

RESEARCH ARTICLE

Data-driven analysis of biomedical literature suggests broad-spectrum benefits of culinary herbs and spices

N. K. Rakhi^{1,2}*, Rudraksh Tuwani¹*, Jagriti Mukherjee¹, Ganesh Bagler^{1*}

1 Center for Computational Biology, Indraprastha Institute of Information Technology (IIIT-Delhi), New Delhi, India, **2** Department of Bioscience and Bioengineering, Indian Institute of Technology Jodhpur, Jodhpur, India

* These authors contributed equally to this work.

* ganesh.bagler@gmail.com, bagler@iiitd.ac.in



OPEN ACCESS

Citation: Rakhi NK, Tuwani R, Mukherjee J, Bagler G (2018) Data-driven analysis of biomedical literature suggests broad-spectrum benefits of culinary herbs and spices. PLoS ONE 13(5): e0198030. <https://doi.org/10.1371/journal.pone.0198030>

Editor: Mahesh Narayan, The University of Texas at El Paso, UNITED STATES

Received: January 22, 2018

Accepted: May 11, 2018

Published: May 29, 2018

Copyright: © 2018 Rakhi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by a senior research fellowship from the Ministry of Human Resource Development, Government of India and Indian Institute of Technology Jodhpur to NKR. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Spices and herbs are key dietary ingredients used across cultures worldwide. Beyond their use as flavoring and coloring agents, the popularity of these aromatic plant products in culinary preparations has been attributed to their antimicrobial properties. Last few decades have witnessed an exponential growth of biomedical literature investigating the impact of spices and herbs on health, presenting an opportunity to mine for patterns from empirical evidence. Systematic investigation of empirical evidence to enumerate the health consequences of culinary herbs and spices can provide valuable insights into their therapeutic utility. We implemented a text mining protocol to assess the health impact of spices by assimilating, both, their positive and negative effects. We conclude that spices show broad-spectrum benevolence across a range of disease categories in contrast to negative effects that are comparatively narrow-spectrum. We also implement a strategy for disease-specific culinary recommendations of spices based on their therapeutic tradeoff against adverse effects. Further by integrating spice-phytochemical-disease associations, we identify bioactive spice phytochemicals potentially involved in their therapeutic effects. Our study provides a systems perspective on health effects of culinary spices and herbs with applications for dietary recommendations as well as identification of phytochemicals potentially involved in underlying molecular mechanisms.

Introduction

Culinary practices across cultures around the world have evolved to incorporate spices and herbs in them. The potential utility of these aromatic plant products in recipes has received a lot of attention leading to multiple rationales for their wide-spread use in food preparations [1,2]. Apart from their use as flavoring agents, spices have been suggested to be of value for their ability to inhibit or kill food-spoilage microorganisms [2]. Beyond their antimicrobial properties, the diverse therapeutic values of spices have been highlighted through *in vivo* and *in vitro* studies. Spices have been reported to possess therapeutic potential for their hypolipidemic [3], anti-diabetic [4], anti-lithogenic [5], antioxidant [6], anti-inflammatory and anti-carcinogenic [7] activity.

Competing interests: The authors have declared that no competing interests exist.

Scientific investigations into the health effects of spices have resulted in a large body of biomedical literature mentioning their direct or indirect connections to health and diseases. With focus on a specific spice/herb, such studies have discussed their health consequences to report heterogeneous results. While some of the surveys have attempted to collate and summarize this knowledge [3,6,8], a comprehensive picture of health impacts of culinary herbs and spices based on empirical evidence still evades us. Data from MEDLINE suggests an exponential increase in scientific reports associating culinary spices and herbs with diseases since 1990's. Given their importance in food preparations, it is imperative to systematically investigate these empirical data to investigate health consequences of culinary herbs and spices.

Beyond their culinary use, traditional medicinal systems have also advocated the role of spices as therapeutic agents [8,9]. Apart from obtaining a coherent picture of the impact of these exceptional culinary ingredients on health, it would also be of value to probe the molecular mechanisms behind their action which remain largely unknown. A framework that integrates data on spice-disease associations and their phytochemicals to explore their underlying connections will help unravel molecular mechanisms behind the health impact of culinary spices and herbs (Fig 1). Towards this end, we set out to find associations between spices and diseases from biomedical abstracts available from MEDLINE using a text mining approach.

One of the earliest attempts in linking diet and diseases from literature was by Swanson who suggested the utility of dietary fish oil for the treatment of Reynaud's syndrome from indirect associations manually inferred from literature survey [10]. Biomedical literature has expanded by many folds since this pilot study making it impossible to manually concatenate the information available from research articles to infer relationships between different entities or to formulate a hypothesis. Computational approaches to text mining and natural language processing are potent tools in this pursuit [11] and many studies in recent years have contributed to efforts in this direction [12–15]. NutriChem [15,16] database relates plant-based foods, their phytochemicals, and diseases by using a text mining approach. HerDing [17] is another resource which links herbs to diseases by indirectly connecting constituent chemicals of the former to genes associated with the latter.

We investigated the impact of culinary spices and herbs for their role as regulators of health by text mining biomedical literature to assimilate, both, positive and negative associations. We observed that in general, the benevolent effects of spices span a broader spectrum of disorders than their adverse effects. Thus by exhaustively integrating evidence for beneficial and harmful effects of spices, we provide a framework for identification of spices whose benefits far outweigh their harms. We also suggest ways for their informed culinary use as well as for identification of phytochemicals with potential therapeutic value. In summary, our study offers a systems perspective of health effects of spices and herbs to provide informed culinary recommendations and insights into underlying molecular mechanisms behind their therapeutic utility.

Results

Protocol for integration of spice-phytochemical-disease data

We text mined spice-disease associations from abstracts available in MEDLINE, the largest database of biomedical literature containing more than 28 million references to research articles in biomedicine. First, a comprehensive dictionary of 188 species of culinary spices and herbs was manually compiled from various sources such as FooDB (<http://foodb.ca>), Wikipedia (https://en.wikipedia.org/wiki/List_of_culinary_herbs_and_spices), PFAF (Plants For A Future, <http://www.pfaf.org/user/Default.aspx>), FPI (Food Plants International, <http://foodplantsinternational.com>) and FlavorDB [18] (<http://cosylab.iitd.edu.in/flavordb>). This dictionary was then used to retrieve relevant abstracts from MEDLINE database. We then

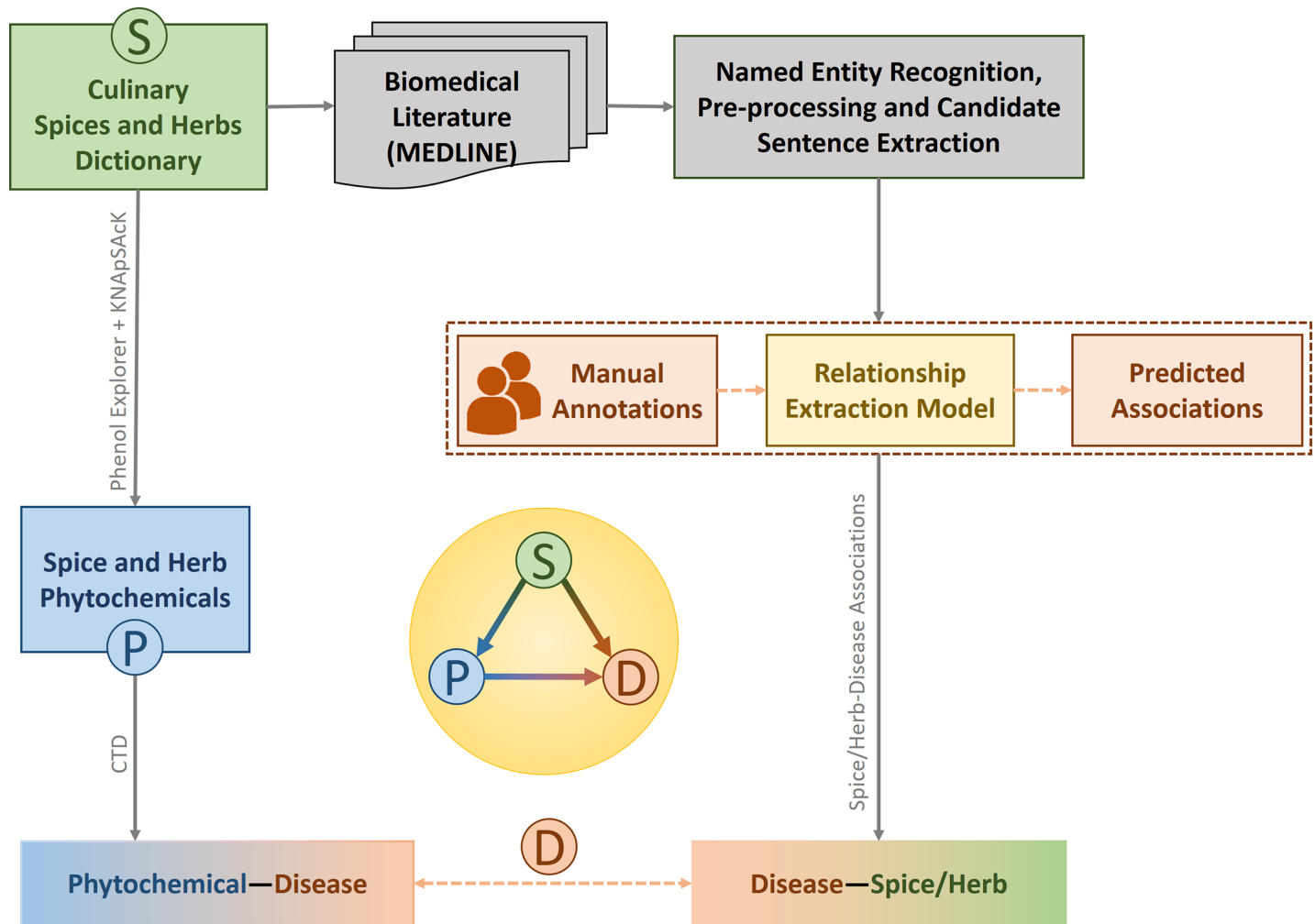


Fig 1. Workflow implemented for data-driven analysis of biomedical literature associating culinary spices and herbs to diseases. Starting with compilation of an exhaustive dictionary of culinary spices and herbs, towards identification of spice-disease associations, one thread of investigation involved implementation of a computational protocol for text mining of biomedical literature including named entity recognition of herbs/spices as well as diseases, pre-processing, extraction of candidate sentences, manual annotations followed by predictions of associations with a machine learning based model. The other thread involved identification of bioactive spice phytochemicals and linking them to diseases. By integrating tripartite information of spices-phytochemicals-diseases, this study establishes the broad-spectrum benevolence of spices, suggests ways for their disease-specific culinary recommendations and probes potential molecular mechanisms underlying their therapeutic properties. Thus it provides a systems perspective to health effects of spices with potential culinary and medicinal applications.

<https://doi.org/10.1371/journal.pone.0198030.g001>

carried out Named Entity Recognition (NER) and normalization of spice and disease entities using a dictionary matching approach for the former and NCBI's TaggerOne [19] tool the latter. For extracting relations, we only considered sentences that mention at least one spice/herb and disease and manually labeled a subset of these for positive, negative and neutral associations between the spice-disease pairs. The labeled sentences were then used to train a machine learning classifier to categorize the associations between the spice-disease pairs in the unlabeled sentences. To further probe putative molecular mechanisms for benevolent effects of spices, we identified spice phytochemicals from PhenolExplorer [20] and KNApSack [21] and found their therapeutic associations with diseases using Comparative Toxicogenomic Database [22] (CTD). Fig 1 depicts the computational framework implemented for integrating and extracting tripartite spice-phytochemical-disease associations. These information are made available through an interactive resource, SpiceRx [23].

Spices disease associations

By combining manually annotated and predicted associations, we obtained a total of 8957 spice-disease associations from 5769 abstracts. Among these 8172 were positive spice-disease associations and 783 were negative. Out of 188 spices present in the dictionary, we obtained associations for 152 spices linking them to 848 unique disease-specific MeSH [24] (Medical Subject Headings) IDs (S1 Dataset). We used a Convolutional Neural Network (CNN) classifier with word, position, part of speech and chunk embedding [25–27] to predict positive, negative or neutral association in a spice-disease pair. It was evaluated on an external test set and found to have an accuracy of 86.7% and macro-averaged precision, recall and F1 score of 90.7%, 80% and 84.2% respectively. The class-wise performance metrics for the model are provided in Table 1.

Disease entities were recognized and normalized to their corresponding MeSH IDs using TaggerOne [19]. MeSH [24] is a controlled vocabulary of biomedical terms curated and developed by National Library of Medicine. It organizes terms hierarchically from general to more specific (S1 Fig). In this hierarchical structure, a spice may have associations with a disease at multiple levels of specificity. For example, Endocrine System Diseases (C19) present at the first level of MeSH hierarchy constitutes disease sub-categories such as Adrenal Gland Diseases (C19.053), Diabetes Mellitus (C19.246) at the second level. Further, specific types of Diabetes Mellitus such as ‘Diabetes Mellitus, Type 1 (C19.246.267)’, ‘Diabetes Mellitus, Type 2 (C19.246.300)’ appear at the third level. To conduct a multi-level analysis, we associated spices with disease terms at three levels of MeSH hierarchy labeled as ‘category’, ‘sub-category’ and a ‘disease’.

We observed an exponential increase in articles reporting therapeutic properties of spices after 1995 (Fig 2), with the number of abstracts reporting positive associations of spices with diseases far out-numbering those reporting negative associations (Fig 3). A large number of spices such as ginger (*Zingiber officinale*) and turmeric (*Curcuma longa*) have very few negative associations reported in MEDLINE whereas a few others like liquorice (*Glycyrrhiza glabra*) and celery (*Apium graveolens*), have almost an equal number of abstracts reporting positive and negative associations. The complete list of associations for spices is provided in S2 Dataset. These data suggest that, in general, beneficial effects of spices have been reported more widely than their adverse effects in biomedical literature.

On analyzing individual diseases (third level of MeSH hierarchy) associated with spices, we found that diabetes mellitus, inflammation, and carcinogenesis have the highest number of positive associations (Fig 4) (S3 Dataset). Spices were also shown to have a preventive role in various cancers including breast, colorectal, prostatic and liver neoplasms. Among the diseases adversely affected by spices were hypersensitivity, dermatitis, rhinitis, hypertension and allergic rhinitis, (Fig 5) (S3 Dataset). It is worth noting that majority of these diseases are autoimmune in nature and are subjective to certain individuals sensitive to that spice. In such cases, spices may act as triggering factors rather than causal agents.

Table 1. The class-wise performance metrics for the best CNN model, implementing word, position, part of speech and chunk embedding features, used for spice-disease relationship extraction. All negative associations were cleaned manually.

Class	Precision	Recall	F1-Score
No-Association	88.06%	89.83%	88.9%
Negative	1*	65.96%	79.49%
Positive	83.98%	84.32%	84.15%
Macro averaged	90.68%	80.03%	84.20%

*cleaned manually.

<https://doi.org/10.1371/journal.pone.0198030.t001>

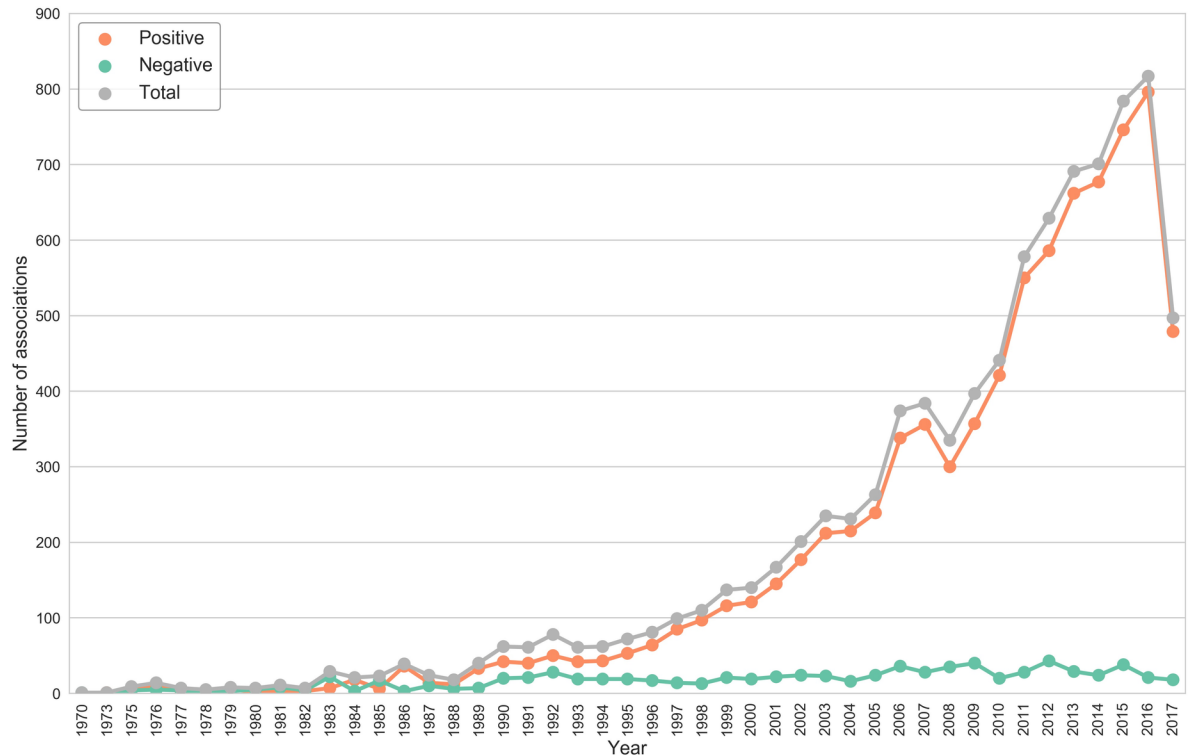


Fig 2. Statistics of spice-disease associations. Historical trend in biomedical literature reporting spice-disease associations. There is an exponential increase in articles reporting the therapeutic effects of spices in last few decades. Data of research articles archived in MEDLINE till July 2017 is represented in the illustration.

<https://doi.org/10.1371/journal.pone.0198030.g002>

Broad-spectrum benevolence of herbs/spices

To probe for the effects of spice/herb across a spectrum of disorders, we analyzed its associations with disease ‘sub-categories’ at the second level of MeSH hierarchy (S1 Fig). Analyzing associations at this level provides a balance between specificity and generality of disease terms. Among the disease sub-categories positively associated with spices, pathologic processes, signs and symptoms, metabolic diseases, diabetes mellitus, vascular diseases as well as central nervous system diseases were found to be dominant (Fig 6), S4 Dataset). Top disease categories which were negatively associated with spices included vascular diseases, skin diseases, hypersensitivity and respiratory hypersensitivity (Fig 7), S4 Dataset).

To quantify the broad impact a spice may have across diverse disease categories as well as sub-categories, we devised a ‘spectrum score (Ω_s)’. This metric computes the sum of proportion of disease terms associated with a spice at the second level of MeSH hierarchy (sub-categories), multiplied by the number of disease terms associated at the first level (categories). (See Materials and Methods). With 27 disease categories, the lower and upper bound for the spectrum score is 0 and 729 respectively. To elucidate further, let us consider a spice that is associated with all diseases in exactly half of the MeSH disease categories versus another spice that has associations with half of the diseases in every disease category. In such a case, the latter would have a higher spectrum score than the former. We computed the spectrum score for both positive (benevolence spectrum score, Ω_s^+) as well as negative associations (adverse spectrum score, Ω_s^-).

The spices with highest ‘benevolence spectrum score’ according to our analysis were garlic (*Allium sativum*), ginger (*Zingiber officinale*), turmeric (*Curcuma longa*), liquorice (*Glycyrrhiza*

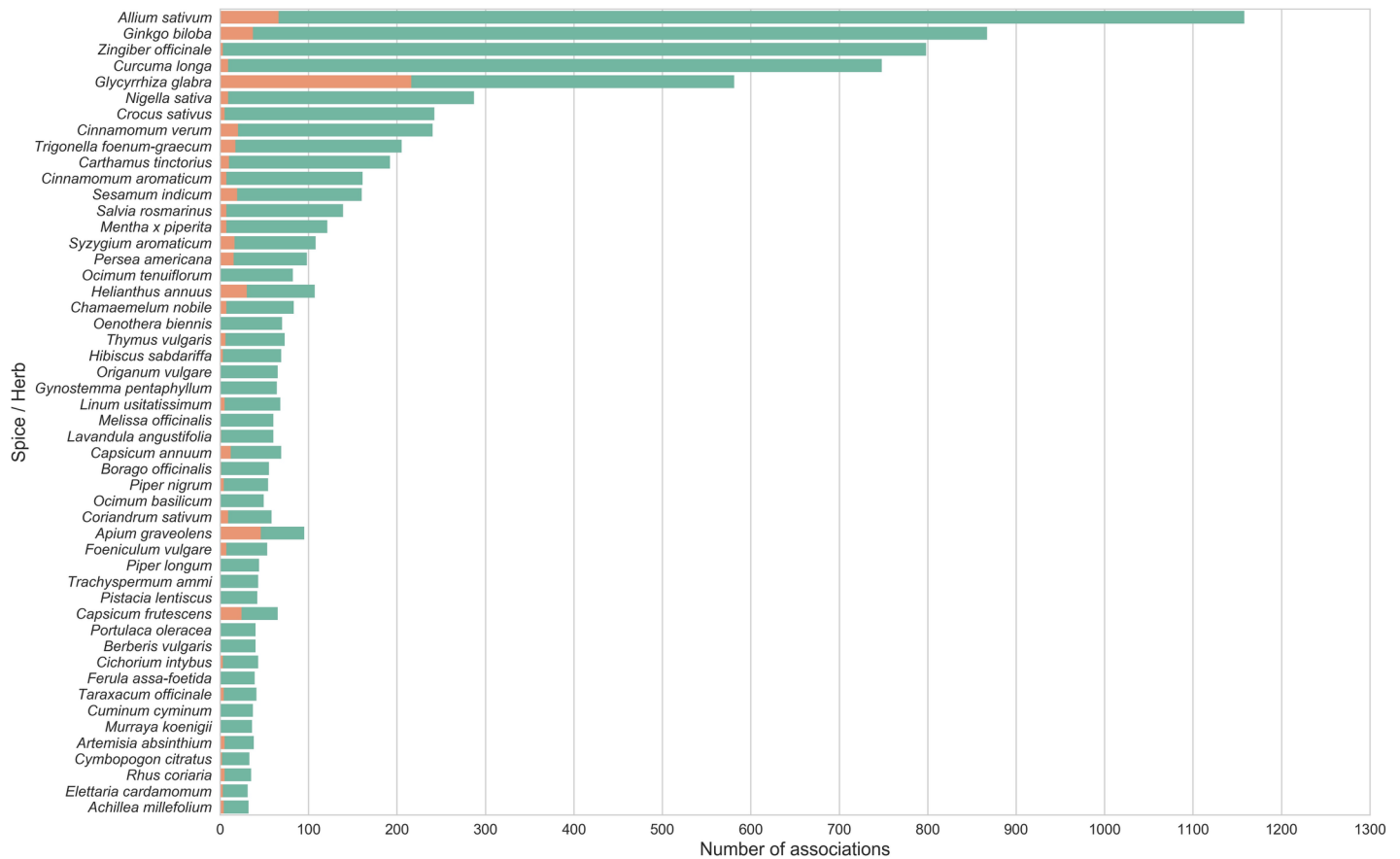


Fig 3. Statistics of positive and negative disease associations for the top 50 spices with most number of associations. Notice that certain spices like liquorice (*Glycyrrhiza glabra*) and celery (*Apium graveolens*) had equal number of positive as well as negative associations. The bias in number of associations may also indicate the inherent biases in scientific literature suggesting that certain spices are studied more than others.

<https://doi.org/10.1371/journal.pone.0198030.g003>

glabra), ginkgo (*Ginkgo biloba*), black cumin (*Nigella sativa*), cinnamon (*Cinnamomum verum*) and saffron (*Crocus sativus*) whereas the top adverse spectrum spices were liquorice (*Glycyrrhiza glabra*), ginger (*Zingiber officinale*), fenugreek (*Trigonella foenum-graecum*), ginkgo (*Ginkgo biloba*), sunflower (*Helianthus annuus*) and Celery (*Apium graveolens*). Spices such as garlic, liquorice and ginkgo have a high benevolence as well as adverse spectrum scores.

We found that for 150 out of 152 spices, the ‘benevolence spectrum score’ exceeded the ‘adverse spectrum score’, with almost 50 spices having ‘relative benevolence’ ($\Delta\Omega_r$) greater than 50 (Fig 8). Hence, it may be concluded that in general spices have positive effects with a broad spectrum of diseases in contrast to their negative effects which are comparatively narrow-spectrum. In line with our analysis, spices have been reported to be effective against a range of disorders [3–5,7]. Details of benevolent, adverse as well as relative benevolence scores for all spices are provided in the S5 Dataset.

Culinary recommendations

Each of the MeSH disease categories refers to a class of disorders such as nutritional and metabolic disorders, cardiovascular diseases, nervous systems diseases, digestive system diseases, immune system diseases, neoplasms, bacterial infections and mycoses, virus diseases and such.

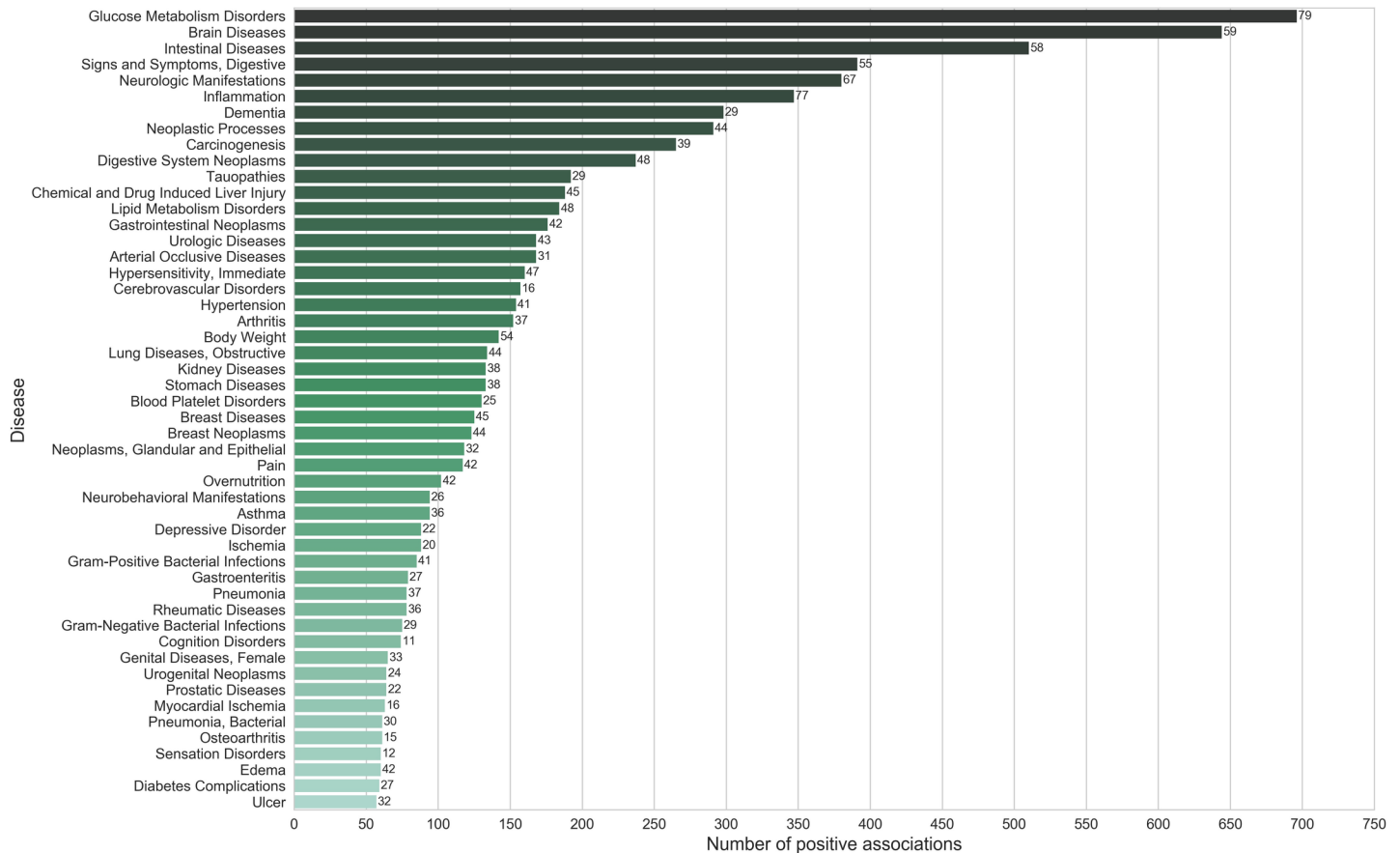


Fig 4. Top diseases (Third level of MeSH hierarchy) ranked according to their total number of positive associations. Numbers shown against the bars indicate the ‘number of spices’ involved in the associations. The number of positive disease associations for spices outnumber the number of negative associations (Fig 5) indicating that spices, in general, have been reported with beneficial health effects.

<https://doi.org/10.1371/journal.pone.0198030.g004>

The spectrum score forms the basis to prioritize spices for culinary intervention against a MeSH disease category. We computed the category-specific ‘benevolence spectrum’ and ‘adverse spectrum’ scores to enumerate the ‘trade-off score’ that represents the therapeutic value of a spice against a class of disorders. [S6 Dataset](#) provides a list of culinary recommendations intended as a dietary intervention against various disease categories.

There is ample amount of empirical evidence for the recommendations provided by our study. Our data suggest that spices show therapeutic effects against most of the viral diseases. Among them, turmeric (*Curcuma longa*) is the most broad-spectrum antiviral spice and is reported with inhibitory properties against various viruses including HIV, influenza, and coxsackievirus [28]. Studies in human and animal models have shown that dietary spices significantly stimulate the activities of digestive enzymes of the pancreas and small intestines such as pancreatic lipase, amylase and proteases thereby acting as digestive stimulants. Spices like ginger and garlic stimulate TRPV1, a sensor in the digestive system which has implications for gastrointestinal tract pathology and physiology [29,30]. Prominent spices recommended for cardiovascular diseases, such as tulsi (*Ocimum tenuiflorum*), mint (*Mentha X piperita*), ginkgo (*Ginkgo biloba*) and ginger (*Zingiber officinale*), have been reported with beneficial effects against cardiovascular disorders. Epidemiological studies suggest that these spices lower cholesterol level, decrease platelet aggregation, reduce blood pressure, and increases antioxidant status which in turn decreases the progression of cardiovascular diseases [31]. Black cumin

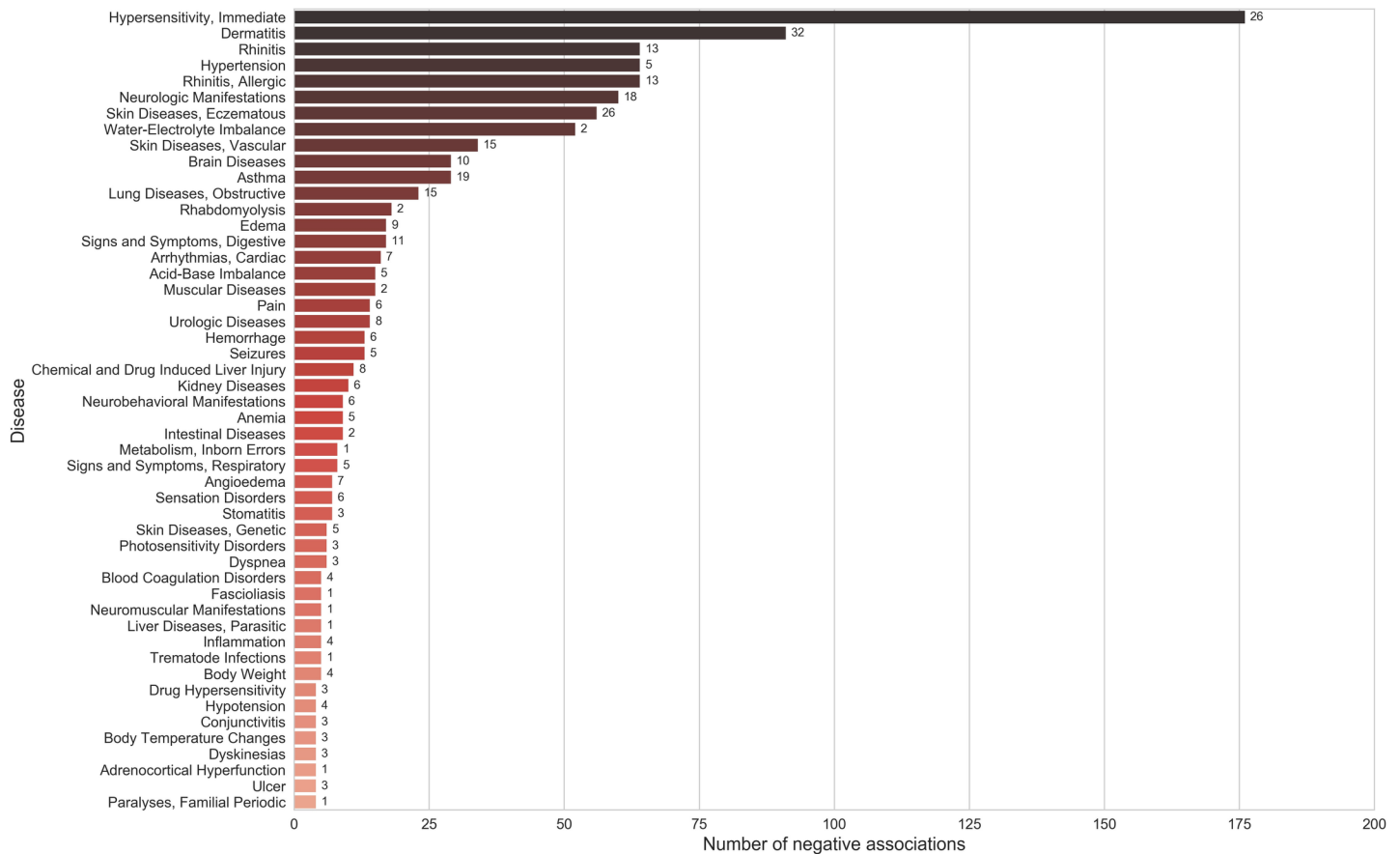


Fig 5. Top diseases (Third level of MeSH hierarchy) ranked according to their total number of negative associations. The numbers mentioned on the bars indicate the number of spices associated negatively with each disease.

<https://doi.org/10.1371/journal.pone.0198030.g005>

(*Nigella sativa*), turmeric (*Curcuma longa*) and garlic (*Allium sativum*) are prominent spices recommended for diabetes, a major metabolic disorder. Evidence from animal studies and human trials have indicated that these spices modulate hyperglycemia and lipid profile function. Their antioxidant characteristics and effects on insulin secretion, glucose absorption, and gluconeogenesis make them potent candidates towards treating diabetes [32,33]. Similarly, the anti-diabetic property of ginkgo (*Ginkgo Biloba*) may be linked to the ability of its extract to reduce insulin resistance.

Incidentally, the spices that frequent in the culinary recommendations are among those used for culinary and medicinal preparations across cultures. Curcumin (*Curcuma longa*) and tulsi (*Ocimum tenuiflorum*), widely used in Indian culinary and medicinal preparations, were present across recommendations made throughout the spectrum of MeSH disease categories. Similarly garlic, used in Southern European especially Italian cuisine, also appeared in culinary recommendations across all categories of diseases. Some of the other most potent spices include ginger (*Zingiber officinale*), black cumin (*Nigella sativa*) and ginkgo (*Ginkgo biloba*) (see S1 Table).

Linking spices to diseases through phytochemicals

Our analysis suggests that beyond their utility as flavoring, coloring, and food preserving (anti-microbial [2]) agents, spices may have been incorporated in traditional culinary practices due

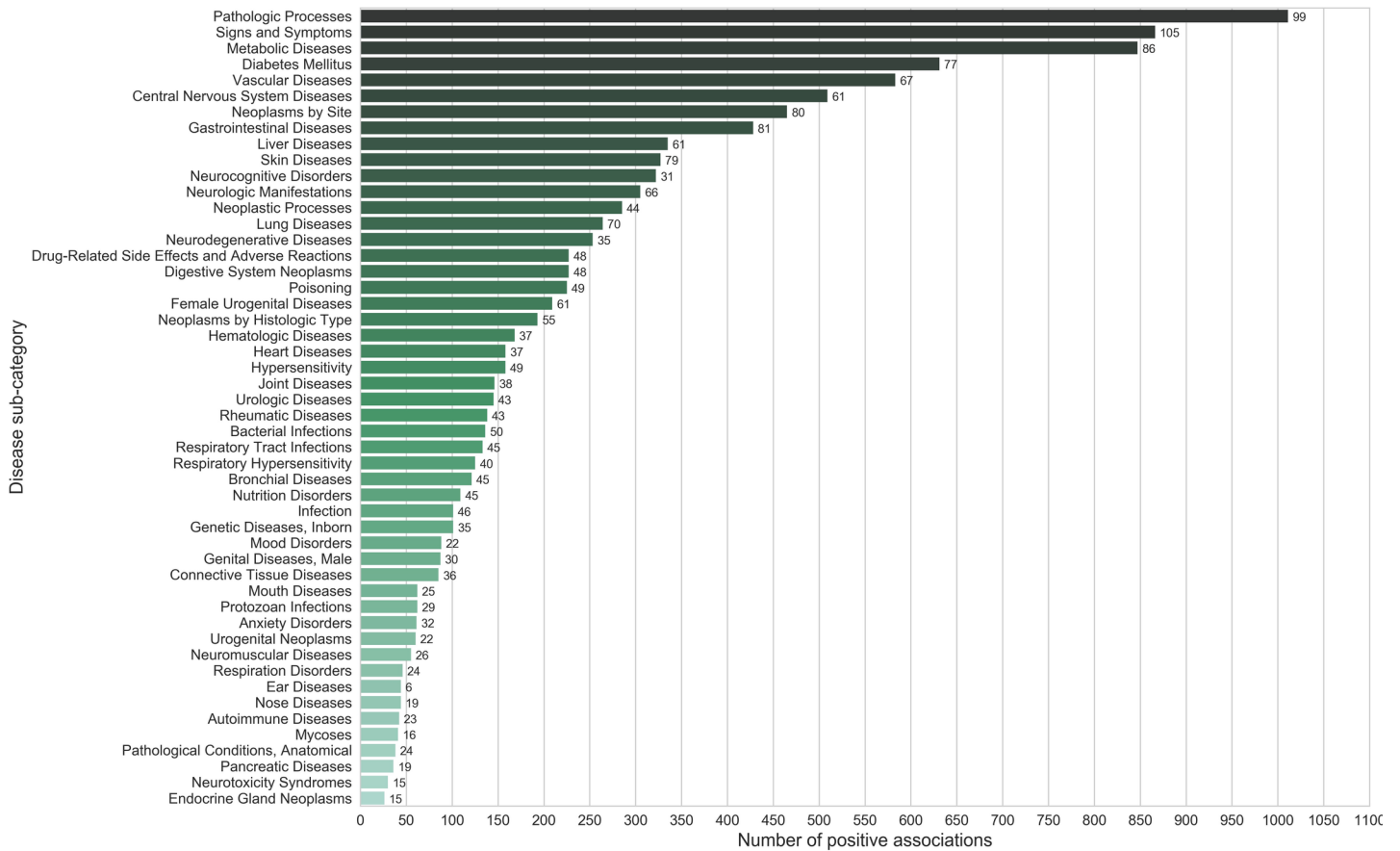


Fig 6. Disease categories (First level of MeSH hierarchy) ranked according to the number of positive associations with spices. Numbers shown against the bars indicate the ‘number of spices’ linked with each of the associations. The number of positive disease category associations for spices outnumber those with negative associations (Fig 7) further confirming the benevolent health effects of spices.

<https://doi.org/10.1371/journal.pone.0198030.g006>

to their beneficial health effects across a spectrum of disorders. Given that the therapeutic properties of plants are mediated by their phytochemicals [34–36] we hypothesize that the broad spectrum benevolence of spices can be attributed to the presence of bioactive phytochemicals such as polyphenols [36]. For example, curcumin, a polyphenol from turmeric is known to have a wide range of health benefits including antioxidant, anti-inflammatory, and anticancer effects [37]. Ajoene, a polyphenol compound derived from garlic, has been shown to induce apoptosis in leukemic cells [38]. Similarly, eugenol present in clove is reported to have antifungal property [39]. The antioxidant activity of black pepper has been attributed to the presence of β -caryophyllene, limonene, β -pinene, piperine and piperolein in its essential oil and oleoresins [39]. The anticancer properties of ginger are attributed to the presence of certain pungent vallinoids, gingerol, and paradol, as well as some other constituents like shogaols, zingerone, amongst others [39]. Going beyond the investigation of spice-disease associations, we linked spices to their constituent bioactive molecules and further connected them to diseases to obtain potential evidence of therapeutic associations (Fig 1).

We obtained 866 chemical compounds corresponding to 142 culinary spices in our dictionary from PhenolExplorer [20] and KNApSACk [21], and consisted of 2042 spice-phytochemical associations. These data were filtered using PubChem [40] to keep only 570 bioactive phytochemicals, as they are known to react with tissues or cells. Further, we associated spice phytochemicals to diseases with the help of CTD [22], a public database of curated and

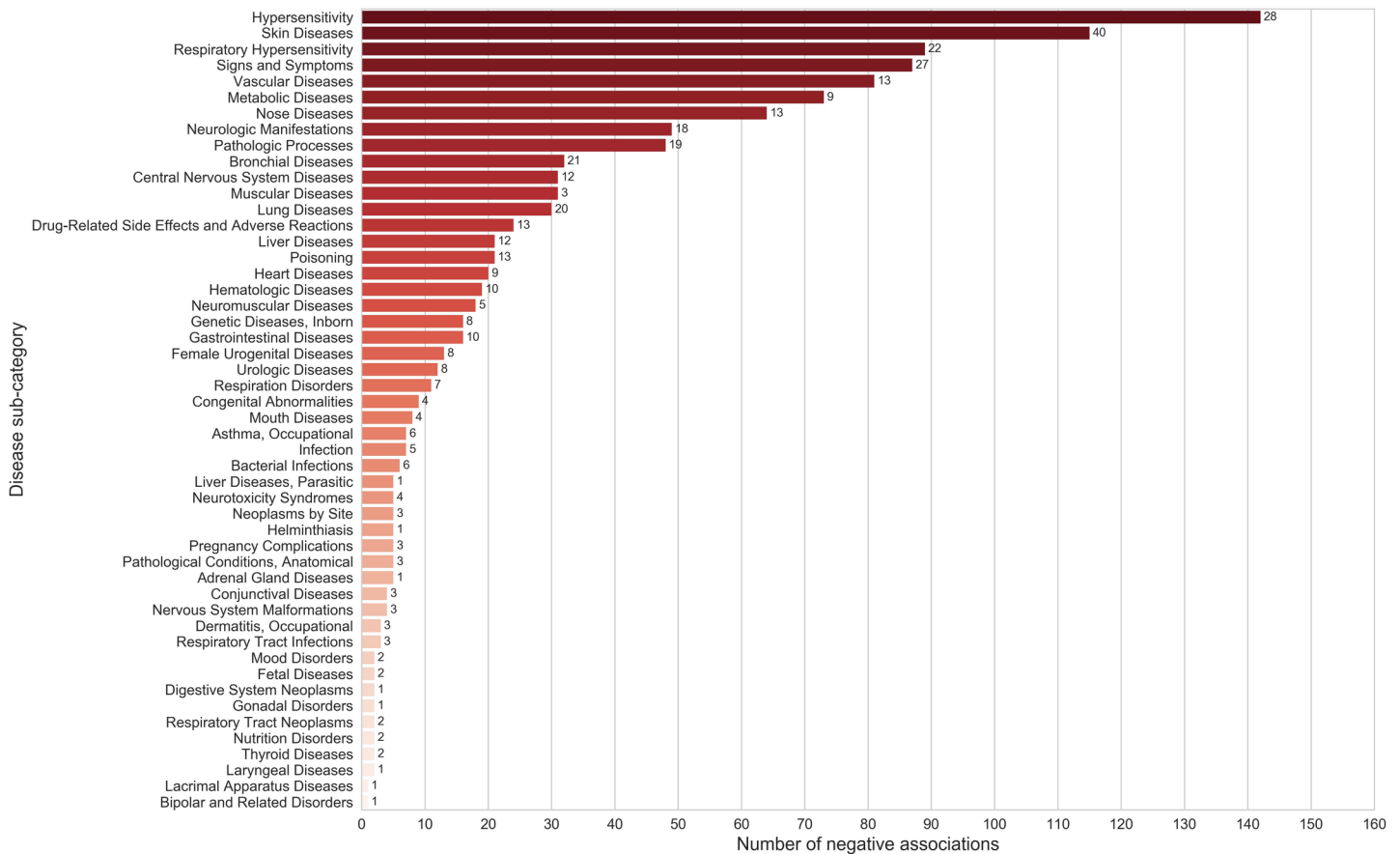


Fig 7. Disease categories (First level of MeSH hierarchy) ranked according to the number of negative associations with spices. Numbers shown against the bars indicate the ‘number of spices’ linked with each of the associations.

<https://doi.org/10.1371/journal.pone.0198030.g007>

inferred chemical-disease associations from the literature. CTD [22] classifies chemical-disease associations into therapeutic, inferred or marker associations. Therapeutic and marker associations are directly curated from the literature, whereas inferred relations are obtained from indirect associations. Therapeutic associations between a phytochemical and disease imply the presence of direct evidence of that phytochemical in alleviating the disease. For our further analysis, we focused only on 211 bioactive chemicals from the spices which were reported to have therapeutic associations.

We integrated the data of spice-disease associations with spice-phytochemical and phytochemical-disease mappings. This tripartite data of spice-phytochemical-disease associations can form the basis for finding putative molecular mechanisms behind the beneficial effects of spices against diseases. Using data of curated phytochemical-disease associations from CTD, we found that out of 4380 positive spice-disease associations (where disease terms were mapped to third level of MeSH), 37% (1619) could be explained through evidence of phytochemical-disease associations. To elucidate, we found empirical evidences supporting anti-carcinogenic effects of garlic (*Allium sativum*) against liver neoplasms. With the help of CTD [22], we found allyl sulfide, a compound in garlic, to be therapeutically associated with liver neoplasms. It can therefore be hypothesized that the anti-carcinogenic effects of garlic can be attributed to the presence of allyl sulfide. Incidentally, this hypothesis is independently supported by the literature [41]. The 63% spice-disease associations which could not be explained

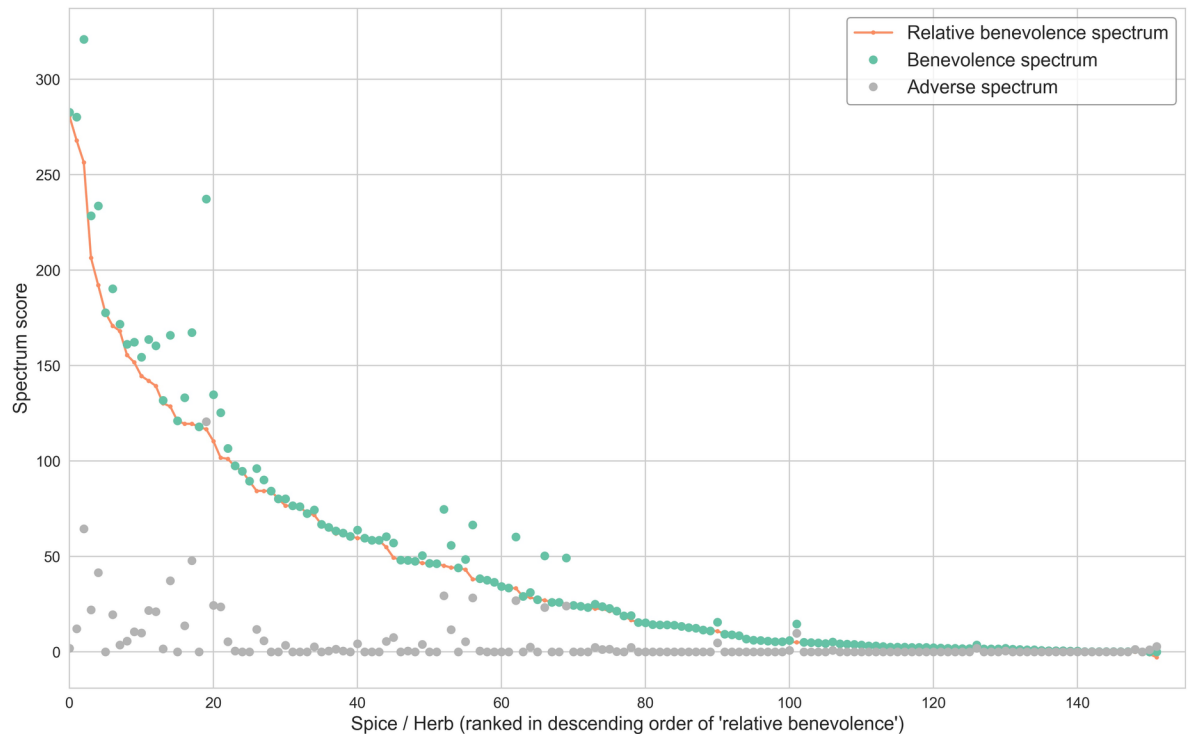


Fig 8. Spices ranked according to their ‘relative benevolence score’ highlighting their broad-spectrum benevolence. This score enumerates the relative health benefits as reflected in the difference between ‘benevolence spectrum’ and ‘adverse spectrum’ scores. Barring two, all spices had positive scores with a large number of them showing significantly larger therapeutic effects compared to their adverse effects.

<https://doi.org/10.1371/journal.pone.0198030.g008>

through evidence of phytochemical-disease relations may serve as hypotheses for unearthing the putative molecular mechanisms by utilizing the data of spice-phytochemical associations.

[S7 Dataset](#) provides an exhaustive list of spice-disease associations and phytochemicals identified from the integration of tripartite data of diseases, spices and their phytochemicals ([Fig 1](#)) and [S8 Dataset](#) provides the list of positive spice-disease associations for which no specific therapeutic phytochemical from a spice could be obtained.

Discussion

Humans are unique in having developed the ability to cook, which has been argued to be critical for the emergence of their large brains [42,43]. While cooked food must have provided with much-needed energy supply, it is intriguing that they flavor the food with nutritionally insignificant quantities of herbs and spices. Going beyond the ability of spices to act as flavoring and antimicrobial agents [2], our analysis of spice-disease associations text-mined from biomedical literature shows the broad-spectrum benefits of spices. Recent studies have shown the potential benefits of consumption of spices such as chillies through cohort studies [44] as well as the role of specific spice phytochemicals in their health effects [45]. Interestingly, the broad-spectrum benevolence score of a spice was not positively correlated with its phytochemical repertoire ([S2 Fig](#)) suggesting that richness in the phytochemical content itself does not explain its therapeutic value.

We also point out negative health effects of spices, largely reflected in allergies, immune system, and skin-related disorders. Few of the negative effects of spices have been linked with their excessive use. For example, licorice, a beneficial herb for hypertension can cause weight

loss, hypokalemia and other related adverse effects if consumed in large doses. Beyond probing the molecular basis of positive associations, it would also be of interest to identify toxic phytochemicals present in spices and assess their effect on specific diseases so as to provide an advisory against their consumption. Negative associations for spices projected by our study can serve as a basis for such investigations.

As opposed to a previous attempt in this direction [15,16] that linked all plant-based foods with diseases and phytochemicals from literature, our study focused on culinary spices and herbs. We investigated an exhaustive dictionary of 188 culinary herbs and spices with far better coverage (99 additional) than that of NutriChem [15,16]. Overall, in terms of the number of disease associations, the depth of our analysis was better than that of NutriChem [15,16] (S3 Fig) and our data comprised a larger set of associations for most spices (S4 Fig). NutriChem [15,16] used dictionary based string matching approach for named entity recognition and normalization of diseases as well as plants. In case of diseases, it is empirically shown that depending on the disease dictionary used, the string matching approach typically leads to a low precision and recall [46]. We used TaggerOne [19], a machine learning based named entity recognition tool which yields state of the art performance. Even though the performance of our relationship extraction model was evaluated on a dataset consisting of positive, negative and neutral associations in contrast to previous studies which evaluated on only positive and negative associations, our model achieves a comparative F1 score. In addition to this, we provide an accurate information of adverse effect of spices by manually correcting all predicted negative associations. Despite our best efforts to ensure accurate extraction of spice-disease associations, our method is constrained by shortcomings inherent to text mining approaches and use of limited information pertaining to biomedical literature, namely, title and abstract. Overall, our analysis serves as a precursor to systematic reviews including meta-analysis as well as hypothesis-driven investigations into the health effects of spices and herbs. The data compiled as part of our study are made available through an interactive resource, SpiceRx [23].

Similar to languages where words are synthesized from the same phonetic repertoire, cuisines around the world have concocted their own unique ingredient combinations, especially those made from spices [47,48]. Interestingly, many cuisines around the world such as those from the Indian subcontinent (*paanch phoron*, *garam masala*, *sambar masala* among a host of others referred to as *masala*), Ethiopia (*berbere*) and Middle East (*baharat*) to mention a few, have ended up developing unique spice combinations of their own. It remains to be critically examined whether these have been deliberately composed with an appreciation of therapeutic properties of spices and herbs, or are accidentally emerged constructs. Spices are frequently used as part of functional foods, for example, the Indian dish *rasam* is a concoction of different spices and has been reported to be hypoglycemic, anti-anemic and antipyretic [49]. *Sambar*, another predominantly spice-based recipe has been shown to work against prostate cancer [50]. Traditional medicinal systems are also known to recommend spices as part of their prescriptions. *Trikatu* [51], a spice concoction made with black pepper, long pepper, and dried ginger has been advised to be of value against rheumatoid arthritis by Ayurveda, a classical traditional medicinal system from India. In Chinese traditional medicine, *Xiaoyao-san*, a combination of various spices, has been recommended for management of stress and depression-related disorders [52].

Cooking typically involves high-temperature processing via heating, boiling, frying and such. It could be argued [53] that heating is a simpler and more effective means of killing microbes, thereby refuting the antimicrobial hypothesis [2]. Other beneficial effects of spices (such as anti-diabetic, anti-carcinogenic and antioxidant and inflammatory), unearthed in this study, could not be argued against with this logic. Ironically, this argument raises another critical question: Whether the therapeutic properties and bioactivity of spice phytochemicals can sustain the intense heating processes typically involved in cooking [54]? Besides that, one of

the ambiguous factors in appreciating the benevolence of spices is the distinction between the effectiveness of individual compounds vis-à-vis their synergistic actions. Apart from these aspects, there is ample scope for improvising the strategy for culinary recommendations as well as for identifications of molecular mechanisms involved in health impact of spices by including the data of quantity and disease-specific potency of their constituent phytochemicals. While raising a host of such critical questions related to dietary intake of herbs and spices, by investigating evidence from biomedical literature reporting health effects of culinary herbs and spices our data-driven analysis suggests their broad-spectrum benevolence.

Materials and methods

Compilation of spices and herbs dictionary

We compiled a dictionary of 188 species of culinary spices and herbs. Scientific names and common names were obtained from Foodb (<http://foodb.ca/>) and Wikipedia (https://en.wikipedia.org/wiki/List_of_culinary_herbs_and_spices). Varieties in scientific names, wherever available, were standardized to their respective species name. For example, *Capsicum baccatum* var. *pendulum*, the scientific name of Peruvian pepper, was standardized to *Capsicum baccatum*. All scientific names were then mapped to their respective NCBI Taxonomy IDs. This dictionary was further enriched by adding common names from FPI (Food Plants International, <http://foodplantsinternational.com/plants/>), NCBI Taxonomy (<https://www.ncbi.nlm.nih.gov/taxonomy>) and PFAF (Plants for a Future, <http://www.pfaf.org>). Singular and plural forms of common names of the spices and herbs were also included. Common names that did not exclusively map to an NCBI Taxonomy ID were removed.

Biomedical literature

We used MEDLINE (Medical Literature Analysis and Retrieval System Online, <https://www.nlm.nih.gov/bsd/mms/medlineelements.html>) as our source of biomedical literature. It includes citations from more than 5600 scholarly journals with over 24 million references to peer-reviewed biomedical and life science research articles from as early as 1946. The data was downloaded in bulk from the FTP server of NCBI (https://www.nlm.nih.gov/databases/download/pubmed_medline.html). A modified version of PubMed parser (https://github.com/titipata/pubmed_parser) was used to extract information of PMID, Date, Title, Abstract, Journal, and Authors from the XML files. Articles for which no abstract text was available were not considered. The modified parser is available at https://github.com/cosylabiiit/pubmed_parser.

Named entity recognition

We adopted a dictionary matching approach for Named Entity Recognition (NER) of spices and herbs. With a large dictionary, the process of dictionary matching becomes a computational bottleneck. Therefore, we used a modified implementation of Aho-Corasick algorithm (NoAho, <https://github.com/JDonner/NoAho>) to efficiently obtain non-overlapping and longest matches at the token level. For disease NER (DNER) and normalization, we used TaggerOne [19] which utilizes semi-Markov models with a rich feature set. It was reported to have a precision of 85% and a recall of 80% on the Biocreative V Chemical Disease Relation test set [46]. We used the pre-trained disease-only model available with TaggerOne [19] on our data.

Preprocessing

Sentence segmentation was carried out on the retrieved abstracts using Stanford CoreNLP package [55]. Only sentences with mention of at least one herb/spice and one disease were

considered for extracting relations. Those with mentions of multiple herbs/spices and/or diseases were simplified by duplicating the sentence while iteratively masking all except a specific spice-disease pair. In all sentences, numbers were replaced by a standard identifier token and, barring some punctuation characters (!,;,), all special characters were removed. The preprocessed sentences were then tokenized using GENIA [56] and the part-of-speech (PoS) tag, as well as the chunk tag of each token were obtained. Further, we also computed the distance of each token from the candidate spice-disease pair and used them as position features.

Labelling associations

Hitherto, to the best of our knowledge, no labeled corpus for associations between plant-based foods and diseases is publicly available. We thus manually annotated a total of 6712 spice-disease pairs to tag positive, negative and neutral associations. Out of all the annotated pairs, 2669 had positive associations, 301 had negative associations, and 3742 had neutral or no associations. This data was used for training as well as evaluating our relationship extraction models.

Relation extraction model

We developed a Machine Learning Classifier to categorize tagged spice-disease pair(s) in a sentence as having positive, negative or neutral associations. The following models were tested: (i) Linear Support Vector Machine (SVM) with unigram and bigram word features; (ii) Convolutional Neural Network (CNN) with word embedding [27] features; and (iii) CNN with word, position, PoS and chunk embedding features.

For the Linear SVM model, we obtained the unigram and bigram word features and scaled their respective weights using Term Frequency-Inverse Document Frequency (TF-IDF) approach. This model was trained using one-versus-all strategy. Following are the equations describing the method for computing TF-IDF weights of features: (i) $tf(t,s) = f_{t,s}$; (ii) $idf(t) = \log \frac{N}{n_t}$; and (iii) $tfidf(t,s) = tf(t,s) * idf(t)$, where $f_{t,s}$ denotes the number of times feature t appears in sentence s , n_t is the number of sentences in which the feature t appears and N is the total number of sentences.

The architecture for our CNN models is based on state-of-the-art models for sentence classification and relation extraction [25–27]. As input, we fed mini-batches of sentence sequences to the models. The two CNNs differ in the representation of the tokens or words present in input sequences. For the first model, we only used the word embedding as the token representation, whereas for the second model we used the PoS, chunk and position embedding in addition to word embedding. The word embedding was initialized using pre-trained weights from Chiu *et. al* [57], with the embedding for unknown words initialized from a uniform $(-\alpha, \alpha)$ distribution. The parameter α was determined on the basis of the variance of the known words [58]. Further, CNN requires all input sequences to have consistent size, thus sentences were zero-padded to equalize their lengths to that of the longest sentence. The input to the CNNs was a $b \times d \times n \times 1$ tensor, where b is the size of the mini-batch, d is the length of the ‘token’ vector of the sentence and n is the length of the longest sentence in the corpus. The architecture of the second CNN is depicted in Fig 9. The first layer is a Convolutional layer with n_f filters of different filter sizes f and rectified linear unit (ReLU) activation. The respective maximum activations of all the filters are then concatenated into a single vector of size n_f and fed to a Dropout layer [59], which randomly sets an activation to zero with probability p . This is followed by a dense layer of h hidden units with ReLU activation and a softmax layer with 3 units. For both the networks, we used categorical cross entropy as our objective function and applied l_2 regularization of 3 on the dense layers only. The networks were trained using mini-batch gradient descent with shuffled batches of size 50 and Adam [60] optimizer. We adopted

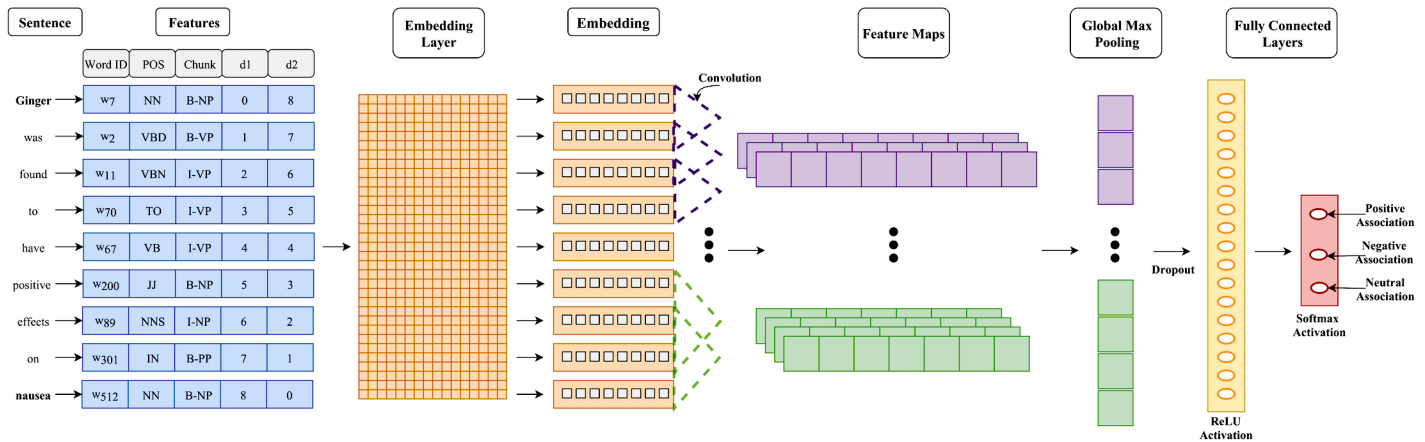


Fig 9. Architecture of the Convolutional Neural Network. Illustration of the convolutional neural network model utilizing word, position, part of speech and chunk embeddings.

<https://doi.org/10.1371/journal.pone.0198030.g009>

an early stopping criterion for the training process and stopped model training if the validation loss did not decrease for 5 epochs. To address the class imbalance problem, we over-sampled the negative class and the positive class by a factor of 12 and 1.35 respectively. The hyper-parameters of both the neural networks were determined using 5-fold cross validation and are available in [S2 Table](#). The code as well as the data used for the CNNs is available at the Complex System Laboratory, IIIT-Delhi’s GitHub page: <https://github.com/cosylabiit/spice-disease-associations>.

Evaluation metrics

We evaluated the performance of our model based on its precision, recall, F1 score and accuracy: $Precision = TP / (TP + FP)$; $Recall = TP / (TP + FN)$; $F1 - score = 2 \cdot Precision \cdot Recall / (Precision + Recall)$; $Accuracy = TP + TN / (TP + TN + FP + FN)$, where TP , FP , TN , FN are True Positives, False Positives, True Negatives and False Negatives respectively.

MeSH hierarchy

MeSH is a controlled vocabulary of biomedical terms curated and developed by National Library of Medicine. The terms are hierarchically organized from generic to more specific. The DNER tool used in this study (TaggerOne [19]) normalizes the tagged entities to MeSH IDs. The hierarchical structure of MeSH results in situations where a spice is typically associated with a disease at multiple levels of specificity. For example, in the first level of MeSH hierarchy a spice may be linked with the disease category Endocrine System Diseases (C19) and at the second level C19 may be associated with sub-categories such as Adrenal Gland Diseases (C19.053) or Diabetes Mellitus (C19.246). Further, it may be linked to the specific type of Diabetes Mellitus, say, ‘Diabetes Mellitus, Type 1 (C19.246.267)’ or ‘Diabetes Mellitus, Type 2 (C19.246.300)’ appearing at the third level. We conducted a multi-level analysis by associating spices with disease terms at top three levels of MeSH hierarchy which were referred to as category, sub-category, and a disease ([S1 Fig](#)).

Adverse and benevolent spectrum scores

The ‘spectrum score of a spice (Ω_s)’ encodes diversity of adverse (Ω_s^-) or therapeutic (Ω_s^+) effects of a spice s across the MeSH disease categories as well as their constituent subcategories,

and is defined as $\Omega_s = \hat{D}_s \cdot \sum_i^D \hat{d}_s^i / d^i$. Here, D is total number of MeSH disease categories, \hat{D}_s represents the number of disease categories with which spice s has therapeutic association with, d^i represents the total number of disease sub-categories in the i th disease category, and \hat{d}_s^i represents the number of disease subcategories in the i th disease category with which the spice s is associated. When calculating the ‘spectrum scores’ across all 27 categories, the ‘adverse spectrum score’ and ‘benevolent spectrum score’ vary between 0 and 729. Further, for each spice the ‘relative benevolence’ ($\Delta\Omega_s = \Omega_s^+ - \Omega_s^-$) that encodes its residual therapeutic benefit was computed.

‘Therapeutic tradeoff score’ for culinary recommendations

Category-specific (benevolence and adverse) spectrum score was defined as $\Omega_s^i = \hat{d}_s^i \cdot \sum_k^{d^i} \hat{\alpha}_s^k / \alpha^k$. Here, \hat{d}_s^i represents the number of disease sub-categories in the i th disease category with which spice s has therapeutic association with, α^k represents the total number of diseases in the k th disease sub-category, and $\hat{\alpha}_s^k$ represents the number of disease subcategories in the k th disease sub-category with which the spice s is associated. The ‘therapeutic tradeoff score’, $\Delta\Omega_s^i$, represents the difference between the ‘benevolence spectrum’ and ‘adverse spectrum’ of spice s for category i ; the higher the tradeoff score of a spice the better is its therapeutic value against the spectrum of diseases represented by this category. Thus, tradeoff score of a spice serves as a basis for its recommendation against a MeSH disease category.

Linking phytochemicals from spices/herbs to diseases

We obtained the phytochemical information for spices/herbs using KNApSACk [21] and CTD [22]. The different compound identifiers were standardized to PubChem IDs and further PubChem BioAssay [40] was used for ascertaining their bioactive status. Therapeutic associations of a compound were obtained from CTD [22] after mapping its PubChem ID to corresponding MeSH ID.

Supporting information

S1 Fig. Hierarchical structure of MeSH disease headers. For the purpose of multi-level analysis, spices were associated with disease terms at three levels of MeSH hierarchy—‘category’, ‘sub-category’ and a ‘disease’.

(TIF)

S2 Fig. Correlation between the number of phytochemicals in spices and their broad-spectrum benevolence. The data indicate that the broad-spectrum benevolence score of spices and their phytochemical repertoire are not correlated.

(TIFF)

S3 Fig. Comparison of the number of associations obtained for spices reported by our study with that of NutriChem [15,16] indicating richer associations in our data.

(TIFF)

S4 Fig. Comparison of associations retrieved for ‘individual spices’ by NutriChem[15,16] to those from our study, suggesting better depth/coverage in the latter.

(TIFF)

S1 Table. Top ten broad spectrum spices and number of MeSH disease categories and sub-categories with which they are positively associated.

(DOCX)

S2 Table. Hyper-parameters selected for the convolutional neural network Model 2 and Model 3.

(DOCX)

S1 Dataset. Statistics of positive and negative spice-disease associations for each spice.

(XLSX)

S2 Dataset. Statistics of positive and negative associations as well as number of spices, at the third level of MeSH.

(XLSX)

S3 Dataset. Statistics of positive and negative associations as well as the number of spices at the third level of MeSH disease hierarchy.

(XLSX)

S4 Dataset. Statistics of positive and negative associations as well as the number of spices at the second level (sub-category) of MeSH disease hierarchy.

(XLSX)

S5 Dataset. Benevolent, adverse as well as relative benevolence scores for all spices.

(XLSX)

S6 Dataset. List of culinary recommendations against various disease categories.

(XLSX)

S7 Dataset. Tripartite associations for a spice and a disease along with specific phytochemicals reported to be involved in the therapeutic action.

(XLSX)

S8 Dataset. Statistics of spice-disease associations for which two specific phytochemicals were ascertained.

(XLSX)

Acknowledgments

We thank the Indraprastha Institute of Information Technology (IIIT-Delhi) for providing computational facilities to GB and RT. This work was supported by a senior research fellowship from the Ministry of Human Resource Development, Government of India and Indian Institute of Technology Jodhpur to NKR. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

Conceptualization: Ganesh Bagler.

Data curation: N. K. Rakhi, Jagriti Mukherjee, Ganesh Bagler.

Formal analysis: N. K. Rakhi, Rudraksh Tuwani, Ganesh Bagler.

Investigation: N. K. Rakhi, Rudraksh Tuwani, Ganesh Bagler.

Methodology: N. K. Rakhi, Rudraksh Tuwani, Ganesh Bagler.

Project administration: Ganesh Bagler.

Resources: Ganesh Bagler.

Supervision: Ganesh Bagler.

Validation: Rudraksh Tuwani, Ganesh Bagler.

Visualization: N. K. Rakhi, Rudraksh Tuwani, Ganesh Bagler.

Writing – original draft: N. K. Rakhi, Rudraksh Tuwani, Ganesh Bagler.

Writing – review & editing: N. K. Rakhi, Rudraksh Tuwani, Ganesh Bagler.

References

1. Sherman PW, Billing J. Darwinian Gastronomy: Spices taste good because they are good for us. *Bioscience*. 1999; 49: 453–463. <https://doi.org/10.2307/1313553>
2. Billing J, Sherman PW. Antimicrobial functions of spices: why some like it hot. *Q Rev Biol*. 1998; 73: 3–49. PMID: [9586227](https://pubmed.ncbi.nlm.nih.gov/9586227/)
3. Srinivasan K. Spices as influencers of body metabolism: An overview of three decades of research. *Food Res Int*. Elsevier; 2005; 38: 77–86. <https://doi.org/10.1016/j.foodres.2004.09.001>
4. Srinivasan K. Plant foods in the management of diabetes mellitus: Spices as beneficial antidiabetic food adjuncts. *Int J Food Sci Nutr*. Taylor & Francis; 2005; 56: 399–414. <https://doi.org/10.1080/09637480500512872> PMID: [16361181](https://pubmed.ncbi.nlm.nih.gov/16361181/)
5. Srinivasan K. Anti-cholelithogenic potential of dietary spices and their bioactives. *Crit Rev Food Sci Nutr*. 2017; 57: 1749–1758. <https://doi.org/10.1080/10408398.2014.1003783> PMID: [26147513](https://pubmed.ncbi.nlm.nih.gov/26147513/)
6. Yashin A, Yashin Y, Xia X, Nemzer B. Antioxidant Activity of Spices and Their Impact on Human Health: A Review. *Antioxidants*. Multidisciplinary Digital Publishing Institute (MDPI); 2017; 6: 70. <https://doi.org/10.3390/antiox6030070> PMID: [28914764](https://pubmed.ncbi.nlm.nih.gov/28914764/)
7. Kaefer CM, Milner JA. The role of herbs and spices in cancer prevention. *Journal of Nutritional Biochemistry*. Elsevier; 2008. pp. 347–361. <https://doi.org/10.1016/j.jnutbio.2007.11.003> PMID: [18499033](https://pubmed.ncbi.nlm.nih.gov/18499033/)
8. Srinivasan K. Role of Spices Beyond Food Flavoring: Nutraceuticals with Multiple Health Effects. *Food Rev Int*. 2005; 21: 167–188. <https://doi.org/10.1081/FRI-200051872>
9. Johri RK, Zutshi U. An Ayurvedic formulation “Trikatu” and its constituents. *Journal of Ethnopharmacology*. Elsevier; 1992. pp. 85–91. [https://doi.org/10.1016/0378-8741\(92\)90067-2](https://doi.org/10.1016/0378-8741(92)90067-2) PMID: [1434692](https://pubmed.ncbi.nlm.nih.gov/1434692/)
10. Swanson D. Fish Oil, Raynaud’s Syndrome, and Undiscovered Public Knowledge. *Perspectives in Biology and Medicine*. 1986. pp. 7–18. PMID: [3797213](https://pubmed.ncbi.nlm.nih.gov/3797213/)
11. Jensen LJ, Saric J, Bork P. Literature mining for the biologist: from information retrieval to biological discovery. *Nat Rev Genet*. Nature Publishing Group; 2006; 7: 119–129. <https://doi.org/10.1038/nrg1768> PMID: [16418747](https://pubmed.ncbi.nlm.nih.gov/16418747/)
12. Lee S, Choi J, Park K, Song M, Lee D. Discovering context-specific relationships from biological literature by using multi-level context terms. *BMC Med Inform Decis Mak*. 2012; 12. <https://doi.org/10.1186/1472-6947-12-S1-S1> PMID: [22595086](https://pubmed.ncbi.nlm.nih.gov/22595086/)
13. Yang H, Swaminathan R, Sharma A, Ketkar V, D’Silva J. Mining biomedical text towards building a quantitative food-disease-gene network. *Stud Comput Intell*. 2011; 375: 205–225. https://doi.org/10.1007/978-3-642-22913-8_10
14. Srinivasan P, Libbus B. Mining MEDLINE for implicit links between dietary substances and diseases. *Bioinformatics*. 2004; 20: 290–296. <https://doi.org/10.1093/bioinformatics/bth914> PMID: [15262811](https://pubmed.ncbi.nlm.nih.gov/15262811/)
15. Jensen K, Panagiotou G, Kouskoumvekaki I. Integrated text mining and chemoinformatics analysis associates diet to health benefit at molecular level. *PLoS Comput Biol*. 2014; 10: e1003432. <https://doi.org/10.1371/journal.pcbi.1003432> PMID: [24453957](https://pubmed.ncbi.nlm.nih.gov/24453957/)
16. Jensen K, Panagiotou G, Kouskoumvekaki I. NutriChem: A systems chemical biology resource to explore the medicinal value of plant-based foods. *Nucleic Acids Res*. 2015; 43: D940–D945. <https://doi.org/10.1093/nar/gku724> PMID: [25106869](https://pubmed.ncbi.nlm.nih.gov/25106869/)
17. Choi W, Choi C-H, Kim YR, Kim S-J, Na C-S, Lee H. HerDing: herb recommendation system to treat diseases using genes and chemicals. *Database*. Oxford University Press; 2016; baw011: 1–7. <https://doi.org/10.1093/database/baw011> PMID: [26980517](https://pubmed.ncbi.nlm.nih.gov/26980517/)
18. Garg N, Sethupathy A, Tuwani R, NK R, Dokania S, Iyer A, et al. FlavorDB: a database of flavor molecules. *Nucleic Acids Res*. Oxford University Press; 2017; 46: D1210–D1216. <https://doi.org/10.1093/nar/gkx957> PMID: [29059383](https://pubmed.ncbi.nlm.nih.gov/29059383/)

19. Leaman R, Lu Z. TaggerOne: Joint named entity recognition and normalization with semi-Markov Models. *Bioinformatics*. 2016; 32: 2839–2846. <https://doi.org/10.1093/bioinformatics/btw343> PMID: [27283952](https://pubmed.ncbi.nlm.nih.gov/27283952/)
20. Neveu V, Perez-Jimenez J, Vos F, Crespy V, du Chaffaut L, Mennen L, et al. Phenol-Explorer: an online comprehensive database on polyphenol contents in foods. *Database*. 2010; 2010: bap024. <https://doi.org/10.1093/database/bap024> PMID: [20428313](https://pubmed.ncbi.nlm.nih.gov/20428313/)
21. Afendi FM, Okada T, Yamazaki M, Hirai-Morita A, Nakamura Y, Nakamura K, et al. KNApSACk family databases: Integrated metabolite-plant species databases for multifaceted plant research. *Plant Cell Physiol*. 2012; 53: e1. <https://doi.org/10.1093/pcp/pcr165> PMID: [22123792](https://pubmed.ncbi.nlm.nih.gov/22123792/)
22. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, King BL, McMorran R, et al. The Comparative Toxicogenomics Database: Update 2017. *Nucleic Acids Res*. 2017; 45: D972–D978. <https://doi.org/10.1093/nar/gkw838> PMID: [27651457](https://pubmed.ncbi.nlm.nih.gov/27651457/)
23. Rakhi N, Tuwani R, Garg N, Mukherjee J, Bagler G. SpiceRx: an integrated resource for the health impacts of culinary spices and herbs. *bioRxiv* 273599. Cold Spring Harbor Laboratory; 2018; 1–24. <https://doi.org/10.1101/273599>
24. Lipscomb CE. Medical Subject Headings (MeSH). *Bull Med Libr Assoc*. 2000; 88: 265–266. PMID: [10928714](https://pubmed.ncbi.nlm.nih.gov/10928714/)
25. Kumar Sahu S, Anand A, Oruganty K, Gattu M. Relation extraction from clinical texts using domain invariant convolutional neural network. *Proc 15th Work Biomed Nat Lang Process*. 2016; 71. <https://doi.org/10.18653/v1/W16-2928>
26. Nguyen TH, Grishman R. Relation Extraction: Perspective from Convolutional Neural Networks. *Work Vector Model NLP*. 2015; 39–48. <https://doi.org/10.3115/v1/W15-1506>
27. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. *arXiv Prepr arXiv13013781*. 2013; 1–12. <https://doi.org/10.1162/153244303322533223>
28. Zorofchian Moghadamtousi S, Abdul Kadir H, Hassandarvish P, Tajik H, Abubakar S, Zandi K. A review on antibacterial, antiviral, and antifungal activity of curcumin. *BioMed Research International*. Hindawi; 2014. p. 186864.
29. Holzer P. Transient receptor potential (TRP) channels as drug targets for diseases of the digestive system. *Pharmacol Ther*. Pergamon; 2011; 131: 142–170. <https://doi.org/10.1016/j.pharmthera.2011.03.006> PMID: [21420431](https://pubmed.ncbi.nlm.nih.gov/21420431/)
30. Platel K, Srinivasan K. Digestive stimulant action of spices: a myth or reality? *Indian J Med Res*. 2004; 119: 167–79. PMID: [15218978](https://pubmed.ncbi.nlm.nih.gov/15218978/)
31. Rahman K, Lowe GM. Garlic and Cardiovascular Disease: A Critical Review. *J Nutr*. American Society for Nutrition; 2006; 136: 736–740.
32. Bi X, Lim J, Jeyakumar Henry C. Spices in the management of diabetes mellitus. *Food Chem*. Elsevier; 2017; 217: 281–293. <https://doi.org/10.1016/j.foodchem.2016.08.111> PMID: [27664636](https://pubmed.ncbi.nlm.nih.gov/27664636/)
33. Heshmati J, Namazi N. Effects of black seed (*Nigella sativa*) on metabolic parameters in diabetes mellitus: A systematic review. *Complement Ther Med*. Churchill Livingstone; 2015; 23: 275–282. <https://doi.org/10.1016/j.ctim.2015.01.013> PMID: [25847566](https://pubmed.ncbi.nlm.nih.gov/25847566/)
34. Visioli F, Borsani L, Galli C. Diet and prevention of coronary heart disease: The potential role of phytochemicals. *Cardiovascular Research*. Oxford University Press; 2000. pp. 419–425. [https://doi.org/10.1016/S0008-6363\(00\)00053-5](https://doi.org/10.1016/S0008-6363(00)00053-5) PMID: [10963715](https://pubmed.ncbi.nlm.nih.gov/10963715/)
35. Mishra BB, Tiwari VK. Natural products: An evolving role in future drug discovery. *Eur J Med Chem*. 2011; 46: 4769–4807. <https://doi.org/10.1016/j.ejmech.2011.07.057> PMID: [21889825](https://pubmed.ncbi.nlm.nih.gov/21889825/)
36. Manach C, Scalbert A, Morand C, Rémésy C, Jiménez L. Polyphenols: Food sources and bioavailability. *American Journal of Clinical Nutrition*. American Society for Nutrition; 2004. pp. 727–747. <https://doi.org/10.1093/ajcn/79.5.727> PMID: [15113710](https://pubmed.ncbi.nlm.nih.gov/15113710/)
37. Kocaadam B, Şanlıer N. Curcumin, an active component of turmeric (*Curcuma longa*), and its effects on health. *Crit Rev Food Sci Nutr*. Taylor & Francis; 2017; 57: 2889–2895. <https://doi.org/10.1080/10408398.2015.1077195> PMID: [26528921](https://pubmed.ncbi.nlm.nih.gov/26528921/)
38. Bayan L, Koulivand PH, Gorji A. Garlic: a review of potential therapeutic effects. *Avicenna J phytomedicine*. Mashhad University of Medical Sciences; 2014; 4: 1–14.
39. Shahidi F, Ambigaipalan P. Phenolics and polyphenolics in foods, beverages and spices: Antioxidant activity and health effects—A review. *J Funct Foods*. Elsevier Ltd; 2015; 18: 820–897. <https://doi.org/10.1016/j.jff.2015.06.018>
40. Wang Y, Bryant SH, Cheng T, Wang J, Gindulyte A, Shoemaker BA, et al. PubChem BioAssay: 2017 update. *Nucleic Acids Res*. Oxford University Press; 2017; 45: D955–D963. <https://doi.org/10.1093/nar/gkw1118> PMID: [27899599](https://pubmed.ncbi.nlm.nih.gov/27899599/)

41. Wu C-C, Chung JG, Tsai S-J, Yang JH, Sheen LY. Differential effects of allyl sulfides from garlic essential oil on cell cycle regulation in human liver tumor cells. *Food Chem Toxicol*. Pergamon; 2004; 42: 1937–1947. <https://doi.org/10.1016/j.fct.2004.07.008> PMID: 15500931
42. Wrangham R. *Catching Fire: How Cooking Made Us Human*. Basic Books; 2009.
43. Fonseca-Azevedo K, Herculano-Houzel S. Metabolic constraint imposes tradeoff between body size and number of brain neurons in human evolution. *Proc Natl Acad Sci U S A*. 2012; 109: 18571–18576. <https://doi.org/10.1073/pnas.1206390109> PMID: 23090991
44. Chopan M, Littenberg B. The association of hot red chili pepper consumption and mortality: A large population-based cohort study. Gualillo O, editor. *PLoS One*. Public Library of Science; 2017; 12: e0169876. <https://doi.org/10.1371/journal.pone.0169876> PMID: 28068423
45. Jiang J, Emont MP, Jun H, Qiao X, Liao J, Kim D il, et al. Cinnamaldehyde induces fat cell-autonomous thermogenesis and metabolic reprogramming. *Metabolism*. W.B. Saunders; 2017; 77: 58–64. <https://doi.org/10.1016/j.metabol.2017.08.006> PMID: 29046261
46. Li J, Sun Y, Johnson RJ, Sciaky D, Wei CH, Leaman R, et al. BioCreative V CDR task corpus: a resource for chemical disease relation extraction. *Database (Oxford)*. Oxford University Press; 2016; 2016: baw608. <https://doi.org/10.1093/database/baw068> PMID: 27161011
47. Jain A, Rakhi NK, Bagler G. Analysis of food pairing in regional cuisines of India. *PLoS One*. 2015; 10. <https://doi.org/10.1371/journal.pone.0139539> PMID: 26430895
48. Jain A, Rakhi NK, Bagler G. Spices form the basis of food pairing in Indian cuisine. arXiv:150203815. 2015; 30.
49. Devarajan A, Mohanmarugaraja MK. *A Comprehensive Review on Rasam: A South Indian Traditional Functional Food*. *Pharmacogn Rev*. Wolters Kluwer—Medknow Publications; 2017; 11: 73–82. https://doi.org/10.4103/phrev.phrev_13_17 PMID: 28989243
50. Prasad V, Reddy N, Francis A, Nayak P, Kishore A, Nandakumar K, et al. Sambar, an Indian dish prevents the development of dimethyl hydrazine-induced colon cancer: A preclinical study. *Pharmacogn Mag*. Wolters Kluwer—Medknow Publications; 2016; 12: 441. <https://doi.org/10.4103/0973-1296.191454> PMID: 27761072
51. Doss HM, Ganesan R, Rasool M. Trikatu, an herbal compound ameliorates rheumatoid arthritis by the suppression of inflammatory immune responses in rats with adjuvant-induced arthritis and on cultured fibroblast like synoviocytes via the inhibition of the NFκB signaling pathway. *Chem Biol Interact*. 2016; 258: 175–186. <https://doi.org/10.1016/j.cbi.2016.09.003> PMID: 27613480
52. Liu C-T, Wu B-Y, Hung Y-C, Wang L-Y, Lee Y-Y, Lin T-K, et al. Decreased risk of dementia in migraine patients with traditional Chinese medicine use: a population-based cohort study. *Oncotarget*. 2017; 8: 79680–79692. <https://doi.org/10.18632/oncotarget.19094> PMID: 29108348
53. McGee H. In victu veritas. *Nature*. 1998; 392: 649–650. <https://doi.org/10.1038/33528> PMID: 9565025
54. Suresh D, Gurudutt KN, Srinivasan K. Degradation of bioactive spice compound: Curcumin during domestic cooking. *Eur Food Res Technol*. 2009; 228: 807–812. <https://doi.org/10.1007/s00217-008-0993-9>
55. Manning CD, Surdeanu M, Bauer J, Finkel J, Bethard SJ, McClosky D. The Stanford CoreNLP Natural Language Processing Toolkit. Association for Computational Linguistics (ACL) System Demonstrations. 2014. pp. 55–60.
56. Kim J-D, Ohta T, Tateisi Y, Tsujii J. GENIA corpus—a semantically annotated corpus for bio-textmining. *Bioinformatics*. 2003; 19: i180–i182.
57. Chiu B, Crichton G, Korhonen A, Pyysalo S. How to Train good Word Embeddings for Biomedical NLP. *Proceedings of the 15th Workshop on Biomedical Natural Language Processing*. 2016. pp. 166–174.
58. Kim Y. Convolutional Neural Networks for Sentence Classification. arXiv:14085882. 2014;
59. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J Mach Learn Res*. 2014; 15: 1929–1958. <https://doi.org/10.1214/12-AOS1000>
60. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. arXiv:14126980. 2014;