



Using Log-Linear Models and Odd Ratios to Determine Outcome of Admissions in a Hospital

Patrick Kwabena Amoakoh

Department of Mathematics & ICT, OLA College of Education, Cape Coast, Ghana

ARTICLE INFO	ABSTRACT
Published Online: 07 September 2018	<p>The purpose of this study is to determine the outcome of patients' admission in the medical ward (male, female and paediatrics) in the Central Regional Hospital</p> <p>The specific objectives of the study are as follows: to determine association between Outcome of admission (alive and death), Year of admission conditioning on the type of Medical Ward, to determine association between Medical Ward and Outcome of admission (alive and death) conditioning Year of admission, to determine association between Year of admission, Medical Ward and Outcome of admission (alive and death) and determine associations among variables by means of log-linear model Statistics and likelihood ratios</p> <p>The statistical tools used in analyzing the data were SPSS Micro soft excel and SAS. The Outcome of admission variable was categories into two levels; the medical ward had three levels and year variable had five levels and is ordinal. The model approach based on log-linear models was used. In this case, the homogeneous association was tested by comparing the saturated model (SM) and a model assuming homogeneity (HAM). The conditional independence was checked by comparing the HAM model with different models assuming conditional independence. The best model was chosen to conduct further analysis on the effect of medical ward on the outcome of admission while controlling by year.</p> <p>It realized that total number of patients admitted in the hospital was twelve thousand four hundred and twenty six (12426) for the 5 year period. Ten thousand seven hundred and sixty one (10761) were discharged (alive); representing 86.6%. Also one thousand six hundred and sixty five died; representing 13.4% for the same five year period.</p> <p>We observed that the female medical ward recorded highest patients' death in the hospital from 2006 to 2010.</p>
Corresponding Author: Patrick Kwabena Amoakoh	<p>In this model, the conditional likelihood ratio between any two variable are identical. We therefore concluded that the best model is the saturated log-linear model.</p>
KEYWORDS: Log-linear models, Odd ratios, Admissions, Medical Wards.	

Introduction

Admission to Hospital Medical Wards

"The decision to admit patients to the medical wards is determined by age, co-existing illness (co morbidity), physical and laboratory findings, the ability of the patient to comply reliably with an oral medication, and the resources available to the patient outside the hospital" (Ali, Woldie and Mirkuzie, 2010). In studies done in developed countries, medical admissions accounted for 22.2%, 33.0%, and 13.0% of total hospital admissions in U.S.A, Western Australia and Hong Kong, respectively (Oregon Health & Science University Pharmacy).

In a South African study, admissions to the medical wards constituted 40% of the total hospital admissions. In developed countries non-communicable diseases namely cardiovascular diseases are the main reasons for medical admissions. (Elias Ali, 2010) For instance the Australian study, the most common reason for admissions to the medical wards (29% of patients) was cardiovascular disease. In another study, admissions to medical wards at a hospital in Hong Kong were most frequently associated with the cardiovascular system, which made up 30.3% of all medical cases. However, in cities and towns of developing countries, the increasing urbanization and westernization of the population is changing the morbidity pattern of diseases.

It is becoming widely accepted that non-communicable or chronic diseases are also now the major causes of death and disability in low and middle-income countries, (Elias Ali, 2010).

In particular, smoking is increased in underdeveloped countries. The annual cigarette consumption per adult (in cigarettes) has increased from 860 in the early 1970s 1410 in 1995. The reason was aggressive marketing of tobacco companies, delay in implementing antismoking regulations, and because of the public perception of the risk of smoking is still low. This has been supported by studies from South Africa, (Elias Ali, 2010). Circulatory disorders (22%) and infectious diseases (19%) were shown to be the main causes for admission to medical wards of GF Jooste Hospital. In another study from the same country medical admissions were mostly associated with the circulatory system (27.9%) followed by respiratory (15.9%) and infectious diseases (11.9%). Unlike the findings in developed and some Sub-Saharan African countries, the leading reasons for admission to the medical wards in Ethiopia were found to be communicable diseases. Particularly, communicable diseases such as acute febrile illness of infectious origin, pneumonia and tuberculosis were the reasons for admission to the medical wards.

. All patients who were admitted to the medical wards of JUSH during the study period (January.1, 2008 to December. 31, 2008) and whose case notes were available in the hospital registration room archive was included. A case note was classified according to year of admission. The recent case notes of patients admitted to the medical wards in the year 2008 was retrieved. Hence, all cases found during the study period was included in the study and no sampling technique was used. The dependent variables in the study were reasons for admission and outcome of admission, while the explanatory variables included socio-demographic characteristics of the patients in the case notes reviewed, co-morbidities, complications, duration of hospital stay and month of admission.

Outcome of Admission in Medical wards

Leng et al. (1999) stated that in recent years have seen rises in readmission both in the UK and the US. In Oxfordshire, readmission rates almost doubled between 1968 and 1985, with 75% increase in emergency readmissions. A similar rise was also seen in Scotland, where the overall readmission rate rose from 7.1% in 1982 to 11.4% in 1994. Unfortunately, the definition of readmission has varied considerably between many studies, including the time since discharge before readmission, and the type of readmission (elective or emergency). This focuses on unplanned, or emergency, readmissions. The rising trend in emergency readmissions is worrying partly because of implications about quality of care but also because of the burden placed on provision of hospital services. Reasons for the increase are less clear and are likely to be complex. Possible

explanations include changes in the social and demographic structure of the population, falling lengths of stay, and the medical condition itself. The increasing number of elderly people in the population may be particularly important in generating the rise because admission rates increase dramatically with age, especially in those living alone. It has also been suggested that readmissions are related to recurring medical problems, indicated by a higher than expected number of admissions in the period before the readmission.

Huffman (1990) stated that the medical record “must contain sufficient data to identify the patient, support the diagnosis or reason for attendance at the health care facility, justify the treatment and accurately document the results of that treatment”

The main purpose of the medical record is: To record the facts about a patient's health with emphasis on events affecting the patient during the current admission or attendance at the health care facility, and for the continuing care of the patient when they require health care in the future.

A patient's medical record should provide accurate information on: who the patient is and who provided health care; what, when, why and how services were provided; and the outcome of care and treatment.

The medical record has four major sections: Administrative, which includes demographic and socio-economic data such as the name of the patient (identification), sex, date of birth, place of birth, patient's permanent address, and medical record number; legal data including a signed consent for treatment by appointed doctors and authorization for the release of information; financial data relating to the payment of fees for medical services and hospital accommodation; and clinical data on the patient whether admitted to the hospital or treated as an outpatient or an emergency patient.

Statement of the Problem

If critically ill patients are admitted and do not receive proper care, death may occur at any time. The medical teams must fully utilize their knowledge and skills to reduce mortality in the hospital. As the role of the hospital unit in admission outcomes at Central Regional Hospital has yet to be researched, this study explores the situation in the hospital.

Objective of the Study

The broad goal of health services delivery is to improve the health of all people living in the municipality regardless of age, ethnicity, religious conviction, political affiliation or socio-economic standing. This broad goal encompasses many specific objectives, among them, an increase in life expectancy, reduction in morbidity and fertility rates, and improvement in quality of life. The specific objectives of the study are as follows:

- To determine association between Outcome of admission and Year of admission conditioning on the type of Medical Ward.
- To determine association between Medical Ward and Outcome of admission conditioning on Year of admission
- To determine association between Year of admission, Medical Ward and Outcome of admission.
- Determine associations among variables by means of log-linear models.

Methodology

The research method used for this work is quantitative research method. This will explain the phenomena ‘outcome of admission a hospital’ by collecting numerical data that are analysed using mathematically based methods. According to Aliaga and Gunderson (2000), describes what we mean by quantitative research methods very well: Quantitative research is ‘Explaining phenomena by collecting numerical data that are analysed using mathematically based methods (in particular statistics)’.

In order to be able to use mathematically based methods, our data have to be in numerical form. This is not the case for qualitative research. Qualitative data are not necessarily or usually numerical, and therefore cannot be analysed by using statistics. Therefore, as quantitative research is essentially about collecting numerical data to explain a particular phenomenon, particular questions seem immediately suited to being answered using quantitative methods.

Population and Sampling Technique

The population of the study consisted of the patients admitted in the medical wards (Male, Female and Pediatric wards) from 2005 to 2010 in Central Regional Hospital (Interburtin), Cape Coast – Ghana. The total number of patients admitted in the hospital was Twelve Thousand Four Hundred and Twenty Six (12,426) for the five (5) year period. The patients and medical personnel were interviewed. A purposeful sampling technique was used.

Instrumentation

The instruments used for the study were observation, interview and secondary data from biostatistics department in the hospital. Interviews were granted to the medical personnel and some patients. Table records of all patients’ admissions in Medical wards from 2006 to 2010 were collected and retrospectively analyzed using a model based on log-linear model. The statistical software’s employed in this study are SAS and Microsoft Excel 2007. Information regarding sex, admission, and the discharge or death outcome was recorded.

Basic Strategy and Key Concepts

The basic strategy in log-linear modeling involves fitting models to the observed frequencies in the cross-tabulation of categorical variables. There can then be models represented by a set of expected frequencies that may or may not resemble the observed frequencies. Models will vary in terms of the marginal’s they fit, and can be described in terms of the constraints they place on the associations or interactions that are present in the data. The pattern of association among variables can be described by a set of odds and by one or more odds ratios derived from them. Once expected frequencies are obtained, we then compare models that are hierarchical to one another and choose a preferred model, which is the most parsimonious model that fits the data. It’s important to note that a model is not chosen if it bears no resemblance to the observed data. The choice of a preferred model is typically based on a formal comparison of goodness-of-fit statistics associated with models that are related hierarchically (models containing higher also implicitly include all lower order terms). Ultimately, the preferred model should distinguish between the pattern of the variables in the data and sampling variability, thus providing a defensible interpretation.

Theory and Calculations

The Log-linear Model

The following model refers to the traditional chi-square test where two variables, each with two levels (2 x 2 tables), are evaluated to see if an association exists between the variables:

$$\text{Log}(m_{ij}) = \mu + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY} \dots\dots\dots (1)$$

Log(m_{ij}) = is the log of the expected cell frequency of the cases for cell ij in the Contingency table.

μ = is the overall mean of the natural log of the expected frequencies

λ = terms each represent “effects” which the variables have on the cell frequencies

X and Y = are treated as outcome variables

i and j refer to the categories or level within the variables X and Y

Therefore:

λ_i^X = the main effect for variable X

λ_j^Y = the main effect for variable Y

λ_{ij}^{XY} = the interaction effect for variables X and Y

The above model in equation (1.1) is considered a Saturated Model because it includes all possible one way and two-way effects. Given that the saturated model has the same amount of cells in the contingency table as it does effects, the expected cell frequencies will always exactly match the observed frequencies, with no degrees of freedom remaining.

.For example, in a 2 x 2 table there are four cells and in a saturated model involving two variables there are four effects, $\mu, \lambda_i^X, \lambda_j^Y, \lambda_{ij}^{XY}$

XY , therefore the expected cell frequencies will exactly match the observed frequencies. Thus, in order to find a

more parsimonious model that will isolate the effects best demonstrating the data.

Patterns, a non-saturated model must be sought. This can be achieved by setting some of the effect parameters to zero. For instance, if we set the effects parameter λ_{ij}^{XY} to zero (i.e. we assume that variable X has no effect on variable Y or vice versa) we are left with the unsaturated model:

$$\text{Log}(m_{ij}) = \lambda_i^X + \lambda_j^Y + \dots (2)$$

This particular unsaturated model is titled the Independence Model because it lacks an interaction effect parameter between X and Y . Implicitly, this model holds that the variables are unassociated. Note that the independence model is analogous to the chi square analysis, testing the hypothesis of independence.

Table 1: Cell Counts in a 2×2 Contingency Table

Level of X	Level of Y		Total
	1	2	
1	n_{11}	n_{12}	n_{1+}
3	n_{21}	n_{22}	n_{2+}
Total	n_{+1}	n_{+2}	N

Table 2: Cell Probabilities in a 2×2 Contingency Table

Level of X	Level of Y		Total
	1	2	
1	π_{11}	π_{12}	π_{1+}
3	π_{21}	π_{22}	π_{2+}
Total	π_{+1}	π_{+2}	1

The motivation for the use of log-linear models is that statistical independence can be express in terms of a linear combination of the logarithms of the cell probabilities. In particular if the variables X and Y in a 2×2 table are statistically independent, then the probability of individuals being in the first row (level 1 of X) among those in the first column (level 1 of Y) would be the same as the probability for the first row among those in the second column (level 2 of Y).

Therefore, $\frac{\pi_{11}}{\pi_{+1}} = \pi_{1+}$ (3)

And $\pi_{11} = \pi_1 * \pi_{+1}$. Similar arguments lead to the general result that if the row and column variables are independent, then $\pi_{ij} = \pi_{i+} * \pi_{+j}$, for $i, j = 1, 2$.

You can then express independence as a general relation involving all four cell probabilities. First, if X and Y are statistically independent.

$$\frac{\pi_{11}}{\pi_{+1}} = \frac{\pi_{12}}{\pi_{+2}} \dots (4)$$

Table 3: Log-linear Model Expected cell counts

Level of X	Level Y	
	1	2
1	$\exp(\mu + \lambda_1^X + \lambda_1^Y + \lambda_{11}^{XY})$	$\exp(\mu + \lambda_1^X - \lambda_1^Y - \lambda_{11}^{XY})$
2	$\exp(\mu - \lambda_1^{XY} + \lambda_1^X - \lambda_{11}^{XY})$	$\exp(\mu + \lambda_1 - \lambda_1) + \lambda_{11}$

Since $\pi_{+1} = \pi_{11} + \pi_{21}$ and $\pi_{+2} = \pi_{12} + \pi_{22}$, the relationship is

$$\frac{\pi_{11}}{\pi_{11} + \pi_{21}} = \frac{\pi_{12}}{\pi_{12} + \pi_{22}} \dots (5)$$

so that $\pi_{11} (\pi_{12} + \pi_{22}) = \pi_{12} (\pi_{11} + \pi_{21})$. This simplifies to $\pi_{11}\pi_{22} = \pi_{12}\pi_{21}$. Therefore, the row and column variables are independent if

$$\psi = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}} = 1 \dots (6)$$

Where ψ is called the cross-product ratio, or the odds ratio. Taking logarithms of both sides expresses statistical independence as a linear combination of the logarithms of the cell probabilities:

$$\text{Log } \psi = \text{log } \pi_{11} - \text{log } \pi_{12} - \text{log } \pi_{21} + \text{log } \pi_{22} = 0 \dots (7)$$

Log-linear models for 2×2 contingency tables involve the logarithm of the cross-product ratio in a special way.

The saturated log-linear model for a 2×2 table is:

$$\text{log}(m_{ij}) = \mu + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}; i, j = 1, 2 \dots (8)$$

Where $m_{ij} = n\pi_{ij}$ is the expected frequency in the (i, j) cell. This model is similar to the two-way analysis of variance model for a continuous response y :

$$E(y_{ij}) = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} \dots (9)$$

With overall mean μ , main effects α_i and β_j and interaction effects $(\alpha\beta)_{ij}$. The use of the terms λ_i^X , λ_j^Y , and λ_{ij}^{XY} instead of α_i , β_j and $(\alpha\beta)_{ij}$ is the common log-linear model notation and is especially convenient when considering tables of higher dimensions.

Since there is $1 + 2 + 2 + 4 = 9$ parameters in the saturated log-linear model, but only four observations, the model is over parameterized. Imposing the usual sum-to-zero constraints yields three non-redundant λ parameters (λ_i^X , λ_j^Y , λ_{ij}^{XY}). The fourth parameter, μ , is fixed by the total sample size n . table (1) displays the expected cell frequencies m_{ij} in terms of the model parameters μ , λ_i^X , λ_j^Y , and λ_{ij}^{XY}

$$\sum_{i=1}^2 \lambda_i^X = 0 \quad \sum_{j=1}^2 \lambda_j^Y = 0 \quad \sum_{i=1}^2 \lambda_{ij}^{XY} = 0 \quad \sum_{j=1}^2 \lambda_{ij}^{XY} = 0$$

The odds ratio can also be expressed as a function of the expected frequencies:

$$\Psi = \frac{m_{11}m_{22}}{m_{12}m_{21}}$$

So that: $\log \psi = \log m_{11} - \log m_{12} - \log m_{21} + \log m_{22} = 4\lambda_{11}^{XY} \dots s \dots (10)$

Therefore, the hypothesis of independence of X and Y is equivalent to:

$H_0: \lambda_{11}^{XY} = 0$. The corresponding independence log-linear model is given by

$$\log(m_{ij}) = \mu + \lambda_i^X + \lambda_j^Y \quad i, j = 1, 2 \dots (11)$$

This model has one degree of freedom for testing lack of fit. The Pearson chi-square test of independence for a 2 x 2 contingency table. An alternative approach is the test $H_0: \lambda_{11}^{XY} = 0$ using the likelihood ratio test to compare the fit of the independence and saturated log-linear models.

The likelihood ratio test of independence can be derived directly from the multinomial likelihood.

$$f(n_{11}, n_{12}, n_{21}, n_{22}) = \frac{n! \pi_{11}^{n_{11}} \pi_{12}^{n_{12}} \pi_{21}^{n_{21}} \pi_{22}^{n_{22}}}{n_{11}! n_{12}! n_{21}! n_{22}!} \dots (12)$$

The unrestricted maximum likelihood estimates (MLEs) of the π_{ij} values are given

By $P_{ij} = \frac{n_{ij}}{n}$, the maximized likelihood is then:

$$\max L = \frac{n!}{n_{11}! n_{12}! n_{21}! n_{22}!} \prod_{i=1}^2 \prod_{j=1}^2 (n_{ij})^{n_{ij}} \dots (13)$$

Under the independence hypothesis $H_0: \pi_{ij} = \pi_i+ \pi_{.j}$, the likelihood is:

$$L_0 = \frac{n!}{n_{11}! n_{12}! n_{21}! n_{22}!} \pi_{1+}^{n_{1+}} \pi_{2+}^{n_{2+}} \pi_{.1}^{n_{.1}} \pi_{.2}^{n_{.2}} \dots (14)$$

The MLEs for the π_{ij} under this model are $P_{ij} = n_i+ n_{.j} / n$ and the maximized log-likelihood is

$$\max L_0 = \frac{n!}{n_{11}! n_{12}! n_{21}! n_{22}!} \prod_{i=1}^2 \prod_{j=1}^2 \left(\frac{n_{i+} n_{.j}}{n} \right)^{n_{ij}} \dots (15)$$

The likelihood ratio is:

$$\text{Max } L_0 = \lambda \frac{\max L_0}{\max L} = \prod_{i=1}^2 \prod_{j=1}^2 \frac{(m_{ij})^{n_{ij}}}{n_{ij}} \dots (16)$$

$$G^2 = -2 \log \lambda = 2 \sum_{i=1}^2 \sum_{j=1}^2 n_{ij} \log \left(\frac{n_{ij}}{m_{ij}} \right) \dots (17)$$

Where $m_{ij} = n_{i+} n_{.j} / n$ and the likelihood ratio statistics G^2 has an asymptotic chi-square distribution with 1 df if H_0 is true, and it is asymptotically equivalent to the Pearson chi-square statistic.

Log-linear Model for the s x r Table

When a sample of n observations is classified with respect to two categorical variables, one having s levels and the other having r levels, the resulting frequencies can be displayed in an s x r contingency table as shown in table 4 below. The corresponding cell probabilities are π_{ij} , with row and column marginal probabilities $\{\pi_{.j}+\}$ and $\{\pi_{i+}\}$, respectively. The generalization of the log-linear models from the table to the s x r table straightforward. The saturated model is

$$\text{Log}(m_{ij}) = \mu + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY} \quad i=1, \dots, s, \quad j=1, \dots, r \dots (18)$$

Where $m_{ij} = n\pi_{ij}$ is the expected frequency in the (i, j) cell. The parameters μ is fixed by the sample size n and the model has s + r + sr parameters λ_j^X, λ_j^Y , and λ_{ij}^{XY} . The sum-to-zero constraints

$$\sum_i \lambda_i^X = 0 \quad \sum_j \lambda_j^Y = 0 \quad \sum_{ij} \lambda_{ij}^{XY} = 0 \quad \sum_{ij} \lambda_{ij}^{XY} = 0 \dots (19)$$

Fitting Log-linear Models

Once a model has been chosen for investigation the expected frequencies need to be tabulated. For two variable models, the following formula can be used to compute the direct estimates for non-saturated models. (Column total) * (row total)/grand total for larger tables, an iterative proportional fitting algorithm (Deming-Stephan algorithm) is used to generate expected frequencies. This procedure uses marginal tables fitted by the model to insure that the expected frequencies sum across the other variables to equals the corresponding observed marginal tables

Testing for goodness of fit

Once the model has been fitted, it is necessary to decide which model provides the best fit. The overall goodness-of-fit of a model is assessed by comparing the expected frequencies to the observed cell frequencies for each model. The Pearson Chi-square statistic or the likelihood ratio (G^2) can be used to test a models fit. However, the (G^2) is more commonly used because it is the statistic that is minimized in maximum likelihood estimation and can be partitioned uniquely for more powerful test of conditional independence in multi way tables:

$$G^2 = 2 \sum n \log(n/m_{ij}) \dots (20)$$

where n and m denote the observed and fitted cell frequencies. The corresponding Pearson chi-square statistics is equal to;

$$Qp = \sum(n - m) \dots (21)$$

G^2 follows a chi-square distribution with the degrees of freedom (df) equal to the number of lambda terms set equal to zero. Therefore, the G^2 statistic tests the residual frequency that is not accounted for by the effects in the model (the λ parameters set equal to zero). The larger the G^2 relative to the available degrees of freedom, the more the expected frequencies depart from the actual cell entries. Therefore, the larger G^2 values indicate that the model does not fit the data well and thus, the model should be rejected. It is often found that more than one model provides an adequate fit to the data as indicated by the non-significance of the likelihood ratio. At this point, the likelihood ratio can be used to compare an overall model within a smaller, nested model (i.e. comparing a saturated model with one interaction or main effect dropped to assess the importance of that term). The equation is as follows:

$$G^2 \text{ comparison} = G^2 \text{ model 1} - G^2 \text{ model 2}$$

Model 1 is the model nested within model 2. The degrees of freedom (df) are calculated by subtracting the df of model 2 from the df of model 1. If the L^2 comparison statistic is not

significant, then the nested model (1) is not significantly worse than the saturated model (2). Therefore, choose the more parsimonious (nested) model.

Log-linear Residuals

In order to further investigate the quality of fit of a model, one could evaluate the Individual cell residuals. Residual frequencies can show why a model fits poorly or can point out the cells that display a lack of fit in a generally good-fitting model. The process involves standardizing the residuals for each cell by dividing the difference between frequencies observed and frequencies expected by the square root of the frequencies expected ($(F_{obs} - F_{exp}) / \sqrt{F_{exp}}$). The cells with the largest residuals show where the model is least appropriate. Therefore, if the model is appropriate for the data, the residual frequencies should consist of both negative and positive values of approximately the same magnitude that are distributed evenly across the cells of the table.

Results and Discussion

Preliminary Analysis

Table 4: Cross Tabulation between Outcome of admission and medical ward

Medical Ward (X)	Outcome of Admission (Y)				Total	
	Alive		Death			
	Freq	%	Freq	%	Freq	
Male	3589	87.77	500	12.23	4089	33
Female	2279	75.59	736	24.71	3015	24
Paedics	4893	91.19	429	8.06	5322	43
Total	10761	88.6	1665	13.4	12426	100

Table 4 deals with Cross tabulation Outcome of Admission and Medical Ward.

It was observed that more patients were discharged / alive in the paediatric ward

Representing 4893 (91.19%) than those discharge in the male and female medical ward representing 3389 (87, 77%) and 2279 (75.58%) respectively.

On the other hand, it was also noted that female medical ward recorded highest number of death representing 736 (24,71%) than paediatrics and female medical wards representing 429 (8.06) and 500 (12.23) respectively.

Table 5: Cross Tabulation between Outcome of admission and year of admission

YEAR (Z)	Outcome of Admission (Y)				Total	
	Alive		Death			
	Freq	%	Freq	%	Freq	
2006	1905	82.29	410	17.71	2315	
2007	1843	85.52	312	14.48	2155	
2008	2215	86.83	336	13.17	2551	
2009	2337	88.96	290	11.03	2627	
2010	2461	88.58	317	11.41	2778	
Total	10761	86.60	1665	13.39	12426	

In table 5, we realized that the year 2009 recorded highest number of patients discharged/ alive representing 2337 (88.96%). This is followed by year 2010 also representing 2461 (88.58). Also year 2006 recorded least number of patients discharge.

On the other hand, year 2006 recorded highest of death representing 401 (17.71%). Meanwhile, year 2009 recorded

least number of patients death representing 290 (11.03%). It was noted that, values obtained for both patients alive and death are evenly distributed within the five year period.

Finally, patients admitted in the medical wards from 2006-2010 were 12426. Of this patients discharge was 10761(86.60%) and 10605 (13.39%) representing death rate during the same period.

Saturated model for X, Y, Z $\log(m_{ijk}) = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ}$..(3.2)

Table 6: Maximum Likelihood Analysis of Variance

Source	DF	Chi-Square	Pr > Chi- Square
Year (Z)	4	11.67	0.0199
Medical (X)	2	3.67	0.1599
year*medical (ZX)	8	56.29	<.0001
Outcome (Y)	1	3855.54	<.0001
year*outcome(ZY)	4	37.98	<.0001
medical*outcome (XY)	2	414.80	<.0001
year*medical*outcome(ZXY)	8	553.25	<.0001
Likelihood Ratio	0	.	.

The Likelihood ratio test in the analysis of variance table compares this model to the saturated model and thus tests the null hypothesis of no three- factor interaction. Since the model is saturated, the likelihood ratio test is equal to zero

(table 6) Thewald test of the three – factor interaction is ($QW = 553.25, 8 \text{ df}, p = 0.0001$) is significant.

$\log(m_{ijk}) = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z$ (22.)

Table 7: Maximum Likelihood Analysis of Variance

Source	DF	Chi-Square	Pr > Square
Year (Z)	4	99.54	<.0001
Medical (X)	2	630.73	<.0001
Outcome (Y)	1	5021.19	<.0001
Likelihood Ratio	22	1814.87	<.0001

We could evaluate the difference in the Chi-square statistics, based on the difference in the degrees of freedom; if the differential Chi-square statistic is significant, then we would conclude that the three-way interaction model provides a significantly better fit to the observed table than the model without this interaction. Therefore, the three-way interaction is statistically significant.

Year (Z), Medical ward (X), Outcome(Y) Variables

$\log(m_{ijk}) = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ}$ (3.2) Saturated model

$\log(m_{ijk}) = \mu + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}$ (4.6) Homogeneous model

$\log(m_{ijk}) = \mu + \lambda_i^X + \lambda_j^Y + \lambda_k^Z$ (4.5) Mutually independent model

The log-linear models were used to check for associations. In this approach, based on fitted models, it was checked whether the homogeneity assumption holds and thus, whether conditional independence models, joint independence models or the mutually independent model gave a good fit to the data. First of all, the goodness of fit

was tested for each model. The table 7 summarizes the results. The likelihood ratio statistic (G^2) was used when testing goodness of fit. The choice of G^2 was suggested mainly by the fact that it can be partitioned.

This property allows the comparison of nested models and the saturated model. The values of G^2 in 14 are differences of likelihood ratio for a given model and the likelihood ratio

for the saturated model. A large P-value indicates that there is no significant difference of the related model and the saturated model. It can be seen that the homogeneity assumption holds since the model ((ZX, ZY, XY) is significantly different of the saturated model (P-value < 0.0001), thus there is a homogeneous association between the three variables. The homogeneous association model can therefore be considered as baseline to check conditional independence among the variables. Table 8 gives results on comparison of nested models to the HAM. The model (ZY, XY) (P-value< 0.0001) is significantly different from the HAM. Thus, year and outcome are conditionally independent given the medical ward and outcome of

“Using Log-Linear Models and Odd Ratios to Determine Outcome of Admissions in a Hospital”

admission. The model (ZX, XY), assuming conditional independence of Medical ward and the outcome of admission given year, showed a p-value=0.0001. This means that interaction ZY cannot be removed from the HAM. It is deduced that year and the outcome of admission are not independent. Thus, the three variables are not mutually independent. The last association checked is the

joint independence of medical ward to year and outcome of admission. By taking the conditional independence model (ZY, ZX) as baseline, it is shown in table 15 that joint independence model (X, ZX) is not significantly different from the model (ZY,ZX) (P-value=0.0001) which means that medical ward is jointly independent to therapy and response-to-chemotherapy.

Table 8: Goodness of fit for log-linear models relating Year (Z), Medical ward (X), and Outcome (Y)

Model	G ²	Df	P-value
(ZX,ZY,XY)	661	8	< 0.0001
(ZY,XY)	1337.86	16	< 0.0001
(ZX,XY)	724.09	12	< 0.0001
(ZY,ZX)	1081.89	10	< 0.0001
(X,ZY)	1755.26	18	<0.0001
(Y, ZX)	1141.50	14	<0.0001
(Z, XY)	1397.47	20	< 0.0001

From table 13, the models (ZY, XY), (ZX, XY) and (ZY, ZX) are conditionally independent.

Also the models (X, ZY), (Y, ZX) and (Z, XY) are jointly independent.

Table 9: Test of nested models

Models Compared	Deviance	DF	P-value
(ZX,ZY,XY) –Saturated	661-0	8	<0.0001
(ZY,XY) - (ZX,ZY,XY)	676.86	8	<0.0001
(ZX,XY) - (ZX,ZY,XY)	63.01	4	<0.0001
(ZY,ZX) - (ZX,ZY,XY)	420.89	2	<0.0001
(X,ZY) - (ZX,ZY,XY)	1094.26	10	<0.0001
(Y, ZY) - (ZX,ZY,XY)	480.5	6	<0.0001
(Z,XY) - (ZX,ZY,XY)	736.47	12	<0.0001

The model with no three – factor interaction provides a good fit to the observed data. Thus, no pair of variables is conditionally independent. In this model, the conditional likelihood ratio between any two variable are identical. We therefore concluded that the model is three- factor interaction model.

Conclusion

Based on the results of the research, we observed that the use of log-linear models is successfully employed to determine the associations between Year of admission (Z), Medical Ward(Y) and Outcome of admission(X). From the results, we can infer that the female medical ward recorded more death than the male and the paediatrics ward. Again the results reveal that more patients died in 2006 and year 2009 recorded least number of patients death. We father conclude that female medical ward recorded the highest number of death and paediatric ward recorded least number of death from 2006 to 2010

We can then conclude that there is association between medical ward and the outcome of admission, medical ward and the year, and year and outcome of admission. The three-way association tested was between category of

medical ward, outcome of admission and year of admission (model 2d), and it was found to be significant. After that, a two-way association between medical ward, outcome of admission and year (models 1a, and 2a) were tested and all of them were found to be significant. This shows that there is an association between the three variables. Therefore, the best model is the saturated model.

Recommendations

Based on the findings from the log-linear model for outcome on admissions in Hospital from 2006 – 2010, the following recommendations have been made:

- The Hospital Authorities could use the log-linear modeling and likelihood ratio methods to analyze data periodically.
- Medical personnel should make maximum use of available facilities to help reduced mortality rate in the hospital. This is because current rate is not very good.
- Doctor to patient ratio must be improved for store further death in the hospital.
- Further studies should be conducted in other hospitals to see whether or not same problem exist.

References

1. Agresti, A (1996). An Introduction to Categorical Data Analysis. John Wiley & Sons, Inc. New York, New York, USA. and Outcome in Acute Myocardial Infarction and Outcome in Acute Myocardial Infarction - IMAJ . Vol 5 . Angela Jeansonne (2002) log-linear models
2. Antwi, K.B. (2008) Stakeholders Perception of Current Health Education Situation
3. Ashton, C. M. and Wray, N. P. (1996). A conceptual framework for the study of early readmission as an indicator of quality of care. *Soc.Sci.Med.* **43**, 1533-1541.
4. Beck, E. J., Ehrmann, F. D., Thivierge, C., Guérard, L., & Déry, V. (2000) *The need and feasibility of a prospective monitoring system for chronic heart failure in Montréal : SurvICe*. Montreal: Direction de la santé publique, Régierégionale de la santé et des services sociaux de Montréal-Centre.
5. Brook, R.H., Kahn, K. L., & Kosecoff, J. (1992). Assessing clinical instability at discharge: The clinician's responsibility. *JAMA* **268**, 1321-1322.
6. Burns, R.B., McCarthy, E.P., Moskowitz, M.A., Ash, A., Kane, R.L., & Finch, M. (1997) Outcomes for older men and women with congestive heart failure. *J Am GeriatrSoc* 1997; 45:276-80.
7. Cleland, J. G. F., Gemmell, I., Khand, A., & Boddy, A. (1999). Is the prognosis of heart failure improving? *Eur. J. Heart Fail*; 1:229-41.
8. Cordle, F., & Tyroler, H. A. (1974). The use of Hospital medical records for epidemiologic research. Differences in Hospital Utilization and in hospital mortality by age- race-sex- place of residence and socioeconomic status in a defined community population. *Medical care*; 12: 596-610.
9. Dean, M. P.G., Gear, J. S. S. Medical admissions to Hillbrow Hospital, Johannesburg, by discharge diagnosis. *S Afr Med J.* 1986;69: 672-3.
10. Debbie, E., Feldman, C. T., Lise Guérard,(2001) Trends in Mortality and Admissions in Hospitals
11. Denise, C. (2009). Admission and Discharge of Patients from Hospital and Residential Settings Policy October No: CLIN/0012/v4
12. Feldman, D.E. (2001) changing trends in mortality and admissions to hospital
13. Leng, G.C., Walsh, D., Fowkes, F.G.R., & Swainson, C.P. (1999). *Quality in Health Care*, 8:234–238
14. G. Amit et al. (2003) the Association between Hospital Department, Medical Treatment .IMAJ Vol 5.
15. Garko, S.B., Ekweani, C.N., & Anyiam, C.A. (2003). Duration of Hospital stay and Mortality in the medical wards of Ahmadu Bello University Teaching Hospital, Kaduna, Nigera. *Annals of African Medicine*, 2(2), 68-71.
16. Glasgow, Weissberg, et al. (1991). Readmissions of children with diabetes mellitus to a children's hospital. *Pediatrics* **88**, 98-104.
17. Goff, D. C., Pandey, D. K. K., Chan, F. A., Ortiz, C., & Nichaman, M. Z. (2000) Congestive heart failure in the United States : Is there more than meets the I(CD Code)? The Corpus Christi Heart Project. *Arch Intern Med*;160:197-202.
18. Graves, E. J. Detailed diagnoses and procedures: National Hospital Discharge Survey, 1989. *Vital Health Stat 13* 1991; Sept(108):1-236.
19. Gregg, C., Fonarow, M. D., & Ahmanson (2008). Day of Admission and Clinical Outcomes for Patients Hospitalized for Heart Failure.
20. Guy Amit, M. D., Sylvie Goldman, M. D., Liora Ore, M. D., Marcelo Low, M. P. H., Jeremy, D., & Kark, M. D. (2003). The Association between Hospital Department, Medical Treatment.
21. Habte –Gabr, E., Girma, B., Mehrete, M., Mehari, A., Tekle, E., & Belachew, K. (1976) Analysis of Admissions to Gondar Hospital in NorthWestern Ethiopia, *Ethiop Med J*; 14:49-59.
22. Hayward, R. A., Bernard, A. M., Rosevear, J. S., Anderson, J. E., & McMahan-L. F. J. (1993). An evaluation of generic screens for poor quality of hospital care on a general medicine service *.Med.Care*31, 394-402.
23. Herlitz, J., Bengtson, A., Hjalmarson, A., & Karlson, B. W. (1988). Morbidity during five years after myocardial infarction and its relation to infarct size. *Clin.Cardiol.* 11, 672-677.
24. http://asio.jde.aca.mmu.ac.uk/new_gis/analysis/loglin.htm- A brief tutorial on log-linear models. (retrieved April 15, 2018).
25. <http://wizard.ucr.edu/~rhannema/soc203a/loglin.html>- A comprehensive article on hierarchical log-linear models. Retrieved April 15, 2018
26. <http://www.math.yorku.ca/SCS/Courses/grcat/grc8.html>- A brief discussion on log linear models and how to use the statistical software, SAS, for log linear modeling. (retrieved April15, 2018).
27. <http://www.statsoftinc.com/textbook/stloglin.html>- A tutorial on log-linear analysis of frequency tables. Retrieved April 15, 2018
28. <http://www2.chass.ncsu.edu/garson/pa765/logit.html>- A brief tutorial on log-linear, logit and probit models. (retrieved April 15, 2018).
29. Huffman (1990) Medical Records Manual. A Guide for Developing Countries
30. Hussein, K. (1997). Pattern of Medical Admissions to Jimma Hospital. *Bull Inst Health Sci.*;7(1): 51-62.

“Using Log-Linear Models and Odd Ratios to Determine Outcome of Admissions in a Hospital”

31. Knoke, D., & Burke, P. J. (1980). *Log-Linear Models*. Sage Publications, Inc. Newberry Park, California, USA.
32. Kolo, P. M., & Chijioke Adindu. (2010). Gender Disparities among Medical Admissions of a tertiary health facility Iloroin. Nigeria.
33. Medical Records Manual A Guide for Developing Countries (World Health Organization 2002)
34. Rich MW. (1999) Heart failure disease management: a critical review. *J Card Fail*; 5:64-75.
35. Taylor, I. C., & McConnell, J. G. (1995) Patterns of admission and discharge in an acute geriatric user www.sfsu.edu/~efc/.../Log%20Linear%20Models.htm