# Exploring Common Symptoms in Patients with Respiratory Allergies Using K-Means Algorithm in the North-East of Iran in 2012-2015 


#### Abstract

Somaye Norouzi ${ }^{1}$, Samane Sistani ${ }^{2}$, Maryam Khoshkhui ${ }^{3}$, Reza Faridhosseini ${ }^{3}$, Payam Payandeh ${ }^{4}$, Fahimeh Ghasemian ${ }^{5}$, Leila Ahmadian ${ }^{6}$, Mohammadhossein Pourasad ${ }^{7}$, Farahzad Jabbari Azad ${ }^{3}$ ${ }^{1}$ Student Research Committee, Faculty of Management and Medical Information Sciences, Kerman University of Medical Sciences, Kerman, Iran, ${ }^{2}$ Department of Medical Informatics, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran, ${ }^{3}$ Allergy Research Center, Mashhad University of Medical Sciences, Mashhad, Iran, ${ }^{4}$ Department of Pediatrics, School of Medicine, Non-Communicable Pediatric Disease Research Center, Health Research Institute, Amirkola Hospital, Babol University of Medical Sciences, Babol, Iran, ${ }^{5}$ Department of Computer Engineering, Faculty of Engineering, Shahid Bahonar University, Kerman, Iran, ${ }^{6}$ Department of Health Information Sciences, Faculty of Management and Medical Information Sciences, Kerman University of Medical Sciences, Kerman, Iran, ${ }^{7}$ Department of Health Information Technology, School of Paramedical, Kermanshah University of Medical Sciences, Kermanshah, Iran


Background: As a common disease among people of almost any age, allergic rhinitis has many adverse effects such as lowering the quality of life and efficiency at work or school. Considering these conditions and the collection of large amounts of data, the present research was conducted on allergic rhinitis and asthma patients' data to extract the common symptoms of these diseases using cluster analysis and the k-means algorithm.
Materials and Methods: The present cross-sectional research was conducted in Mashhad city. The inclusion criteria were affliction with one or two respiratory allergy diseases diagnosed by an allergy specialist through clinical history taking and physical examination. A researcher-made checklist was used in the present study for data collection. Then, the K-means algorithm's cluster analysis model was conducted to extract clusters (WEKA software (3, 6, 9)).
Results: Overall, 1,231 patients met the inclusion criteria. The result of the Cluster analysis consisted of
1: Cluster 1 in allergic rhinitis consisted of 702 patients, and cluster 2 consisted of 382 patients.
2: 46 asthma patients were assigned to cluster 1 and 23 to cluster 2.
3: Also, 60 asthma and allergic rhinitis patients were assigned to cluster 1 and 19 to cluster 2 . The most common symptoms in all patients were rhinorrhea, sneezing, nasal congestion, and itchy nose.
Conclusion: Overall, Salsola kali was the most common allergen in allergic rhinitis and asthma patients. Also, the most common symptoms in patients are rhinorrhea, sneezing, itchy nose, and nasal congestion. This study can help physicians diagnose allergic rhinitis and asthma in geographical areas with a high prevalence of Salsola kali.

Keywords: Allergic rhinitis; Asthma; Data mining; Cluster analysis

## INTRODUCTION

Allergic rhinitis is a common disease among people of almost any age. It is the most prevalent among adolescents. Though allergic rhinitis is not considered a serious disease, it is clinically important as it is associated with many adverse effects, which can lower the quality of life and
efficiency at work or school(1). Basic scientific and epidemiologic findings show that allergic rhinitis is a part of a systematic inflammatory process and is followed by other inflammatory disorders in mucous membrane such as asthma, rhinosinusitis, and allergic conjunctivitis(2).

This disease is the main risk factor for controlling mild asthma, as most people afflicted with allergic and nonallergic asthma are affected by the disease (3).

Annually, 20-40 million people are afflicted with allergic rhinitis in the US. $10-30 \%$ of these people are adults, and about $40 \%$ of this population are children (4). According to the existing research in Iran, the prevalence of allergic rhinitis is $11.9 \%$ in $6-7$-year-old children and $21.2 \%$ among the 13-14-year-old population (5). Allergic rhinitis is not only accompanied by physical symptoms but also by social and psychological adverse effects. Thus, it affects other aspects of patients' lives too. In light of the existing literature, those afflicted with allergic rhinitis are constantly reporting a low quality of life. This disease imposes high medical costs on patients and indirectly affects their social and psychological health (6).

As a complicated disease, allergic rhinitis is associated with asthma. Thus, it is necessary to explore allergic rhinitis and asthma phenotypes and how they are correlated (7). This disease is often neglected, underdiagnosed, or mistaken for another disease. It is detrimental to health and imposes high costs on society.

In recent years, the healthcare industry has focused on data science, which is a domain that includes specific rules and regulations, statements of problems, algorithms, and processes for extracting applicable models from panel data. Cluster analysis is one of the best-known artificial intelligence methods used to investigate allergic rhinitis (8). Cluster analysis is a data mining method based on defining a group of data (cluster) according to the properties of the existing data and finding the representative points of the cluster. Each cluster consists of a group of data with similar features. One of the wellknown algorithms for cluster analysis is the k-means algorithm $(8,9)$.

The present research was conducted on the allergic rhinitis and asthma patients' data to extract the common
symptoms of these diseases using cluster analysis and the k-means algorithm.

## MATERIALS AND METHODS

## Setting, Sample Size, and Eligibility Criteria

The present cross-sectional research was a retrospective descriptive study which was conducted in Mashhad, a major medical destination in northeast Iran. The participants were patients in the age range of 1-74 years diagnosed with respiratory allergy (allergic rhinitis and asthma) who were visited in the allergy and asthma clinic of Ghaem Hospital in Mashhad. Ghaem Hospital is a general hospital affiliated with Mashhad University of Medical Sciences. There are three allergy and clinical immunology subspecialists in this hospital. This center admits two subspecialist assistants annually to pass a 3year sub-specialization course taught by professional instructors.

The inclusion criteria were affliction with one or two respiratory allergy diseases diagnosed by an allergy specialist through the clinical history taking and physical examination. A prick test was also done on the patient's skin by an allergy specialist when needed. The exclusion criterion was affliction with a background disease such as a cold, flu, chronic sinusitis, or any other rhinitis infectious symptoms.

## Data collection instrument

A researcher-made checklist was used in the present study for data collection. This checklist was validated by allergy and clinical immunology subspecialists. It consisted of three parts:

1. Patients' demographic information (age, gender, and place of residence and occupation).
2. Clinical symptoms of disease and family history (runny nose, itchy nose, itchy eye, etc.).
3. The prick test result (allergens such as Salsola kali, Ash, etc.).

## Data preprocessing

The main step before data mining is data preprocessing (9). Thus, before the final analysis, data preprocessing was done in the present research to solve problems such as low-quality data, redundant, recurrent, unclear, heterogeneous, and missing data. The data need to be precise, valid, and impeccable for the final analysis. Also, sometimes, it is necessary to homogenize the size of data. The age and year/season variables were normalized in the clustering algorithm via the $\mathrm{x}=(\mathrm{x}-\mathrm{min}) /($ max-min) formula.

## Cluster analysis model

The next step in data preprocessing is determining an algorithm for the data set. K-means is among the main clustering algorithms (8). This method is used for clustering the dataset to separate groups when there is no information available about the shape of the clusters. The K-means algorithm works by defining an initial set of cluster centers extracted from data. Then it assigns each record to the most similar cluster based on the value in the entry fields of the record. Finally, after each clustering, a percentage is estimated for each symptom, indicating the number of people with the same symptom (9). In this algorithm, variables are assumed to be of equal weight, and the Euclidean distance is used for clustering. The WEKA software $(3,6,9)$ was used in the present research for cluster extraction. After administering the k-means algorithm to the existing data, different types of clustering were tested on the data. These two criteria were used to assess and select clusters: 1. reliance on an allergy and asthma specialist's comments, 2. the highest intra-cluster similarity and inter-cluster distinction.

## Ethical Consideration

The current study was based on retrospective data, and patients had a free chance to contribute to the project. However, informed consent for participation was obtained from the participants. Furthermore, ethical approval for this study was obtained from the Ethics Committee of Kerman University of Medical Sciences (IR.KMU.REC.1400.433).

## RESULTS

## Descriptive statistics

During the three years of conducting this research, 1,881 patients were visited in the allergy clinic. Among them, 1,445 had at least one positive prick test result. Overall, 1,231 patients met the inclusion criteria. A total number of 1,083 patients ( $88 \%$ ) were afflicted with allergic rhinitis, among whom $54 \%$ were male. Moreover, $5.6 \%$ ( $\mathrm{n}=69$ ) were afflicted with asthma, among whom $53.6 \%$ were male. Also, $6.4 \% ~(n=79)$ were afflicted with asthma and allergic rhinitis, among whom $60.8 \%$ were male. Overall, the frequency of males was higher than female participants. Regarding the risk factors, 718 patients had a family history of atopia, which was prevalent in $58.5 \%$ of allergic rhinitis patients, $59.4 \%$ in asthma patients, and $54.4 \%$ in the group with asthma and allergic rhinitis.

## Cluster analysis results

The clustering technique was applied for each disease separately. In the following, the results are presented for each disease.

## Cluster analysis results across demographic groups Allergic rhinitis

Patients afflicted with allergic rhinitis were divided into two clusters. Cluster 1 consisted of 702 patients, and Cluster 2 consisted of 382 patients. Cluster 1 included $50 \%$ males, $82 \%$ of $>15$ years old, $18 \%$ of $<15$ years years old, $61 \%$ with a family history of atopic, $84 \%$ with prevalent symptoms in urban areas, and $16 \%$ of those living in rural areas. The symptoms were exacerbated for $90 \%$ in spring, $92 \%$ in summer, $70 \%$ in autumn, and $64 \%$ in winter (Table 1).

Cluster 2 included $58 \%$ males, $72 \%$ of $>15$ years old, $28 \%$ of < 15 years old, $57 \%$ with a family history of atopia, $76 \%$ with prevalent symptoms in urban areas, and $24 \%$ in rural areas. The symptoms were exacerbated for $80 \%$ in spring, $80 \%$ in summer, $64 \%$ in autumn, and $63 \%$ in winter (Table 1).

## Asthma

Patients who have asthma were divided into two clusters. Forty-six patients were assigned to Cluster 1 and 23 to Cluster 2. Cluster 1 included $47 \%$ males, $56 \%$ of $>15$ years old, $44 \%$ of $<15$ years old, $21 \%$ with a family history of atopia, $89 \%$ with prevalent symptoms in urban areas,
and $11 \%$ in rural areas. The symptoms were exacerbated for $84 \%$ in spring, $82 \%$ in summer, and $56 \%$ in autumn and winter (Table 1).

Cluster 2 included $65 \%$ males, $35 \%$ of $>15$ years old, $65 \%$ of $<15$ years old, $39 \%$ with a family history of atopia, $56 \%$ with prevalent symptoms in urban areas, and $44 \%$ in rural areas. The symptoms were exacerbated for $69 \%$ in spring, $65 \%$ in summer, $82 \%$ in autumn, and $78 \%$ in winter (Table 1).

## Asthma and allergic rhinitis

Patients who have asthma and allergic rhinitis at the same time were divided into two clusters. Sixty patients were assigned to Cluster 1 and 19 to Cluster 2. Cluster 1 included $57 \%$ males, $48 \%$ of $>15$ years old, $52 \%$ of < 15 years old, $36 \%$ with a family history of atopia, $84 \%$ with prevalent symptoms in urban areas, and $16 \%$ in rural areas. The symptoms were exacerbated for $78 \%$ in spring, $84 \%$ in summer, and $73 \%$ in autumn and winter (Table 1).

Cluster 2 included $61 \%$ males, $70 \%$ of $>15$ years old, $30 \%$ of $<15$ years old, $60 \%$ with a family history of atopia, $76 \%$ with prevalent symptoms in urban areas, and $24 \%$ in rural areas. The symptoms were exacerbated for $76 \%$ in spring, $73 \%$ in summer, $70 \%$ in autumn, and $61 \%$ in winter (Table 1).

## Cluster analysis results across clinical symptoms

 Allergic rhinitisThe most common symptoms in patients afflicted with allergic rhinitis in cluster 1 were rhinorrhea ( $97 \%$ ),
sneezing ( $96 \%$ ), itchy eyes ( $91 \%$ ), itchy nose ( $96 \%$ ), nasal congestion ( $81 \%$ ), itchy throat ( $81 \%$ ), red eyes ( $71 \%$ ), itchy skin ( $52 \%$ ), burning eyes ( $43 \%$ ) and others as summarized in Table 2.

The most common symptoms in patients afflicted with allergic rhinitis in cluster 2 were rhinorrhea ( $74 \%$ ), sneezing ( $65 \%$ ), itchy eyes ( $42 \%$ ), and itchy nose ( $33 \%$ ), nasal congestion ( $36 \%$ ), and others as summarized in Table 2.

## Asthma

The most common symptoms in patients who have asthma in cluster 1 were rhinorrhea ( $100 \%$ ), sneezing ( $93 \%$ ), itchy eyes ( $84 \%$ ), itchy nose ( $86 \%$ ), nasal congestion ( $96 \%$ ), itchy throat ( $50 \%$ ), red eyes ( $41 \%$ ), itchy skin ( $41 \%$ ) and others as summarized in Table 2.

The most common symptoms in patients who have asthma in cluster 2 were rhinorrhea ( $30 \%$ ), sneezing ( $17 \%$ ), itchy eyes ( $13 \%$ ), itchy nose ( $8 \%$ ), nasal congestion ( $47 \%$ ), and others as summarized in Table 2.

## Asthma and allergic rhinitis

The most common symptoms among patients afflicted with asthma and allergic rhinitis in cluster 1 were rhinorrhea ( $91 \%$ ), sneezing ( $83 \%$ ), itchy eyes ( $68 \%$ ), itchy nose $(68 \%)$, nasal congestion ( $71 \%$ ), itchy throat ( $58 \%$ ), red eyes ( $51 \%$ ), itchy skin ( $43 \%$ ), burning eyes ( $25 \%$ ) and others as summarized in Table 2.

Table 1. Demographic information in the different clusters in respiratory allergies

| Characteristics |  | Allergic rhinitis |  | Asthma |  | Allergic rhinitis \&asthma |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 702 | 382 | 46 | 23 | 60 | 19 |
|  |  | cluster1 | cluster2 | cluster1 | cluster2 | cluster1 | cluster2 |
| Gender | Male | 50\% | 58\% | 47\% | 65\% | 61\% | 57\% |
|  | Female | 50\% | 42\% | 53\% | 35\% | 39\% | 43\% |
| age>15 |  | 82\% | 72\% | 56\% | 35\% | 70\% | 48\% |
| age<15 |  | 18\% | 28\% | 44\% | 65\% | 30\% | 52\% |
| Family History of Atopia |  | 61\% | 57\% | 21\% | 39\% | 60\% | 36\% |
| Place of residence | Urban | 84\% | 76\% | 89\% | 56\% | 76\% | 84\% |
|  | Rural | 16\% | 24\% | 11\% | 44\% | 24\% | 16\% |
| Exacerbation Month | Spring | 90\% | 80\% | 84\% | 69\% | 76\% | 78\% |
|  | Summer | 92\% | 80\% | 82\% | 65\% | 73\% | 84\% |
|  | Fall | 70\% | 64\% | 56\% | 82\% | 70\% | 73\% |
|  | Winter | 64\% | 63\% | 56\% | 78\% | 61\% | 73\% |

The most common symptoms in patients afflicted with allergic rhinitis in cluster 2 were rhinorrhea (63\%), sneezing ( $63 \%$ ), itchy eyes ( $42 \%$ ), itchy nose ( $47 \%$ ), nasal congestion ( $31 \%$ ), itchy throat ( $36 \%$ ), red eyes ( $26 \%$ ), itchy skin ( $21 \%$ ), burning eyes ( $36 \%$ ) and others as summarized in Table 2.

## Main clinical symptoms and allergens in Respiratory

## Allergies

Various clinical symptoms and allergens extracted in the first cluster of diseases are presented in Table 3.

## Cluster analysis results for allergens

## Allergic rhinitis:

The most common allergens in patients afflicted with allergic rhinitis in cluster 1 were Salsola kali (86\%), pigweed
mix ( $71 \%$ ), tree mix ( $58 \%$ ), ash ( $56 \%$ ), and grass mix ( $53 \%$ ). Less common allergens are summarized in Table 3.

The most common allergens in patients afflicted with allergic rhinitis in cluster 1 were Salsola kali ( $82 \%$ ), pigweed mix ( $62 \%$ ), tree mix ( $45 \%$ ), ash ( $48 \%$ ), and grass mix ( $41 \%$ ). Less common allergens are summarized in Table 3.

## Asthma

The most common allergens in patients with asthma in cluster 1 were Salsola kali (56\%), pigweed mix (34\%), tree mix ( $41 \%$ ), D. pteronyssinus ( $52 \%$ ), and D. farinae ( $41 \%$ ). The rest are summarized in Table 3.

The most common allergens in patients with asthma in cluster 2 were Salsola kali (47\%), pigweed mix (34\%), tree mix (34\%), D. pteronyssinus (34\%), and D. farinae (21\%). The rest are summarized in Table 3.

Table 2. Frequency of Clinical symptoms in the different clusters of respiratory allergies

| Symptoms | Allergic rhinitis |  | Asthma |  | Allergic rhinitis \&asthma |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 702 | 382 | 46 | 23 | 60 | 19 |
| Symptoms | Cluster 1 | Cluster 2 | Cluster 1 | Cluster 2 | Cluster 1 | Cluster 2 |
| Rhinorrhea | $97 \%$ | $74 \%$ | $100 \%$ | $30 \%$ | $91 \%$ | $63 \%$ |
| Sneezing | $96 \%$ | $65 \%$ | $93 \%$ | $17 \%$ | $83 \%$ | $63 \%$ |
| Itchy eyes | $91 \%$ | $42 \%$ | $84 \%$ | $13 \%$ | $68 \%$ | $42 \%$ |
| Itchy nose | $96 \%$ | $33 \%$ | $86 \%$ | $8 \%$ | $68 \%$ | $47 \%$ |
| Nasal Congestion | $81 \%$ | $36 \%$ | $69 \%$ | $47 \%$ | $71 \%$ | $31 \%$ |
| Itchy throat | $81 \%$ | $19 \%$ | $50 \%$ | $17 \%$ | $58 \%$ | $36 \%$ |
| Red eyes | $71 \%$ | $11 \%$ | $41 \%$ | $13 \%$ | $51 \%$ | $26 \%$ |
| Itchy Skin | $52 \%$ | $15 \%$ | $41 \%$ | $4 \%$ | $43 \%$ | $21 \%$ |
| Sore throat | $35 \%$ | $29 \%$ | $5 \%$ | $13 \%$ | $40 \%$ | $42 \%$ |
| Sinus infection | $45 \%$ | $6 \%$ | $39 \%$ | $4 \%$ | $40 \%$ | $10 \%$ |
| Burning eyes | $43 \%$ | $8 \%$ | $23 \%$ | $0 \%$ | $25 \%$ | $36 \%$ |
| Sputum | $37 \%$ | $11 \%$ | $23 \%$ | $21 \%$ | $36 \%$ | $31 \%$ |
| Shortness of Breath | $32 \%$ | $12 \%$ | $52 \%$ | $15 \%$ | $57 \%$ | $33 \%$ |
| Swollen eyelids | $25 \%$ | $13 \%$ | $30 \%$ | $4 \%$ | $20 \%$ | $10 \%$ |
| Wheezing | $21 \%$ | $9 \%$ | $10 \%$ | $56 \%$ | $35 \%$ | $47 \%$ |
| Red spot | $21 \%$ | $7 \%$ | $23 \%$ | $4 \%$ | $16 \%$ | $21 \%$ |
| Anorexia | $19 \%$ | $8 \%$ | $8 \%$ | $8 \%$ | $20 \%$ | $31 \%$ |
| Heartburn | $17 \%$ | $3 \%$ | $10 \%$ | $8 \%$ | $20 \%$ | $15 \%$ |
| Regurgitation | $7 \%$ | $12 \%$ | $17 \%$ | $0 \%$ | $10 \%$ | $31 \%$ |
| Nausea | $13 \%$ | $3 \%$ | $6 \%$ | $13 \%$ | $11 \%$ | $26 \%$ |
| Chest pain | $12 \%$ | $2 \%$ | $6 \%$ | $4 \%$ | $15 \%$ | $36 \%$ |
| Stomachache | $10 \%$ | $3 \%$ | $6 \%$ | $13 \%$ | $8 \%$ | $26 \%$ |
| Vomit | $7 \%$ | $1 \%$ | $6 \%$ | $4 \%$ | $8 \%$ | $21 \%$ |
| Diarrhea | $6 \%$ | $1 \%$ | $10 \%$ | $4 \%$ | $8 \%$ | $5 \%$ |
| Blisters | $4 \%$ | $1 \%$ | $6 \%$ | $0 \%$ | $1 \%$ | $10 \%$ |

Table 3. The various clinical symptoms and allergens in the first cluster of respiratory allergies

| Clinical symptom/Allergen | Frequency |  |
| :--- | :---: | :---: |
|  | Allergic rhinitis | Asthma |
| Red eyes | $81 \%$ | $50 \%$ |
| Sore throat | $71 \%$ | $41 \%$ |
| Burning eyes | $35 \%$ | $5 \%$ |
| Sputum | $43 \%$ | $23 \%$ |
| Shortness of breath | $37 \%$ | $23 \%$ |
| Salsola kali | $32 \%$ | $52 \%$ |
| Pigweed mix | $86 \%$ | $56 \%$ |
| Ash | $71 \%$ | $34 \%$ |
| Grass mix | $56 \%$ | $19 \%$ |
| D. pteronyssinus | $53 \%$ | $15 \%$ |
| D.farinae | $24 \%$ | $52 \%$ |

## Asthma and allergic rhinitis

The most common allergens in patients with asthma in cluster 1 were Salsola kali ( $71 \%$ ), pigweed mix ( $48 \%$ ), tree mix (36\%), Ash (21\%), grass mix (26\%), D. pteronyssinus $(40 \%)$, and D. farinae ( $30 \%$ ). The rest are summarized in Table 3.

The most common allergens in patients with asthma in cluster 2 were Salsola kali ( $100 \%$ ), pigweed mix ( $94 \%$ ), tree mix (94\%), Ash (78\%), grass mix ( $84 \%$ ), D. pteronyssinus $(52 \%)$, and D. farinae ( $36 \%$ ). The rest are summarized in Table 4.

The most common allergens and clinical symptoms in different clusters are illustrated in Figure 1.


Figure 1. Frequency of the common symptoms and allergens in the different clusters of respiratory Allergies

Table 4. Frequency of Allergens in the different clusters of respiratory allergies

| Allergens |  | Allergic rhinitis |  | Asthma |  | Allergic rhinitis \& asthma |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 702 | 382 | 46 | 23 | 60 | 19 |
|  |  | Cluster 1 | Cluster 2 | Cluster 1 | Cluster 2 | Cluster 1 | Cluster 2 |
| Outdoor Allergens | Salsola kali | 86\% | 82\% | 56\% | 47\% | 71\% | 100\% |
|  | Pigweed mix | 71\% | 62\% | 34\% | 34\% | 48\% | 94\% |
|  | Tree mix | 58\% | 45\% | 41\% | 34\% | 36\% | 94\% |
|  | Ash | 56\% | 48\% | 19\% | 21\% | 21\% | 78\% |
|  | Grass mix | 53\% | 41\% | 15\% | 26\% | 26\% | 84\% |
|  | Birch tree | 9\% | 7\% | 2\% | 0\% | 8\% | 0\% |
| Indoor Allergens | D. pteronyssinus | 24\% | 25\% | 52\% | 34\% | 40\% | 52\% |
|  | D. farinae | 25\% | 26\% | 41\% | 21\% | 30\% | 36\% |
|  | Alternaria alternate | 17\% | 14\% | 21\% | 13\% | 16\% | 36\% |
|  | Aspergillus mix | 11\% | 8\% | 21\% | 0\% | 10\% | 31\% |
|  | Cockroach | 6\% | 7\% | 0\% | 8\% | 1\% | 5\% |
|  | Candida Albicans | 0\% | 0\% | 4\% | 0\% | 0\% | 0\% |
|  | Feather | 0\% | 0\% | 0\% | 0\% | 0\% | 0\% |

## DISCUSSION

The present research aimed to extract the common symptoms of allergic rhinitis and asthma through cluster analysis and the k -means algorithm. There were 1,881 patient participants in this research who visited the asthma and allergy clinic of Mashhad University of Medical Sciences. Among them, 1,231 (65.4\%) were diagnosed with allergic rhinitis, asthma, or both, thus meeting the inclusion criterion. Most participants were diagnosed with allergic rhinitis ( $\mathrm{n}=1,083,88 \%$ ). In the three disease groups (allergic rhinitis, asthma, and both), most patients belonged to cluster 1 . Those in cluster 1 were the most homogenous (similar to each other) in terms of the symptoms and allergens. The most different symptoms and allergens in the two groups of patients included itchy throat, red eyes, sore throat, Salsola kali, Pigweed mix, Ash, D. pteronyssinus, and D. farinae.

Epidemiologic and pathophysiologic studies showed that allergic rhinitis and asthma are very similar and even occur often at the same time. Also, allergic rhinitis increases the risk of affliction with asthma $(5,10)$. Seventynine patients were simultaneously afflicted with allergic rhinitis and asthma in this study. After the cluster analysis, most of these patients were assigned to cluster 1. The most common symptoms were running nose, sneezing, itchy nose, and eyes. Moreover, patients in this cluster were mostly allergic to the Salsola kali allergen.

Amizadeh et al. and Kashef et al. $(11,12)$ reported no significant correlation between gender and the effect of allergens. However, in several other studies $(10,13,14)$, allergy was significantly more prevalent in men than women. Yet, Klossek et al. reported a higher prevalence of allergic rhinitis in women than men (15). However, in the present study, respiratory allergic diseases prevailed more in men.

The incidence rate of allergic rhinitis has been increasing within the past few decades, influenced by the global rising incidence of atopia (5). The frequency of the family history of atopia was found to be $58.5 \%$ in patients with allergic rhinitis, $59.4 \%$ in patients with asthma, and
$54.4 \%$ in those with asthma and allergic rhinitis at the same time. The present findings were not consistent with the rate of $85 \%$ reported by Farrokhi et al. in the west-south of Iran (10). These divergent findings can probably be due to the different geographical areas.

According to Farrokhi et al., the prevalence of allergens is significantly higher in hot and humid areas. Thus, people are at a higher risk of internal allergies, especially in summer and winter (10). In this study, the frequency of respiratory allergic symptoms in summer was consistent with Farrokhi et al.'s study, yet different in winter. This divergence can potentially be due to different climates. Also, in the present study, the frequency of allergic rhinitis was higher in spring due to pollination.

In the present research, in the three groups of diseases, in cluster 1 , the most frequent symptoms were runny nose, sneezing, and itchy eyes and nose. The least frequent symptoms were metabolic symptoms such as diarrhea, nausea, and vomiting. Similarly, Amizadeh et al. and Farrokhi et al. mentioned runny nose as the most common symptom in allergic rhinitis patients $(10,11)$. Besides, Klossek et al. and Sapsaprang et al. reported eye symptoms as the most common symptom in allergic patients $(15,16)$. The findings of these four studies are consistent with the present research. Yet, in the study conducted by Ghaffari et al., coughing, shortness of breath, and wheezing were reported as the most common symptoms of asthma (17).

The prevalence of air allergens differs across the weather types (10). In this study, in both clusters of asthma and allergic rhinitis groups, the most common allergens were Salsola kali, weed mix, and tree mix. Similarly, many studies reported Salsola kali as the most common allergen (18-20). In a study in Shiraz, Kashef et al. found weed mix as the most common allergen (12), which is consistent with the present finding. However, Farrokhi et al. reported mold as the most common factor, probably related to humidity in that geographic area (10). Also, Khazaei et al. found mites and fungi as the most common symptoms in Sistan Baluchestan (21).

In the present research, in the group of patients with asthma and allergic rhinitis together, though the fewest patients belonged to cluster2, the most commonality was that of Salsola, weed mix, and tree mix allergens.

## Limitations of study and suggestions for further research

Though the present descriptive research was conducted through a cluster analysis, diseases such as allergic rhinitis and asthma are geography-dependent. Their prevalence in a geographic area may differ across cities. Therefore, further research is required in more geographic areas to identify the most common symptoms and allergens.

For future research, models can be developed using algorithms such as artificial neural networks or decision trees to speed up the diagnostic process.

## CONCLUSION

Overall, Salsola kali was the most common allergen in allergic rhinitis and asthma patients. Also, the most common symptoms in patients were rhinorrhea, sneezing, itchy nose, and nasal congestion. This study can help physicians diagnose allergic rhinitis and asthma in geographical areas with a high prevalence of Salsola kali.

## Conflict of Interest

The authors have no conflicts of interest.

## Funding

This study was supported by the Vice Chancellor for Research, the Student Research Committee of Kerman University of Medical Sciences (grant No.99000605).

## Acknowledgments

The authors would like to appreciate the Clinical Research Development Unit, Ghaem Hospital, Mashhad University of Medical Sciences, Mashhad, Iran for their assistance in the present manuscript. Also, we would like to gratefully acknowledge the contribution of all patients who voluntarily participated in this study.

## REFERENCES

1. Mir E, Panjabi C, Shah A. Impact of allergic rhinitis in school going children. Asia Pac Allergy 2012;2(2):93-100.
2. Min YG. The pathophysiology, diagnosis and treatment of allergic rhinitis. Allergy Asthma Immunol Res 2010;2(2):65-76.
3. Kim H, Bouchard J, Renzi PM. The link between allergic rhinitis and asthma: a role for antileukotrienes? Can Respir J 2008;15(2):91-8.
4. Dykewicz MS, Fineman S. Executive Summary of Joint Task Force Practice Parameters on Diagnosis and Management of Rhinitis. Ann Allergy Asthma Immunol 1998;81(5 Pt 2):463-8.
5. Mohammadzadeh I, Barari-Savadkoohi R, Alizadeh-Navaei R. The prevalence of allergic rhinitis in Iranian children: A systematic review and descriptive meta-analysis. Journal of Pediatrics Review 2013;1(2):19-24.
6. Payandeh P, Fadaee J, Jabbari Azad F, Bakhshaii M, Sistani S. Allergens Prevalence among Patients with Respiratory Allergies in Mashhad, Iran. Tanaffos 2019;18(2):133-41.
7. Burte E, Bousquet J, Varraso R, Gormand F, Just J, Matran R, et al. Characterization of Rhinitis According to the Asthma Status in Adults Using an Unsupervised Approach in the EGEA Study. PLoS One 2015;10(8):e0136191.
8. Jun YJ, Jung J, Lee HM. Medical data science in rhinology: Background and implications for clinicians. Am J Otolaryngol 2020;41(6):102627.
9. Sistani S, Norouzi S, Hassibian MR, Tara M, Tabesh H, Hasibian S, et al. The discovery of major heart risk factors among young patients with ischemic heart disease using kmeans techniques. Int Cardiovasc Res J 2019;13(3):85-90.
10. Farrokhi S, Gheybi MK, Movahed A, Tahmasebi R, Iranpour D, Fatemi A, et al. Common aeroallergens in patients with asthma and allergic rhinitis living in southwestern part of Iran: based on skin prick test reactivity. Iran J Allergy Asthma Immunol 2015;14(2):133-8.
11. Amizadeh M, Safizadeh H, Bazargan N, Farrokhdoost Z. Survey on the prevalence of allergic rhinitis and its effect on the quality of high school students' life. Iran J Otorhinolaryngol 2013;25(71):79-84.
12. Kashef S, Kashef MA, Eghtedari F. Prevalence of aeroallergens in allergic rhinitis in shiraz. Iran J Allergy Asthma Immunol 2003;2(4):185-8.
13. Abbasi Ranjbar Z. Prevalence of allergic rhinitis among children in Rasht. Journal of Guilan University of Medical Sciences 2005;14(53):56-62.
14. Sattar HA, Mobayed H, al-Mohammed AA, Ibrahim AS, Jufairi AA, Balamurugan P, et al. The pattern of indoor and outdoor respiratory allergens in asthmatic adult patients in a humid and desert newly developed country. Eur Ann Allergy Clin Immunol 2003;35(8):300-5.
15. Klossek JM, Annesi-Maesano I, Pribil C, Didier A. Un tiers des adultes ont une rhinite allergique en France (enquête INSTANT) [INSTANT: national survey of allergic rhinitis in a French adult population based-sample]. Presse Med 2009;38(9):1220-9.
16. Sapsaprang S, Setabutr D, Kulalert P, Temboonnark P, Poachanukoon O. Evaluating the impact of allergic rhinitis on quality of life among Thai students. Int Forum Allergy Rhinol 2015;5(9):801-7.
17. Ghaffari J, Khademloo M, Saffar MJ, Rafiei A, Masiha F. Hypersensitivity to house dust mite and cockroach is the most common allergy in north of iran. Iran J Immunol 2010;7(4): 234-9.
18. Fereidouni M, Hossini RF, Azad FJ, Assarehzadegan MA, Varasteh A. Skin prick test reactivity to common aeroallergens among allergic rhinitis patients in Iran. Allergol Immunopathol (Madr) 2009;37(2):73-9.
19. Mahram M, Barikani A, Nejatian N. The frequency of common allergens in allergic rhinitis among the patients referred to the allergy clinic of Qods Hospital in Qazvin during 2007-2010. J Aller Ther 2013;4(130):2.
20. Mahboubi Oskouei Y, Farid Hosseini R, Ahanchian H, Jarahi L, Ariaee N, Jabbari Azad F. Report of Common Aeroallergens among Allergic Patients in Northeastern Iran. Iran J Otorhinolaryngol 2017;29(91):89-94.
21. Khazaei HA, Hashemi SR, Aghamohammadi A, Farhoudi F, Rezaei N. The study of type 1 allergy prevalence among people of South-East of iran by skin prick test using common allergens. Iran J Allergy Asthma Immunol 2003;2(3):165-8.
