Türkiye Klinikleri Biyoistatistik Dergisi
Turkiye Klinikleri Journal of Biostatistics
Turkiye Klinikleri J Biostat. 2022;14(1):1-13

ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

DOI: 10.5336/biostatic.2021-84929

A Spatial Analysis of Air Quality and Food Consumption Surrogates in Relation to Asthma Deaths: An Analytical Research

Hava Kalite ve Gıda Tüketim Belirteçlerinin Astıma Bağlı Ölümlerin İlişkisinin Mekânsal Bir Analizi: Analitik Bir Araştırma

Begüm TÜZÜNER[a], Mahmut MERGEN[b], Selman AKTAŞ[c,d], Mehmet KOÇAK[b,e]

[a]Department of Industrial Engineering, Boğaziçi University Faculty of Engineering, İstanbul, Türkiye
[b]Department of Biostatistics, İstanbul Medipol University International Faculty of Medicine, İstanbul, Türkiye
[c]Department of Biostatistics, University of Health Sciences Faculty of Medicine, İstanbul, Türkiye
[d]Department of Biostatistics, İstanbul University-Cerrahpaşa, Cerrahpaşa Faculty of Medicine, İstanbul, Türkiye
[e]Department of Preventive Medicine, College of Medicine, The University of Tennessee Health Science Center, Memphis, Tennessee, USA

ABSTRACT Objective: Asthma is among the most common diseases affecting more than 300 million people globally overall. Despite growing understanding of the association between the environmental and food intake markers and the etiology and progression of asthma, there is still unmet knowledge gap which requires further investigations to be done in a spatial and temporal manner. We carried out such a study to explore the association of air-quality markers, fruit and vegetable consumption and drinking water source trajectories with asthma death rate in spatial models. **Material and Methods:** Province-level data on asthma deaths, air-quality markers, namely, particular matter 10 and 2.5, sulfur-dioxide, carbon monoxide (CO), nitrogen-dioxide, and ozone, drinking water source data from rivers, dams, wells, and springs as well as fruits and vegetables sales data were obtained for 81 provinces of Turkey for years 2018 and 2019. Mixed modelling approach taking into consideration the spatial autocorrelation was used to investigate the associations of these environmental and food consumption variables with asthma deaths. **Results:** These models revealed decreased asthma deaths with increased consumption of apple, banana, cabbage, lemon, onion, pineapple, spinach and zucchini, and increased asthma deaths with increased CO concentration and increased broad bean consumption. **Conclusion:** As a conclusion, we showed that spatial and temporal analyses have premise to offer much needed information to help close the knowledge gap in understanding the association of environmental and food intake markers with asthma deaths in granular spatial models. We have also showed that such efforts are possible by extracting publicly available data as well.

Keywords: Asthma deaths; air quality markers; drinking water sources; fruit consumption; vegetable consumption

ÖZET Amaç: Astım, tüm dünyada 300 milyondan fazla insanı etkileyen en yaygın hastalıklar arasındadır. Çevresel belirteçler ve gıda tüketim belirteçleri ile astımın etiyolojisi ve ilerlemesi arasındaki ilişkiye dair artan bilgiye rağmen hâlâ mekânsal ve zamansal daha fazla araştırmanın yapılmasını gerektiren doldurulamamış bir bilgi boşluğu bulunmaktadır. Biz de hava kalitesi belirteçleri, meyve ve sebze tüketimi ve içme suyu kaynakları ile astım ölüm oranı arasındaki ilişkiyi mekânsal modellemeler kullanarak incelemek amacıyla böyle bir çalışma yürüttük. **Gereç ve Yöntemler:** İl düzeyinde astım ölümleri, hava kalitesi belirteçleri yani partikül maddesi 10 ve 2,5, kükürtdioksit, karbonmonoksit (CO), nitrojen dioksit ve ozon; nehirlerden, barajlardan, kuyulardan ve kaynaklardan sağlanan içme suyu miktarları ile Türkiye'nin 81 ilinden 2018 ve 2019 yıllarına ait meyve-sebze satış verileri elde edilmiştir. Mekânsal otokorelasyonu dikkate alan karma modelleme yaklaşımı (MIXED models), astım ölümlerinin çevresel ve gıda tüketim değişkenleri ile ilişkisini araştırmak için kullanılmıştır. **Bulgular:** Bu modeller; elma, muz, lahana, limon, soğan, ananas, ıspanak ve kabak tüketiminde artışla birlikte azalan astım ölümlerini ve CO konsantrasyonu ve bakla tüketimi ile artan astım ölümlerini ortaya çıkarmıştır. **Sonuç:** Sonuç olarak mekânsal ve zamansal analizlerin, çevresel faktörler ve gıda tüketimi değişkenleri ile astıma dayalı ölümler arasındaki ilişkiyi anlamadaki bilgi boşluğunu kapamadaki faydasını gösterdik. Bunun yanı sıra bu tür çabaların, kamuya açık verilerin çıkarılmasıyla mümkün olduğunu da ortaya koyduk.

Anahtar kelimeler: Astım ölümleri; hava kalitesi belirteçleri; içme suyu kaynakları; meyve tüketimi; sebze tüketimi

Correspondence: Mehmet KOÇAK
Department of Biostatistics, İstanbul Medipol University International Faculty of Medicine, İstanbul, Türkiye
E-mail: mehmetkocak@medipol.edu.tr

Peer review under responsibility of Turkiye Klinikleri Journal of Biostatistics.

Received: 09 Jun 2021 ***Received in revised form:*** 18 Oct 2021 ***Accepted:*** 19 Oct 2021 ***Available online:*** 27 Oct 2021

According to the World Health Organization (WHO), asthma can be defined as a major noncommunicable disease characterized by recurrent attacks of breathlessness and wheezing, which varies in severity and frequency from person to person. It was estimated that more than 339 million people suffer from asthma worldwide.[1]

Environmental factors, in association with genetic susceptibility, play a critical role in asthma pathophysiology.[2] von Mutius suggests that gene-environment interactions play an important role in asthma etiology and progression and, contrary to common expectation, such interactions generate subtle effects on many genes rather than strong effect on few genes.[3] Among environmental risk factors, air pollutants are considered the most important especially for respiratory diseases including asthma as suggested by WHO. Guarnieri and Balmes demonstrated in a review article that air quality markers such as particulate matter (PM), ozone (O_3), nitrogen-dioxide (NO_2), and sulphur-dioxide (SO_2) are among the specific risk factors potentially increasing the oxidative injury to the airways, causing inflammation, and magnified sensitization, thus triggering new-onset asthma or existing asthma progression.[4]

In addition, as eating is an indispensable part of life, what we consume would directly affect our health positively or negatively. Since Ancient Greece, the question has been asked repeatedly. For example, Hippocrates's quote "let food be thy medicine" gives the clue to investigate the association between diet and health in general. It has been also investigated that analysis of dietary patterns such as Mediterranean diet (Med-diet) and the Western diet provides several advantages in examining the relationship between dietary intake and health outcomes: foods or nutrients are not eaten in isolation, and considerable interactions and synergistic effects exist with common nutrients in foods.[5] McKeever and Britton summarized the literature evidence that shows potential impact of fruit and vegetable consumption on asthma and concluded that while cross-sectional studies showed some evidence, case-control and longitudinal studies showed limited evidence.[6] Therefore, these potential associations between diet and asthma need to be more specifically explored in a spatiotemporal manner.

Gray et al. took a comprehensive approach by generating an Environment Quality Index composed of air, water, land, build, and sociodemographic markers and investigated the association of this index with the asthma prevalence using the medical claims data in the United States.[7] There is still a need to carry out such explorations towards a hard-endpoint like deaths due to asthma.

In this study, we aim to investigate the association of fruit and vegetable consumption, air quality markers, and drinkable water sources with asthma related deaths using spatiotemporal data from 81 provinces of Turkey. We hypothesize that there will be fruit and vegetable consumption and air quality markers significantly associated with the asthma death rates positively or negatively. We will describe our data and analytical approach in Material and Methods section, will present the study findings in the Results section, and discuss these findings with comparisons and contrasts with the current literature in the Discussion section.

MATERIAL AND METHODS

In this work, the data were congregated from three main data sources in a spatiotemporal manner. The spatial dimension is defined by Turkey's 81 provinces and the temporal dimension is defined by monthly data for years 2018 and 2019. The first data source aimed to cover the DIET dimension. To do this, fruit and vegetable sales data in local markets within each province was requested from the Turkish Ministry of Commerce for years 2018 and 2019. This data included 49 fruit and vegetable types that are sold in kilograms, bunches or as pieces locally; these local sales variables will be uses as surrogates for the food consumption signatures of these provinces. The second data source aimed to help generate ENVIRONMENT dimension. To achieve this aim, air quality data including PM_{10} and $PM_{2.5}$, SO_2, carbon monoxide (CO), nitrogen-monoxide, NO_2 as well as NO_X, and O_3 were manually extracted from the periodic Environmental Impact Evaluation (EIE) reports generated by the Turkish Ministry of Environment and Urbanization for each state in a monthly basis. In addition to air quality data, as a third dimension of data, drinking water sources data including the amount in the municipal water supply coming from stream, dam, lake, pond, spring and well were manually extracted for each of the 81 provinces from the periodic EIE reports generated by the Turkish Ministry of Environment and Urbanization for each state in an annual basis between years 2000-2018. Finally, asthma related deaths were considered as the primary outcome of this study and formed the fourth dimension of the data. The annual number of asthma deaths within each province for years 2018 and 2019 on monthly basis were requested from the Turkish Statistical Institute.

Data from these four dimensions were merged by the local and time identifiers. For data normalization, population size of each province for years 2018 and 2019 were also added. As the outcome variable is a hard endpoint, not an acute condition, each data item was aggregated for each year within each province as the monthly trend was not of primary interest in this particular study. Asthma deaths were expressed in a normalized manner as 'asthma deaths per 100,000 population'. For the fruit and vegetable types, the annual sum of consumption within each province has been divided by the population of each province in the corresponding years so to obtain a per capita among of sales as a surrogate marker of per capita consumption measure. The air quality measures were expressed as annual averages within each province. The water resources data is used to generate the profiles of the states via creating trends on the municipal water supply in over the years. Representative maps generated based on the above definitions of our primary outcome variables and predictors have been provided in Figure 1 and Figure 2.

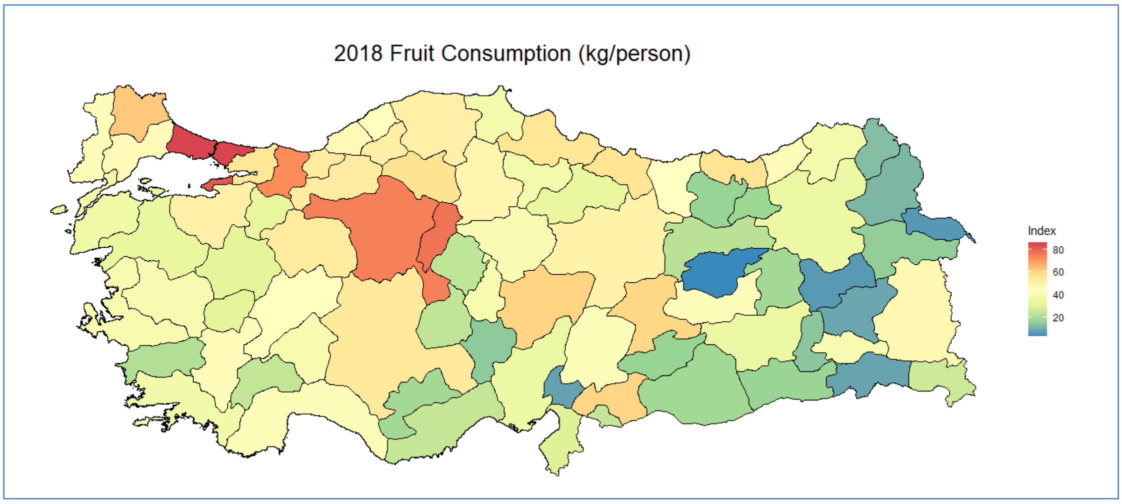


FIGURE 1: 2018 total fruit consumption (kg) per person.

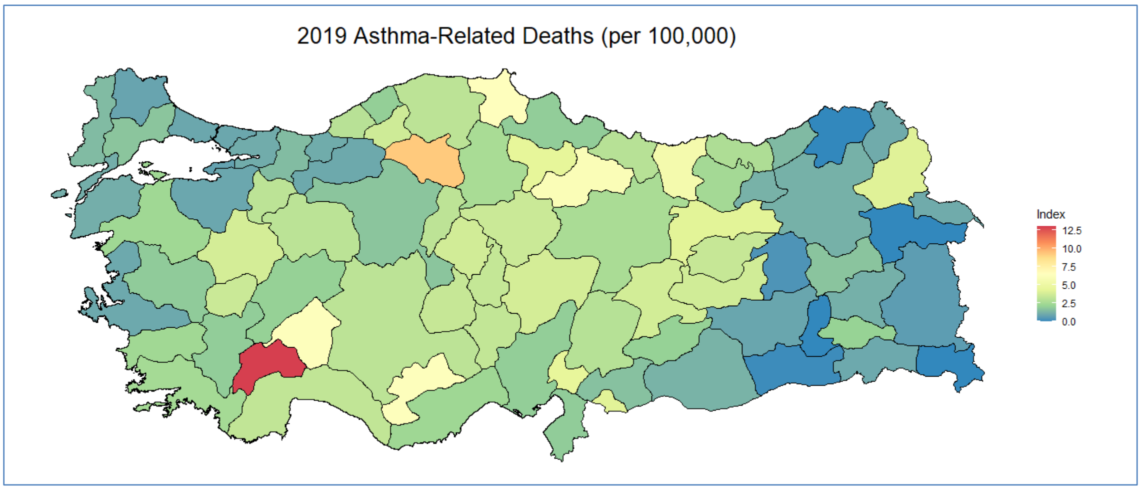


FIGURE 2: Asthma-related deaths per 100,000 population.

As the outcome variable is of a spatial nature, we tested for the spatial autocorrelation for each year of data using the Moran's I statistic.[8] For 2018, spatial autocorrelation was not significant (p=0.37), while it was significant for year 2019 (p=0.002), which is also apparent from Figure 2. To address this significant spatial autocorrelation, Mixed Modelling framework in SAS Version 9.4 (®, Cary, North Carolina, USA) through the MIXED procedure was utilized with spatial power covariance, SP(POW) structure with the central latitude and longitude values of each province since provinces cannot be treated as independent subjects any longer due to spatial autocorrelation in any regression model. Spatial-power autocorrelation approach estimates the underlying auto-correlation coefficient, ρ, and forms the covariance matrix such that each cell of the matrix is multiple of the variance estimates and power of the spatial-autocorrelation to the distance between the geographic centroids of any 2 pairs of the provinces as shown below for K spatial-locations:

$$SP(POW)(lat, long) = \sigma^2 \begin{bmatrix} 1 & \rho^{d_{12}} & \rho^{d_{13}} & \rho^{d_{14}} & \cdots & \rho^{d_{1K}} \\ \rho^{d_{21}} & 1 & \rho^{d_{23}} & \rho^{d_{24}} & \cdots & \rho^{d_{2K}} \\ \rho^{d_{31}} & \rho^{d_{32}} & 1 & \rho^{d_{34}} & \cdots & \rho^{d_{3K}} \\ \rho^{d_{41}} & \rho^{d_{42}} & \rho^{d_{43}} & 1 & \cdots & \rho^{d_{4K}} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{d_{K1}} & \rho^{d_{K2}} & \rho^{d_{K3}} & \rho^{d_{K4}} & \cdots & 1 \end{bmatrix}$$

where d_{ij} represents the distance between i^{th} and j^{th} locations using the latitude and longitude of these locations.

The general structure of such a spatial model can be given as follows:

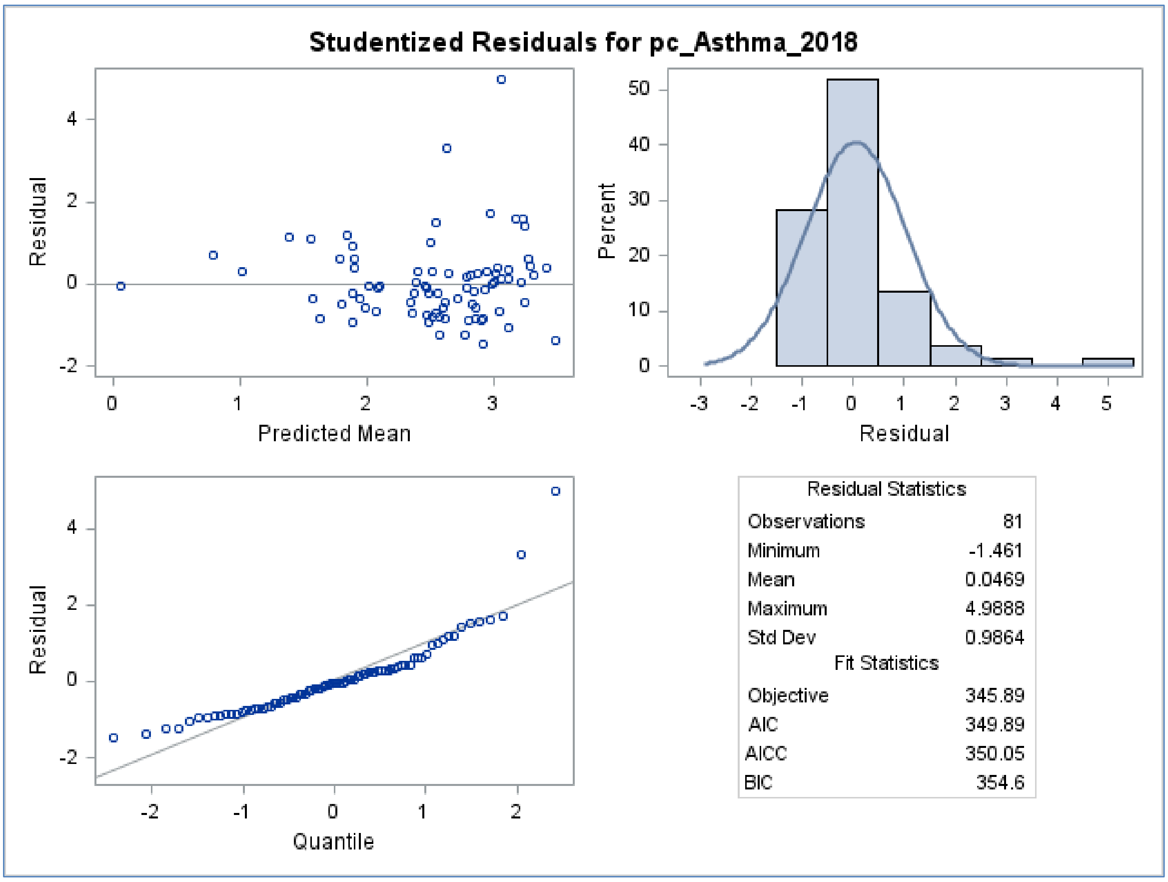
$$y_i = \mu + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \varepsilon_i$$

where $\varepsilon_i's$ are correlated errors following N(0, Σ) with Σ being the spatial-correlation matrix as shown in the matrix above. So, this model is identical to least squares regression models excepting that the residuals are not independent but follow a spatial correlation structure.

Model diagnostic plots were assessed for model adequacy, existence of outliers and influential observations. Overall, a few possible outliers were not beyond what would be expected in such spatial models and residual distribution was unimodal without significant departure from normality. An example of such diagnostic graphs was provided in Appendix 1.

To generate a food consumption and air quality signature of each province, data across the years of 2018 and 2019 were combined and categorized as quartiles for each food and air quality markers; thus, each province's signature was either lower than the first quartile, between first quartile and median, between median and third quartile, or higher than the third quartile. This approach helped to eliminate the impact of potential outliers and help with interpretations of the modeling results. An example of such signatures is shown for NO_2 in Figure 3.

APPENDIX 1: An example of model-diagnostic plots.



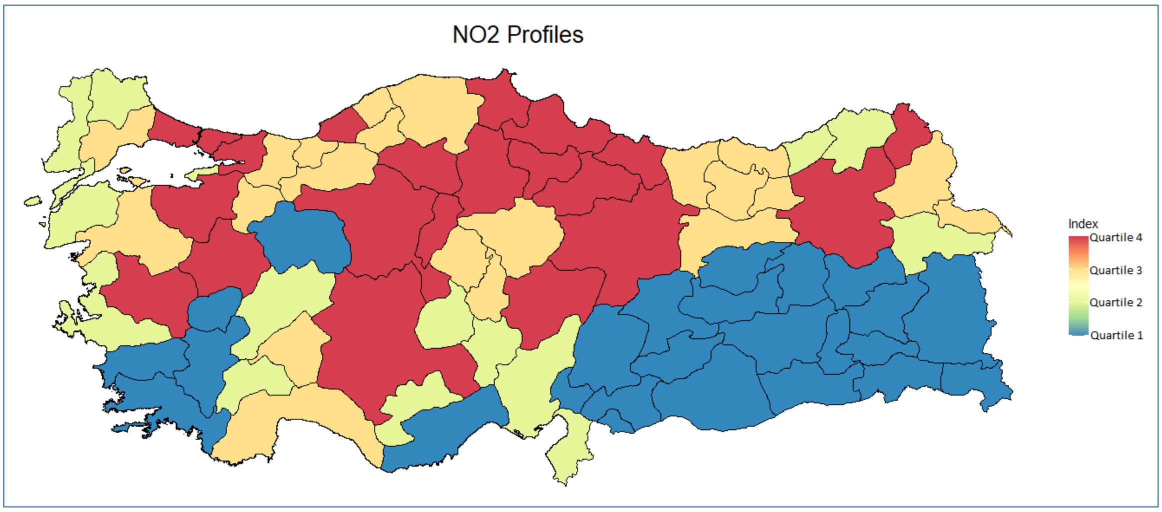


FIGURE 3: NO_2 profiles in quartiles.

In addition to the above "quartile" approach, whenever that are sufficient (i.e., data from at least three timepoints) longitudinal data, we built trajectories for a given marker using the TRAJ procedure developed for SAS by Jones and Nagin.[9] This approach was feasible for PM_{10}, SO_2, NO_2, O_3 and $PM_{2.5}$ for the air quality markers, and among the drinking water source data, it was feasible for Dam, Spring, Well, and River trajectories. We present an example of such trajectories in Figure 4.

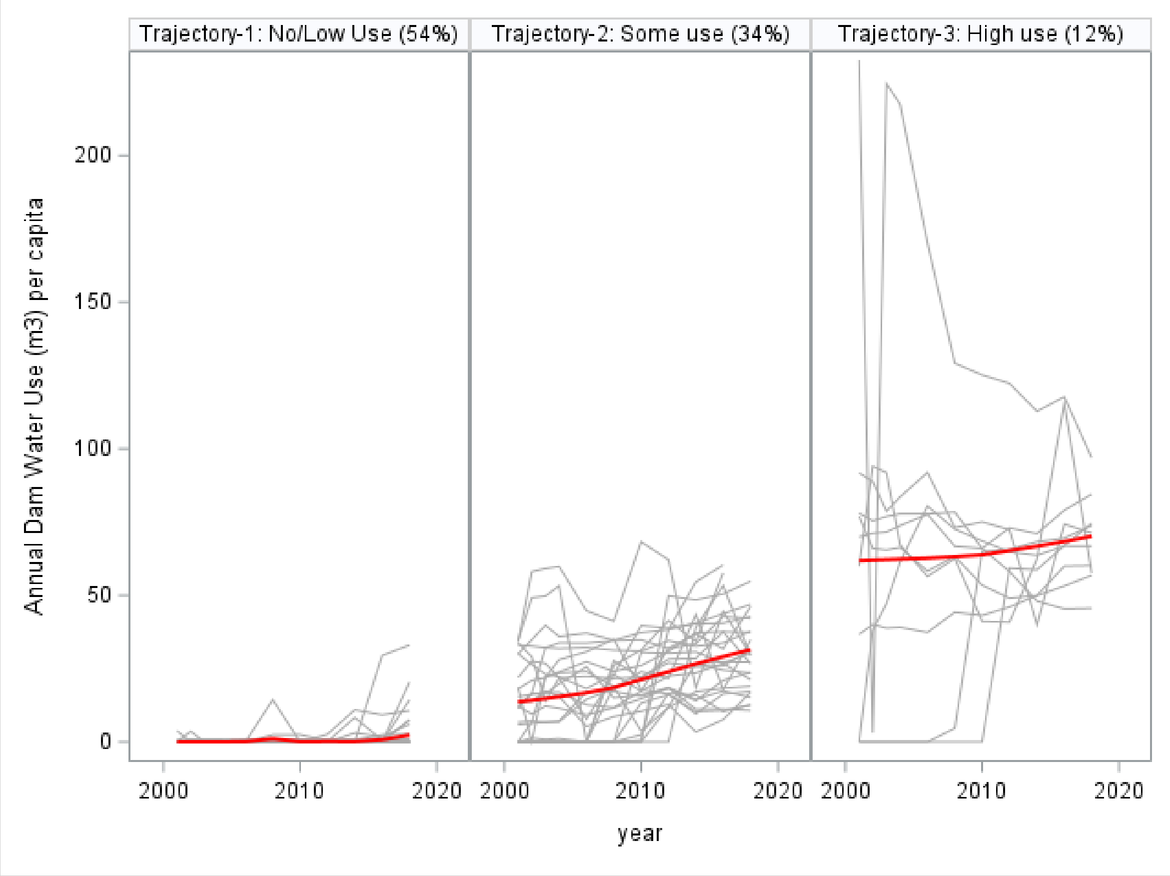


FIGURE 4: Annual dam water use trajectories. Numbers in parentheses represent the percentages of provinces falling into that trajectory.

Overall, 49 fruit and vegetable consumption markers including the total fruit and vegetable consumption, 8 air quality markers with limited data for some markers, and 4 drinking water source profiles. Despite these large number of predictors, we did not utilize a multiplicity-correction approach such as Benjamini-Hochberg False Discovery Rate approach due to the fact that this analysis is of exploratory nature and aims to generate hypotheses for further investigation.[10] Therefore, we used the traditional Type 1 error rate of 0.05 to indicate significance in our results from both univariable models as well bivariable models.

In our analytic approach, we first built univariable models with asthma-related deaths per 100,000 population as the dependent variable and each of the food, air-quality and water-source markers as predictors. Furthermore, bivariable models were built using a limited list of markers that were significant at unadjusted p value=0.1 to be more inclusive in this exploration stage. The above models were built independently for year 2018 and 2019 so that consistent significant results become more meaningful in generating hypotheses for future prospective studies.

RESULTS

Univariable model results significant at alpha=0.05 level at least in one of the analysis years are presented in Table 1. Overall vegetable consumption as well as 4 fruit (apple, banana, pear, and pineapple) and 5 vegetable consumption (cabbage, eggplant, lettuce, onion, and pepper) were found to be negatively associated with the deaths due to asthma. For example, every two additional kilograms of apple or onion consumption corresponds to 1 less death in 100,000 due to asthma. The list of environmental and food consumption markers significant at alpha=0.1 (Appendix 2) were included in our bivariable analysis. Interestingly, none of the air quality metrics were found to be associated with the asthma deaths and only CO was included in the bivariable models with its suggestive positive association with asthma deaths. Among the water source trajectories, only the river water consumption trajectory was found to be associated with our outcome variable. The model results of all air quality markers and water source trajectories were presented in Appendix 3 regardless of significance.

TABLE 1: Univariable results significant at α=0.05 in at least one year.

	Year 2018		Year 2019	
Predictor	**Beta**	**p value**	**Beta**	**p value**
Apple (kg)	-0.54	0.011	-0.38	0.038
Banana (kg)	-0.46	0.045	-0.52	0.0078
Cabbage (kg)	-0.41	0.067	-0.59	0.0017
Eggplant (kg)	-0.21	0.33	-0.39	0.030
Lettuce (bunch/piece)	-0.20	0.33	-0.37	0.032
Onion (kg)	-0.49	0.021	-0.33	0.065
Pear (kg)	-0.50	0.017	-0.46	0.011
Pepper (kg)	-0.37	0.089	-0.37	0.044
Pineapple (kg)	-0.45	0.073	-0.43	0.042
Total vegetable (bunch/piece)	-0.27	0.20	-0.39	0.030
River water trajectories	-0.55	0.13	-0.63	0.041

APPENDIX 2: Univariable results significant at α=0.10 in at least one year to be included in the bivariable analyses.

	Year 2018		Year 2019	
Predictor	**Beta**	**p value**	**Beta**	**p value**
Cabbage (kg)	-0.41	0.067	-0.59	0.0017
Banana (kg)	-0.46	0.045	-0.52	0.0078
Apple (kg)	-0.54	0.011	-0.38	0.038
Pear (kg)	-0.50	0.017	-0.46	0.011
Onion (kg)	-0.49	0.021	-0.33	0.065
Eggplant (kg)	-0.21	0.33	-0.39	0.030
Total vegetable (bunch/piece)	-0.27	0.20	-0.39	0.030
Lettuce (bunch/piece)	-0.20	0.33	-0.37	0.032
River water trajectories	-0.55	0.13	-0.63	0.041
Pineapple (kg)	-0.45	0.073	-0.43	0.042
Pepper (kg)	-0.37	0.089	-0.37	0.044
Broad bean (kg)	-0.26	0.27	-0.39	0.056
Coliflower (kg)	-0.20	0.37	-0.37	0.057
Spinach (kg)	-0.36	0.071	-0.34	0.062
Tomato (kg)	-0.38	0.063	-0.26	0.15
Parsley (bunch/piece)	-0.30	0.19	-0.36	0.064
Pomegranate (kg)	-0.39	0.065	-0.19	0.29
Lemon (kg)	-0.41	0.067	-0.26	0.15
Total vegetable (kg)	-0.37	0.069	-0.29	0.10
Fruit (kg)	-0.39	0.072	-0.34	0.072
Kiwi (kg)	-0.14	0.50	-0.32	0.074
Zuchini (kg)	-0.38	0.074	-0.25	0.16
Fresh garlic (kg)	-0.32	0.14	-0.34	0.076
Beets (kg)	-0.36	0.14	-0.36	0.089
Carbon monoxide	0.33	0.090	0.04	0.81
Dry garlic (kg)	-0.36	0.090	-0.15	0.41
Grape (kg)	-0.30	0.20	-0.33	0.096

APPENDIX 3: Univariable results for air-quality and water-source markers regardless of significant.

	Year 2018		Year 2019	
Predictor	**Beta**	**p value**	**Beta**	**p value**
Carbon monoxide	0.33	0.090	0.04	0.81
Nitrogen-dioxide	0.26	0.27	0.06	0.77
Nitrogen-dioxide trajectories	-0.73	0.45	0.21	0.79
Ozone	0.12	0.61	-0.03	0.90
Ozone trajectories	0.29	0.67	-0.18	0.75
Particular matter-10 trajectories	-0.17	0.50	-0.20	0.34
Particular matter-2.5	0.00	1.00	0.02	0.94
Particular matter-2.5 trajectories	-0.90	0.30	-0.48	0.52
Sulphur-dioxide	-0.05	0.82	-0.16	0.37
Dam water trajectories	-0.29	0.35	-0.19	0.46
River water trajectories	-0.55	0.13	-0.63	0.041
Spring water trajectories	0.04	0.91	0.13	0.62
Well water trajectories	0.86	0.13	0.48	0.33

The results of the bivariable models were presented in Table 2. When considered together with fruit and vegetable consumption variables, CO was consistently positively associated with asthma deaths in both years. Similarly, onion consumption was consistently negatively associated with asthma deaths in both years adjusted for another significant predictor. Other than these two predictors, apple, banana, cabbage, lemon, pineapple, spinach and zucchini were found to be negatively correlated, and broad bean showed a suggestive positive association with asthma deaths. Interestingly, the river water usage trajectories were associated with asthma deaths negatively as well.

TABLE 2: Bivariable results significant at α=0.05 in at least one marker for at least one year.

Predictor-1	**Predictor-2**	Year 2018				Year 2019			
		Predictor-1		Predictor-2		Predictor-1		Predictor-2	
		Beta	**p value**	**Beta**	**p value**	**Beta**	**p value**	**Beta**	**p value**
Apple (kg)	Carbon monoxide	-0.66	0.0021	0.46	0.016	-0.38	0.034	0.32	0.040
Banana (kg)	Carbon monoxide	-0.43	0.048	0.58	0.0027	-0.29	0.13	0.37	0.032
Broad bean (kg)	Onion (kg)	0.13	0.66	-0.56	0.043	0.52	0.031	-0.68	0.0049
Cabbage (kg)	Carbon monoxide	-0.46	0.037	0.38	0.050	-0.38	0.037	0.33	0.037
Lemon (kg)	Carbon monoxide	-0.47	0.022	0.53	0.0033	-0.36	0.051	0.34	0.14
Onion (kg)	Carbon monoxide	-0.43	0.032	0.51	0.0048	-0.41	0.0034	0.34	0.064
Pineapple (kg)	River water trajectories	-0.52	0.038	-0.63	0.082	-0.01	0.95	-0.63	0.043
Spinach (kg)	Carbon monoxide	-0.50	0.014	0.52	0.0033	-0.42	0.017	0.35	0.048
Zuchini (kg)	Carbon monoxide	-0.42	0.049	0.36	0.066	-0.32	0.078	0.32	0.046

Asthma death distributions by the quartiles of the above-mentioned significant variables are presented in Figure 5 and Figure 6.

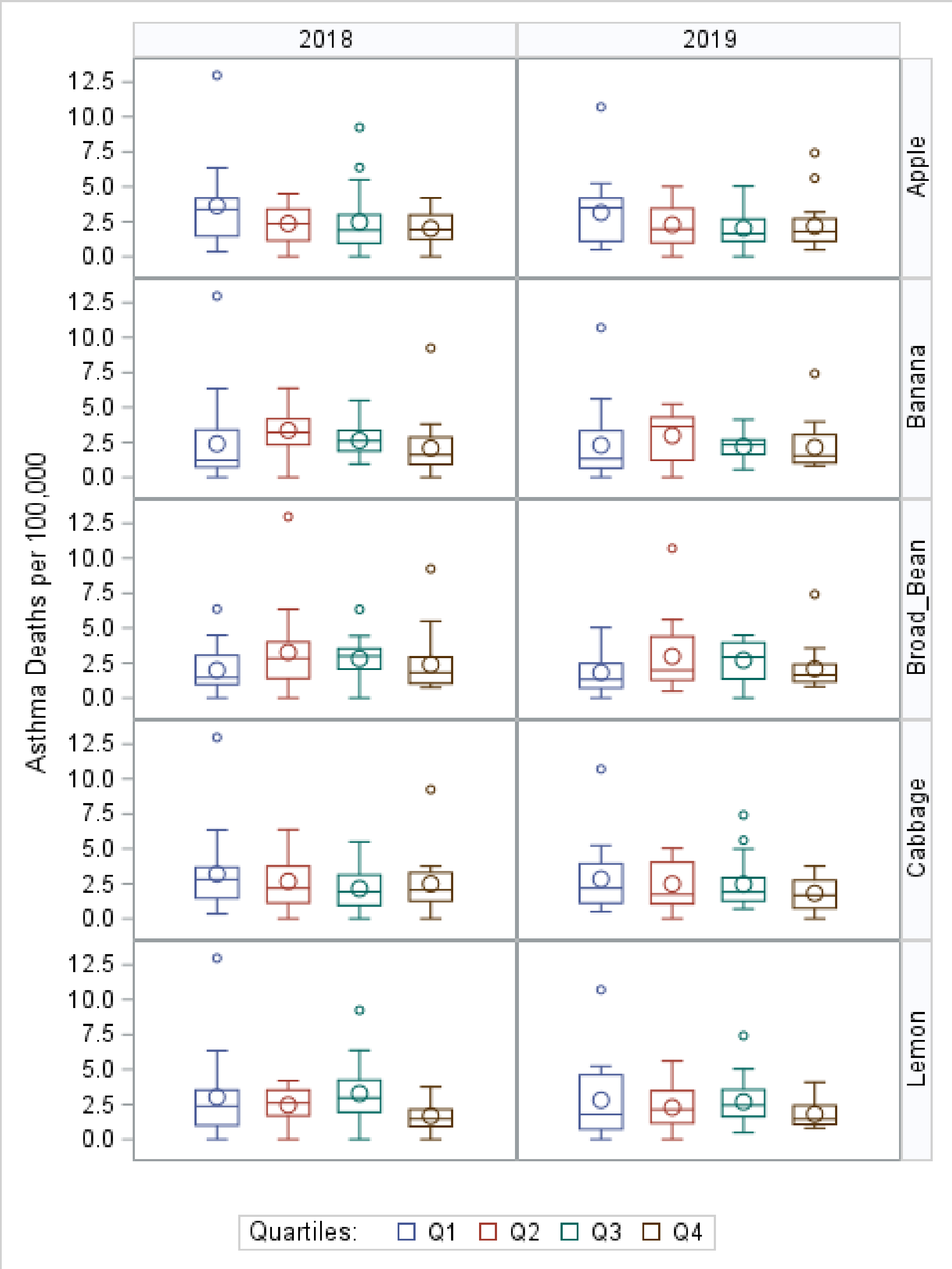


FIGURE 5: Asthma death distributions by the quartiles of apple, banana, broad bean, cabbage, and lemon.

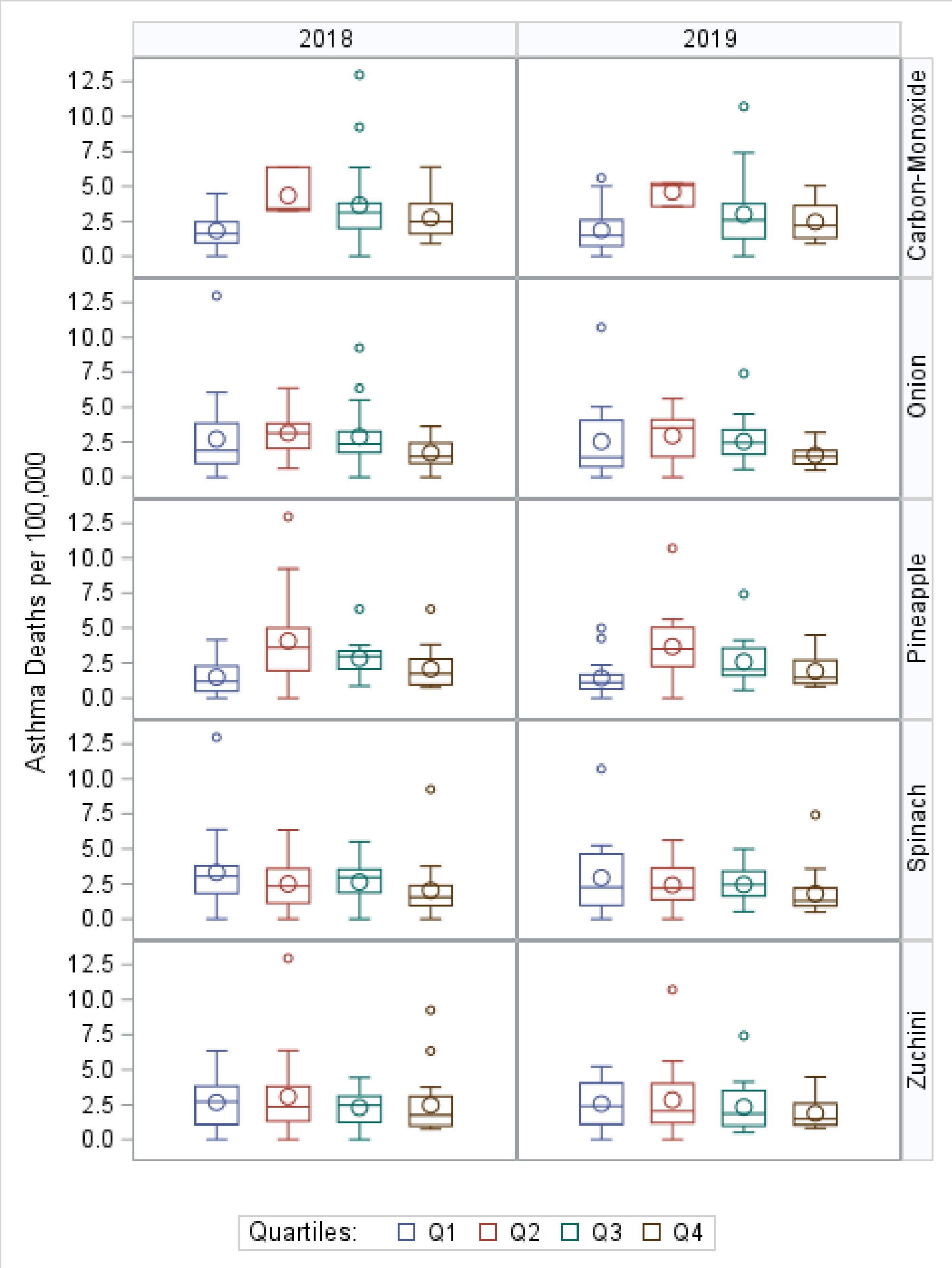


FIGURE 6: Asthma death distributions by the quartiles of onion, pineapple, spinach, zucchini, and carbon monoxide.

DISCUSSION

In this study, we have investigated the association of fruit and vegetable sales data as the surrogates of food consumption, air quality markers, and drinking water sources with asthma deaths per 100,000 population in spatial models. In doing so, we have utilized the distributional quarters for the entire country for each potential predictor, which represented the profile of a given province for that specific marker compared to other provinces as quartiles so that low, intermediate, and high consumption or concentration levels can be compared; we were also able to construct trajectories for air-quality and drinking water source predictors, which had sufficient data to produce such trajectories. One immediate advantage of such an approach is to reduce the potential impact of outliers and influential observations on model results where each province can be represented with a single group assignment based on longitudinal data, thus leading to more robust findings.

As it is expected that provinces closer to each other would be expected to have similar air-quality and dietary intake, as also shown through formal statistical tests, the existence of spatial autocorrelation is inescapable. We addressed the spatial autocorrelation by utilizing Mixed Modeling framework with Spatial Power covariance structure, SP(POW) (latitude, longitude), specified in the REPEATED statement in the MIXED procedure of SAS. We have approached the temporal component of our analyses by building separate models for year 2018 and year 2019 so that we can assess the consistency of significant findings over two years, an approach somewhat similar to training-validation type comparison in model building. In summarization of our model results, although we recognize the multiplicity issue, we did not apply a multiplicity correction approach such as Benjamini-Hochberg False Discovery Rate, to control the familywise error rate, mainly because this study is an exploratory study where we generate hypotheses for further investigation; therefore, we wanted to be as liberal as possible in providing at least promising results without any claim of causality.

One of the main findings of our analyses is the not-surprising positive association of CO with asthma deaths. This finding is confirmatory to the work by Evans et al.[11] They also showed an association of asthma exacerbation of children with $PM_{2.5}$ while our study did not find any significant association perhaps due to the fact that the $PM_{2.5}$ data were not complete for Turkey with only 48 of 81 provinces having the relevant data. The association of CO with asthma deaths also displayed a quadratic behavior (Figure 6, first row) where provinces in Quartile-1 having relative lower asthma deaths compared to provinces in Quartile-2 while provinces in Quartile-3 and -4 also having lower asthma deaths compared to Quartile-2 as well on average. Such a quadratic pattern was statistically significant for both years 2018 and 2019. Such a pattern existed for broad bean and pineapple as well, which makes the interpretation a bit harder. The next consistently significant finding is the potential beneficial effect of onion consumption, which is in line with the suggestions by a review paper by Bystrická et al.[12] This potential impact may be due to the antioxidant properties of onion as shown by Marefati and Boskabady in animal models.[13] Shaheen et al. showed in a case-control study the potential benefit of consuming apples and onions for asthma in adults and our study also showed a negative significant association of apple consumption with asthma death.[14] Romieu et al. analyzed the French branch (E3N, Etude Epide´miologique aupre`s des femmes de la MGEN) of the European Prospective Investigation into Cancer and Nutrition study in a similar manner as we did in our study and found reduced odds of asthma diagnosis by increased consumption of leafy vegetables such as lettuce and spinach; our study also showed negative association of cabbage and spinach with asthma death rate.[15] These findings are also in line with those by Woods et al. who showed a negative association of apple, pear and green-leaf vegetable consumption with asthma in young adults.[16]

A surprising association among the major findings of our analyses was the negative association of river water consumption trajectories. This association may be a spurious one as the correlation of the river water consumption trajectories with the quartiles of total fruit and vegetable consumption was significantly positively correlated, which may indicate that the regions with rivers are already in advantage in terms of fruit and vegetable production, thus consumption, while the rivers are the main source of drinking waters supplied by municipalities after proper filtering in those regions. One final association we have shown in our analyses was the consistent protective effect of lemon for asthma deaths. Choi et al. showed anti-asthma and anti-inflammation effects of lemon oil in animal studies, which seems to be in line with our findings.[17]

Our study has several weaknesses to mention. Perhaps the most important of them is that our predictors of asthma death rate are not an exact measure but an average for geographic regions (i.e., 81 provinces in our case) utilizing limited amount of data over few years. Secondly, our food and vegetable consumption markers are just surrogates and not ideal nutritional intake variables as they are simply the per-capita sales data of these individual fruits and vegetables and the reason for lower or higher consumption of certain food items may be due to multiple convoluted factors such as the geographic and cultural characteristics of a given region like its climate, elevation, local cuisine, etc., as well as its demographic structure such as urbanization and concentration of different age-segmentation of its population, main income source such as agriculture and farming. Eyles et al. showed a significant correlation between the market sales data with actual household nutritional intake.[18] Therefore, we still believe that the local market sales data can be reasonable surrogates when individual or household level consumption data are not available. However, we still lack in our data sources in terms of some critical food types such as meat, dairy, and flour-based items. The ideal study design to investigate the association between fruit and vegetable consumption would require granular data including individual level of disease follow-up and individual level of food consumption profiles; however, such data are close to impossible to obtain in a large spatial spectrum like ours and therefore, we were limited by representing each province by a signature of fruit and vegetable consumption using the sales data as surrogates as well as air quality markers with an average behavior over few years. Additionally, we used a hard-endpoint for the asthma profile for the provinces as deaths due to asthma rather than actual prevalence of asthma diagnoses.

Another critical weakness of our study is that the death due to asthma is a rare event. In fact, in four provinces out of 81, there were no deaths due to asthma in 2018 and in two provinces in 2019; similarly, in 11 and 17 provinces respectively in 2018 and 2019, asthma death rate for every 100,000 population was less than 1.0. The range of the asthma death rate was 0 to 13 per 100,000 population, with a median of 2.3 and 1.9 respectively for 2018 and 2019. This rare event nature of our outcome variable may have impacted the statistical power of our models negatively. Xue et al. discussed the modelling challenges especially around model convergence for correlated longitudinal rare events.[19] We did not encounter any convergence issues with our spatial models perhaps because our response variable is a population-size adjusted variable, thus a continuous nature of an outcome variable, which reduces the issue of rare-event problems. For example, when we combine the asthma deaths for years 2018 and 2019, the number of significant variables went up considerably from 11 markers to 19 markers; since we are mainly interested in the consistency of our finding across years, we chose not to employ this data aggregation approach for our outcome variable.

As future work, we plan to obtain direct measures of water quality markers currently missing from our models. We are in the process of requesting such data from the municipalities and the Ministry of Health, and we are also working on extracting publicly available water-quality data from local authorities. Similarly, we are expanding the environmental markers to include markers such as air temperature, air pressure, humidity, rain fall, etc., to be included in our array of predictors. Finally, to expand the food consumption dimension of our data beyond fruit and vegetable consumption, our team is working on identifying the proper sources of new dietary markers such as meat, poultry, and fish consumption and their derivatives, as well as dairy products and flour-based products such as bread, pastries, desserts, etc. In addition, our death records data are available only for years 2018 and 2019, which limits the generalizability of our findings. As the digitalization of data flow at the governmental level as well as at the local non-governmental organization level improves and the above-mentioned data sources become more accessible, the array of our analyses will naturally expand to include more data dimensions and longer time periods to catch associations in a more generalizable manner.

CONCLUSION

In this work, we have identified several fruit and vegetable consumption, and one air quality marker associated with the respiratory disease death rate in our spatial models consistently for years 2018 and 2019. Among these, apple, banana, cabbage, lemon, onion, pineapple, spinach, and zucchini consumption as well as river water consumption trajectories as drinking water were found to be negatively associated with, thus protective against asthma deaths while CO concentration and broad bean consumption were found to be positively associated with asthma deaths, thus detrimental. As these associations do not necessarily mean causation, they must be viewed as new hypotheses and new well-designed prospective studies must be developed to investigate these new hypotheses.

Availability of Data and Material

As the death records and fruit and vegetable consumption data used in this report were granted access only to the corresponding author, we do not have permission to share these data components; however, we can share the air quality data upon request.

Acknowledgement

We also thank the Turkish Republic Ministry of Commerce and Turkish Statistical Institute for data sharing. The opinions raised in this article solely belong to its authors, and does not represent the position of TUBITAK, Turkish Republic Ministry of Commerce and Turkish Statistical Institute in any shape of form.

Source of Finance

Partial financial support was received from TUBITAK Directorate of Science Fellowships and Grant Programmes (BIDEB)-2232 International Fellowship for Outstanding Researchers.

Conflict of Interest

No conflicts of interest between the authors and/or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Authorship Contributions

Idea/Concept: *Mehmet Koçak;* ***Design:*** *Mehmet Koçak;* ***Control/Supervision:*** *Mehmet Koçak, Selman Aktaş;* ***Data Collection and/or Processing:*** *Mehmet Koçak, Selman Aktaş, Begüm Tüzüner, Mahmut Mergen;* ***Analysis and/or Interpretation:*** *Mehmet Koçak, Begüm Tüzüner;* ***Literature Review:*** *Begüm Tüzüner, Mehmet Koçak;* ***Writing the Article:*** *Begüm Tüzüner, Mehmet Koçak;* ***Critical Review:*** *Begüm Tüzüner, Mehmet Koçak, Mahmut Mergen;* ***References and Fundings:*** *Mehmet Koçak.*

REFERENCES

1. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2017;390(10100):1211-59. Erratum in: Lancet. 2017;390(10106):e38. [PubMed] [PMC]
2. Guilleminault L, Williams EJ, Scott HA, Berthon BS, Jensen M, Wood LG. Diet and asthma: is it time to adapt our message? Nutrients. 2017;9(11):1227. [Crossref] [PubMed] [PMC]
3. von Mutius E. Gene-environment interactions in asthma. J Allergy Clin Immunol. 2009;123(1):3-11; quiz 12-3. [Crossref] [PubMed]
4. Guarnieri M, Balmes JR. Outdoor air pollution and asthma. Lancet. 2014;383(9928):1581-92. [Crossref] [PubMed] [PMC]
5. Tapsell LC, Neale EP, Satija A, Hu FB. Foods, nutrients, and dietary patterns: interconnections and implications for dietary guidelines. Adv Nutr. 2016;7(3):445-54. [Crossref] [PubMed] [PMC]
6. McKeever TM, Britton J. Diet and asthma. Am J Respir Crit Care Med. 2004;170(7):725-9. [Crossref] [PubMed]
7. Gray CL, Lobdell DT, Rappazzo KM, Jian Y, Jagai JS, Messer LC, et al. Associations between environmental quality and adult asthma prevalence in medical claims data. Environ Res. 2018;166:529-36. [Crossref] [PubMed] [PMC]
8. Moran PA. Notes on continuous stochastic phenomena. Biometrika. 1950;37(1-2):17-23. [Crossref] [PubMed]
9. Jones BL, Nagin DS. Proc TRAJ: A SAS Procedure for Group-Based Modeling of Longitudinal Data. In: Annual Meeting Home. 2007. [Link]
10. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological). 1995;57(1):289-300. [Crossref]
11. Evans KA, Halterman JS, Hopke PK, Fagnano M, Rich DQ. Increased ultrafine particles and carbon monoxide concentrations are associated with asthma exacerbation among urban children. Environ Res. 2014;129:11-9. [Crossref] [PubMed] [PMC]
12. Bystrická J, Musilová J, Vollmannová A, Timoracká M, Kavalcová P. Bioactive components of onion (Allium cepa L.)--a review. Acta Alimentaria. 2013;42(1):11-22. [Crossref]
13. Marefati N, Boskabady MH. The effect of onion (Allium cepa) extract on serum oxidant, antioxidant biomarkers in rat model of asthma. Iranian Journal of Allergy, Asthma & Immunology. 2018;17:164-5. [Link]
14. Shaheen SO, Sterne JA, Thompson RL, Songhurst CE, Margetts BM, Burney PG. Dietary antioxidants and asthma in adults: population-based case-control study. Am J Respir Crit Care Med. 2001;164(10 Pt 1):1823-8. [Crossref] [PubMed]
15. Romieu I, Varraso R, Avenel V, Leynaert B, Kauffmann F, Clavel-Chapelon F. Fruit and vegetable intakes and asthma in the E3N study. Thorax. 2006;61(3):209-15. [Crossref] [PubMed] [PMC]
16. Woods RK, Walters EH, Raven JM, Wolfe R, Ireland PD, Thien FC, et al. Food and nutrient intakes and asthma risk in young adults. Am J Clin Nutr. 2003;78(3):414-21. [Crossref] [PubMed]
17. Choi GG, Chung KJ, Cheong KJ. Anti-asthma and anti-inflammation effects of lemon oil in OVA-induced allergic asthma mouse model. Journal of Digital Convergence. 2014;12(10):577-85. [Crossref]
18. Eyles H, Jiang Y, Ni Mhurchu C. Use of household supermarket sales data to estimate nutrient intakes: a comparison with repeat 24-hour dietary recalls. J Am Diet Assoc. 2010;110(1):106-10. [Crossref] [PubMed]
19. Xue X, Kim MY, Wang T, Kuniholm MH, Strickler HD. A statistical method for studying correlated rare events and their risk factors. Stat Methods Med Res. 2017;26(3):1416-28. [Crossref] [PubMed] [PMC]