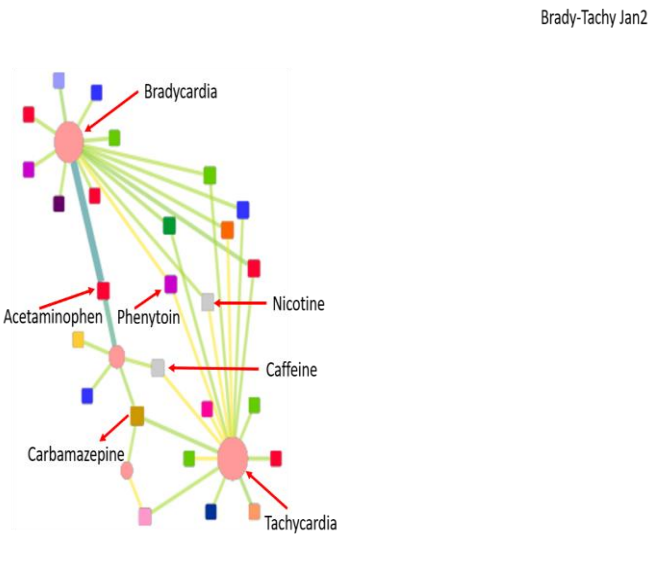
Figure 7c shows Hypotension, hypertension network and its associated chemicals. Hypotension or abnormally low blood pressure is linked to 16 pharmaceuticals whereas, hypertension (or persistently high systemic blood pressure) is linked to 19 pharmaceuticals. While the impact of most of the drugs are mild, the most harmful pharmaceutical in this network appears to be hydrocortisone. Hydrocortisone linked hypertension has been linked to 42 genes (ctd.org). The main mechanism has been reported to be high fluid retention (Hulisz, 2008), due to the overstimulation of the mineralocorticoid receptor. This leads to sodium retention in the kidneys and subsequent effects. There are many reports in literature detailing the effects of hydrocortisone on regulation of blood pressure (Sudhir (1989), Buning (2016)).

Figure 7d show the chemicals linked to heart block, heart failure and heart arrest. Heart arrest is associated with 13 chemicals, heart failure with 7 chemicals and hear block with 2 from our list of 63 pharmaceuticals. Carbamazepine is linked to all 3 diseases and phenytoin to heart arrest and failure. Toxicity arising from use of anticonvulsants like carbamazepine and phenytoin, has been reported in literature to be linked to cardiac arrest (Durelli 1985). Hydrocortisone has the strongest evidence for heart failure (dark, thick edge interaction), followed by methotrexate and phenytoin. As discussed earlier, steroids such as hydrocortisone has been linked to higher risk of cardiovascular diseases (Hitti 2004).

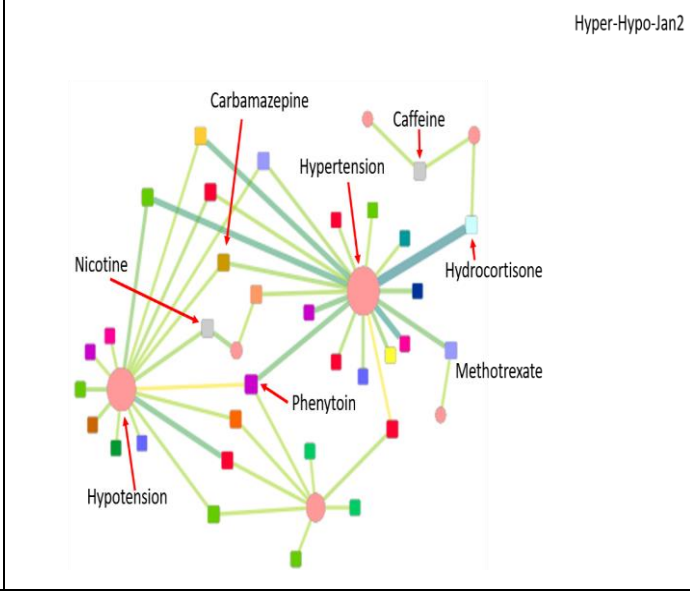
Next, Figure 7e shows the subnetwork for Stroke and Myocardial diseases. Myocardial infraction is linked to 12 chemicals, followed by 9 chemicals each for cardiomyopathies and stroke. The worst offenders in this network are amphetamine (linked to 6 diseases in this class), followed by nicotine (4 diseases), caffeine and phenytoin. The evidence for the link between nicotine, caffeine and phenytoin is also very strong as evidenced by the thicker and darker edge interactions (in figure 7e). Several studies have reported the link between amphetamine use and myocardial infraction (Zukkor 2015).

In summary, for the cardiovascular network, carbamazepine (linked to 31 CV diseases), nicotine (linked to 30 CV diseases), amphetamine (linked to 25 CV diseases), caffeine to 22 and methotrexate to 20 CV diseases. Of all the subclasses of cardiovascular diseases that we discussed in this section; hypertension has the highest degree as it is linked to 19 pharmaceuticals. This means that hypertension is linked to 19 out of the 63 pharmaceuticals that we have analyzed in our study. This is a concerning observation as systemic arterial hypertension is the most crucial precursor for all-cause morbidity and mortality worldwide and is associated with increased risk of many other cardiovascular diseases (Oparil, 2018). The common cardiovascular diseases that develop from hypertension include chronic heart disease, stroke, and coronary heart disease. Elevated blood pressure is positively correlated to the risk of stroke and coronary heart disease. Unfortunately, in worst cases complications include heart failure, peripheral vascular disease, renal impairment, retinal hemorrhage, and even visual impairment (Singh, 2017). Hence, there is an urgent need to focus attention on evaluating the profoundness of the impact that these contaminating pharmaceuticals can have on consumers, in triggering adverse conditions like hypertension.

7b



7c



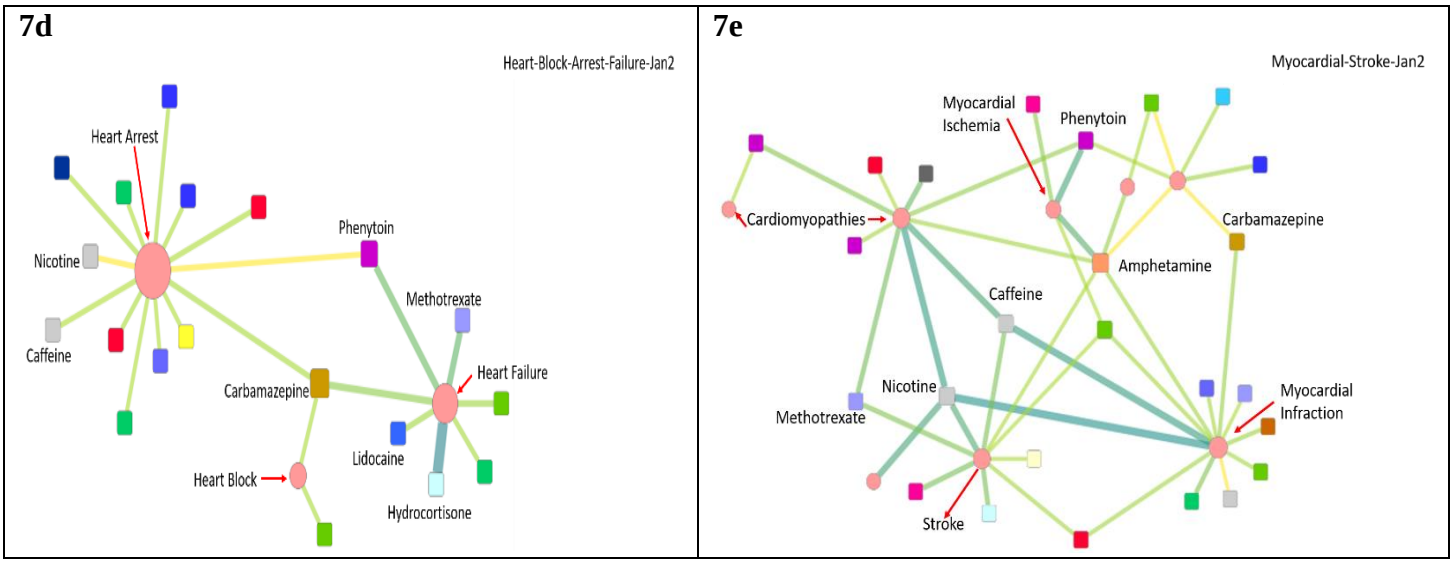


Fig 7b-7e: Cardiovascular subnetworks (b) Bradycardia-Tachycardia (c) Hypertension-Hypotension (d) Heart block, heart arrest, heart failure (e) Stroke and Myocardial diseases. Squares represent pharmaceuticals and circles represent diseases. The lines represent chemical-disease association. The stronger the association the darker and thicker the line

8. Cancer-pharma network

Figure 8 shows the drug-disease data network specifically for Cancer. Squares represent pharmaceuticals, color coded according to their drug category and circles represent the different cancer diseases. Again, the stronger the association between a chemical and a disease, the darker and thicker the line or edge interaction. 24 of the 63 chemicals studied in this work are linked to 44 types of cancer. The worst affecter drugs are Nicotine, linked to 17 types of cancer, and phenytoin linked to 12 types of cancer. Centers for Disease Control and Prevention (CDC) reports that nicotine can cause cancer in any part of the body (https://www.cdc.gov/cancer/tobacco/index.htm). This is in agreement with our mapping studies where we see nicotine is linked to colon, urinary bladder, lung, nose, pancreatic, stomach and other types of neoplasms. Several articles in literature (Grado 2014) details the direct contribution of nicotine to cancer and the carcinogenic mechanisms of such actions. The regulation of nicotine containing drugs or products need proper and strict regulation as nicotine is known to affect the cellular processes like proliferation, oxidative stress, apoptosis and DNA mutation causing cancer. It also affects the tumor proliferation and metastasis and causes resistance to cancer therapies to existing patients relying on chemo and radio therapeutic agents.

Phenytoin, which is an anticonvulsant used for the treatment of seizures. As per our mapping studies, this drug has been linked to many adverse effects, including 12 different types of cancer. There have been reports of cancer risk associated with phenytoin including cancer risk in people with epilepsy: the role of antiepileptic drugs (Singh 2005). The most concerning research finding indicates that phenytoin has been causally implicated in three human cancers: lymphoma, myeloma and neuroblastoma, the latter specifically in the setting of foetal hydantoin syndrome

(Adelöw et al., 2006), a similar observation was also derived from our data network. Methotrexate has the strongest association with hepatocellular carcinoma, due to the thicker edge lines for the interaction. In total, 176 genes have been reported to be linked to this association (ctd.org). Previous studies have shown that rheumatoid arthritis patients treated with methotrexate led to lymphoproliferative disorders in patients (Xavier et al., 2002). Fortunately, these studies included only several hundred patients and are not very reliable to detect an excessive risk of lymphoma upon the drug consumption. Methotrexate is still the most commonly prescribed drug for Rheumatoid arthritis.

Some of the other strong associations seen in caffeine with breast neoplasm and nicotine with lung neoplasm. Caffeine-rich foods such as tea, coffee, and chocolate were suggested to be carcinogenic in 1970s and 1980s (James, 1983). But the validation of such claims are unsettled to this day. The reason behind that is the probability of associations between coffee consumption and breast cancer may be confounded by other aspects of diet or by the lack of appropriate control for nondietary confounding factors. Liver and urinary bladder neoplasms are the major diseases triggered by this group of compounds. Liver neoplasm is linked to 10 chemicals, while urinary bladder neoplasm is linked to 8.

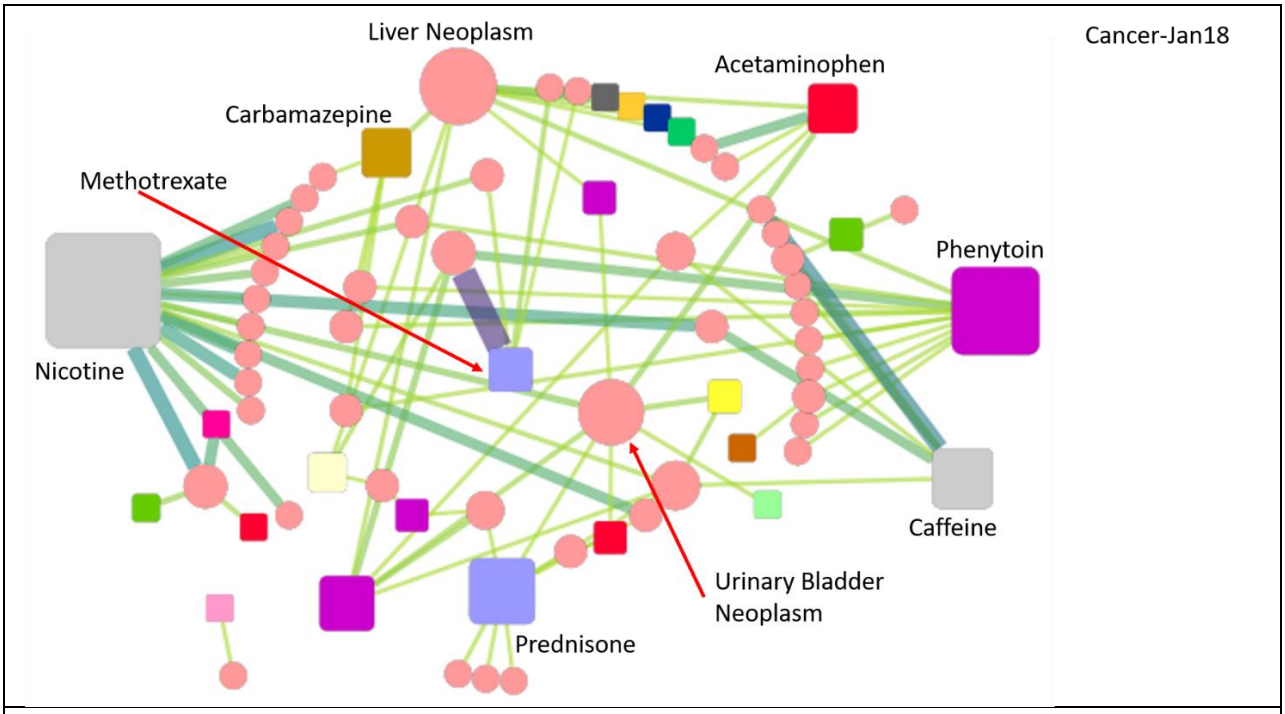


Fig 8: Cancer network. Squares represent pharmaceuticals and circles represent diseases. The size of each node corresponds to the number of nodes linked to it. The lines represent chemical-disease association. The stronger the association the darker and thicker the line.

The Pharmaceutical-disease interaction map and its statistical significance

The chemical disease interaction network of pharmaceutical contaminants polluting water sources constructed in Cytoscape is shown in Figures 4-8. The network is comprised of 2208

interactions (or edges) between 63 pharmaceutical contaminants in water and 555 diseases or illnesses. Highly associated chemicals can be seen to form hubs in the map. These are Methotrexate, linked to 170 diseases, Nicotine linked to 142 diseases, Carbamazepine linked to 132 diseases, Phenytoin linked to 128 diseases, Acetaminophen linked to 106 diseases. The interaction network is sparsely connected as indicated by the average number of interactions per node (Avg. Degree): 7.100, meaning that on average every node is connected to approximately 7 other disease nodes. The network density is also low for the same reason. Further analyzing the network via Cystoscape's Network Analyzer tool had uncovered various significant feature of the Chemical-Disease network including, but not limited, to characteristic path length, clustering coefficient, and network density. Table 1 shows the most relevant measures of the network as calculated by Cystoscape's Network Analyzer tool.

Table 1: Cystoscape's Network Analyzer tool.

Attributes	Values
Number of Nodes	618
Number of Edges	2208
Avg. Degree	7.100
High Score interacting Nodes	Methotrexate (170), Nicotine (142), Carbamazepine (132), Phenytoin (128), Acetaminophen (106)
Clustering Coefficient	0.000
Network Density	0.012
Characteristic Path Length	3.231

This is better shown in the bar plot below (Figure 9)

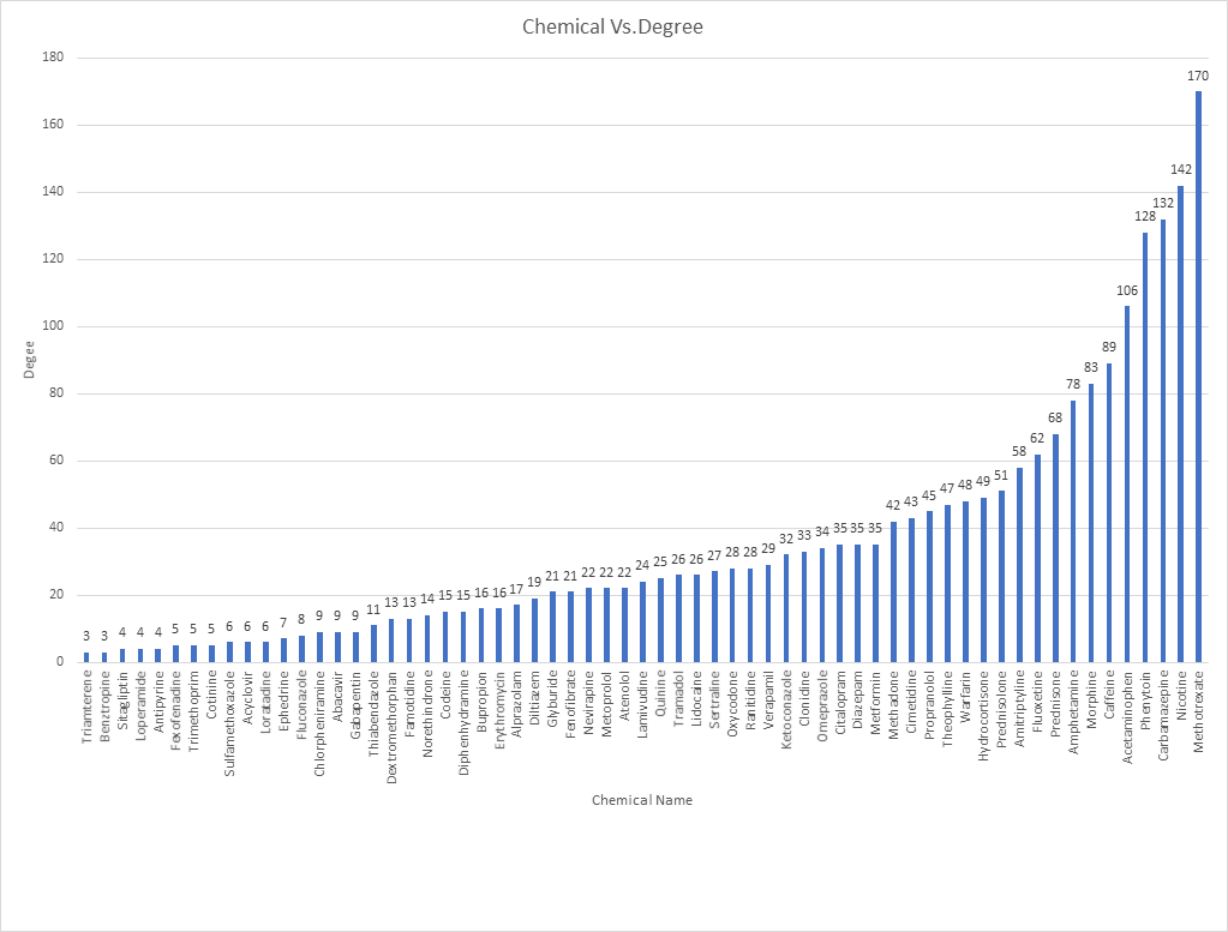


Figure 9: Bar plot showing the distribution of degree amongst the analyzed chemicals.

Additionally, for the 63 chemicals, their respective clustering coefficient or the edge density among neighboring nodes would be 0. This is because the network is unidirectional, meaning interactions go from chemical to disease and there are no interactions between diseases. Figure 10 provides insight into the relationship between the different chemicals and their respective betweenness centrality, which is a measure of centrality based on shortest path lengths. Particularly, betweenness centrality “describes a walker's movement from one node to another via the shortest path; therefore, nodes with a large number of visits by the walker shows high betweenness centrality” and nodes with a small number of visits show a low betweenness centrality (Iida 2018).

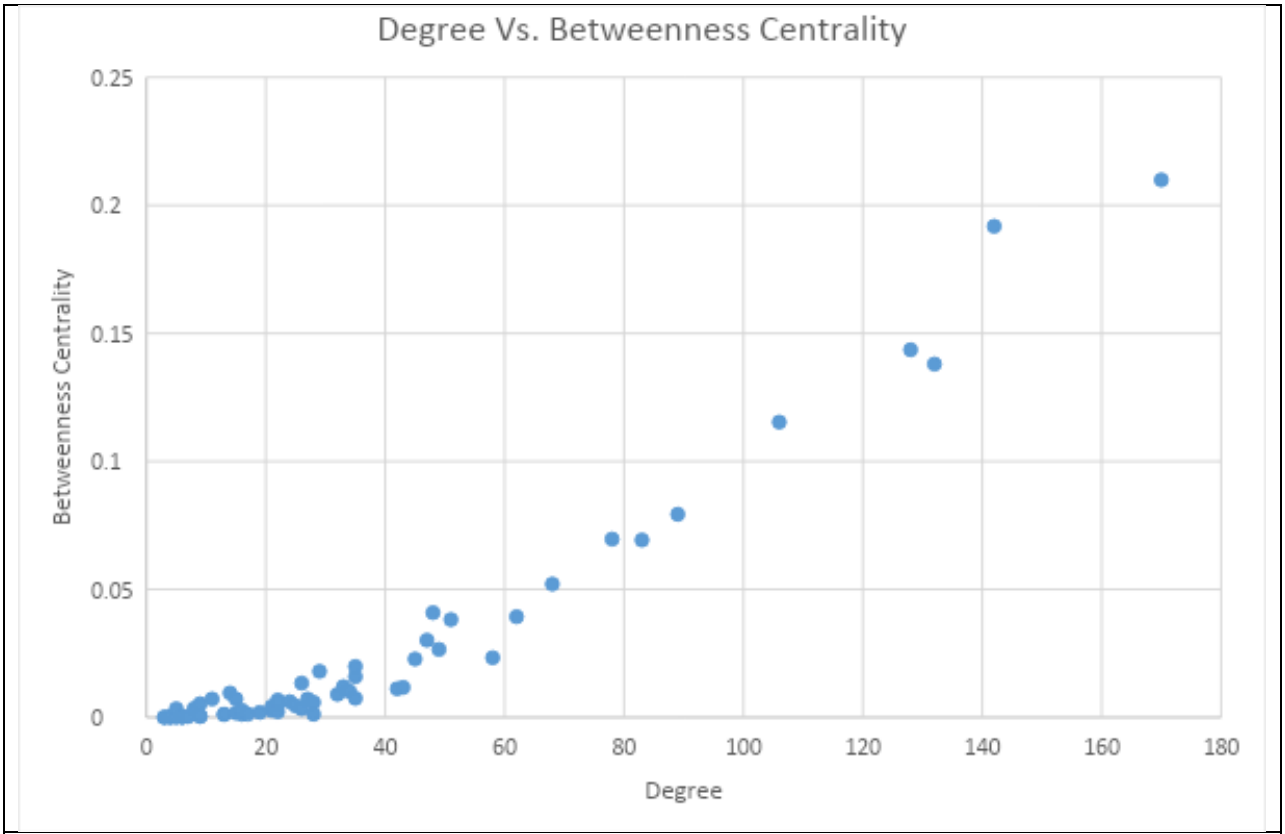


Figure 10: Degree vs. Betweenness Centrality

CONCLUSION

A recent study published by the USGS, earlier this year, analyzed the contamination levels of commonly prescribed pharmaceuticals in the waterways throughout the USA. Starting with the 88 pharmaceuticals reported in that study, we developed a chemical-disease interaction map of 63 out of 88 pharmaceuticals in that report. This approach helped us to identify the harmful effects of 63 water-contaminating pharmaceuticals. The 63 chemicals were linked to 562 unique diseases. The diseases can be broadly classified into 18 classes. Of these cardiovascular and nervous system diseases were most prevalent. The 63 pharmaceuticals are distributed amongst 28 categories. The USGS study had identified Nicotine, Metformin, Caffeine, Cotinine and Lidocaine to be present in the highest concentrations in the waterways. These compounds collectively are linked to many cardiovascular, nervous system, urogenital and mental disorder diseases. Our analysis indicate that the top 5 harmful contaminants are Methotrexate (linked to 170 diseases), Nicotine (linked to 142 diseases), Carbamazepine (linked to 132 diseases), Phenytoin (linked to 128 diseases), Acetaminophen (linked to 106 diseases). Analysis of the cardiovascular sub-network shows that, carbamazepine (linked to 31 CV diseases), nicotine (linked to 30 CV diseases), amphetamine (linked to 25 CV diseases), caffeine to 22 and methotrexate to 20 CV diseases respectively. For the cancer sub-network, the worst offenders are Nicotine, linked to 17 types of cancer, and phenytoin linked to 12 types of cancer. It is hard to pinpoint the exact source of such pharmaceutical contamination of our waterways, but it would not be far off to state that we as consumers are the largest contributors to such pollution, starting with dumping unused and expired medications in toilets. Moreover, our bodies

only metabolize a fraction of the doses, and the remainder are excreted in the wastewaters. Drug manufacturing and agriculture are other sources of this environmental contamination. It is possible that there may be a cumulative effect on humans from such exposures. As protectors of our planet we can take steps to reduce such contamination, by using drug take-back programs, and taking measured steps to get rid of unused medicines. In summary, our studies highlight the long term impact of the contaminating water-borne pharmaceuticals on human health & point to the need for contamination mitigating approaches to improve human and aquatic health.

REFERENCES

1. Bradley, P.M., Journey, C.A., Button, D.T., Carlisle, D.M., Huffman, B.J., Qi, S.L., Romanok, K.M., Van Metre, P.C. (2020) Multi-region assessment of pharmaceutical exposures and predicted effects in USA wadeable urban-gradient streams. *PLoS One* 15(1), e0228214. doi: 10.1371/journal.pone.0228214.

2. National Research Council (US) Safe Drinking Water Committee. Drinking Water and Health: Volume 4. Washington (DC): National Academies Press (US); 1982. VI, Toxicity of Selected Inorganic Contaminants in Drinking Water. Available from: https://www.ncbi.nlm.nih.gov/books/NBK216600/

3.Tchounwou, P.B., Yedjou, C.G., Patlolla, A.K.,Sutton, D.J., 2012. Molecular, clinical and environmental toxicology . 101, 1-30. http://dx.doi.org/10.1007/978-3-7643-8340-4.

4.Hulsen, T., Jamuar, S. S., Moody, A. R., Karnes, J. H., Varga, O., Hedensted, S., Spreafico, R., Hafler, D. A., & McKinney, E. F. (2019). From Big Data to Precision Medicine. Frontiers in medicine, 6, 34. https://doi.org/10.3389/fmed.2019.00034

5. King BL, Davis AP, Rosenstein MC, Wiegers TC, Mattingly CJ (2012) Ranking Transitive Chemical-Disease Inferences Using Local Network Topology in the Comparative Toxicogenomics Database. PLoS ONE 7(11): e46524. https://doi.org/10.1371/journal.pone.0046524

6. Briggs D. (2003). Environmental pollution and the global burden of disease. *British medical bulletin*, *68*, 1–24. https://doi.org/10.1093/bmb/ldg019

7. Su, G., Morris, J. H., Demchak, B., & Bader, G. D. (2014). Biological network exploration with Cytoscape 3. *Current protocols in bioinformatics*, *47*, 8.13.1–8.13.24. https://doi.org/10.1002/0471250953.bi0813s47

8. Davis, A.P., Grondin, C.J., Johnson, R.J., Sciaky, D., Wiegers, J., Wiegers, T.C., Mattingly, C.J. (2020). The Comparative Toxicogenomics Database: Update 2021. Nucleic Acids Res. 2020

9. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T. *Cytoscape: a software environment for integrated models of biomolecular interaction networks.* Genome Res, 13:11 (2498-504). 2003 Nov. PubMed ID: 14597658

10. Cregg, R. Russo,, G., Gubbay, A., Branford, R., Sato, H. (2013) Pharmacogenetics of Analgesic Drugs. Br J Pain 7(4):189-208. doi: 10.1177/2049463713507439.

11. Cazacu, I., Mogosan, A., Loghin, F. (2015) Safety issues of current analgesics: an update. Clujul Med. 88(2): 128–136. doi: 10.15386/cjmed-413

12. Hatzis,G. Karatolios, K. (2016). Cardiovascular Diseases: Genetic Susceptibility, Environmental Factors and their Interaction. Pages 119-136. doi: 10.1016/B978-0-12-803312-8.00006-9

13. Drugs in Water. (2011) Harvard Health Letter. https://www.health.harvard.edu/newsletter_article/drugs-in-the-water

14. Boffetta P, Nyberg, F. (2003). Contribution of environmental factors to cancer risk. Br Med Bull 68:71-94. doi: 10.1093/bmp/ldg023.

15. Lewandowska A.M., Rudzki, M., Rudzki, S. Tomasz Lewandowski, T., Laskowska, B. (2019). Environmental risk factors for cancer - review paper. Ann Agric Environ Med. 22;26(1):1-7. doi: 10.26444/aaem/94299

16. Yee, S.W., Gong, L., Badagnani, I., Giacomini, K.M., Klein, T.E., Altman. R.B. (2011). SLC19A1 Pharmacogenomics Summary. Pharmacogenet Genomics 20(11): 708–715. doi: 10.1097/FPC.0b013e32833eca92

17. Kishi, S., Cheng, C., French, D.., Pei D., Das, S., Cook, E.H., Hijiya, N., Rizzari, C., Rosner, G.L., Frudakis, T., Pui, C.H., Evans, W.E., Relling, M.V. (2007). Ancestry and pharmacogenetics of antileukemic drug toxicity. Blood.109:4151–4157. doi: 10.1182/blood-2006-10-054528

18. Drugs.com https://www.drugs.com/

19. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C, Wilson M. (2017). DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 46(D1):D1074-D1082 doi: 10.1093/nar/gkx1037

20. Uniprot. https://www.uniprot.org

21. Mishra, A., Chaturvedi, P., Datta, S., Sinukumar, S., Joshi, P., Garg, A. (2015). Harmful effects of nicotine. Indian J Med Paediatr Oncol. 36(1): 24–31.

22. Jones, K.L., Lacro, R.V., Johnson, K.A., Adams, J. (1989). Pattern of malformations in the children of women treated with carbamazepine during pregnancy. N Engl J Med (25):1661-6. doi: 10.1056/NEJM198906223202505.

23. Mintzer, S., Skidmore, C.T., Abidin, C.J., Morales, M.C., Chervoneva, I., Capuzzi, D.M., Sperling, M.R. (2009). Effects of antiepileptic drugs on lipids, homocysteine, and C-reactive protein**.** Ann Neurol 65(4):448-56. doi: 10.1002/ana.21615.

24. Hitti, M. (2004). Steroids Linked to Higher Heart Disease Risk. WebMD https://www.webmd.com/asthma/news/20041115/steroids-linked-to-higher-heart-disease-ris

25. Sholter, D.E., Armstrong, P.W. (2000). Adverse effects of corticosteroids on the cardiovascular system. Can J Cardiol 16(4):505-11.

26. Heaney, R.M. (1983). Left bundle branch block associated with propoxyphene hydrochloride poisoning. Ann Emerg Med 12, 780-2. DOI: 10.1016/s0196-0644(83)80259-5

27. Chan, A.T., Manson, J.E., Albert, C.M., Chae, C.U., Kathryn M Rexrode, K.M., Gary C Curhan, G.C., Rimm, E.B., Willett, W.C., Fuchs, C.S. (2006). Nonsteroidal antiinflammatory

drugs, acetaminophen, and the risk of cardiovascular event**.** s Circulation (12):1578-87. doi: 10.1161/CIRCULATIONAHA.105.595793

28. Hulisz, D., Lagzdins, M. (2008). Drug-Induced Hypertension. US Pharmacist. 33(9):HS11-HS20.

29. Sudhir, K., Jennings, G.L., Esler, M.D., Korner, P.I., Blombery, P.A., Lambert, G.W., Scoggins, B., Whitworth, J.A. (1989). Hydrocortisone-induced hypertension in humans: pressor responsiveness and sympathetic function. Hypertension. 13(5):416-21. doi: 10.1161/01.hyp.13.5.416.

30. Buning, J.W., Faassen, M., Brummelman, P., Dullaart, R.P.F., Berg, G., Klauw, M.M., Kerstens , M.N., Stegeman , C.A., Muller Kobold, A.C., Kema, I.P., Bruce, H.R. Wolffenbuttel, André P van Beek (2016). Effects of Hydrocortisone on the Regulation of Blood Pressure: Results From a Randomized Controlled Trial J Clin Endocrinol Metab 101(10):3691-3699. DOI: 10.1210/jc.2016-2216

31. Durelli, L., Mutani, R., Sechi, G.P., Monaco, F., Glorioso, N., Gusmaroli, G. (1985). Cardiac side effects of phenytoin and carbamazepine. A dose-related phenomenon? Arch Neurol 42(11):1067-8. DOI: 10.1001/archneur.1985.04060100049020

32. Zukkor, S. (2015). The Safety of Stimulant Medication Use in Cardiovascular and Arrhythmia Patients. American College of Cardiology https://www.acc.org/latest-in-cardiology/articles/2015/04/28/10/06/the-safety-of-stimulant-medication-use-in-cardiovascular-and-arrhythmia-patients

33. Grando, S.A. (2014). Connections of nicotine to cance**r.** Nat Rev Cancer ;14(6):419-29. doi: 10.1038/nrc3725.

34. Singh, G., Driever, P.H., Sander, J.W. (2005). Cancer risk in people with epilepsy: the role of antiepileptic drugs**.** Brain 128(Pt 1):7-17. doi: 10.1093/brain/awh363.

35. Iida, M., and K. Takemoto. (2018). A Network Biology-Based Approach to Evaluating the Effect of Environmental Contaminants on Human Interactome and Diseases. *Ecotoxicology and Environmental Safety*, 160:316-327. doi: 10.1016/j.ecoenv.2018.05.065.

36. Temple, Jennifer L et al. “The Safety of Ingested Caffeine: A Comprehensive Review.” *Frontiers in psychiatry* vol. 8 80. 26 May. 2017, doi:10.3389/fpsyt.2017.00080

37. Paik, Jin Hui et al. “Symptomatic bradycardia due to nicotine intoxication.” “Bradicardia sintomática causada por intoxicação por nicotina.” *Revista Brasileira de terapia intensiva* vol. 30,1 (2018): 121-126. doi:10.5935/0103-507x.20180018

38. Oparil, Suzanne et al. “Hypertension.” *Nature reviews. Disease primers* vol. 4 18014. 22 Mar. 2018, doi:10.1038/nrdp.2018.14

39. Singh, Shikha et al. “Prevalence and Associated Risk Factors of Hypertension: A Cross-Sectional Study in Urban Varanasi.” *International journal of hypertension* vol. 2017 (2017): 5491838. doi:10.1155/2017/5491838

40. Adelöw, C et al. “Epilepsy as a risk factor for cancer.” *Journal of neurology, neurosurgery, and psychiatry* vol. 77,6 (2006): 784-6. doi:10.1136/jnnp.2005.083931

41. Xavier Mariette, Dominique Cazals-Hatem, Josiane Warszawki, Frédéric Liote, Nathalie Balandraud, Jean Sibilia,

Lymphomas in rheumatoid arthritis patients treated with methotrexate: a 3-year prospective study in France,

Blood,

Volume 99, Issue 11,

2002,

Pages 3909-3915,

ISSN 0006-4971,

https://doi.org/10.1182/blood.V99.11.3909.

(https://www.sciencedirect.com/science/article/pii/S0006497120607601)