American Journal of
Biometrics & Biostatistics

Research Article

Evaluation of Inter Related Biological and Envioronmental Risk Factors of Asthma (onset/ Severity): Using Structural Equation Modeling -

Noreen Akhtar*

Govt postgraduate college, satellite town, Gujranwala, Pakistan

***Address for Correspondence:** Noreen Akhtar,Govt postgraduate college, satellite town, Gujranwala, Pakistan,Tel: +923-348-196-142; ORCid: 0000-0001-9936-6456; E-mail: noreenali.stats@gmail.com

Submitted: 22 July 2020; **Approved:** 16 November 2020; **Published:** 17 November 2020

Cite this article: Akhtar N. Evaluation of Inter Related Biological and Envioronmental Risk Factors of Asthma (onset/ Severity): Using Structural Equation Modeling . American J Biom Biostat. 2020 Nov 17;4(1): 013-024.

Copyright: © 2020 Akhtar N. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

The aim of this study was to evaluate some inter-connected environmental and biological factors which contribute to onset of asthma and its severity level, duration between asthmatic attacks and response of therapy. A cross-sectional study of asthmatic patient's from gujrat (Pakistan) is conducted who were visiting respiratory disease specialists at different hospitals of city. Sample of size 200 is selected. Multivariate analysis technique i-e structural equation modeling is applied to find out the interrelationship of all the factors. Before applying structure equation modeling, factor analysis is used to confirm factors to be used in path diagram as constructs. Goodness of fit index was obtained (GFI = 0.908, NNFI = 0.901, $\chi 2/d.f$ = 1.607, CFI = 0.921, RMSEA = 0.053). The main conclusion got by applying SEM is that there is direct effect of some factors as well as, there are a lot of factors that are attached behind them.

INTRODUCTION

Many studies have been conducted to define asthma for its impact on lung function i.e., airway inflammation, its reversibility, and airway hyper-responsiveness. However their attempts to define asthma have been discouraged by the lack of understanding of mechanisms involved. If we look at the functional loss of airway inflammation then we can give more detailed and simple description of asthma [1]. In area where this study has been conducted most of inhabitants whether they are patients of asthma or healthy ones, don't have knowledge about how to take Asthma management steps. This research will provide information and awareness about those inter-connected risk factors; they are facing in routine life. Being asthmatics these factors may contribute to worsen their condition e.g. increased attacks and/or may affect response of therapy/medication. This research has been conducted to provide awareness to asthmatics and their carers as well.

A single framework to explain asthma disparities among people would be very complex and complicated if all the dimensions are to be studied to make a single framework to explore their connection with each other, but a few of them are combined here in this paper and a framework is developed which defines that asthma is not just a hereditary disorder but also it prevails by economic well being associated with the social context of people's lives [2] in different countries as well as in Pakistan. Because there is direct effect of some factors as well as, there are lot of factors that are attached behind them [3]. These factors include socio-economic effects [4], genetic tendency to the development of asthma [5], airway hyper-responsiveness [6], Gender disparities and some other biological factors [7]. Also environmental factors contribute to prevalence of asthma. Some environmental factors modify the chances to happening that asthma will develop in prone individuals and some environmental factors can also lead towards making asthma worse; these are also called sudden/ precipitate factors**.** Allergens and occupational factors (tobacco, secondhand smoking, air pollution etc) have been considered the most important causes of asthma; in fact they could first sensitize the airways and then keep asthma severe by sudden asthma attacks or leading to long term continuance of symptoms, even after its effect have been controlled. Exposure to an allergen is an important risk factor for sensitization to that particular allergen like pollens, dust mites etc [8] and exposure to allergens in sensitized individuals is a risk factor for making already bad or problematic situation (i-e. asthma) worse and leads to long term continuation of symptoms [9].

In study area it has been observed and according to the physicians, asthma is basically a hereditary disorder but a lot of other factors that are discussed in this study, contributes to make it bitter/severe. In United States, patterns of hospital admissions and emergency room visits show that low income of asthmatics was associated with an increase in mortality and morbidity rates [10]. Even the precision of relationship between socio-economic factors and asthma may vary from country to country especially in developing countries like Pakistan, India and etc. Environmental disadvantages, Indoor/ Outdoor allergens, hyperventilation, extreme emotional expression and family size [3]. So this study would prove a valuable source of awareness to people. This may help them to reduce exposure to asthma risk factors in home environment as well as to reduce exposure to asthma risk factors in outdoor environment. And will result in increased access to asthma care.

The interaction between exposure to psychological & environmental disadvantageous conditions as well as molecular and genetic disorders i.e. biomedical framing is studied, with influence of socioeconomic status i.e. ecological framing [2]. To consider this, data is collected on biological factors as well as on socio-economic status and contributing factor i.e. stress and violence, which are rising important risk factors [11].Multivariate approach, structural equation modeling is used for model fitting. Descriptive statistics of the responses of patients on several indicators of factors is also shown in first section. Second section is about comparison of all the factors with respect to demographic characteristics like age, gender, area, because these characteristics had been considered as being associated [12] with variation in asthma expression among patient to patient [13,14].

STUDY POPULATION

It is cross sectional study. A questionnaire was used to take responses of patients. First half of the questionnaire was on demographic factors and 2nd half is on risk factors that are included in path diagram. Since whole population of asthmatics in gujrat city can't be dealt easily. It's a time consuming and expensive process to deal with whole target population. So sample of 200 patients is selected. A non-probability sampling technique is used. Because there was no record of registered total asthmatic patients in city found. Data of asthma patients is collected by visiting different hospitals including urban and rural areas. Sample of those patients is included who visits regularly for treatment/therapy

CONCEPTUAL FRAMEWORK

Family history of asthma contributes in developing asthma at early age. Also environmental factors, allergens along with genetics tendency are contributing risk factors [15]. It is observed that there is association between exposure to domestic violence and asthma prevalence [16]. Sensitization to allergens may lead to development of asthma [17]. There is great role of psychosocial stressors that influence biological mechanism involved in asthma through which stress can contribute to asthma causation and also that these risk factors are involved interdependently to cause asthma or to trigger its severity, to increase number of minor/major attacks or to slower down the response of therapy [2]. The interaction between genetics and environment to which an individual is exposed is to be studied which helps explaining asthma variations among individuals in society with respect to different inequalities.

A proposed framework of inter-related risk factors supported by Wright et al, [2].

ANALYTICAL APPROACH

The strategies used are given below.

Factor analysis

Factor analysis is a broad term for multivariate statistical methods used to identify common underlying variables called factors within a larger set of measures. Basically, factor analysis determines which group of variables goes together. A factor is a group of related variables representing an underlying domain or theme. Factors are indicated by covariances among two or more variables.

Factor analysis empirically explores the interrelationships and dimensions among variables to cluster inter-correlated variables into smaller sets of basic factors. It reduces the number of variables and also classifies variables by exploring the underlying theoretical structures.

Confirmatory Factor Analysis (CFA) and Exploratory Factor Analysis (EFA) are powerful statistical techniques. The questionnaire identifies the factor structure or what we think it is. However, some questions may not measure what we thought they should. If the factor structure is not confirmed, EFA is the next step [18]. EFA helps us determine what the factor structure looks like according to how participant responses. Exploratory factor analysis is essential to determine underlying constructs for a set of measured variables.

Structural equation modeling

The structural equation modeling techniques are used to study relations among variables. The relations are typically assumed to be linear. Structural equation modeling is a multivariate statistical analysis technique that is used to analyze structural relationships. It is the combination of factor analysis and multiple regression analysis, and it is used to analyze the structural relationship between measured variables and latent constructs. The structural regression models are similar to confirmatory factor analysis models as it is used to test whether the hypothesized relationship between constructs in path diagram is significant. Since regression models are one-way i.e. pointing arrow from one construct to dependent construct. In structural regression models, the sum of all the responses of that specific constructs are involved whose relation is to be tested. And using simple linear regression analysis significance of relationship is tested by relying on *p*-value. Structural regression models are made by breaking two-way or other forms of relationships between constructs. As in path model type of relations that exist are likewise that at a time a construct is dependent but on the other turn it might be possible that it is playing role of independent construct/variable.

RESULTS AND DISCUSSIONS

Factors which are considered in proposed framework includes the interaction between exposure to psychological & environmental disadvantageous conditions as well as molecular and genetic disorders i.e. biomedical framing is studied, with influence of socioeconomic status i.e. ecological framing [2].

Data is collected on biological factors as well as social-status variables which are cause of variation in asthma expression and risks among individuals and also cause variation in social exposures i.e. stress and violence, which are rising as important risk factors [11]. To analyze it is used to apply statistical techniques, which are factor analysis and structural equation modeling. Descriptive statistics of the responses of patients on several indicators of those factors are also shown in first section.

Second section is about comparison of all the factors with respect to demographic characteristics like age, gender, area, because these characteristics had been considered as being associated with variation in asthma expression from patient to patient [12-15]. Since other factors also contribute to cause asthma (onset, severity, duration and response to therapy), So tests are applied for those factors also, to check whether these varies for different age groups, gender differences and area of residence of patients in Gujrat city (i.e. rural, urban areas) of patients whose sample has been taken for data analysis.

Descriptive statistics & normality test

Patients of each gender category are included in data. Among total respondents there are 40 percent males and 60 percent female patients. Among 200 patients, there are 39 percent patients who are residents of urban areas and 61 percent are rural area residents.

People are living with poor socio-economic conditions. We can test normality by using various descriptive statistics which are (i) the relative values of the mean, median (middle value) and mode (most frequent value) (ii) The skewness (iii) Kurtosis (iv) A test for normality (Kolmogorov-Smirnov).

It is rare to get perfectly normally distributed data. We want to fulfill the criteria of our assumptions fairly. For example, if data look almost bell-shaped and have low levels of kurtosis and skewness, we wish to accept the alternative hypothesis. As we can see in table 1, for age variable, mean age is found to be 40.74, skewness of the data is .567, which indicates a positive skewness, which is moderate. The kurtosis is -.176. SPSS measures kurtosis around 0, where 0 relates to normal (mesokurtic) position [16-18].

In order to find out if the distribution of data is normal, kolmogrov-simrnov test is used. The null hypothesis here would be that test distribution is normal. It is shown in Table 2 for age the *p*-value 0.017 is less than significance level of 0.05, so accepting this confidence level, the null hypothesis is rejected in favor of the alternative hypothesis and the data is therefore not normally distributed w.r.t age variable. Furthermore normality of total score of asthma expression is being checked and also *p*-value (0.015) is less than 0.05 so the null hypothesis is rejected with the conclusion that test distribution is not normal. Genetics and physiological disruption is being tested for normality where *p*-value (0.006) is less than 0.05 so the null hypothesis is rejected and concluded that the test variable is not normally distributed. Exposure to Indoor-outdoor allergens, childhood status is not normally distributed with p-value of 0.036. Exposure to psychological stress, total score of socioeconomic status, total score of current family status, total score of neighborhood, total score of collective efficacy, total score of behavioral factors, and total score of physical condition is normally distributed.

Overall table 2 shows that Genetics and physiological disruption, total score of asthma expression Exposure to Indoor-outdoor allergens, childhood status are showing non normality. Exposure to psychological stress, total score of socioeconomic status, total score of current family status, total score of neighborhoods, total score of collective efficacy, total score of behavioral factors, total score of physical condition are normally distributed.

Table 1: Descriptive statistics.

N Valid	**Missing**	200 0
Mean		40.74
Std. Error of Mean		1.235
Median		40.00
Mode		50
Std. Deviation		17.465
Variance		305.028
Skewness		.567
Std. Error of Skewness		.172
Kurtosis		-.176
Std. Error of Kurtosis		.342
Range		86
Minimum		10
Maximum		96
Percentiles **50** **75**	**25**	25.00 40.00 53.00

Table 2: Checking normality by Kolmogrov-simrnov test method variables.

Variables	Kolmo-grov-sim-rnov	Asympt. sig(2-tailed)
Age	1.545	0.017
Total Score of Socio-economic Status	1.339	0.056
Total Score of Asthma	1.568	0.015
Total Score of Behavioral Factors	1.072	0.201
Total Score of Genetics & physiological Disrup	1.694	0.006
Total Score of Current Family Status	0.814	0.522
Total Score of in/Out Door Allergens	1.416	0.036
Total Score of Childhood Status	2.495	0.000
Total Score of Exposure to Psychological Stress	1.234	0.095
Total Score of Physical Condition	1.062	0.209
Total Score of Collective Efficacy	1.059	0.212
Total Score of Neighborhood Context	0.939	0.341

CONFIRMATORY FACTOR ANALYSIS

CFA is used to test the hypothesis that a relationship between the observed variables and their underlying latent constructs exists. The researcher uses knowledge of the theory, empirical research, or both, hypothesized the relationship pattern a priori and then tests the hypothesis statistically.

Traditional statistical methods normally utilize one statistical test to determine the significance of the analysis. However, Structural Equation Modeling (SEM), CFA specifically, relies on several statistical tests to determine the adequacy of model fit to the data.

In my study, proposed factors which are considered are:

- Asthma (onset, severity, duration b/w attacks)
- Behavioral Factors
- Genetics and Physiological Disruption
- Current Family Status
- Childhood Status
- Exposure to psychological stress
- Exposure to In/Out door Allergens
- Neighborhood context
- Socio-economic Status
- Physical Condition
- Collective Efficacy

Using Statistica 7.0 version, Confirmatory factor analysis is being performed whose results are given in next section.

Construct validity

The chi-square test indicates the amount of difference between expected and observed covariance matrices. A chi-square value close to zero indicates little difference between the expected and observed covariance matrices. In addition, the probability level must be greater than 0.05 when chi-square is close to Zero.

The Comparative Fit Index (CFI) is equal to the discrepancy function adjusted for sample size. It is also known as Bentler Comparative Fit Index. CFI compares the existing model fit with a null model. Where null model contains the assumption that the latent variables in the model are uncorrelated (the "assumption of independence"). CFI ranges from 0 to 1 with a larger value indicating better model fit. Acceptable model fit is indicated for a CFI value of 0.90 or greater [19].

Root Mean Square Error of Approximation (RMSEA) is related to residual in the model. RMSEA value ranges from 0 to 1 with a smaller RMSEA value indicating better model fit. If its value is 0.08 then it means model is just moderately fit. Acceptable model fit is indicated by an RMSEA value of 0.06 or less [19].

The Goodness of Fit Index (GFI) is a measure of fit between the hypothesized model and the observed covariance matrix. The Adjusted Goodness of Fit Index (AGFI) corrects the GFI, which is affected by the number of indicators of each latent variable.

AGFI is a variant of GFI which adjusts GFI for degrees of freedom: the quantity (1 - GFI) is multiplied by the ratio of your model's df divided by df for the baseline model, and then AGFI is 1 minus this result. AGFI's use has been declining and it is no longer considered a preferred measure of goodness of fit.

With indication of model fit, the parameter estimates are examined. The ratio of each parameter estimate to its standard error is distributed as a z statistic and is significant at the 0.05 level if its value exceeds 1.96 and at the 0.01 level it its value exceeds 2.56.

After performing analysis values which are to be considered normally are shown in table 3 & table 4.

Factor 1: Asthma (AE)

GFI = 0.887, CFI = 0.768, RMSEA = 0.143, Chi-square = 95.6073, D.F = 20, *p*-value = 0.00

We see that the chi-square is rejected. The RMSEA is 0.1 which indicates mediocre fit. The CFI is 0.768 and the GFI is 0.887, not at the threshold of 0.95 and 0.90.It is not confirming to the general criteria because to get confirmed, all criteria's given before should be fulfilled.

Note: In CFA, if unacceptable model fit is found, an EFA can be performed [18].

Factor2: Behavioral Factors (BF)

GFI = 0.917, CFI = 0.698, RMSEA = 0.133, Chi-square = 68.4782, D.F = 14, p-value = 0.00

Factor2 i.e. behavioral factors is also not confirming to the general criteria. So we will move towards EFA.

Factor3: Genetics and Physiological Disruption (G&PD)

GFI = 0.982, CFI = 0.997, RMSEA = 0.023, Chi-square = 11.1858,D.F = 10, p-value = 0.3432

Factor3 is confirming to all general criteria of confirmatory factor analysis.

Factor4: Current Family Status (CFS)

GFI = 0.909, CFI = 0.913, RMSEA = 0.102,Chi-square = 82.2668,D.F = 29, p -value = 0.000

Factor4 i.e. current family status is not confirming to the general criteria. So we will need to move towards EFA.

Factor5: Childhood Status (CS)

GFI = 0.998, CFI = 1.000, RMSEA = 0.000, Chi-square = 0.64556, D.F = 2, p -value = 0.724

Factor5 is confirming to all general criteria of confirmatory factor analysis so there is no need to check through exploratory factor analysis.

Factor6: Exposure to Psychological Stress (EPS)

GFI = 0.975, CFI = 0.964, RMSEA = 0.059,Chi-square = 15.288,D.F = 9, p -value = 0.083

Factor6 is confirming to all general criteria of confirmatory factor analysis so there is no need to check through exploratory factor analysis.

Factor 7: Exposure to In\out door Allergens (I/O)

GFI = 0.940, CFI = 0.929, RMSEA = 0.075, Chi-square = 54.265,D.F = 27, p -value = 0.001

Factor7 i.e. In/Out door Allergens is not confirming to the general criteria. So we will need to move towards EFA.

Factor 8: Neighborhood context (NC)

GFI = 0.933, CFI = 0.822, RMSEA = 0.078, Chi-square = 66.7592, D.F = 2, p -value = 0.000

Factor8 i.e. Neighborhood context is not confirming to the general criteria. So we will need to move towards EFA.

Factor 9: Socio-economic Status (SES)

Table 3: single sample fit indices for factors.

Factors	AE	BF	G&PD	CFS	CS	I/O	EPS	SES	CE	PC	NBH
Fit indices	0.887	0.917	0.984	0.906	0.998	0.940	0.975	0.903	0.748	0.976	0.957
Joreskog GFI	0.797	0.834	0.953	0.811	0.992	0.900	0.942	0.839	0.411	0.928	0.914
Joreskog AGFI	0.641	0.485	0.140	0.471	0.084	0.454	0.197	0.661	1.118	0.159	0.316
Akaike Info- Criterion	0.906	0.717	0.305	0.703	0.216	0.752	0.396	0.959	1.317	0.325	0.548
Schwarz's Bayesian Criterion	0.649	0.491	0.143	0.477	0.086	0.463	0.202	0.670	1.122	0.162	0.322
Browne Cudeck Cross Validation	353.20	201.12	456.22	630.19	218.00	417.56	187.6	377.04	557.99	116.89	216.28
Independence Model Chi-Square	28.000	21.000	10.000	21.000	6.000	36.000	15.00	36.00	15.000	10.000	21.000
Independence Model df	0.729	0.660	0.983	0.896	0.997	0.870	0.919	0.747	0.644	0.901	0.838
Bentler-Bonett Normed Fit	0.674	0.546	0.987	0.872	1.019	0.904	0.939	0.732	0.418	0.876	0.839
Bentler-Bonett Non-Normed Fit Index	0.768	0.698	0.994	0.915	1.000	0.929	0.964	0.799	0.651	0.938	0.893
Bentler Com-parative Fit	0.521	0.440	0.491	0.597	0.332	0.653	0.551	0.560	0.387	0.450	0.559
James-Mulaik-Brett Parsi-monious Fit	0.621	0.489	0.966	0.843	0.991	0.827	0.864	0.662	0.407	0.801	0.758
Bollen's Rho	0.773	0.709	0.994	0.916	1.006	0.930	0.965	0.804	0.655	0.941	0.896

Table 4: Basic summary statistics.

Factors Fit indices	AE	BF	G&PD	CFS	CS	IO	EPS	SES	CE	PC	NBH
Discrepancy Function	0.480	0.344	0.039	0.331	0.003	0.273	0.077	0.480	0.997	0.058	0.176
Maximum Residual Cosine	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Maximum Absolute Gradient	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ICSF Criterion	0.000	0.000	0.000	-0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ICS Criterion	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ML Chi Square	95.607	68.478	7.803	65.784	0.646	54.265	15.29	95.45	198.47	11.627	34.955
Degrees of Freedom	20.000	14.000	5.000	14.000	2.000	27.000	9.000	27.00	9.000	5.000	14.000
p-level	0.000	0.000	0.167	0.000	0.724	0.001	0.083	0.000	0.000	0.040	0.001
RMS Standar-dized Residual	0.104	0.097	0.024	0.056	0.011	0.060	0.050	0.081	0.200	0.057	0.061

GFI = 0.903, CFI = 0.799, RMSEA = 0.113, Chi-square = 95.45, D.F = 27, p -value = 0.000

Factor9 i.e. Socio-economic Status is not confirming to the general criteria. So we will need to move towards EFA.

Factor 10: Physical Condition (PC)

GFI = 0.976,CFI = 0.938, RMSEA = 0.08,Chi-square = 11.627,D.F = 5, p -value = 0.040

Factor10 i.e. physical condition is confirming to the general criteria. So we will need to move towards EFA.

Factor 11: Collective Efficacy (CE)

GFI = 0.784, CFI = 0.651, RMSEA = 0.328, Chi-square = 198.469,D.F = 9, p -value = 0.000

Factor11 i.e. collective efficacy is not confirming to the general criteria. So we will need to move towards EFA.

Note: those factors that are not confirmed for them we will run exploratory factor analysis.

In CFA, if unacceptable model fit is found, an EFA can be performed [18].

The factors which are not confirmed for them we will run exploratory factor analysis.

Exploratory factor analysis

Exploratory Factor Analysis (EFA) could be described as orderly simplification of interrelated measures. EFA has been used to explore the possible underlying factor structure of a set of observed variables without imposing a preconceived structure on the outcome [20]. By performing EFA, the underlying factor structure is identified.

Two tests tell us that whether data is suitable for applying factor analysis. The Kaiser-Meyer Olkin Measure of Sampling Adequacy is a statistic which indicates the proportion of variance in your variables which is common variance, i.e. which might be caused by underlying factors. It should contain high value say near 1.0 to ensure that a factor analysis may be useful for our data. A value of .6 is a suggested minimum. Also an identity correlation matrix shows that our variables are independent. And Bartlett's test of sphericity is used for testing whether our correlation matrix is identity or not. An identity matrix is matrix in which all of the diagonal elements are 1 and all off diagonal elements are 0. You want to reject this null hypothesis. The significance level indicates that whether there is significance relation between variables or not. If p value is less than specified level, it means there is significant relationship between variables. If significance value is 0.10 or near, then it indicates that our data is not suitable for factor analysis. Communalities table shows the proportion of each variable's variance that can be explained by the factors. It is also noted as h^2 and can be defined as the sum of squared factor loadings for the variables. Initial with principal factor axis factoring, the initial values on the diagonal of the correlation matrix are determined by the squared multiple correlation of the variable with the other variables. Extraction – The values in this column indicate the proportion of each variable's variance that can be explained by the retained factors. Variables with high values are well represented in the common factor space, while variables with low values are not well represented. These are the reproduced variances from the factors that you have extracted. We can find these values on the diagonal of the reproduced correlation matrix. Initial Eigenvalues eigenvalues are the variances of the factors. Because we conducted our factor analysis on the correlation matrix, the variables are standardized, which means that the each variable has a variance of 1, and the total variance is equal to the number of variables used in the analysis. Total this column contains the eigenvalues. The first factor will always account for the most variance (and hence have the highest eigenvalue), and the next factor will account for as much of the left over variance as it can, and so on. Hence, each successive factor will account for less and less variance. Percentage of Variance this column contains the percent of total variance accounted for by each factor. Cumulative percentage this column contains the cumulative percentage of variance accounted for by the current and all preceding factors. Extraction Sums of Squared Loadings the number of rows in this panel of the table correspond to the number of factors retained. Rotation Sums of Squared Loadings the values in this panel of the table represent the distribution of the variance after the varimax rotation. Varimax rotation tries to maximize the variance of each of the factors. The scree plot graphs the eigenvalues against the factor number. Factor Matrix this table contains the unrotated factor loadings, which are the correlations between the variable and the factor. Because these are correlations, possible values range from -1 to +1. Reproduced Correlations this table contains two tables, the reproduced correlations in the top part of the table, and the residuals in the bottom part of the table. Reproduced Correlation the reproduced correlation matrix is the correlation matrix based on the extracted factors. You want the values in the reproduced matrix to be as close to the values in the original correlation matrix as possible. This means that the residual matrix, which contains the differences between the original and the reproduced matrix to be close to zero. If the reproduced matrix is very similar to the original correlation matrix, then you know that the factors that were extracted accounted for a great deal of the variance in the original correlation matrix, and these few factors do a good job of representing the original data. The numbers on the diagonal of the reproduced correlation matrix are presented in the communalities table in the column labeled Extracted. Residual represents the differences between original correlations (shown in the correlation table at the beginning of the output) and the reproduced correlations, which are shown in the top part of this table. Rotated Factor Matrix this table contains the rotated factor loadings, which represent both how the variables are weighted for each factor but also the correlation between the variables and the factor. Because these are correlations, possible values range from -1 to +1.

Results of EFA for different factors are discussed above.

Factor 1: Asthma expression

For Kaiser-Meyer-Olkin measure in table 5, value of 0.610, which is greater than 0.50 indicates that factor analyses can be applied to this data. Since our P-value by Bartlett's test is 0.000 that is less than 0.05 and indicates that the relationship between variables which are wheezing, shortness of breath, coughing, tightness of chest etc in asthma factor is significant.

Table 6 of Communalities shows that we have only one low value in the extraction column, which is of variable 'coughing a lot', is 0.545 and this is not going to be well represented in the common factor because only that variable will be well represented in the common factor whose values are high. These are the reproduced variances from the number of components that we have shown in this table which are 10. These values are on the diagonal of the reproduced correlation matrix. Next are given eigenvalues in table 7 are showing that each

Table 5: Exploratory Factor Analysis of Asthma expression: KMO and Bartlett's Test.

Factors	Kaiser-Meyer-olkin Measure of Sampling Adequacy	Bartlett's Test of Sphericity		
		Approx. χ^2	d.f	Sig.
Asthma Expression	0.610	380.831	45	0.000
Behavioral Factors	0.550	357.102	105	0.000
Current Family Status	0.820	634.219	36	0.000
In/Outdoor Allergens	0.789	458.467	66	0.000
Neighborhood Context	0.587	243.313	36	0.000
Socioeconomic Status	0.746	399.206	55	0.000
Collective Efficacy	0.687	565.467	28	0.000

Table 6: Communalities for Asthma Expression.

Communalities	Initial	Extraction
Wheezing	1.000	.649
Coughing a lot	1.000	.545
Chest tightness	1.000	.646
Breath shortness	1.000	.749
Attacks with change in season	1.000	.707
Worse symptoms at night	1.000	.786
Need emergency treatment	1.000	.869
Increasing severity	1.000	.878
Satisfied with treatment	1.000	.684
Excessive use of inhaler	1.000	.779

Extraction Method: Principal component analysis.

Table 7: Eigen-values and factor loadings.
Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.554	25.543	25.543	2.554	25.543	25.543	1.926	19.260	19.260
2	1.372	13.720	39.264	1.372	13.720	39.264	1.573	15.726	34.986
3	1.277	12.767	52.031	1.277	12.767	52.031	1.340	13.405	48.391
4	1.082	10.817	62.848	1.082	10.817	62.848	1.247	12.474	60.864
5	1.007	10.069	72.917	1.007	10.069	72.917	1.205	12.053	72.917
6	.764	10.06	80.555						
7	.652	8.522	87.076						
8	.581	5.805	92.882						
9	.518	5.177	98.059						
10	.194	1.941	100.000						

Extraction Method: Principal component analysis.©

variable has a variance of 1, and the total variance is equal to the number of variables used in the analysis i.e. 10 variables. The column named as total, contains the first component accounting the most variance (and hence have the highest eigenvalue) which is 2.554 and the next component accounts for as much of the remaining variance, and so on. The "% of Variance" column shows that one factor is explaining 25.543 percent of variance. Next column contains the cumulative percentage of variance accounted for by the current and all preceding principal components. For example, the third row shows a value of 12.767. This means that the first three components together account for 12.767% of the total variance. (Remember that because this is principal components analysis, all variance is considered to be true and common variance. In other words, the variables are assumed to be measured without error, so there is no error variance.) . Here we have first '5' components whose eigenvalues are greater than 1 so only 5 factors are been extracted and rotated.

Table 8: Indices for fit path model.

Single Sample Fit Indices	
	Value
Joreskog GFI	0.908
Joreskog AGFI	0.872
Akaike Information Criterion	1.321
Schwarz's Bayesian Criterion	2.033
Browne-Cudeck Cross Validation Index	1.364
Independence Model Chi-Square	980.555
Independence Model df	136.000
Bentler-Bonett Normed Fit Index	0.820
Bentler-Bonett Non-Normed Fit Index	0.901
Bentler Comparative Fit Index	0.921
James-Mulaik-Brett Parsimonious Fit Index	0.663
Bollen's Rho	0.777
Bollen's Delta	0.923

Table 9: Noncentrality fit indices.

Noncentrality Fit Indices			
	Lower 90%	**Point**	**Upper 90%**
Population Noncentrality Parameter	0.149	0.308	0.507
Steiger-Lind RMSEA Index	0.037	0.053	0.068
McDonald Noncentrality Index	0.776	0.857	0.928
Population Gamma Index	0.944	0.965	0.983
Adjusted Population Gamma Index	0.922	0.951	0.976

Table 10: Basic Summary statistics.

Basic Summary Statistics	
Basic Summary Statistics	**Value**
Discrepancy Function	0.889
Maximum Residual Cosine	0.000
Maximum Absolute Gradient	0.001
ICSF Criterion	-0.000
ICS Criterion	0.000
ML Chi-Square	176.826
Degrees of Freedom	110.000
p-level	0.000
RMS Standardized Residual	0.085

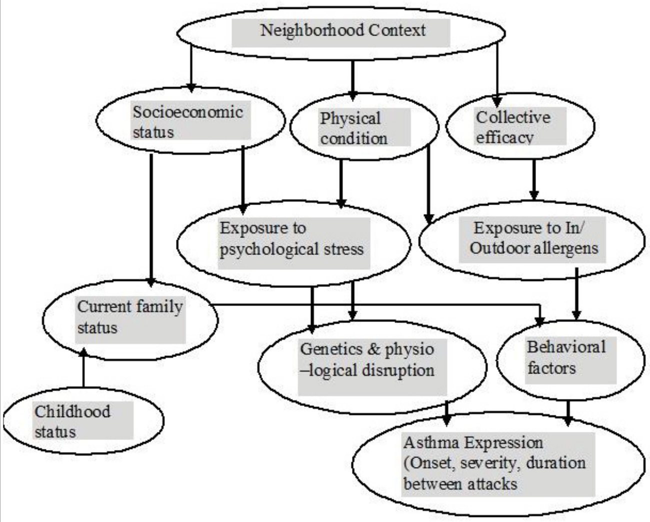
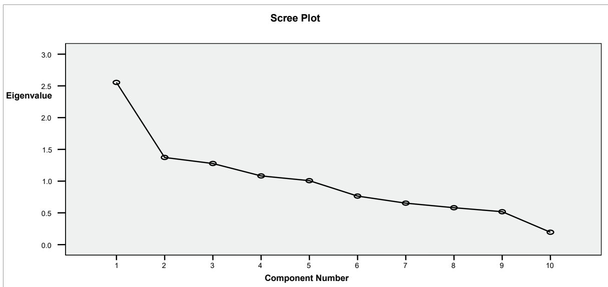


Figure 1: Proposed conceptual framework



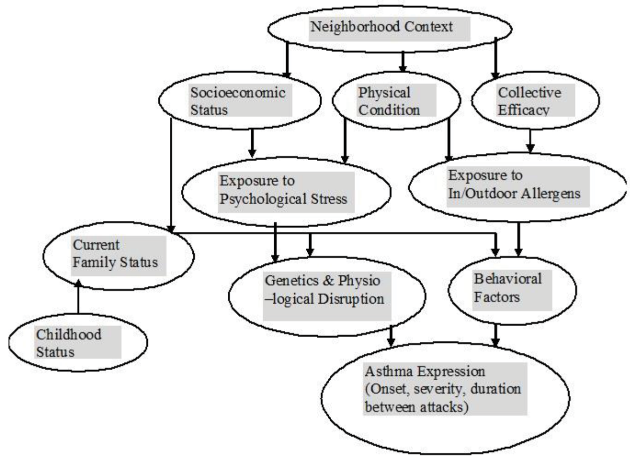


Figure 2: Proposed modal.

The scree plot shows the eigenvalues for initial components or factors. It is used to help determine the optimal number of factors or components to maintain in the solution. For a good factor analysis, this chart will look roughly like the intersection of two lines. Generally, the factors you want to keep are the ones on the vertical slope. The

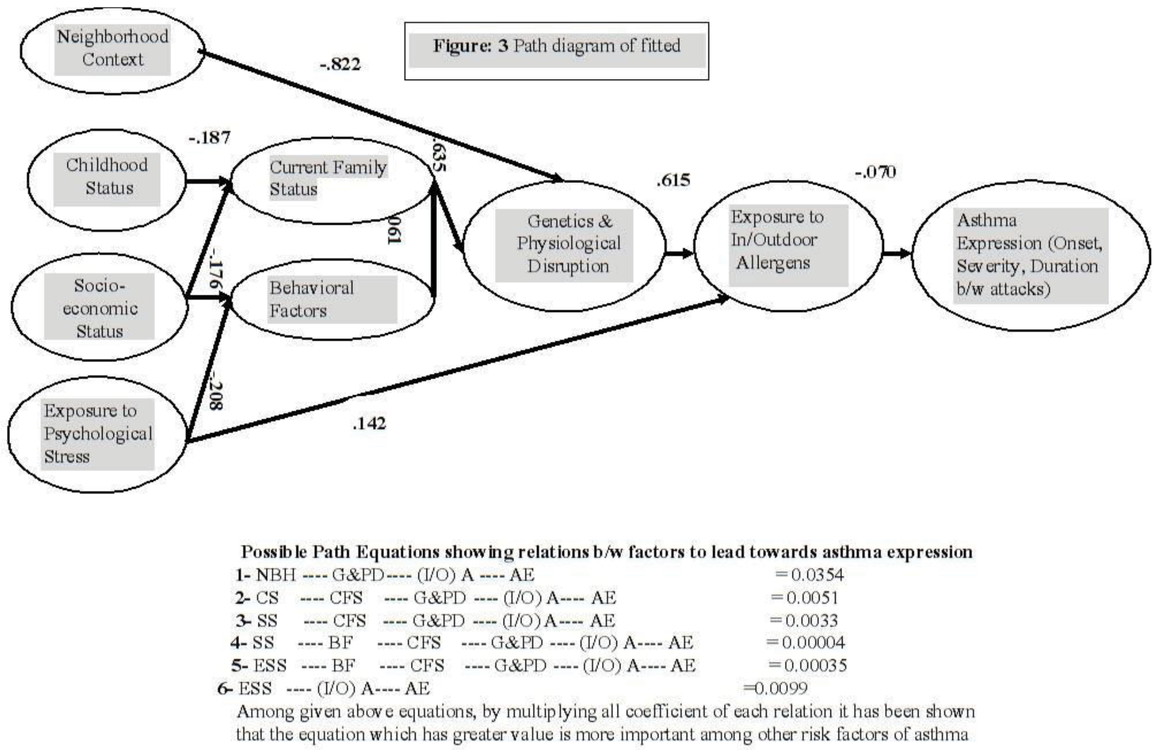


Figure 3: Path diagram of fitted.

Relationship Statements:

1. Neighborhood Context (NBH) → Genetics & physiological disruption (G&PD) → Exposure to In/outdoor Allergens(I/O)A → Asthma Expression(AE) = 0.0354
2. Childhood status → Current family Status(CFS) → Genetics & physiological disruption (G&PD) → Exposure to In/outdoor Allergens(I/O)A → Asthma Expression(AE) = 0.0051
3. Socio-economic Status → Current family Status → Genetics & physiological disruption (G&PD) → Exposure to In/outdoor Allergens(I/O)A → Asthma Expression(AE) =0.0033
4. Socio-economic Status → Behavioral Factors → Current family Status → Genetics & physiological disruption (G&PD) → Exposure to In/outdoor Allergens(I/O)A → Asthma Expression(AE) =0.00004
5. Exposure to Psychological stress → Behavioral Factors → Current family Status → Genetics & physiological disruption (G&PD) → Exposure to In/outdoor Allergens(I/O)A → Asthma Expression(AE) =0.00035
6. Exposure to psychological stress → Exposure to In/outdoor allergens(I/O)A → Asthma Expression(AE) =0.0099

ones on the one-dimensional slope contribute relatively little to the solution, and can be excluded. From the 2nd component the line is almost flat, meaning the each successive component is accounting for smaller and smaller amounts of the total variance. In general, we are interested in keeping only those principal components whose eigenvalues are greater than 1. Components with an eigenvalue of less than 1 account for less variance and so are of little use.

The table of Component Matrix contains component loadings, which are the correlations between the variable and the component. Because these are correlations, possible values range from -1 to +1. This table reports the factor loadings for each variable on the unrotated components or factors. Each number represents the correlation between the item and the unrotated factor. These correlations can help you formulate an interpretation of the factors or components. This is done by looking for a common line among the variables that have large loadings for a particular factor or component. It is common to see items with large loadings on several of the unrelated factors, which can make interpretation difficult. In these cases, it can be helpful to examine a rotated solution. If there would be only 1 component than rotated component matrix will not appear automatically.

EFA is confirming three variables i.e. need of emergency treatment, increasing severity, and excessive use of inhaler, to be included in asthma expression.

Factor2: Behavioral Factors

Criteria of different measures found by performing analysis and graph suggests that EFA is confirming three variables i.e. use medicines properly, prescribed diet, and home cooked meals, to be included in behavioral factor.

Factor3: Current family status

EFA is confirming five variables i.e. hopeless future of family, Absence of pleasure and joy, change in routine meal, difficulty in decision making, and difficulty to concentrate things, body paralyzed to be included in this factor named as current family status.

Factor4: Exposure to in/out door allergens

EFA is confirming five variables i.e. Dust, Pollution, stress/tension, smoke to be included in this factor named as indoor outdoor allergens.

Factor5: Neighborhood context

EFA is confirming four variables i.e. undoubted cleanliness of neighborhood, well caring about health, disease observed (T.B), disease observed (Asthma) to be included in this factor named as neighborhood context.

Factor6: Socio-economic Status

EFA is confirming five variables i.e. highest grade of education, motorcycle, refrigerator, mobile, telephone, and computer to be included in this factor named as socio-economic status.

Factor7: Collective Efficacy

EFA is confirming three variables i.e. close relationship with community, Trusted NBH, and communication at to be included in this factor and for use of these variables in SEM.

Path analysis

To check whether there exists causality among relationship between constructs path analysis is done.

To check that there is cause and effect relationship between constructs, simple regression is applied. Hence the estimated model is

Asthma expression = 8.002 - 0.070(Exposure to in/out allergens)

The t statistics can help to determine the relative importance of each variable in the model. Value of t-statistic should be below -2 and above +2. Here it is less than -2 which shows relative importance of in/outdoor allergens variable in the model. Also *p*-value is less than significance value of 0.05 which shows dependence of asthma expression on In/out door allergens.

Next estimated model is

In/out door allergens = -0.128+0.615(Genetics and physiological disruption) +0.142(Exposure to psychological stress)

It is found that value of t-statistic is greater than +2 for both independent variables which show relative importance of Genetics and physiological disruption & Exposure to psychological stress in the model. Also *p*-value is less than significance value of 0.05 it means in/out door allergens depends on Genetics and physiological disruption & Exposure to psychological stress.

Now another estimated model is shown

Genetics and physiological disruption = 6.600-0.822(Neighborhood context) +0.635(Current family status)

It is found that t-statistic is less than -2 of neighborhood context and greater than +2 of current family status, which shows relative importance of both the independent variables in the model. Also *p*-value is less than significance value of 0.05.

Next equation is

Current family status = 5.729 -0.187(Childhood status) + 0.120(Socio-economic status)-0.061(Behavioral factors)

It is found that t-statistic is less than -2 of all independent variables in this model, which show relative importance of all three of the independent variables in the model. Also *p*-value is less than significance value of 0.05.

Another estimated model is

Behavioral factors = 13.849+0.176(Socio-economic status) + 0.208(Exposure to psychological stress).

Here value of t-statistic is less than -2 of both variables that show the relative importance of both the independent variables in the model. Also *p*-value is less than significance value of 0.05 it means

As linear regression has confirmed the casual relationship between constructs, a path diagram has been formed between those constructs whose relation is confirmed.

Path diagram of fitted model

A path diagram in structural equation modeling with effective goodness of fit index was obtained including factors

- Asthma expression (onset, severity, duration b/w attacks)
- Behavioral factors
- Genetics and physiological disruption
- Current family status
- Childhood status
- Exposure to In/out door Allergens
- Exposure to psychological stress
- Neighborhood context including one variable
- Socio-economic status

And our fitted model with fit-indices according to general criteria is given above.

Using Statistica7.0 version, results obtained by performing structural equation modeling are as above.

General criteria of model fitting (by Siu Loon Hoe): There are several indicators of goodness-of-fit and most SEM scholars recommend evaluating the models by observing more than one of these indicators [21]. Based on this stated criteria, Garver and Mentzer recommended the Non-Normed Fit Index (NNFI); the Comparative Fit Index (CFI), and the Root Mean Squared Approximation of Error (RMSEA). Therefore, the commonly applied fit indices are given above.

Chi-square statistic: chi-square test is used to assess actual and predicted matrices. Thus, non-significance means that there is no considerable difference between the actual and predicted matrices [21]. Therefore, low $\chi 2$ values, which result in significance levels greater than 0.05 or 0.01, indicate that actual and predicted inputs are not statistically different. *p*-values indicate whether the model is significantly different than the null model. The null hypothesis is the hypothesized model in which the parameters were set up for the hypothesized model, indicating whether a path should exist or not between variables. A high *p*-value, or a value larger than zero, would mean that the null hypothesis is rejected leading to a high probability that it would be wrong in doing so. There is a limitation to the chi-square test. The $\chi 2$ is highly sensitive to sample size especially if the observations are greater than 200. An alternate evaluation of the $\chi 2$ statistic is to examine the ratio of $\chi 2$ to the degrees of freedom (d.f.) for the model [22]. Klin suggested that a $\chi 2$/ d.f. ratio of 3 or less is a reasonably good indicator of model fit.

Composite Fit index (CFI): The CFI is a test statistic that indicates the overall proportion of variance explained by the model; good fit is indicated by a CFI > 0.9.

Root Mean Square Error of Approximation (RMSEA): The RMSEA is a model fit index that considers the model's residuals; an RMSEA < 0.08 indicates a satisfactory model representation. Values less than 0.05 indicate good fit, values up to 0.08 reasonable fit and ones between 0.08 and 0.10 indicate mediocre fit.

Non-Normal Fit index (NNFI): The NNFI, also known as the Tucker Lewis index, compares a proposed model's fit to a null model. Additionally, NNFI measures parsimony by assessing the degrees of freedom from the proposed model to the degrees of freedom of the null model. NNFI also seems flexible against variations in sample size and, thus, is highly recommended. An acceptable threshold for this index is 0.90 or greater.

Goodness of Fit index (GFI): GFI cannot be interpreted as percent of error explained by the model. Rather it is the percent of observed covariances explained by the covariances implied by the model. That is, R^2 in multiple regression deals with error variance whereas GFI deals with error in reproducing the variance-covariance matrix.

Hence values of fit indices are given

GFI = 0.908 NNFI = 0.901 $\chi 2$/ d.f = 1.607

CFI = 0.921 RMSEA = 0.053

As the general criterion is being fulfilled so it is indication of good model fit.

CONCLUSION

Fitted model we obtained by applying SEM, describes that, there is direct effects of some factors as well as, there are a lot of factors that are attached behind them also. The main conclusion got by applying SEM is that our fitted model shows that firstly Indoor/Outdoor allergens sensitize person's air-tubes that leads to airway inflammation and hyper responsiveness towards some inhaled stimuli's/ allergens. In other words it would be convenient to say that some individuals are sensitive to some irritants in our environment i-e allergic. But as regression coefficient of relation between indoor/outdoor allergens is negative. This could be defended by the statement that exposure to some allergens may actually be negatively associated with current expression of asthma as it had found in some cross sectional and prospective studies [23]. Genetics factors are base of development of asthma among some people. During data collection most of the patients had family history of asthma. So it means, asthma may be developed by genetic disorder and indoor-outdoor allergens making it severe or worsening its symptoms.

Some papers presented in past contains that asthma could be developed more commonly among those who have genetically disorder of airways, lungs and other organic mechanism. In this study, very deep mechanism of genetics and organism is not of main because this can only be well described by basic sciences researcher. Extreme psychological stress can increase severity of asthma by narrowing airways. Interrelationship between other factors i.e. socioeconomic status, neighborhood context, childhood status, behavioral factors and current family status also support genetic and allergen factors to make Asthma severe and also leads to make asthmatic patient's to no response to treatment and that cause to increased duration.

SUGGESTIONS

It was estimated that in 2016 more than 339 million people had Asthma globally [24]. It is a common disease among children. According to WHO estimates, in 2016, there were 417,918 deaths due to asthma at the global level and 24.8 million DALYS attributable to Asthma [25]. Rate of morbidity and deaths associated with asthma are high, even sufficient treatment facility was available to them. To explore worldwide prevalence of asthma, Global Initiative for Asthma obtained data on asthma prevalence from 20 different regions worldwide, from literature primarily published through the European Community Respiratory Health Survey and International Study on childhood Asthma and Allergens. According to this report, it is estimated that in Pakistan, 4.3% children suffer from asthma. There is no proper record of most of the asthma patients in Pakistan or record is not poorly managed, so this estimate may not be exact estimate of prevalence of asthma.

Due to limited resources it was difficult to take large sample for analysis so future research with large sample would provide firm grounds to results of this study.

REFERENCES

1. National heart lung and blood institute (NHLBI).
2. Wright RJ, Suglia SF, Levy J, Fortun K, Shields A, Subramanian S, Wright, R. Transdisciplinary research strategies for understanding socially patterned disease: the Asthma Coalition on Community, Environment, and Social Stress (ACCESS) project as a case study. Ciencia & saude coletiva.2008; 13(6):1729–1742. doi: 10.1590/s1413-81232008000600008.
3. Matthew M, Denise F, Shaun, H, and Richard, B. Global burden of asthma. A report developed for the global initiative for asthma (GINA). 2004.
4. Wisconsin Asthma Plan 2009-2014.
5. Marjorie L, Slankard MD. What Causes Asthma? 2009 Jul27.
6. O'Byrne PM, Inman MD. Airway Hyperresponsiveness. Journal of the American College of Chest Physicians. 2003 Mar: 123; 411-416.
7. de Marco R, Locatelli F, Sunyer J, Burney P. Differences in incidence of reported asthma related to age in men and women. A retrospective analysis of the data of the European Respiratory Health Survey. Am J Respir Crit Care Med. 2000 Jul;162(1):68-74. doi: 10.1164/ajrccm.162.1.9907008. PMID: 10903222.
8. von Mutius E. The environmental predictors of allergic disease. J Allergy Clin Immunol. 2000 Jan;105(1 Pt 1):9-19. doi: 10.1016/s0091-6749(00)90171-4.

PMID: 10629447.

9. Gold DR. Environmental tobacco smoke, indoor allergens, and childhood asthma. Environ Health Perspect. 2000 Aug;108 Suppl 4(Suppl 4):643-51. doi: 10.1289/ehp.00108s4643. PMID: 10931782; PMCID: PMC1637671.
10. Weiss KB, Gergen PJ, Hodgson TA. An economic evaluation of asthma in the United States. N Engl J Med. 1992 Mar 26;326(13):862-6. doi: 10.1056/NEJM199203263261304. PMID: 1542323.
11. Williams DR, Sternthal M, Wright RJ. Social determinants: taking the social context of asthma seriously. Pediatrics. 2009 Mar;123 Suppl 3(Suppl 3):S174-84. doi: 10.1542/peds.2008-2233H. PMID: 19221161; PMCID: PMC3489274.
12. Gissler M, Järvelin MR, Louhiala P, Hemminki E. Boys have more health problems in childhood than girls: follow-up of the 1987 Finnish birth cohort. Acta Paediatr. 1999 Mar;88(3):310-4. doi: 10.1080/08035259950170088. PMID: 10229043.
13. Litonjua AA, Carey VJ, Weiss ST, Gold DR. Race, socioeconomic factors, and area of residence are associated with asthma prevalence. Pediatric Pulmonology. 1999 Dec o6; 28:394-401.doi: 10.1002/(SICI)1099-0496(199912)28
14. Liang PH, Shyur SD, Huang LH, Wen DC, Chiang YC, Lin MT, Yang HC. Risk factors and characteristics of early-onset asthma in Taiwanese children. J Microbiol Immunol Infect. 2006 Oct;39(5):414-21. PMID: 17066205.
15. Sears MR, Burrows B, Flannery EM, Herbison GP, Holdaway MD. Atopy in childhood. I. Gender and allergen related risks for development of hay fever and asthma. Clin Exp Allergy. 1993 Nov;23(11):941-8. doi: 10.1111/j.1365-2222.1993.tb00279.x. PMID: 10779282.
16. Subramanian, S.V, Ackerson, L.K, Subramanyam, M.A, and Wright, R.J. (2007) Domestic violence is associated with adult and childhood asthma prevalence in India. International Journal of Epidemiology. 2007 Feb 28: 36(3) 569–579. doi: 10.1093/ije/dym007.
17. Lin CK, Hung MS, Fang CP, Shen TY, Hsiao SH, Huang TJ, Yang CT. Survey of asthmatic patients' sensitization to house dust mites in southwestern Taiwan. Chang Gung Med J. 2006 Nov-Dec;29(6):568-75. PMID: 17302220.
18. Diana DS. Statistical Analyst in Institutional Research at the University of Northern Colorado. In 1999 she earned a Ph.D. in Educational Psychology at UNC. She has been a SAS programmer since 1984.
19. Hu, LT, Bentler PM. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives," Structural Equation Modeling.1999; 2009 03 Nov: 6 (1), 1-55. doi: 10.1080/10705519909540118
20. Child D. *The essentials of factor analysis.* 2nd ed. Cassell Educational Ltd. 1990.
21. Hair JF, Anderson RE, Tatham RL, Black WC. Multivariate Data Analysis. 5th ed. Upper Saddle River, NJ: Prentice Hall. 1998.
22. Jöreskog KG. Testing Structural Equation Models. In KA. Bollen, JS. Long, Testing Structural Equation Models. Thousand Oaks, CA: Sage. 1993 March 23: 294-316
23. Wieringa MH, Weyler JJ, Van Bastelaer FJ, Nelen VJ, Van Sprundel MP, Vermeire PA. Higher asthma occurrence in an urban than a suburban area: role of house dust mite skin allergy. Eur Respir J. 1997 Jul;10(7):1460-6. doi: 10.1183/09031936.97.10071460. PMID: 9230231.
24. 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 Sep 16;390(10100):1211-1259. doi: 10.1016/S0140-6736(17)32154-2. Erratum in: Lancet. 2017 Oct 28;390(10106):e38. PMID: 28919117; PMCID: PMC5605509.
25. Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016. Geneva, Global Health Estimates 2016: World Health Organization; 2018.