DEVELOPING A RISK-ADJUSTED CAPITATION REGIME FOR NIGERIA HEALTH INSURANCE SCHEME (NHIS)

Ismaila Adeleke
University of Lagos, Nigeria
adeleke22000@gmail.com

Abstract

The National Health Insurance Scheme (NHIS) is set up to operate as Public Private Partnership and directed at providing accessible, affordable and qualitative healthcare for all Nigerians. The Nigerian health insurance industry is faced with the challenge of developing a sustainable fees structure owing to serious data deficits pervading the Nigerian healthcare system and unreliable critical health statistics. Using the capitation payment mechanism the health care provider assumes that for a given insured population, the provider will cover all health care services for a fixed payment per member per month. However, in this arrangement, payments assumed equal risk level for all subscribers and this may encourage risk selection. This article proposes the development of risk-adjusted capitation framework for the Nigeria National Health Insurance Scheme using generalized linear models. Findings show that the usual normal-based risk adjustment models could lead to further risk selection for highly skewed and heavy tailed data. The use of generalized gamma-based regression model for risk adjustment and determining fair capitation rates is suggested. The study relied on healthcare cost of various diagnoses based on international classification of diseases and information on enrollment to facilities with regards to enrollees characteristics. This framework would allow the healthcare management system to consolidate on the profile of its present and historical data contained in the production system and provides pathways for clinical and administrative information, actuarial valuation and in-depth statistical analysis. This is in tandem with the healthcare delivery agenda of the Millennium Development Goals (MDGs) and sustainable development. Policy implications and recommendation are discussed.

Keywords: NHIS**,** Risk Adjustment, Capitation, Healthcare, Generalized linear model, Risk-Adjusted models

1. INTRODUCTION

The distinctive factor in the National Health Insurance Scheme (NHIS) is the way the health care consumer pays for medical service. In contrast to the previous fee-for-service system, payments are made by capitation to a primary healthcare provider by the HMOs on behalf of a contributor for services to be rendered by the healthcare provider. The capitation is regularly made in advance whether the enrollee utilizes the facility or not (NHIS, 2012). Using the capitation payment mechanism the health care provider assumes that for a given insured population, the provider will cover all health care services for a fixed payment per member per month (Toso and Farmer, 1994; Ibiwoye and Adeleke, 2008; Adeleke *et. al.*, 2016; Adeleke *et. al.*, 2012). However, in this arrangement, payments assumed equal risk level for all subscribers and this may encourage risk selection. A risk-based capitation arrangement has not been explored for the current NHIS. From a financial perspective, risk selection looks like an 'arbitrage' strategy in which the health plan takes a short position in the poor risks and an offsetting long position in the good risks. To eliminate the arbitrage opportunity, and discourage the health plan from engaging in risk selection, the sponsor must eliminate predictable sources of priced risk from the pricing scheme. If the sponsor can do so, successfully, then the health plan's expected return from serving the sponsor's beneficiaries behaves as a random variable and the incentive diminishes. Most of the recent growth in enrollment has been fostered by a belief in both the public and private sectors that capitated systems offer the potential for lower costs and more comprehensive coverage than is available in the fee-for-service sector.

The goal of the National Health Insurance Scheme is to protect families from the financial burden of huge medical bills, guarantee equitable distribution of health care costs among different income groups, maintain high standard of health care services delivery within the scheme and promote private sector participation in the provision of health care services. The Scheme's objectives are to: ensure that every Nigerian has access to good healthcare services; protect families from the financial hardship of huge medical bills; limit the rise in the cost of healthcare services; ensure equitable distribution of healthcare costs among different income groups; maintain high standard of healthcare delivery services within the scheme; ensure efficiency in healthcare services; improve and harness private sector participation in the provision of healthcare services; ensure adequate distribution of health facilities within the Federation; ensure equitable patronage of all levels of healthcare; and ensure the availability of funds to the health sector for improved services (NHIS, 2012).

However, these goals might not be achievable if adequate attention is not paid to the pricing policies as this could lead to adverse and favorable risk selection, and ultimately increased medical care expenditures for the healthcare programs. In addition, these problems bring windfall profits for certain healthcare providers, and limited access to comprehensive health insurance for individuals who are poor health risks. Uniform capitation rates benefit Health Maintenance Organizations (HMOs) that enroll a healthier case mix. Usually, capitation rates were set based on the expected health care costs of the average beneficiary without adjustments for health status. Because risk adjustment pays insurers more for high-risk individuals and less for low-risk individuals, it reduces the risk selection incentives in the market (Toso and Farmer, 1994). The objective of this study is to explore the development of a risk-adjusted capitation mechanism for the Nigeria National Health Insurance Scheme. This mechanism would account for varying levels of risk using demographic strata (such as age, gender and race) as well as health risk factors such as case mix. The remainder of this article is arranged as follows: section 2 presents the theoretical and conceptual frameworks of the

study; the materials and methods used in the study are discussed in section 3; section 4 deals with data analysis and discussion of results, while section 5 concludes the write-up.

2. THEORETICAL AND CONCEPTUAL FRAMEWORKS

2.1 Economic Theory and Healthcare Expenditures Economics

Health care goods or services are characterized by two essential features that differentiate them from other goods; that is, the demand for health care goods is largely a random phenomenon and, it changes immensely depending on the type of individual or group demanding them (see Alberta Health, 1998; Varde, and Diderichsen, 2000; Porell, Gruenberg, Sawitz, and Beiser, 1989). Generally, the demand curve relates the quantities of goods or services that would be purchased to possible prices; and the prices in the demand analysis are presumed to be paid by the consumer without the option of having the costs being covered by a third party. However, the healthcare market is unique due to the factor of third-party payments usually by public or private health insurance payers. These third-party payments cover significant shares of health costs limiting what health consumers pay out for health goods as a results, the consumer becomes less cost conscious and the quantity purchased at every price tends to increase (Babazono, Weiner, Tsuda, Mino, and Hillman, 1998; Beck, 1998; Mehmud, and Yi, 2012). Thus, third-party payment will shift both the health costs demand curve to the right and produce a more inelastic demand curve for health care, by lowering the price which the individual pays for healthcare. The increased demand for health care generated as a result of having health insurance coverage sometimes lead to moral hazard, since insurance reduces the healthcare costs of the consumer and a rational insured is expected to consume larger quantities of care than if they were paying the entire price. They will expand utilization to the point where perceived marginal benefits of additional care purchased are just equal to the marginal costs incurred for that care. However, these marginal costs paid by the consumer are not the entire marginal costs, so that demand is increased. It is therefore, expected that as health insurance coverage increases, both consumer and physician-induced demand will increase, leading to increases in total expenditures in healthcare. Insurers are expected to continue to respond with cost control and cost-sharing mechanisms. Current and forecasted increases in health expenditures will affect consumer and physician demand for health care services, and also the potential moral hazard impact on demand of third-party payers (Brown, and Amelung, 1999; Carr-Hill, Rice, and Smith. 2000; Chinitz, 1994).

2.2 Conceptual Framework

The conceptual framework adopted for this study is derived from the health care demand and health care expenditures model. The factors captured in the model are: the prices of health care, economic resources, access to health care, life cycle stage, and socio-demographic factors. The price charged by medical provider is often not what individual consumers pay as part might be paid by a third-party payer or by a scheme instituted by a government. Subscription to an insurance scheme or purchase of health insurance play important role in reducing health care costs of consumer. When insured individuals accessed a facility(ies) health insurance may cover a substantial proportion of the amount and as a result, insurance status significantly influences the household demand for health care, which results in an increase in the quantity of services demanded. The demand for health care is constrained by the financial resources available. Relying on Bojanic's (1992), life cycle stages are also adopted for this study. The life cycle variables are created using the age of the enrollee, marital

status, and number of dependents under age eighteen, as these provide useful measures for identifying the level of health care need at each stage of the household life cycle. Age is closely related to the incidence of illness and chronic disease and the desire for health care (Feldstein 1993), and in particular, elderly tend to have higher expenditures than other age groups because health status declines with age (Gregory and Sabelhaus 1995). Married persons tend to be healthier and are less likely to utilize hospital care than single persons, because married persons are more likely to have health insurance coverage than single persons. The number of children by a subscriber is expected to affect the demand for health care. Thus, each life cycle stage, based upon the age and marital status of a household head and the presence of young children, may have varying effects on the need for health care goods and services. Finally, health care utilization is likely to be influenced by socio demographic factors, such as education and household size (Clark, Saunders, Baluch, and Simon, 1995; Varde, and Whitehead, 1997; Ellis, and Pope, 1996; Kahn, Parke, and Yi, 2014). These characteristics account for enrollee preferences with respect to health care utilization, which affects healthcare costs.

3. MATERIALS AND METHODS

3.1 Research Design

The study adopted the exploratory study design. The main purpose of such design is formulating a problem for more precise investigation and developing the working hypotheses from an operational point of view (Kothari, 2004). The major emphasis is on the discovery of ideas and insights, which in this case, trying to understand how various characteristics of enrollees can help develop appropriate rate that is commensurate with risk they bring into the pool. As such the research design appropriate for this kind of study must be flexible enough to provide opportunity for considering different aspects of a problem under study. Inbuilt flexibility in research design is needed because the research problem, broadly defined initially, is transformed into one with more precise meaning in exploratory studies, which fact may necessitate changes in the research procedure for gathering relevant data. The target population of the study was the enrollees of Health Maintenance Organizations (HMOs) under the National Health Insurance Scheme (NHIS) in Nigeria.

3.2 Data Collection Methods

The study used secondary data representing healthcare claims from record of healthcare providers and health maintenance organizations. It is made up of close to sixty thousand utilization payments on behalf of the beneficiaries of health insurance. The data obtained shows dates of birth of beneficiaries, utilization dates, ailments/diagnosis billed, categories of ailments/investigations, provider payments and HMO payments. The different ailments have been categorized using the International Classification of Primary Care 2nd edition (ICPC-2).

3.3 Method of Data Analysis

Data analysis techniques are employed to give meaning to the collected data for the research. It is the systematic maneuvering of available data in order to answer the research questions, test hypotheses and give meaning to data collected during research. The data analysis processes begins with checking of the data for consistencies, completeness, relevance and reliability. Descriptive statistics will be generated to describe the concentration and spread of

the data, after which generalized models are fitted to the data. The fitted models are used to compute the risk scores of the enrollees.

3.4 Risk-Adjustment Modeling

Risk-adjusted models that incorporate the rudimentary demographic data are investigated. A number of different determinants of enrollees' healthcare costs are captured. In this study, the generalized linear model (GLM) was used in developing the risk-adjusted model. A feature of this model is that it expresses the mean response as a function of linear combinations of explanatory variables. By considering regression in the GLM context, it will be possible to handle dependent variables that are not normally distributed and allow us to introduce new applications, such as gamma regressions that are useful for fat-tailed distributions. Estimation procedures for calibrating GLM models, significance tests and goodness-of-fit statistics for documenting the usefulness of the model, and residuals for assessing the robustness of the model fit have been extensively discussed by (Frees, 2010; Jong, and Heller, 2008; Kahn, Parke, and Yi, 2014; Mehmud, and Yi, 2012).

For the test of significant difference in risk scores among the different categories of the members, the analysis of variance technique will be employed. This statistical method is an extremely useful technique concerning researches in the fields of economics, biology, education, psychology, sociology, business/industry and in researches of several other disciplines. It technique is used when multiple sample cases are involved and enables us to perform this simultaneous test and as such is considered to be an important tool of analysis in this study. Using this technique, one can draw inferences about whether the samples have been drawn from populations having the same mean.

4. DATA ANALYSIS AND DISCUSSION OF FINDINGS

Descriptive statistical analysis was used to identify frequencies and percentages of various categories and classifications of the enrollees. It describes the distributions of the healthcare claims costs presented by the providers and the actual claims settlement by the HMO. Tables 1 and 2 give the description of the claims data regarding the age and ailment distribution.

Table 1
Frequency distribution

	Frequency	Percentage
Infancy (0 - 1 year)	2546	4.3
Early childhood (2 - 4 years)	8570	14.4
Childhood (5 - 12 years)	5960	10.0
Adolescence (13 - 16 years)	818	1.4
Late teenage (17 - 20 years)	459	.8
Early adulthood (21 - 40 years)	32457	54.5
Middle age (41 - 70 years)	8267	13.9
Senescence (70+ years)	441	.7
Total	59518	100.0

Table 2
Distribution of data by diagnostics

	Frequency	Percentage
General and Unspecified	24795	41.7
Blood, Blood Formation and immune mechanism	457	.8
Digestive	5917	9.9
Eye	2361	4.0
Ear	427	.7
Cardiovascular	3080	5.2
Musculoskeletal	2028	3.4
Neurological	579	1.0
Psychological	320	.5
Respiratory	7463	12.5
Skin	2630	4.4
Endocrine/Metabolic and Nutritional	755	1.3
Urological	921	1.5
Pregnancy, Child Bearing, Family Planning	5439	9.1
Female Genital	2346	3.9
Total	59518	100.0

Over 50 per cent of the utilization of facilities as are those in the age bracket 21 – 40, i.e. the early adulthood categories while those in the age ranges 17 – 20 and 70+ years account for a little over 1.5 percent utilization. The data also revealed that majority (41.7%) of the utilization is as a result of ailments categorized as general and unspecified disease followed at

a distance by respiratory which account for 12.5 percent. A further scrutiny of the data revealed that majority of the early adulthood group is diagnosed for general and unspecified disease, especially malaria related as presented in Table 3. A chi-square of independence (Table 4) suggests association between age group of members and the type of disease diagnosed at time of utilization.

Table 3
Classification of data by age group and diagnostic

	Age group								
	Infancy (0 - 1 year)	Early childhood (2 - 4 years)	Childhood (5 - 12 years)	Adolescence (13 - 16 years)	Late teenage (17 - 20 years)	Early childhood (21 - 40 years)	Middle age (41 - 70 years)	Senescence (70+ years)	Total
General and Unspecified	1104	3484	2330	348	183	13768	3416	162	24795
Blood, Blood Formation and immune mechanism	14	48	40	2	5	280	68	0	457
Digestive	164	858	627	100	47	3210	855	56	5917
Eye	71	322	215	41	23	1364	309	16	2361
Ear	32	66	45	1	6	209	67	1	427
Cardiovascular	87	487	328	49	37	1609	446	37	3080
Musculoskeletal	85	297	239	20	19	1059	290	19	2028
Neurological	30	112	61	7	2	290	76	1	579
Psychological	18	50	33	3	4	167	45	0	320
Respiratory	474	1048	832	102	55	3927	971	54	7463
Skin	80	371	264	30	17	1445	411	12	2630
Endocrine/Metabolic and Nutritional	23	113	69	15	5	430	97	3	755
Urological	88	148	90	7	5	448	132	3	921
Pregnancy, Child Bearing, Family Planning	155	838	542	62	32	2978	770	62	5439
Female Genital	121	328	245	31	19	1273	314	15	2346
Total	2546	8570	5960	818	459	32457	8267	441	59518

Table 4
A test of significance of the dependence of diagnostic on age

Chi-Square Tests

	Value	Df	Asymp. Sig. (2-sided)
Pearson Chi-Square	450.313	98	.000
Likelihood Ratio	441.598	98	.000
Linear-by-Linear Association	13.375	1	.000
N of Valid Cases	59518		

Analyses of the data by healthcare cost are presented in Tables 6 and 7. There is evidence that the claims data is heavily tailed and highly peaked which suggest that the data is significantly non normal.

Table 5
Descriptive analysis of healthcare claims cost by age

Age group	Mean	N	Std. Deviation	Median	Kurtosis	Skewness
Infancy (0 - 1 year)	6724.149	2546	14988.16	2650.82	343.933	14.327
Early childhood (2 - 4 years)	5801.135	8568	13134.12	2373.55	565.57	16.519
Childhood (5 - 12 years)	6438.945	5958	12696.71	2458.14	182.733	9.323
Adolescence (13 - 16 years)	8397.263	818	14459.73	3000	32.154	4.756
Late teenage (17 - 20 years)	7857.907	459	14833.7	2538.2	35.248	5.059
Early adulthood (21 - 40 years)	11519.06	32448	41556.52	2963.225	1056.507	23.094
Middle age (41 - 70 years)	12356.6	8266	58448.37	3689.157	1365.606	32.91
Senescence (70+ years)	22672.63	441	97362.79	7170	334.285	17.328
Total	10109.77	59504	39357.44	2862	1740.774	32.033

Although the utilization of those in age 70+ is very low, the cost of healthcare for this category of members is the highest compared to other age groups. The average claims by provider and the claims settled by the HMO is the highest for this group.

Table 6
Descriptive analysis of healthcare claims cost by ailments

Classification group	Mean	N	Std. Deviation	Median	Kurtosis	Skewness
General and Unspecified	8678.279	24786	29065.03	2548.526	973.409	22.036
Blood, Blood Formation and immune mechanism	8316.454	457	18516.75	2500.85	27.733	4.795
Digestive	9778.252	5917	28379.51	2665.7	323.959	14.744
Eye	10362.73	2360	41717.15	2995.7	1015.61	27.528
Ear	5175.232	427	10635.39	2035.5	62.141	6.408
Cardiovascular	9058.324	3077	20299	2535.25	87.641	7.442
Musculoskeletal	13043.49	2028	45983.11	3158.375	259.203	14.168
Neurological	8319.032	579	24087.21	2765.55	123.016	9.79
Psychological	12132.21	320	33693.17	3270.5	58.687	7.023
Respiratory	12654.89	7463	47614.31	4546	1688.117	32.271
Skin	10363.11	2629	46805.66	2750	793.558	24.852
Endocrine/Metabolic and Nutritional	8310.45	755	28458.39	2250	95.362	8.861
Urological	14526.88	921	40535.36	5995.72	187.495	11.868

Classification group	Mean	N	Std. Deviation	Median	Kurtosis	Skewness
Pregnancy, Child Bearing, Family Planning	13756.11	5439	74037.98	3403	880.292	26.625
Female Genital	8082.877	2346	21290.11	2640.5	142.303	9.549
Total	10109.77	59504	39357.44	2862	1740.774	32.033

From the exploratory data analysis results displayed in Tables 5 and 6, very positive skewness and heavy tailed kurtosis are observed for the all the age groups and ailment types. For the combined age group the healthcare claims has skewness of 28.959 and 32.033 respectively for claims filed in by provider and claims actually settled by the HMO, with kurtosis of 1487.645 and 1740.774; for ailment classification groups, the minimum skewness computed was 4.795 and a value as high as 27.528 can be seen. The computed kurtoses for all the ailments are also very high. The preliminary exploratory data analysis findings are that the healthcare claims costs are heavily tailed and highly peaked suggesting the suitability of generalized linear modeling.

4.1 Healthcare Cost Modelling

Tables 7 and 8 present the estimates members characteristics based on the normal and gamma regression analysis respectively. These show that the different age categories are significant in determining healthcare costs. Regarding classification of ailments, musculoskeletal, respiratory and diseases related to child bearing are common factors to the other models.

The results presented suggest that the fitted models are significant going by the goodness of fit tests. The regression models fitted considered the two possible scenarios for the claims costs. The first assumes that the normal regression model is adequate in modelling the costs while the second considers the case when the claims data is heavily tailed and highly peaked where the gamma distribution can be used. For these scenarios, three different models are fitted depending on the predictor variables captured. Model 1 covers age characteristics as the only predictors; model 2 considers ailment types as the predictors of claims costs, while model 3 uses all the age and ailment characteristics in building the models.

Table 7
Normal Regression analysis of healthcare costs

	Model 1		Model 2		Model 3	
Parameter	B	p-value	B	p-value	B	p-value
(Intercept)	24743.142	0.000	10059.339	0.000	22711.715	0.000
Infancy (0 - 1 year)	-16878.948	.000			-16718.929	.000
Early childhood (2 - 4 years)	-17511.206	.000			-17292.293	.000
Childhood (5 - 12 years)	-16897.856	.000			-16728.131	.000
Adolescence (13 - 16 years)	-14833.242	.000			-14465.509	.000
Late teenage (17 - 20 years)	-14660.721	.000			-14302.125	.000
Early adulthood (21 - 40 years)	-11228.862	.000			-10951.304	.000
Middle age (41 - 70 years)	-10538.012	.000			-10280.380	.000
General and Unspecified			410.661	.642	310.305	.725
Blood, Blood Formation and immune mechanism			376.391	.857	-24.141	.991
Digestive			1611.082	.106	1486.911	.135
Eye			2322.908	.051	2117.783	.075
Ear			-2823.268	.189	-2578.808	.229
Cardiovascular			792.407	.479	739.513	.508
Musculoskeletal			4536.163	.000	4553.774	.000
Neurological			-288.006	.879	92.037	.961
Psychological			4383.917	.072	4574.270	.060
Respiratory			4257.619	.000	4363.520	.000
Skin			2032.840	.080	1888.507	.103
Endocrine/Metabolic and Nutritional			-349.642	.838	-435.719	.798
Urological			5425.577	.001	5772.137	.000
Pregnancy, Child Bearing, Family Planning			5672.179	.000	5528.063	.000
		Goodness of fit test				
Log Likelihood	-7.1387E+05		-7.1393E+05		-7.1379E+05	
Akaike's Information Criterion (AIC)	1.4277E+06		1.4279E+06		1.4276E+06	
Finite Sample Corrected AIC (AICC)	1.4277E+06		1.4279E+06		1.4276E+06	
Bayesian Information Criterion (BIC)	1.4278E+06		1.4280E+06		1.4278E+06	
Consistent AIC (CAIC)	1.4278E+06		1.4281E+06		1.4278E+06	
Df	59496		59489		59482	

Table 8
Gamma Regression analysis of healthcare costs

	Model 1		Model 2		Model 3	
Parameter	B	Sig.	B	Sig.	B	Sig.
(Intercept)	22672.633	0.000	8082.877	0.000	20466.337	0.000
Infancy (0 - 1 year)	-15948.485	.000			-15830.145	.000
Early childhood (2 - 4 years)	-16871.499	.000			-16668.739	.000
Childhood (5 - 12 years)	-16233.688	.000			-16080.843	.000
Adolescence (13 - 16 years)	-14275.370	.000			-13918.450	.000
Late teenage (17 - 20 years)	-14814.726	.000			-14461.323	.000
Early adulthood (21 - 40 years)	-11153.572	.000			-10887.803	.000
Middle age (41 - 70 years)	-10316.037	.000			-10071.932	.000
General and Unspecified			595.402	.483	505.504	.551
Blood, Blood Formation and immune mechanism			233.577	.907	-117.991	.953
Digestive			1695.375	.077	1583.925	.098
Eye			2279.855	.047	2101.575	.066
Ear			-2907.645	.160	-2688.304	.193
Cardiovascular			975.447	.365	926.992	.388
Musculoskeletal			4960.611	.000	4975.414	.000
Neurological			236.155	.897	585.020	.748
Psychological			4049.330	.084	4228.832	.070
Respiratory			4572.016	.000	4663.820	.000
Skin			2280.232	.041	2152.688	.053
Endocrine/Metabolic and Nutritional			227.573	.890	158.093	.923
Urological			6444.004	.000	6747.046	.000
Pregnancy, Child Bearing, Family Planning			5673.228	.000	5542.547	.000
	Goodness of fit test					
Log Likelihood	-5.7571E+05		-5.7642E+05		-5.7501E+05	
Akaike's Information Criterion (AIC)	1.1514E+06		1.1529E+06		1.1501E+06	
Finite Sample Corrected AIC (AICC)	1.1514E+06		1.1529E+06		1.1501E+06	
Bayesian Information Criterion (BIC)	1.1515E+06		1.1530E+06		1.1503E+06	
Consistent AIC (CAIC)	1.1515E+06		1.1530E+06		1.1503E+06	
Df	56784		56777		56770	

The regression models fitted are then used to determine coefficients for each characteristic which are ultimately used to compute the risk score for any subscriber. These scores are the sum of coefficient values for present conditions. Once the risk scores are computed for every enrollees, a test for equality of means was carried out using the F-test of analysis of variance (ANOVA). This is needed to investigate if the scores significantly vary across the age and ailments characteristics. In addition, Duncan's multiple range post hoc tests

is then required to explain how enrollees can be differentiated into groups. Tables 9 to 12 present the results of the ANOVA and the subsequent post hoc tests.

Table 9
Analysis of variance tests for risk score using the three models

		Sum of Squares	Df	Mean Square	F	Sig.
Risk score model 1 (normal)	Between Groups	4.49E+11	7	6.42E+10	1.90E+28	0
	Within Groups	2.01E-13	59510	3.38E-18		
	Total	4.49E+11	59517			
Risk score model 2 (normal)	Between Groups	9.59E+06	7	1.37E+06	0.818	0.572
	Within Groups	9.97E+10	59510	1.68E+06		
	Total	9.97E+10	59517			
Risk score model 3 (normal)	Between Groups	4.49E+11	7	6.42E+10	15528.904	0
	Within Groups	2.46E+11	59510	4.13E+06		
	Total	6.95E+11	59517			
Risk score model 1 (gamma)	Between Groups	5105.626	7	729.375	3.48E+26	0
	Within Groups	1.25E-19	59510	2.09E-24		
	Total	5105.626	59517			
Risk score model 2 (gamma)	Between Groups	3.62	7	0.517	5.001	0
	Within Groups	6152.824	59510	0.103		
	Total	6156.444	59517			
Risk score model 3 (gamma)	Between Groups	5070.529	7	724.361	19143.199	0
	Within Groups	2251.804	59510	0.038		
	Total	7322.334	59517			

Table 10
Duncan's Post hoc Test for Risk Score of age categories Using Gamma Model

Age group	N	Subset for alpha = 0.05							
		1	2	3	4	5	6	7	8
Early childhood (2 - 4 years)	8570	8.72581							
Childhood (5 - 12 years)	5960		8.83998						
Infancy (0 - 1 year)	2546			8.84215					
Late teenage (17 - 20 years)	459				9.02759				
Adolescence (13 - 16 years)	818					9.11979			
Early Adulthood (21 - 40 years)	32457						9.39242		
Middle age (41 - 70 years)	8267							9.46433	
Senescence (70+ years)	441								10.05417

Table 11
Duncan's Post hoc Test for Risk Score of ailment Classification Using Gamma Model

Classification group	N	Subset for alpha = 0.05					
		1	2	3	4	5	6
Neurological	579	9.1828					
Urological	921	9.1871					
Ear	427	9.1970	9.1970				
Psychological	320	9.2033	9.2033	9.2033			
Respiratory	7463	9.2100	9.2100	9.2100	9.2100		
Musculoskeletal	2028		9.2170	9.2170	9.2170	9.2170	
Female Genital	2346		9.2204	9.2204	9.2204	9.2204	
Cardiovascular	3080		9.2223	9.2223	9.2223	9.2223	
Endocrine/Metabolic and Nutritional	755			9.2294	9.2294	9.2294	
General and Unspecified	24795			9.2301	9.2301	9.2301	
Digestive	5917			9.2311	9.2311	9.2311	
Pregnancy, Child Bearing, Family Planning	5439			9.2315	9.2315	9.2315	
Skin	2630				9.2350	9.2350	
Eye	2361					9.2403	9.2403
Blood, Blood Formation and immune mechanism	457						9.2627

4.2 Discussion of Findings

The following are the research findings arising from the study:

i. Analysis of the claims data by health risk factors as displayed in Table 2, revealed that "general and unspecified" ailments constitute 41.7 percent of heath care cost claims; this is followed by "respiratory" with 12.5 percent; "digestive" related ailments, "pregnancy, child bearing and family planning" share the same percentage of around 9 percent each.
ii. A test of independence of age group and diagnostic types was conducted using the chi-square goodness of fit as shown in Tables 3 and 4. Results suggest that health costs of ailment type vary significantly by age at 95% confidence level with the likelihood ratio of 441.598 and p-value of 0.000.
iii. The descriptive statistics on healthcare claims displayed in Table 5 suggest that the claims are generally increasing with age on the average. That is, the cost of healthcare increases with age of the individual and enrollees that are over 70 years incur higher healthcare costs than twice the general average cost of health.
iv. In Tables 7 and 8, the regression results of "Model 1" show that age groups are significant determinant of healthcare costs; moreover, "Model 2" which modeled ailment type against claims cost reveals that "eye", "musculoskeletal", "respiratory", "urological" and "pregnancy, child bearing and family" related ailments are significant contributors of healthcare costs.
v. The analyses of the risk scores displayed in Tables 9-12 show a consistent pattern with the descriptive statistics. That is, the risk scores of the "over 70 years" group are the highest followed by that of age group 41-70 while other age groups have similar risk scores.

The implication of these findings is that the healthcare costs increase significantly from age group 41 – 70 and this justify calls for risk adjustment of the capitation by age.

The tests of equality of risk scores for all the predictors using the generalized gamma regression revealed that risk scores from all three models are significantly different. However, with the normal regression, for model 1 and model 3, there are significant differences in the average risk scores across categories, whereas the risk scores computed using model 2 are not significantly different across various ailments.

The significance difference established by the ANOVA led to a further investigation for the source(s) of these differences. The Duncan post hoc was considered. These test groups characteristics that are similar in risk score together and differentiate them from other groups. An illustration of the implementation of the risk adjustment model is described thus:
Consider the risk score presented in Table 10 for the eight age groups. The average of these risk scores is computed as:

$$Average\ risk\ score = \frac{8.72581 + 8.83998 + \cdots + 10.05417}{8} = 9.18328$$

The adjustment coefficient would be computed for all age group by dividing the age group's risk score by the average risk score as follow

Table 12
Adjusted coefficient for age group based on Gamma and Normal models

Age group	1	2	3	4	5	6	7	8
Coefficient using Gamma	0.95	0.96	0.96	0.98	0.99	1.02	1.03	1.09
Coefficient using Normal	0.57	0.63	0.66	0.77	0.82	1.13	1.21	2.22

The adjustment coefficient results displayed in Table 12 show that the adjustment using the gamma regression leaves a maximum gap of 14.7 percent between age group 1 and age group 8. Whereas, the normal regression model produce a gap as wide as 289.5 percent. The implication is that the normal based risk adjustment could lead to further risk selection if applied to adjust highly skewed and heavy tailed data.

To apply the adjusted coefficient, for instance, if the current rate is 100 across all ages and experience data produced the coefficient displayed in Table 12. The gamma model suggests that the future rate for early childhood (2 – 4years) should be 100×0.95 = 95; while that of middle age (41 – 70 years) should be 100×1.03 = 103.

5. CONCLUSION AND POLICY IMPLICATIONS

This study developed a risk-adjustment model for the NHIS capitation regime. The model is expected to account for the varying levels of risk using demographic as well as health risk factors. The main purpose for this model is to provide the scheme with incentives to produce services efficiently by minimizing risk selection so that health facilities in a competitive market can compete on the basis of quality of healthcare service and medical administrative efficiency rather than on the ability to select risk. Reducing the incentive for positive selection is one of the main objectives of implementing health-based risk adjustment. This will be achieved by rewarding facilities equitably and fairly for the risks they assume and protect the financial sustainability of the scheme. Also, the model would facilitate the consolidation of the present and historical data of the healthcare management system in order to provide pathways for clinical and administrative information, actuarial valuation and in-depth statistical analysis. Problems caused by risk selection can cause a number of problems for health authorities, insurers as well as HMOs; more so for health schemes like the Nigeria Health Insurance Scheme, which establishes payment rates using a formula rather than accepting prices determined by the marketplace. Chief amongst these problems is that total healthcare expenditures can increase for facilities offering multiple-choice options when risk selection occurs. Using historical claims data, it was established that the claims data from the NHIS is highly peaked and heavily tailed and vary significantly between age groups. This demonstrated that the usual normal regression based model for risk adjustment might not be adequate for the data coverage and risk adjustment. The use of generalized gamma regression model to fit healthcare claims data and risk adjustment to determine fair capitation rates is suggested.

The risk adjustment process uses the results of risk assessment to determine fair capitation rates. It can be designed to help accomplish severa1 goals: Help reduce the effects of risk selection so that health facilities in a competitive market can compete on the basis of quality of healthcare service and medical administrative efficiency rather than on the ability to select risk; Reward facilities equitably and fairly for the risks they assume; sustain enrollee

choice between multiple health plans based on service rates or employee contributions that reflect relative administrative and medical efficiencies; Protect the financial sustainability of the scheme.

6. REFERENCES

Adeleke, I.A., Dallah, H., & Oghojafor, B.E.A. (2016). Data integration management system for national health Insurance scheme: Statistical and actuarial perspectives. *Nigerian Journal of Risk and Insurance (in print)*.

Adeleke, I.A., Dallah, H., & Ibiwoye, A. (2012). Evaluation of the capitation regime of Nigeria health insurance scheme. *Journal of Academic Research, 4*(5), 23 – 28.

Alberta Health. (1998). Population Based Funding Model 1998-99. Edmonton: Alberta Health.

Babazono, A., Weiner, J. Tsuda, T., Mino, Y., & Hillman, A.L. (1998). The effect of a redistribution system for health care for the elderly on the financial performance of health insurance societies in Japan. *International Journal of Technology Assessment in Health Care, 4*(3), 458- 66.

Beck, K. (1998). Competition under a regime of imperfect risk adjustment: The Swiss experience. *Sozialund Prdventivmedizin 43*, 7-8.

Bojanic, D. (1992). A look at a modernized family life cycle and overseas travel. *Journal of Travel and Tourism Marketing, 1*, 61-79.

Brown, L.D., & Amelung, V.E. (1999). Manacled Competition: Market reforms in German health care. *Health Affairs 18*(3):76-91.

Carr-Hill, R.A., Rice, N. & Smith, P (2000). Towards locally based resource allocation in the NHS in reforming markets in health care: an economic perspective, ed. P. Smith. Buckingham: *Open University Press. Carr-Hill.*

Chinitz, D.P. (1994). Reforming the Israeli Health Care Market. *Social Science and Medicine, 39* (10), 1447-57.

Clark, D.O., Korff, M. V, Saunders, K., Baluch, W.M., & Simon, G.E. (1995). A chronic disease score with empirically derived weights. *Medical Care, 33*(8), 783-95.

Ellis, R.P. & Pope, G.C. (1996). Diagnosis-based risk adjustment for medicare capitation payments. *Health Care Financing Review 17*(3), 101-28.

Feldstein, P.J. (1993). *Health Care Economics, 4th ed.*, Albany, NY: Delmar Publishers.

Frees, E.W. (2010). *Regression modelling with actuarial and financial applications*. Cambridge University Press, UK.

Gegory, A., & Sabelhaus, J. (1995). Trends in Out-of-Pocket Spending on Health Care, 1980 1992. *Monthly Labor Review*, 35-45.

Jong, P., & Heller, G.Z. (2008). *Generalized Linear Models for Insurance Data*. Cambridge University Press, UK.

Ibiwoye, A., & Adeleke, I.A. (2008). Does national health insurance promote access to quality health care? Evidence from Nigeria. Geneva Papers on Risk and Insurance, *33*, 219-233.

Kahn, H., Parke, R., & Yi, R. (2014). *Risk Adjustment for Pediatric populations*. Milliman Healthcare Reform Briefing Paper.

Kothari, C.R. (2004). *Research Methodology: Methods and Techniques, 2nd ed*. New age International Publishers, New Delhi.

Mehmud, S.M., & Yi, R. (2012). *Uncertainty in Risk Adjustment*, Society of Actuaries' Health Section.

NHIS (2012). National health insurance scheme operational guideline. NHIS Garki, Abuja.

Porell, A.A, Gruenberg, L., Sawitz, E., & Beiser, A. (1989). Adjusting medicare payments using prior utilization data. *Health Care Financing Review 10*(4), 17-29.

Toso, M., & Farmer, A. (1994). Using cost accounting data to develop capitation rates. *Journal of Health Care Finance, 21*(1), 1 – 11.

Varde, A.P.A., & Diderichsen, F. (2000). Modelling of resource allocation to health care authorities in Stockholm county. *Health Care Management Science* (2),141-9.

Varde, D.F.E. & Whitehead, M. (1997). Resource Allocation to Health Authorities: The Quest for an Equitable Formula in Britain and Sweden. *British Medical Journal* 315, 875-8.