Awwad & Ismail, *Cogent Business & Management* (2018), 5: 1477499
https://doi.org/10.1080/23311975.2018.1477499

cogent business & management

Received: 04 August 2016
Accepted: 02 November 2016
First Published: 25 May 2018

*Corresponding author: Emad Abdelgalil Ali Ismail, Quantitative Analysis Department, King Saud University, Saudi Arabia; Mathematics & Insurance Department, Cairo University, Egypt
E-mail: emadali@ksu.edu.sa

Reviewing editor:
David McMillan, University of Stirling, UK

Additional information is available at the end of the article

BANKING & FINANCE | RESEARCH ARTICLE

Evaluation of the degree of policyholder's risk for the individual's health insurance coverage

Fuad A. A. Awwad[1] and Emad Abdelgalil Ali Ismail[1,2]*

Abstract: A policyholder's degree of health risk could be classified as normal or better than normal or high or bad. We examine the relationship between the policyholder's degree of health risk and the effect of his demographic factors. A quantitative model is proposed to support decision-underwriting of the insurer by segmenting the health insurance underwriting portfolio to four risk groups, which are different and mutually exclusive (low risk, normal risk, high risk, bad risk) based on some demographic factors affecting the degree of risk. The likelihood of the insured to risk groups has been estimated using polynomial logistic regression analysis, and the degree of risk most likely has been determined to take appropriate underwriting decision. This study is based on experience of one of the insurance companies in Saudi Arabia, and the subjects were selected using a random sample for detailed data on individual health insurance during the period 2013–2015, based on the random numbers generated. We found a relationship between the degree of health risk and the policyholder's demographic factors. Using this result, we were able to calculate the probabilities of affiliation of the insured for various degrees of risk. This paper presents a model for the rationalization of underwriting decisions in

ABOUT THE AUTHORS

Fuad A. Awwad is an associate professor of Applied Statistics in Quantitative Analysis Department, College of Business Administration, King Saud University, Saudi Arabia. He earned his Ph.D. in Applied Statistics from Lancaster University, England in 1999. Research interests are as follows: use of statistics in decision-making, use and misuse of statistics in business, and statistical modeling of health data.

Emad Abdelgalil Ali Ismail is an associate professor of Applied Statistics in Quantitative Analysis Department, College of Business Administration, King Saud University, Saudi Arabia. He is a professor of Risk Management & Insurance in Mathematics and Insurance Department, Faculty of Commerce, Cairo University, Egypt. He earned his Ph.D. in Risk Management & Insurance from Cairo University, Egypt in 2005. He published many papers in the field of risk management and insurance, nearly 20 researches published locally and internationally. His research interests are pricing risks, reinsurance, underwriting in risks, natural disaster losses, merge insurers and health insurance.

PUBLIC INTEREST STATEMENT

It is one of the most important activities carried out by insurance companies underwriting activity in risks, where underwriting activity means evaluating the degree of risk of the policyholders. This activity involves preparation to calculating the appropriate insurance premium. This paper is concerned with evaluating policyholder's degree of risk for the individual health insurance coverage, which could be categorized as follows: (1) low risk, (2) normal risk, (3) high risk, (4) bad risk; according to certain demographic factors (age, residence, nationality, marital status, gender, occupation, and family history) affecting the degree of risk. The results of this paper help to rationalize underwriting decisions in the insurance companies, because the proposed model helps in calculating the probabilities of various policyholders' degrees of risk for the individual health insurance coverage. Furthermore, categorization of the degree of risk enables the insurer to achieve right underwriting decisions.

the individual health insurance, by classifying the policyholder within the appropriate insurance risk group. In addition, this paper enables to determine the appropriate insurance premium for every policyholder according to his degree of risk and this leads to reduction of the possibility of adverse selection of insurer.

Subjects: Social Sciences; Insurance; Risk Management

Keywords: Health insurance; degree of risk; underwriting; demographic factors

1. Introduction

By the end of 2014, the number of insurance and reinsurance companies licensed in the Saudi market totaled 35 companies, 28 of them are qualified by the Cooperative Health Insurance Council to provide medical insurance services. General insurance includes seven sub-activities namely vehicles, marine, aviation, energy, engineering, insurance, accident, and liability insurance, as well as insurance on property and against fire. The risks to insurance companies vary according to the risk of major insurance activities, as well as competition and growth rates for each insurance activity (Saudi Arabian Monetary Agency (SAMA), 2014).

Medical insurance represented 52% of the insurance market at the end of 2014, and vehicle insurance accounted for 26.5%. Consequently, the medical insurance and vehicle insurance represented 78.5% of the size of the insurance market, while protection and savings insurance represented only 2.6% (Saudi Insurance Sector, 2014).

On the other hand, the total paid claims paid rose by 26% to SAR 20.5 billion in 2014 compared to SAR 16.7 billion in 2013. Net claims incurred for insurance companies amounted to SAR 17.6 billion in 2014 growing 11.2% over the previous year where the figure reached SAR 15.8 billion. The claims of medical insurance accounted for 60% of the total claims incurred during the year (Saudi Insurance Sector, 2014).

Table 1 shows the results of net incurred claims for health insurers that displayed a 6% increase to SAR 10.4 billion, thus the loss ratio decreased to 79% compared to 94% in 2013 (Saudi Insurance Sector, 2014). Table 2, which indicates the loss ratio by line of business, reveals that the loss ratio in the health insurance is the second largest in the various lines of business. Table 3 shows the fluctuations in the loss ratio in the health insurance during the period 2009–2015. These statistics indicate that the underwriting process in the health insurance is not a rational or do not contribute to the improvement of health insurance results over time.

The health insurance underwriting cycle reflects the tendency for health insurance premiums and insurer profitability to systematically fluctuate over time (Patricia Born, 2008). Underwriting in risks is the process by which the insurer decides whether or not to accept a proposal of insurance,

Table 1. Results of health insurance processes

Health insurance (SAR million)	2013	2014	Growth
Gross Written Premium	12,778	15,750	23%
Net Written Premium	11,317	14,659	30%
Net Earned Premium	10,553	13,259	26%
Net Incurred Claims	9,900	10,448	6%
Retention	89%	93%	
Loss Ratio	94%	79%	

Table 2. Loss ratio by line of business in Saudi Insurance Market (2015)

Line of Business	Loss Ratio
Health Insurance	77 %
Motor Insurance	88.3%
Motor Property/Fire Insurance	53%
Engineering Insurance	49%
Accident and Liability Insurance	26.8%
Marine Insurance	46.9%
Energy Insurance	7.1%
Aviation Insurance	26.8%

Table 3. Loss ratio of health insurance (2009–2015)

Years	2009	2010	2011	2012	2013	2014	2015
Loss Ratio	74.8%	71.5%	73%	81.4%	94%	79.2%	77%

on what conditions, in what proportion, and at what price (Diacon & Carter, 1998). This process is the most important for the technical operations in the insurance company. It also has an effect on the outcomes of the insurer's business and may, also, lead to losses that the insurance company may not be able to survive. Underwriting of individual health risks are those processes relating to the evaluation of an individual's exposures to dangers and the possibility of coverage, to help make appropriate underwriting decision. These decisions may be to accept or deny the coverage or acceptance with conditions. Then, it is classified risk unit within the appropriate risk group within its risk underwriting insurer portfolio.

In some very exceptional circumstances, an underwriter may have little previous experience to assess potential claims, and he then must base his assessment largely on gut reaction. However, far more commonly, an underwriter has the benefit of experience of many similar previous claims, and this can be analyzed and used. He can then determine the major underwriting factors (that is, the characteristics that are most likely to influence annual claims costs under the contract) and then classify contracts according to those factors. Identifying and measuring these factors or characteristics require detailed statistical analysis (Diacon & Carter, 1998).

There are several procedures performed by the underwriter in underwriting health risks, as follows:

(i) Determine major underwriting factors affecting the degree of health risk, which depends on the underwriter's experience. According to these factors, they are insured and divided into different risk groups from each other, and each risk group of the insured is similar in the degree of health risk.
(ii) Measure the average annual claims for each risk group, using the frequency distribution data for each of the number of claims and the size of claims.
(iii) Evaluation of the proposed health risk, through the study of factors affecting the degree of risk, and classification of the proposed health risk within the appropriate risk group.

Underwriting health risks process aims to minimize the adverse effects that the insurance company may be exposed to, as a result of selection against the company through the new insurance applicants. As well as minimizing the degree of danger in inherent risks within heterogeneous groups. Adverse selection plays a prominent role in the insurance literature owing to its negative implications for the insurer's financial performance and stability. Adverse selection could be a manageable problem for the insurer (Lee Colquitt, Fier, Hoyt, & Liebenberg, 2012; Viswanathan et al., 2007). Therefore, it is the insurance companies that must follow strict underwriting, and that each branch of the insurance branches practiced.

This paper concerning the study of underwriting health risks, as the subscription of this type of insurance is especially important, because the factors affecting the degree of health risks are many, such as age, gender, nationality, marital status, occupation, and place of residence. Underwriting decision on the health risks in this paper is as follows:

- Acceptance of insurance coverage with a discount price.
- Acceptance of insurance coverage at the normal price.
- Acceptance of insurance coverage with the increase in the price.
- Denial of insurance coverage.

2. Literature review

Arrow, Mossin, and Smith have demonstrated that when insurance is priced at actuarially fair rates, the insured prefers policies that offer full coverage. As insurance is not a costless business, insurers sell policies above the actuarially fair premium to cover their expenses. Smith has shown that when health insurance is available at a cost that exceeds the actuarially fair value and the probability of loss is greater than zero, the optimal level of insurance coverage will depend on an individual's degree of risk aversion and the cost of insurance. For a given risk-averse individual, the optimal level of insurance will decrease as the cost of insurance increases. Depending on the shape of the utility function, the optimal level of health insurance may be zero, or it may exceed the value of the asset, human capital, subject to risk (Browne, 1992). At the equilibrium underwriting, low risks obtain greater coverage than they would without underwriting (Brown & Kamiya, 2012). Based on the underwriting behavior of insurance companies in 1988, medical conditions were classified into three categories: conditions that led to denial of coverage, conditions that led to exclusion restrictions, and conditions that led to higher premiums (Kapur, 2004).

There is a paucity of empirical evidence that is consistent with the existence of adverse selection in the U.S. insurance market. Some potential reasons for the lack of evidence include the fact: (i) that insurers effectively use underwriting and pricing to counteract adverse selection; or (ii) that consumers either do not have, or fail to take advantage of, private information. (Lee Colquitt et al., 2012) Discussion about several strategies to prevent or to counteract the observed negative spillover effects of supplementary insurance. Health insurers may have become more inclined to calculate risk-rated premiums and to use medical underwriting to prevent high-risk applicants from enrolling (Roos & Schut, 2012). The U.S. health care reform debate and legislation discussed the potential effects of the mandate that individuals have health insurance in conjunction with proposed premium subsidies and health insurance underwriting and rating restrictions (Harrington, 2010). An indicator of underwriting profitability in property-liability insurance have changed over time. The findings asserted that underwriting profit has worsened in recent years, and combined ratios are non-stationary. The study affirmed that lifestyles and one's health have an important impact on the underwriting process in health care field (Leng, 2006). A number of alternative explanations have been offered for insurance underwriting cycles. However, no study till date has empirically evaluated this tendency in the health insurance industry. The study used national data over the period from 1960 to 2004 to test if various theories pertaining to price movements in the property and casualty insurance industry can also explain premium behavior in the health insurance industry. The empirical results provide strong support for the capacity constraint, fluctuation in interest rate, and rational expectations with institutional intervention hypotheses (Patricia Born, 2008). Underwriters considered the certain background medical information about four pairs of hypothetical applicants. One member of each pair was described as having positive genetic test information. In seven instances, an adverse underwriting action was taken on applicants based on their genetic test result; in two others, participants indicated uncertainty as to how to underwrite an applicant with genetic test information. In seven of these 92 applications, underwriters said that they would deny coverage, place a surcharge on premiums, or limit covered benefits based on an applicant's genetic information (Pollitz, Peshkin, Bangit, & Lucia, 2007).

Jason Brown and Mark Warshawsky use numerous demographic and health characteristics, this allows for analysis of disability and mortality risk across a number of dimensions, and they find that different risk groups at age 65 have similar projected long-term care expenses, but that the level—periodic—premium structure of most long-term care insurance policies creates incentives for individuals to separate into different risk pools according to observable characteristics, justifying the underwriting observed on the market (Brown & Warshawsky, 2013).

2.1. Objective of the study

The objective of this paper is to evaluate the degree of risk of the policyholder for the individual's health insurance coverage, by examining the relationship between the degree of the individual's health risk and demographic factors affecting the insured and then, propose a quantitative model to support decision-underwriting of the insurer. Achieving this aim helps reduces the possibility of adverse selection of insurer, because every policyholder will pay the insurance premium that commensurate with his degree of risk as well as denying coverage to policyholder with bad risk.

3. Methodology

This paper is for measuring the risks associated with the process of individual health insurance underwriting. A random sample of 1658 insured individuals was obtained from one of the Saudi insurance companies during the period 2013–2015, based on the random numbers generated, for detailed data on individual health insurance. The data were analyzed using Cluster Analysis, One-Way ANOVA, and Multinomial Logistic Regression.

3.1. Assumptions of the model

We assume the following:

(i) The degree of individual health risk varies from one person to another depending on the policyholder's demographic factors.

(ii) The degree of health risk of policyholder is one out of four mutual alternatives, which are as follows: low, normal, high, and bad risk.

(iii) Insurer's underwriting decision making for individual health risk of policyholder depends on the category of the degree of risk.

3.2. Mathematical framework

Cluster Analysis is for dividing the data obtained to the risk groups or clusters that are different and mutually exclusive, and each has its own characteristics, which considers all risk groups internally homogeneous and different from the other risk groups. We can perform analysis of variance test in one direction (One-Way ANOVA), to make sure the differences means of various groups of risks, and test the following null hypothesis:

$$H_0: \ \mu_1 = \mu_2 = \mu_3 = \mu_4$$

3.3. Multinomial or polytomous logistic regression

When the dependent variable is qualitative, discrete, and has several limits or responses, and independent variables are a mixture of quantitative types of variables (discrete and continuous) it would be appropriate to use a Multinomial Logistic Regression. This model has many uses in the process of life, especially in the medical field, when a dependent random variable has several responses, such as assessing the prospects for the symptoms of a disease that (no—there is simple

—there is an average—there are chronically), or when it comes to choose the way of one of the ways of the diet, and in all the previous cases are estimated probability of each response from the variable responses, and determine the most probable value, so as to support making the right decision (Cohen, Patricia Cohen, & West, 2003).

The probability of responses is calculated as follows:

- Model for Probability of Low risk group

$$\hat{P}(Y=0/X)=e^{h_0(x)}/\left[1+e^{h_0(x)}+e^{h_1(x)}+e^{h_2(x)}\right]$$

- Model for Probability of Normal risk group

$$\hat{P}(Y=1/X)=e^{h_1(x)}/\left[1+e^{h_0(x)}+e^{h_1(x)}+e^{h_2(x)}\right]$$

- Model for Probability of High risk group

$$\hat{P}(Y=2/X)=e^{h_2(x)}/\left[1+e^{h_0(x)}+e^{h_1(x)}+e^{h_2(x)}\right]$$

- Model for Probability of Bad risk group

$$\hat{P}(Y=3/X)=1/\left[1+e^{h_0(x)}+e^{h_1(x)}+e^{h_2(x)}\right]$$

where

$$h_0(x)=\hat{\alpha}_0+\sum_{i=1}^{n}\hat{\beta}_{0i}X_i$$

$$h_1(x)=\hat{\alpha}_1+\sum_{i=1}^{n}\hat{\beta}_{1i}X_i$$

$$h_2(x)=\hat{\alpha}_2+\sum_{i=1}^{n}\hat{\beta}_{2i}X_i$$

3.3.1. Estimating model parameters

The likelihood can be generalized to include G outcome categories by taking the product of each individual's contribution across the G outcome categories (Hosmer & Lemeshow, 1989; Saudi Insurance Sector, 2014):

$$L(Y)=\prod_{j=1}^{n}\prod_{g=0}^{g-1}P(Y=g/X)^{y_{jg}}$$

$$where\ y_{jg}\begin{cases}1\ if\ the\ jth\ subject\ has\ D=g\\ (g=0,1,\ldots,G-1)\\ 0\ if\ otherwise\end{cases}$$

Estimated $\alpha's$ and $\beta's$ are those that maximize likelihood.

3.3.2. Wald test

We test the significance of interaction term at each level, for example:

$$H_0:\beta_{g1}=0;g=1,2,\ldots g-1$$

$$H_0:\beta_{g2}=0;g=1,2,\ldots g-1$$

Wald test Statistic:

$$Z = \frac{\hat{\beta}_{gi}}{s_{\hat{\beta}_{gi}}} \sim N(0,1)$$

3.4. Data description

3.4.1. Dependent variable

Dependent variable is the degree of risk, assuming that the Y has several response variables (A, B, C, D), where

- C: Low risk group (cluster 0)

- A: Normal risk group (cluster 1)

- B: High risk group (cluster 2)

- D: Bad risk group (cluster 3)

$$y = \begin{cases} 0 \; if \; y = C \\ 1 \; if \; y = A \\ 2 \; if \; y = B \\ 3 \; if \; otherwise \end{cases}$$

3.4.2. Independent variables

Independent variables are health insurance underwriting factors (policyholder's demographic factors) as follows:

X1: Age

Age affects annual claim costs differently, depending on the type of benefit involved, although both frequency and severity generally increase with the advancement in age for all types of benefits. Most individual medical expense policies are limited as to the amount and type of coverage after a certain age, such as 65 or 70, although some companies have made lifetime coverage available (Black & Skipper, 2000).

X2: Residence

This variable is qualitative, and was regarded as a binary classification (inside the city/other), where

$$X_2 = \begin{cases} 1 \; if \; inside \; the \; city \\ 0 \; if \; otherwise \end{cases}$$

X3: Nationality

This variable is qualitative, and was regarded as a binary classification (Saudi/other), where

$$X_3 = \begin{cases} 1 \; if \; Saudi \\ 0 \; if \; otherwise \end{cases}$$

X4: Marital status:

This is a qualitative variable, and was considered a three-category variable (Married/Single/others), where:

$$X_{41} = \begin{cases} 1 \; if \; Married \\ 0 \; if \; otherwise \end{cases}$$

$$X_{42} = \begin{cases} 1 \; if \; Single \\ 0 \; if \; otherwise \end{cases}$$

X5: Gender:

As with life insurance, a person's gender is of considerable significance in health insurance underwriting. Females show higher disability rates than males at all but the upper ages in most studies. This is true even for policies that exclude or limit coverage of pregnancy, miscarriage, abortion, and similar occurrence (Black & Skipper, 2000). This variable is qualitative, and was regarded as a binary classification (Male/other), where

$$X_5 = \begin{cases} 1 \; if \; Male \\ 0 \; if \; otherwise \end{cases}$$

X6: Occupation:

Occupational risk has two offsetting effects on the purchase of personal accident, sickness, and health insurance (Diacon & Carter, 1998). This variable is qualitative, and was regarded as a binary classification (Employee/other), where

$$X_6 = \begin{cases} 1 \; if \; Employee \\ 0 \; if \; otherwise \end{cases}$$

X7: Family History:

There's not much you can do about your gene pool. However, a family history of stroke, cancer, or other serious medical conditions may predispose you to these ailments and lead to higher rates. Carriers are usually interested in any conditions your parents or siblings have experienced, particularly if they contributed to a premature death. Some carriers put more emphasis on your family's health than others. However, it is likely to have some impact on your premium. This is a qualitative variable, and was considered a four-category variable (Fit/Middle/Not fit/etc.), where

$$X_{71} = \begin{cases} 1 \; if \; fit \\ 0 \; if \; otherwise \end{cases}$$

$$X_{72} = \begin{cases} 1 \; if \; middle \\ 0 \; if \; otherwise \end{cases}$$

$$X_{73} = \begin{cases} 1 \; if \; not \; fit \\ 0 \; if \; otherwise \end{cases}$$

Appendix A shows descriptive statistics of various demographic factors. It can be observed that age is only the variable in ratio scale, and the rest of the variables are in nominal scale.

4. Data analysis and findings of the study

The data of individual health insurance claims and policyholder's demographic factors collected were analyzed using IBM SPSS statistics 22.

4.1. Groups of individual health insurance risks

The individual health insurance to four groups or clusters of claims data are divided according to the demographic factors influencing (age, residence, nationality, marital status, gender, occupation, and family history). These groups are internally homogeneous and mutually exclusive, using cluster analysis technique. Table 4 provides the number of claims in each risk group. We assume that clusters are levels outcome of dependent variable. Table 5 shows Descriptive Statistics of group risks mean of claims, standard deviation, standard error of the estimate, and confidence interval of 95% for each risk group or cluster. In addition, we observed from the table that cluster 3

Table 4. Number of cases in each cluster

Cluster	Number of Cases
1	5
2	1,523
3	108
4	22
Total	1,658

is the most dangerous risk groups and cluster 0 is the lowest dangerous risk groups. Thus, the total numbers of claims have been divided into four graded risk groups. One way ANOVA tests the differences between the average amount of claims for the risk groups.

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

Table 6 provides F value and its level significance *p*-value is zero, so we reject the null hypothesis and accept the alternative that there are differences between the means of amount of claims for the four risk groups.

4.2. Underwriting model in the individual health risks

Multinomial Logistic Regression is used to calculate the probabilities of the policyholder's affiliation for different groups of risk, to determine the most likely value. Using the following equations (Appendix B):

$$h_0(x) = 61.907 + .413X_1 - 12.053X_2 - 14.212X_3 - 27.868X_{41} - 14.833X_{42} - 11.640X_5 + 17.55X_6 - 0.971X_{71} - 9.615X_{72} - 12.724X_{73}$$

$$h_1(x) = 52.358 + .173X_1 - 9.58X_2 - 12.065X_3 - 12.037X_{41} + 1.552X_{42} - 11.277X_5 + 17.284X_6 - 2.182X_{71} - 12.002X_{72} - 12.154X_{73}$$

$$h_2(x) = 42.589 + .028X_1 + 1.947X_2 - 10.111X_3 - 9.746X_{41} + 1.201X_{42} - 8.003X_5 + 18.486X_6 - 9.629X_{71} - 18.031X_{72} - 17.607X_{73}$$

4.2.1. Goodness of fit

4.2.1.1. Likelihood ratio test. As with a standard logistic regression, we can use a likelihood ratio test to assess the significance of the independent variable in our model (Kleinbaum & Klein, 2003).

In this paper, we have a four-level outcome variable and *p* independent variables for each of the outcome comparison. We are being by fitting a full model (with the exposure variable in it) and then comparing that to a reduced model containing only the intercept. The null hypothesis is that the beta coefficients corresponding to the exposure variable are both equal to zero. The likelihood ratio test is calculated as negative two times the log likelihood (log L) from the reduced model minus negative two times the log likelihood from the full model. The resulting statistic is distributed approximately chi-square, with degree of freedom (df) equal to the number of parameters set equal to zero under the null hypothesis, as follows:

$$H_0: \beta_{gi} = 0; g = 1, 2, \dots g - 1,\ i = 1, 2, \dots, p$$

Likelihood ratio test statistic:

$$-2\ log\ L_{reduced} - (-2\ log\ L_{full}) \sim x^2$$

Table 7 shows that negative two times the log likelihood for the reduced model is 1,087.161, and the full model is 259.957. The difference is 827.204. The chi-square *p*-value for this test statistic,

Table 5. Descriptive statistics of group risks

Cluster	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
0	1,523	510.25	1,429.613	36.633	438.39	582.10	0	7,509
1	108	14,196.69	6,454.852	621.118	12,965.40	15,427.99	7,588	32,581
2	22	54258.59	13,488.263	2875.707	48,278.23	60,238.95	35,239	81,000
3	5	1,22,310.80	15,747.082	7042.309	1,02,758.22	1,41,863.38	1,05,200	1,48,000

Table 6. One way ANOVA test results

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	1.515E11	3	5.051E10	6,746.891	.000
Within Groups	1.238E10	1654	7,485,672.514		
Total	1.639E11	1657			

Table 7. Model fitting information

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	1,087.161			
Final	259.957	827.204	30	.000

with 30 degrees of freedom, is 0. We conclude that the independent variables (policyholder's demographic factors) are statistically significant at the 0.01 level.

4.2.1.2. Mcfadden R^2. McFadden in multinomial logistic regression model is similar to the coefficient of determination in linear regression, and has the same concept and characteristics. It has been calculated by McFadden in 1974, where (Cohen et al., 2003; Lattin, Douglas Carroll, & Green, 2003)

$$McFadden\ R^2 = 1 - \frac{\log L_{full}}{\log L_{reduced}}$$

It also has other measures similar to the measure, such as the following: $R_L^2,\ R^2_{Cox\ Snell}$

$$R_L^2 = \frac{\log L_{full} - \log L_{reduced}}{\log L_{full} - 1}$$

$$R^2_{Cox\ Snell} = 1 - \left(\frac{\log L_{reduced}}{\log L_{full}}\right)^{2/n}$$

It also has another measure called Nagelkerke, which depends on $R^2_{Cox\ Snell}$ by dividing the largest estimated value. Table 8 according to McFadden shows that 75.4% of the variation in the degree of risk is to interpret variations in policyholder's demographic factors. 39.3%, 81.2% according Cox Snell and Nagelkerke, respectively.

MathCAD version 3.1 was applied for obtaining multiple logistic regression model, attachment 6 applications, describes the different degree of risk, depending on the policyholder's demographic characteristics. Appendix C shows how various demographic factors affect the degree of risk of policyholders. For example, case 1 considers the case of a male policyholder aged 30 years with middle medical history. The resulting degree of risk was high. When some of the above demographic variables are changed, as illustrated in case 2, the degree of risk will also change from high to low risk and hence it can be noted that the demographic factors of the policyholder affect the degree of risk.

Table 8. Pseudo R-square

Cox and Snell	.393
Nagelkerke	.812
McFadden	.754

5. Discussion of the results

The results show that policyholder's demographic characteristics affect the degree of risk. The same has been demonstrated in Appendix C for six different cases. According to Nagelkerke coefficient, 81.2% of the variation in the degree of risk is explained by variations in policyholder's demographic factors.

6. Conclusion

We examined the relationship between the policyholder's degree of individual health risk and the effect of demographic factors. Data of 1,658 insured were obtained from one Saudi insurance company, and we got a detailed data about individual health insurance during the period 2013–2015. We estimated the policyholders' probabilities to risk groups and determined the degree of most likely risks. This helps the insurer in making underwriting decisions in individual health risks. This paper presented a model for the rationalization of underwriting decisions in the individual health insurance, by classifying the policyholder within the appropriate insurance risk group. In addition, this paper contributes to determine the appropriate insurance premium for every policyholder according to his degree of risk as well as denying coverage to policyholder with bad risk. This leads to reduction in the possibility of adverse selection of insurer.

Acknowledgments
The authors would like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University for its funding this Research group NO **[RG-1435-088**].

Funding
The authors would like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University for its funding this Research group NO [RG-1435-088].

Author details
Fuad A. A. Awwad[1]
Emad Abdelgalil Ali Ismail[1,2]
E-mail: emadali@ksu.edu.sa
[1] Quantitative Analysis Department, King Saud University, Saudi Arabia.
[2] Mathematics & Insurance Department, Cairo University, Egypt.

Citation information
Cite this article as: Evaluation of the degree of policyholder's risk for the individual's health insurance coverage, Fuad A. A. Awwad & Emad Abdelgalil Ali Ismail, *Cogent Business & Management* (2018), 5: 1477499.

References

Black, K., & Skipper, H. D. (2000). *Life & health insurance* (13th ed.). New Jersey: Prentice-Hall, Inc.

Brown, J., & Warshawsky, M. (2013). The life care annuity: A new Empirical Examination of an insurance innovation that addresses problems in markets for life annuities and long-term care insurance. *The Journal of Risk and Insurance*, *80*(3), 677–703. doi:10.1111/j.1539-6975.2013.12013.x

Brown, M. J., & Kamiya, S. (2012). A theory of the demand for underwriting. *The Journal of Risk and Insurance*, *79*(2), 335–349. doi:10.1111/jori.2012.79.issue-2

Browne, M. J. (1992). State restrications on health insurance underwriting criteria: The effect on the uninsured population. *Journal of Insurance Regulation*, *10*(4), 585–597.

Cohen, J., Patricia Cohen, S. G., & West, L. S. A. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). New Jersey: Lawrence Erlbaum Association, Inc.

Diacon, S. R., & Carter, R. L. (1998). *Success in insurance* (3rd ed.). Guildford and King's Lynn: Biddles Ltd.

Harrington, S. E. (2010). The health insurance reform debate. *Journal of Risk and Insurance*, *77*(1), 5–38. doi:10.1111/j.1539-6975.2009.01345.x

Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression*. New York, NY: Wiley.

Kapur, K. (2004). The impact of the health insurance market on small firm employment. *Journal of Risk and Insurance*, *71*(1), 63–90. doi:10.1111/j.0022-4367.2004.00079.x

Kleinbaum, D. G., & Klein, M. (2003). *Logistic regression a self-learning text* (2nd ed.). New York, NY: Springer-Veriag, Inc.

Lattin, J. M., Douglas Carroll, J., & Green, P. E. (2003). *Analyzing multivariate data*. Canada: Thomson Brooks/Cole, Inc.

Lee Colquitt, L., Fier, S. G., Hoyt, R. E., & Liebenberg, A. P. (2012, March). Adverse selection in the credit life insurance market, *Journal of Insurance Regulation*, *31*, 157–180. National Association of Insurance Commissioners.

Leng, C. (2006). Stationary and stability of underwriting profits in property-liability insurance: Part I. *The Journal of Risk Finance*, *7*(1), 38–48. doi:10.1108/15265940610637799

Patricia Born, R. E. (2008). Unraveling the health insurance underwriting cycle. *Journal of Insurance Regulation*, *26*(3), 65–84.

Pollitz, K., Peshkin, B., Bangit, E., & Lucia, K. (2007). Genetic discrimination in health insuraiice: Current legal protections and industry practices. *Excellus Health Plan, Inc*, *44*, 530–568.

Roos, A.-F., & Schut, F. T. (2012). Spillover effects of supplementary on basic health insurance: Evidence from the Netherlands. *The European Journal of Health Economics*, *13*, 51–62. doi:10.1007/s10198-010-0279-6

Saudi Arabian Monetary Agency (SAMA). (2014). Retrieved from www.sama.gov.sa

Saudi Insurance Sector. (2014). *Albilad capital*. Retrieved March 2015, from www.**albilad-capital**.com

Viswanathan, K. S., Lemaire, J., Withers, K., Armstrong, K., Baumritter, A., Hershey, J. C., … Asch, D. A. (2007). Adverse selection in term insurance purchasing due to the BRCA 1/2 genetic test and elastic demand. *The Journal of Risk and Insurance*, *74*(1), 65–86. doi:10.1111/j.1539-6975.2007.00202.x

Appendix A. Descriptive statistics of various demographic factors

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Age X1	1,658	18	66	51.70	8.770	76.912

		Frequency	Percent	Cumulative Percent
Residence X2	Inside the city	1,149	69.3	69.3
	Outside the city	509	30.7	100.0
Total		1,658	100.0	
Nationality X3	Saudi	1,128	68.0	68.0
	Non Saudi	530	32.0	100.0
Total		1,658	100.0	
Marital status X4	Married	645	38.9	38.9
	Single	1,013	61.1	100.0
Total		1,658	100.0	
Gender X5	Male	1,137	68.6	68.6
	Female	521	31.4	100.0
Total		1,658	100.0	
Occupation X6	Employee	1,335	80.5	80.5
	Unemployed	323	19.5	100.0
Total		1,658	100.0	
Family History X7	Fit	546	33	33
	Middle	702	42.3	75.3
	Not fit	410	24.7	100
Total		1,658	100.0	

Appendix B. Parameter estimates of Multinomial Logistic Regression

Parameter Estimates									
y[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
								Lower Bound	Upper Bound
.00	Intercept	61.907	475.690	.017	1	.896			
	x1	.413	.114	13.261	1	.000	1.512	1.210	1.889
	x2	−12.053-	176.060	.005	1	.945	5.830E-6	8.003E-156	4.247E144
	x3	−14.212-	186.338	.006	1	.939	6.730E-7	1.647E-165	2.749E152
	x41	−27.868-	349.163	.006	1	.936	7.889E-13	.000	1.274E285
	x42	−14.833-	277.387	.003	1	.957	3.617E-7	2.792E-243	4.685E229
	x5	−11.640-	195.921	.004	1	.953	8.808E-6	1.502E-172	5.167E161
	x6	17.550	2904.206	.000	1	.995	4.186E7	.000	.[b]
	x71	−.971-	261.975	.000	1	.997	.379	3.847E-224	3.731E222
	x72	−9.615-	16.196	.352	1	.553	6.675E-5	1.092E-18	4.078E9
	x73	−12.724-	16.109	.624	1	.430	2.979E-6	5.779E-20	1.536E8
1.00	Intercept	52.358	386.444	.018	1	.892			
	x1	.173	.111	2.440	1	.118	1.189	.957	1.479
	x2	−9.580-	176.059	.003	1	.957	6.911E-5	9.503E-155	5.026E145
	x3	−12.065-	186.337	.004	1	.948	5.757E-6	1.411E-164	2.349E153
	x41	−12.037-	212.069	.003	1	.955	5.921E-6	1.814E-186	1.932E175
	x42	1.552	1.519	1.044	1	.307	4.720	.240	92.621
	x5	−11.277-	195.920	.003	1	.954	1.267E-5	2.162E-172	7.420E161
	x6	17.284	2904.206	.000	1	.995	3.209E7	.000	.[b]
	x71	−2.182-	261.975	.000	1	.993	.113	1.146E-224	1.111E222
	x72	−12.002-	16.188	.550	1	.458	6.134E-6	1.019E-19	3.691E8
	x73	−12.154-	16.109	.569	1	.451	5.267E-6	1.022E-19	2.715E8
2.00	Intercept	42.589	397.735	.011	1	.915			
	x1	.028	.116	.059	1	.809	1.028	.820	1.290
	x2	1.947	200.272	.000	1	.992	7.004	2.361E-170	2.078E171
	x3	−10.111-	186.339	.003	1	.957	4.062E-5	9.933E-164	1.661E154
	x41	−9.746-	212.072	.002	1	.963	5.851E-5	1.785E-185	1.918E176
	x42	1.201	1.566	.588	1	.443	3.324	.154	71.581
	x5	−8.003-	195.924	.002	1	.967	.000	5.672E-171	1.972E163
	x6	18.486	2904.206	.000	1	.995	1.068E8	.000	.[b]
	x71	−9.629-	261.486	.001	1	.971	6.576E-5	1.741E-227	2.484E218
	x72	−18.031-	1.646	119.976	1	.000	1.476E-8	5.859E-10	3.718E-7
	x73	−17.607-	.000	.	1	.	2.257E-8	2.257E-8	2.257E-8

[a]The reference category is 3.00.

[b]Floating point overflow occurred while computing this statistic. Its value is therefore set to system missing.

https://doi.org/10.1080/23311975.2018.1477499

Appendix C. Effect of various demographic factors on the degree of risk of policyholders

Case_1

			Residence		Nationality		Marital status		Gender		Occupation		Medical History		
		Age	Inside	Outside	Saudi	Non	Married	Single	male	female	Employee	non	Fit	Middle	Non
Risk factors		**30**	**√**		**√**		**√**		**√**		**√**			**√**	
h0(x)	*16.459*	Probabilities Degree of Risks			Low-risk		0.104		Degree of Risk				( ) Low-risk		
h1(x)	*17.871*				Normal risk		0.426						( ) Normal risk		
h2(x)	*17.971*				High risk		0.471						(∏) High risk		
					Bad risk		7.377E-9						( ) Bad risk		
					Sum		1								

Case_2

			Residence		Nationality		Marital status		Gender		Occupation		Medical History		
		Age	**Inside**	**Outside**	**Saudi**	**Non**	**Married**	**Single**	**male**	**female**	**Employee**	**non**	**Fit**	**Middle**	**Non**
Risk factors		**40**	**√**		**√**		**√**		**√**		**√**		**√**		
h0(x)	*40.873*	Probabilities Degree of Risks			Low-risk		0.543		Degree of Risk				(∏) Low-risk		
h1(x)	*40.698*				Normal risk		0.426						( ) Normal risk		
h2(x)	*34.656*				High risk		0.001084						( ) High risk		
					Bad risk		0						( ) Bad risk		
					Sum		1								

https://doi.org/10.1080/23311975.2018.1477499

Case_3

			Residence		Nationality		Marital status		Gender		Occupation		Medical History		
		Age	**Inside**	**Outside**	**Saudi**	**Non**	**Married**	**Single**	**male**	**female**	**Employee**	**non**	**Fit**	**Middle**	**Non**
Risk factors		**35**		**√**	**√**		**√**		**√**		**√**		**√**		
h0(x)	*63.896*	Probabilities Degree of Risks			Low-risk		0.71		Degree of Risk				(∏) Low-risk		
h1(x)	*63.002*				Normal risk		0.29						( ) Normal risk		
h2(x)	*43.516*				High risk		1E-9						( ) High risk		
					Bad risk		0						( ) Bad risk		
					Sum		1								

cogent • business & management

Case_4

			Residence		Nationality		Marital status		Gender		Occupation		Medical History		
		Age	Inside	Outside	Saudi	Non	Married	Single	male	female	Employee	non	Fit	Middle	Non
Risk factors		22	√		√		√		√		√		√		√
h0(x)	−7.504	Probabilities Degree of Risks			Low-risk		2.602E-4		Degree of Risk				( ) Low-risk		
h1(x)	−0.949				Normal risk		0.183						( ) Normal risk		
h2(x)	0.315				High risk		0.345						( ) High risk		
					Bad risk		0.472						(∏) Bad risk		
					Sum		1								

https://doi.org/10.1080/23311975.2018.1477499

Case_5

			Residence		Nationality		Marital status		Gender		Occupation		Medical History		
		Age	**Inside**	**Outside**	**Saudi**	**Non**	**Married**	**Single**	**male**	**female**	**Employee**	**non**	**Fit**	**Middle**	**Non**
Risk factors		**33**	**√**		**√**		**√**		**√**		**√**				**√**
h0(x)	*14.589*	Probabilities Degree of Risks			Low-risk		0.011		Degree of Risk				( ) Low-risk		
h1(x)	*18.238*				Normal risk		0.435						( ) Normal risk		
h2(x)	*18.479*				High risk		0.554						(∏) High risk		
					Bad risk		5.223E-9						( ) Bad risk		
					Sum		1								

cogent • business & management

Case_6

			Residence		Nationality		Marital status		Gender		Occupation		Medical History		
		Age	**Inside**	**Outside**	**Saudi**	**Non**	**Married**	**Single**	**male**	**female**	**Employee**	**non**	**Fit**	**Middle**	**Non**
Risk factors		**33**	**√**		**√**		**√**		**√**		**√**				**√**
h0(x)	17.698	Probabilities Degree of Risks			Low-risk		0.226		Degree of Risk				( ) Low-risk		
h1(x)	18.39				Normal risk		0.451						(∏) Normal risk		
h2(x)	18.055				High risk		0.323						( ) High risk		
					Bad risk		4.653E-9						( ) Bad risk		
					Sum		1								

***Cogent Business & Management* (ISSN: 2331-1975) is published by Cogent OA, part of Taylor & Francis Group.**