International Journal on Advanced Science Engineering Information Technology

Vol.10 (2020) No. 4
ISSN: 2088-5334

Using Multivariate Analysis to Study the Spread of Transitional Diseases in Iraq

Aseel Abdul Razzak Rasheed[a,1], Rawaa Salh Al-saffar[a], Husam A. Rasheed[a]

[a] Collage of administration and Economics / Statistics Department Mustansiriyah University, Iraq
E-mail: [1]aseelstat@uomustansiriyah.edu.iq

***Abstract*— Medical and technological development has achieved continuous successes during the last period. However, the statistics received from the World Health Organization (WHO) show the suffering that millions of people are subjected to daily as a result of their exposure to transitional diseases. The most prevalent transitional disease in the Iraqi provinces and for all age groups for both genders is water pox disease, followed by cutaneous leishmaniosis and then mumps. The most affected governorate with transitional diseases is Baghdad Governorate (Rusafa Sector), followed by Dhi Qar Governorate and Baghdad (Al Karkh Sector). The most age-groups affected with transition diseases are the categories (5-9) years for males, followed by (5-9) years for females. The highest total contribution to the first axis was for cutaneous leishmaniasis, followed by mumps, then Basil dysentery, as for the absolute contribution in the first axis, it was Dhi Qar Governorate, followed by Baghdad. There are more than 14,000,000 people who die each year because of these diseases, and most infections are concentrated in developing countries, including Iraq. Hence, this research is vital to examine the extent of transitional diseases in Iraq for different age groups, both male and female gender. The use of the multivariate method, which is the correspondence analyses, it was found through the research that Baghdad (Rusafa Sector) has a high incidence of transitional diseases and the largest age group at risk of transitional diseases is (5-9) years form male gender.**

***Keywords*— diseases; Iraq; statistical method; pandemic; WHO.**

I. INTRODUCTION

The developments in the field of combating transitional diseases, and the availability of antibiotics to eliminate these diseases, did not prevent members of society from developing new diseases, or the return of some rare diseases. The sources of infection and methods of transmission vary from person to another. Moreover, the transmission of infectious agents from the infected person or the carrier of the disease or animals were infected with diseases common to humans and animals.

The diseases may be transmitted from the environment close to humans, including food and drink contaminated with microorganisms that cause the disease. Transitional diseases are classified as serious in that they often cause death as well as permanent or temporary complications such as polio, which causes permanent disability.[1], [2] Disease can be defined as resulting from the transmission of microorganisms, like the viruses, bacteria, fungi, or parasites from an infected person to another healthy person that leads to his illness.[3] Correspondence analysis was used as a tool for analyzing disease data; given that the symmetric analysis is a classification tool from two sides, the research covers all governorates of Iraq except the Kurdistan Region and the age groups and both gender[4], [5]. In this paper, we used a simple corresponding analysis to study (data classified as multivariate and discrete) and represented by transitional diseases and knowing the most prevalent types of transitional diseases in the Iraqi governorates, according to age groups and gender [6].

II. MATERIALS AND METHOD

A. Correspondence Analysis

Correspondence analysis is a multivariate method that has been known for a long time. Furthermore, Correspondence analysis is a special case of correct correlation analysis that concerned with analyzing the metadata and non-meta data. There is a special kind of graphical display that rows and columns are drawn as two-dimensional points. The location of theses points indicates that they are reconciled [6]–[8].

1) Weighted Euclidean distance: If the points B and A in this field are the sum of the squares of the weighted differences for the coordinates as follows [7]:

$$D^2 \text{ (A, B) =(A-B) T Dq (A-B) =}\sum \text{qi (AJ-BJ)} \quad (1)$$

D_q: q: diagonal matrix with dimension J, in which the diameter elements represent weights.

2) Chi-square (X2) unit measure: It is the distance between the profiles by means of the weighted Euclidean unit of measurement, and by using the profiles the square of the area from □b to □(¯b)are [9]:

$$(\underline{b} - \underline{\bar{b}})^T \; D_{\bar{b}}^{-1} \left(\underline{b} - \underline{\bar{b}}\right) \quad (2)$$

3) Profiles: Are the percentage frequencies whose sum is equal to one, and the side section vector (row or column) is the result of dividing each value in (row or column) by the sum of that (row or column) [10], [11].

4) Masses: Each profile has a point, and accordingly, the mass is the weight of that point relative to its frequency, and the masses (rows or columns) equal the sum (rows or columns) of the corresponding matrix, and the percentage frequencies represent its elements. The purpose of using masses is to distinguish between weights for points and dimensions [12]– [14].

B. Algebra of Two-way Correspondence Analysis

Since the correspondence analysis is only related to the percentage values of data, it is constant of (M..); therefore, the matrix (P) resulting from dividing the original data matrix (Z) by its total sum will be relied upon. The sum of the elements of the matrix (P) equals one, and that the sum of the rows and columns of the array (P) is denoted by the symbol(t ,s), respectively, and the diagonal matrices of (t ,s) are, (D_t,D_s), respectively[15], [16].

$$P=(1/(M..)) \quad (3)$$

The profiles of the rows and columns of the matrix (P)are vectors for the rows and columns divided by their corresponding sum and symbolize the profiles of the rows and columns (t ,s), respectively[17], [18] .

C. The application Sides

Transitional diseases are spreading rapidly through contact infected people with healthy people. To create a healthy environment for living and protecting people from the transmission of these diseases, we have made two analyzes as follows [19]:

- The first analysis: a study of the extent of transmission diseases in Iraqi governorates.
- The second analysis: a study of which age groups are more susceptible to the transmission of diseases of both sexes.

Hence, it was relied on the Iraqi Ministry of Health's annual report for 2017 [20], [21].

III. RESULT AND DISCUSSION

A. Coding

1) Governorate coding:

TABLE I
GOVERNORATE CODING OF DIFFERENT PLACES IN IRAQ

	place		**place**		**place**		**place**
1	Baghdad/Al-Karkh	6	Al-Dewaniya	11	Kirkuk	16	Al-Najaf
2	Baghdad/ Al-Rusafa	7	Diala	12	Wasit	17	Erbil
3	Basrah	8	Al-Anbar	13	Thi-Qar	18	Duhouk
4	Ninevah	9	Babylon	14	Al-Muthanna	19	Al-Sulaimaniya
5	Maysan	10	Kerbela	15	Slah-Al-Deen		

- Age coding by gender

TABLE II
AGE CODING BY GENDER OF DIFFERENT PLACES IN IRAQ

Code	Age (years)	Gender	Code	Age (years)	Gender
1	>1	Male	9	>1	Female
2	1-4	Male	10	1-4	Female
3	5-9	Male	11	5-9	Female
4	10-14	Male	12	10-14	Female
5	15-19	Male	13	15-19	Female
6	20-44	Male	14	20-44	Female
7	45-64	Male	15	45-64	Female
8	65 <	Male	16	65 <	Female

- Coding types of transmission diseases

TABLE III
CODING TYPES OF TRANSMISSION DISEASES OF DIFFERENT PLACES IN IRAQ

Code	**Diseases**	**Code**	**Diseases**	**Code**	**Diseases**
A	Acute flaccid paralysis	**H**	Chickenpox	**O**	Rabies
B	Diphtheria	**I**	Typhoid	**P**	Viral Hepatitis A
C	Pertussis	**J**	Cutaneous Leishmaniosis	**Q**	Viral Hepatitis B
D	Measles	**K**	Black fever	**R**	Viral Hepatitis C
E	Rubella	**L**	Toxoplasmosis	**S**	Viral Hepatitis E
F	Mumps	**M**	Hydatid Cyst	**T**	Meningitis
G	Neonatal tetanus	**N**	Brucellosis	**U**	Bacillary Dysentery

B. Analyzing and Discussing the Results of the Corresponding Relationship between Types of Transitional Diseases and Iraqi Governorates

From the analysis of the two way contingency table represents the data of the matrix (z), where (s) describes the types of transitional diseases and their number (21), and (t) describe the Iraqi governorates and their number are (19), Using the program (SPSS V23) [9], Table I represents the summary of the analysis of the contingency table, we notice through the column (inertia) that the sum of the total inertia of the (18) axes has reached (114.8%) of the total variance, Where we note that the first axis has formed (42.5%) of the total variance of the contingency table, for real data, while the second axis has formed (23%) of the total variance of the contingency table, for real data, we can also note from the proportion of inertia for each axis, that the first and second axes have interpreted the ratio of 57.1%, and through the value of x^2 , and we can see a significant sign between the disease variables and the governorates of Iraq in terms of the number of injuries, because the value of sig was 0,000 which is less than 0.05.

TABLE IV
SUMMARY THE TRANSITIONAL DISEASES WITH THE IRAQ GOVERNORATES

Dimension	Singular Value	Inertia	Proportion of Inertia	
			Accounted for	Cumulative
1	.652	.425	.370	.370
2	.480	.230	.200	.571
3	.380	.144	.126	.696
4	.332	.110	.096	.792
5	.304	.092	.080	.873
6	.197	.039	.034	.907
7	.174	.030	.026	.933
8	.171	.029	.025	.958
9	.138	.019	.017	.975
10	.109	.012	.010	.985
11	.087	.008	.007	.992
12	.073	.005	.005	.997
13	.047	.002	.002	.999
14	.032	.001	.001	.999
15	.017	.000	.000	1.000
16	.013	.000	.000	1.000
17	.010	.000	.000	1.000
18		.000	.000	1.000
Total		1.148	1.000	1.000
Chi-Square= 114858.789with Sig. =0.000				

We note from Table V the absolute contributions of diseases (rows) in each axis, where we notice in the mass column that the most prevalent diseases in the Iraqi governorates is the disease (water pox) with the symbol (H) because it has the largest mass of (0.463), and it is followed by the disease (Cutaneous Leishmaniosis with the symbol J), where its mass reached (0.175), followed by the disease (Mumps with the symbol F) whose mass was by (0.130), and in terms of absolute contribution in the first and second axes, we can see that Cutaneous Leishmaniosis has a higher Absolute contribution in the first axis amounted to (0.443), and does not have the largest block, we also note that the mumps disease has the largest absolute contribution to the second axis, as it reached (0.562) compared to other diseases.

TABLE V
THE MASSES AND THE ABSOLUTE CONTRIBUTIONS TO ALL TRANSITIONAL DISEASES

disease	mass	Score in Dimension		contribution		
				Of point to inertia of Dimension		Of Dimension to Inertia of Point
		1	2	1	2	
a	.005	.131	.245	.000	.034	.120
b	.000	.839	.427	.000	.015	.018
c	.006	.489	.101	.002	.065	.067
d	.000	.083	.405	.000	.001	.021
e	.000	.755	.647	.000	.035	.054
f	.130	1.089	1.442	.236	.425	.973
g	.000	.526	.185	.000	.017	.019
h	.463	.288	.363	.059	.298	.645
i	.012	.766	.397	.011	.103	.123
j	.175	1.284	.272	.443	.802	.828
k	.004	1.072	.272	.008	.401	.420
l	.009	.643	.437	.006	.074	.099
m	.006	.584	.339	.003	.149	.187
n	.021	.119	1.400	.000	.002	.194
o	.000	.795	.339	.000	.057	.065
p	.080	.250	.169	.008	.046	.062
q	.031	.346	.994	.006	.065	.458
r	.012	.512	.389	.005	.112	.160
s	.001	.621	.196	.001	.018	.019
t	.016	.069	.786	.000	.002	.185
u	.028	2.215	1.274	.213	.445	.554
Active Total	1.000			1.000		

From the Table VI, we can observe from the mass column that transitional diseases are more prevalent in Baghdad/Rusafa governorate, because they have the largest mass, reaching (0.212), followed by Dhi Qar governorate, where it reached (0.094), followed by Baghdad/Karkh (0.087), and the governorate of Baghdad / Rusafa had the highest absolute contribution in the second axis, amounting to (0.455), while Dhi Qar governorate had the largest contribution in the first axis, reaching (0.390).

TABLE VI
THE MASSES AND THE ABSOLUTE CONTRIBUTIONS TO ALL GOVERNORATES

place	Score in Dimension		Contribution				
			Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
	1	2	1	2	1	2	total
1	.492	.310	.032	.017	.354	.103	.457
2	.946	1.014	.291	.455	.537	.454	.991
3	.354	.597	.010	.039	.137	.287	.423
4	.293	.616	.004	.026	.037	.119	.156
5	.906	.136	.051	.002	.522	.009	.531
6	.588	.182	.021	.003	.209	.015	.224
7	1.081	.238	.127	.008	.523	.019	.542
8	1.310	.165	.002	.000	.356	.004	.360
9	.098	.368	.000	.005	.005	.050	.055
10	.143	.377	.001	.011	.028	.143	.171
11	.269	.702	.006	.051	.122	.610	.732
12	.547	.205	.015	.003	.298	.031	.329
13	1.643	.758	.390	.113	.668	.105	.773
14	.604	.251	.010	.002	.290	.037	.327
15	1.201	.038	.012	.000	.419	.000	.419
16	.140	.431	.001	.019	.011	.076	.087
17	.232	.586	.002	.017	.012	.057	.070
18	.455	.766	.024	.091	.182	.380	.562
19	.098	1.023	.001	.139	.005	.444	.449
total			1.000	1.000			

C. Analyze and Discuss the Results of the Corresponding Relationship between Transmission Diseases and Age Groups with the Gender Factor

The data was arranged by dividing the age groups into (16) categories, the first (8) categories represent the numbers of injured males, and the second (8) categories represent the number of injured females. This research aims to study the effect of gender in the two-way intersection. The table was relied upon in analyzing the data and represents the matrix of the conjugation matrix (z), where (s) describes the types of transition diseases. The number (21), and (t) describe the age groups of males and females and their number (16),.The calculation has been done by conducting (SPSS V23),[9] Table VII represents the summary for the analysis of the contingency of (15) axis was obtained.

TABLE VII
SUMMARY OF THE TRANSITIONAL DISEASES WITH THE AGE GROUP

Dimension	Singular Value	Proportion of Inertia	
		Accounted for	Cumulative
1	.531	.505	.505
2	.372	.248	.753
3	.264	.125	.878
4	.201	.072	.950
5	.110	.022	.972
6	.080	.011	.983
7	.064	.007	.990
8	.053	.005	.996
9	.031	.002	.997
10	.026	.001	.999
11	.019	.001	.999
12	.014	.000	1.000
13	.013	.000	1.000
14	.008	.000	1.000
15	.005	.000	1.000
Total		1.000	1.000

From the observation of the inertia column, we notice that the value of the inertia for the first axis (0.282) represents 50.5% of the total inertia, while the inertia of the second axis was (0.138) and constitutes (24.8%) of the total inertia. Thus, the first and second axes constitute (75.3%) of the total inertia, and therefore they are sufficient to interpret the original data. Moreover, by observing the value of (sig), we find that it is smaller than (0.05), and this means that there is a strong relationship between the types of transmission diseases and the age groups of both genders.

From Table VIII we notice from the mass column that the chickenpox disease has the largest mass, which amounted to (0.463), which is the most prevalent transitional disease followed by Leishmaniosis disease, where its mass reached (0.175), followed by mumps disease (0.130), and so on for the rest of the diseases. From the contribution column, we can notice that Measles fever has the largest contribution in the first axis (0.242) despite having the lowest mass in the mass column (0.021), whereas Meningitis disease was the largest contribution in the second axis, where it reached (0.262) despite having The least mass in a column of masses (0.016).

TABLE VIII
THE MASSES AND THE ABSOLUTE CONTRIBUTIONS TO ALL TRANSITIONAL DISEASES

disease	Score in Dimension		contribution				
			Of point to the inertia of Dimension		Of Dimension to Inertia of Point		
	1	2	1	2	1	2	total
a	.608	1.493	.003	.028	.103	.436	.539
b	1.686	.233	.000	.000	.086	.001	.087
c	.403	3.566	.002	.204	.015	.848	.863
d	.420	.244	.000	.000	.132	.031	.163
e	1.686	.233	.000	.000	.086	.001	.087
f	.202	.626	.010	.137	.072	.480	.552
g	.124	7.273	.000	.007	.000	.650	.651
h	.403	.160	.142	.032	.791	.087	.878
i	1.128	.130	.030	.001	.771	.007	.778
j	.302	.270	.030	.034	.366	.205	.571
k	.437	2.818	.002	.093	.031	.897	.928
l	3.481	.447	.212	.005	.635	.007	.642
m	2.218	.349	.057	.002	.928	.016	.944
n	2.481	.377	.242	.008	.904	.015	.919
o	.675	.523	.000	.000	.126	.053	.179
p	.387	.106	.022	.002	.304	.016	.320
q	1.736	.227	.175	.004	.711	.008	.720
r	1.615	.156	.060	.001	.680	.004	.685
s	1.631	.123	.007	.000	.623	.002	.626
t	.033	2.463	.000	.262	.000	.868	.868
u	.351	1.534	.007	.179	.065	.865	.929
Active Total			1.000	1.000			

TABLE IX
THE MASSES AND THE ABSOLUTE CONTRIBUTIONS FOR ALL AGE GROUPS AND GENDER FACTORS

Place	Score in Dimension		Contribution				
	1	2	Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
			1	2	1	2	total
1	.052	2.744	.000	.345	.000	.854	.855
2	.306	.657	.015	.101	.135	.437	.572
3	.428	.242	.064	.029	.738	.165	.903
4	.366	.510	.035	.097	.282	.385	.667
5	.544	.125	.036	.003	.396	.015	.411
6	1.028	.058	.075	.000	.730	.002	.732
7	1.687	.166	.073	.001	.661	.004	.665
8	1.630	.621	.014	.003	.397	.040	.438
9	.086	2.639	.000	.297	.001	.887	.889
10	.325	.543	.016	.063	.211	.414	.625
11	.398	.105	.043	.004	.677	.033	.710
12	.311	.374	.019	.040	.434	.440	.874
13	.894	.087	.081	.001	.802	.005	.808
14	2.340	.304	.429	.010	.784	.009	.793
15	1.994	.240	.093	.002	.871	.009	.880
16	1.322	.682	.008	.003	.406	.076	.482
total			1.000	1.000			

From Table IX, we can notice that the largest mass was for the age group (5-9)-male, where it reached (0.186), while the second mass was for the age group (5-9)-female as it reached (0.143). However, they did not have the largest percentage in The contribution of the axes, where we note that the age group (20-44)-female got the most significant proportion of the contribution of the first axis as it reached (0.429) while the age group (<1)-meal had the most significant percentage of the contribution of the second axis, where it reached (0.345) Despite having the least mass.

IV. CONCLUSION

The study showed that there is a significant relationship between the types of transitional diseases and the Iraqi governorates, and between these types of diseases and age groups with the of the gender factor. The most prevalent transitional disease in the Iraqi provinces and for all age groups for both genders is water pox disease, followed by cutaneous leishmaniosis and then mumps. The most affected governorate with transitional diseases is Baghdad Governorate (Rusafa Sector), followed by Dhi Qar Governorate and Baghdad (Al Karkh Sector). The most age-groups affected with transition diseases are the categories (5-9) years for males, followed by (5-9) years for females. The highest total contribution to the first axis was for cutaneous leishmaniasis, followed by mumps, then Basil dysentery, as for the total contribution in the first axis, it was Dhi Qar Governorate, followed by Baghdad (Rusafa Sector). The age group (20-44) years for females had the absolute highest contribution from the first axis, while the Malta fever disease was the most common transitional absolute disease contributed to the first axis. Increase the number of health centers in governorates with a high population density to reduce the spread of transitional diseases in Iraq. Spreading a healthy culture among people and encouraging them to review health centers when they feel ill and not to take medications without consulting a specialist.

REFERENCES

[1] Néstor Rivera, Juan Chica, Ivan Zambrano, and Cristian Garc´ıa. Estudio del comportamiento de un motor ciclo otto de inyeccion electronica respecto de la estequiometr´ıa de la mezcla y del adelanto al encendido para la ciudad de cuenca. Revista Polit´ecnica, 40(1):59–67, 2017.

[2] L Hinke, L Pichler, HJ Pradlwarter, BR Mace, and TP Waters. Modelling of spa- tial variations in vibration analysis with application to an automotive windshield. Finite Elements in Analysis and Design, 47(1):55–62, 2011.

[3] Yoshio Kurosawa, Hideki Enomoto, Shuji Matsumura, and Takao Yamaguchi. High frequency vibration analysis of automotive bodies with panels that have attached viscoelastic layers. In ASME 2003 International Mechanical Engineering Congress and Exposition, pages 23–29. American Society of Mechanical Engineers Digital Collection, 2003.

[4] Seyed Hamed Mirafzal, Amir Mahyar Khorasani, and Amir Hossein Ghasemi. Op- timizing time delay feedback for active vibration control of a cantilever beam using a genetic algorithm. Journal of Vibration and Control, 22(19):4047–4061, 2016.

[5] T Yamaguchi, Y Kurosawa, and H Enomoto. Damped vibration analysis using finite element method with approximated modal damping for automotive double walls with a porous material. Journal of Sound and Vibration, 325(1-2):436–450, 2009.

[6] Sze-jung Wu, Nagi Gebraeel, Mark A Lawley, and Yuehwern Yih. A neural net- work integrated decision support system for condition-based optimal predictive maintenance policy. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 37(2):226–236, 2007.

[7] Zhiqiang Huo, Yu Zhang, Pierre Francq, Lei Shu, and Jianfeng Huang. Incipient fault diagnosis of roller bearing using optimized wavelet transform based multi- speed vibration signatures. IEEE Access, 5:19442–19456, 2017.

[8] Sharaf, H. K., Ishak, M. R., Sapuan, S. M., Yidris, N., & Fattahi, A. (2020). Experimental and numerical investigation of the mechanical behavior of full-scale wooden cross arm in the transmission towers in terms of load-deflection test. Journal of Materials Research and Technology, 9(4), 7937-7946.

[9] Sharaf, H. K., Ishak, M. R., Sapuan, S. M., & Yidris, N. (2020). Conceptual design of the cross-arm for the application in the transmission towers by using TRIZ–morphological chart–ANP methods. Journal of Materials Research and Technology, 9(4), 9182-9188.

[10] Johari, A. N., Ishak, M. R., Leman, Z., Yusoff, M. Z. M., Asyraf, M. R. M., Ashraf, W., & Sharaf, H. K. (2019). Fabrication and cut-in speed enhancement of savonius vertical axis wind turbine (SVAWT) with hinged blade using fiberglass composites. In Proceedings of the Seminar Enau Kebangsaan (pp. 978-983).

[11] Asyraf, M. R. M., Ishak, M. R., Sapuan, S. M., Yidris, N., Johari, A. N., Ashraf, W., ... & Mazlan, R. (2019). Creep test rig for full-scale composite cr ossarm: simulation modelling and analysis. In Seminar Enau Kebangsaan (pp. 34-38).

[12] A. Environmentally, F. Tropical, and C. Hazard, "Journal of Advances in Modeling Earth Systems," pp. 223–241, 2018.

[13] L. Metz, N. Maheswaranathan, B. Cheung, and J. Sohl-Dickstein, "Learning Unsupervised Learning Rules," 2018.

[14] Indarto, "Penginderaan Jauh : Metode Analisis & Interpretasi Citra Satelit," no. June, 2017.

[15] A. Peytchev, A. Peytchev, and E. Peytcheva, "Reduction of Measurement Error due to Survey Length: Evaluation of the Split Questionnaire Design Approach," Surv. Res. Methods, vol. 11, no. 4, pp. 361–368, 2017.

[16] A. Zhang and Y. Xie, "Chaos theory-based data-mining technique for image endmember extraction: Laypunov index and correlation dimension (L and D)," IEEE Trans. Geosci. Remote Sens., vol. 52, no. 4, pp. 1935–1947, 2014.

[17] Angus, M. (2003). Development centre studies the world economy historical statistics: historical statistics. OECD Publishing.

[18] Bland, J. M., & Altman, D. G. (2003). Applying the right statistics: analyses of measurement studies. Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology, 22(1), 85-93.

[19] Burton, A., Altman, D. G., Royston, P., & Holder, R. L. (2006). The design of simulation studies in medical statistics. Statistics in medicine, 25(24), 4279-4292.

[20] Critchley, L. A., & Critchley, J. A. (1999). A meta-analysis of studies using bias and precision statistics to compare cardiac output measurement techniques. Journal of clinical monitoring and computing, 15(2), 85-91.

[21] Peto, R., Darby, S., Deo, H., Silcocks, P., Whitley, E., & Doll, R. (2000). Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. Bmj, 321(7257), 323-329.