# Exploring fraud and abuse in National Health Insurance Scheme (NHIS) using data mining technique as a statistical model

J. D. [1]Kittoe & S. K. [2]Asiedu-Addo

## Abstract

This study explored patterns of fraud and abuse that exist in the National Health Insurance Scheme (NHIS) claims in the Awutu-Effutu-Senya District using data mining techniques, with a specific focus on malaria-related claims. The study employed quantitative research approach with survey design as a strategy of enquiry. This survey explores the utility of various data mining techniques such as data collection, data cleaning/extraction, data integration, data selection, data transformation and pattern evaluation in health domain. Samples of 720 clients diagnosed of malaria in the years 2013, 2014 and 2015 from 4 NHIS service providers in the districts were randomly selected for this study. Results from two-way between-subjects Analysis of Variance (ANOVA) revealed that Hospital B Private and Hospital A Private recorded the highest and lowest mean cost of malaria treatment respectively. The study further revealed that repetition of NHIS registration number, over billing of drugs, drug mismatch, excessive prescription of drugs for malaria treatment and duplication of clients records were some of the fraud and abuse at the facility.

Keywords: national health insurance scheme, fraud and abuse, data mining, malaria

## Introduction

The National Health Insurance Scheme (NHIS) is a form of Health Insurance system established by the Government of Ghana in 2004, with the goal to provide equitable access and financial coverage for basic health care services to persons resident in Ghana. The idea for the NHIS in Ghana was conceived nearly two decades ago by government to replace the 'cash and carry' system in the health care delivery system. Under the cash and carry system, the health need of the individual was only attended to after initial payment for the service was made. Even in cases where patients had been brought into emergency, it was required for the casualty to pay money at every point of service delivery. This system was replaced in 2003 by the new equitable insurance scheme with the law, ACT 650 LI 1809 of the Republic of Ghana (National Health Insurance Authority (NHIA), 2010).

The National Health Insurance Act, Act 650 was passed into law in Ghana in 2003, though implementation in terms of access to benefits began in November 2005 (Witter & Garshong, 2009). Since its implementation, unsubstantiated reports indicate increasing health care and

[1]Josiah Doam Kittoe is the Scheme Manager, National Health Insurance Scheme, Agona Swedru, Ghana. Email: paadoam@gmail.com.

[2]Dr. Samuel Asiedu-Addo also an Associate Professor in the Department of Mathematics Education, University of Education, Winneba Winneba, Ghana. Email. skaseidu-addo@uew.edu.gh

*Kittoe, J. D. & Asiedu-Addo, S. K.*

administrative costs of the various District Mutual Health Insurance Scheme (DMHIS) across the country without any corresponding increase in the premium level. Its policy objective is that, within the next few years, every resident of Ghana shall belong to a health insurance scheme that adequately covers him or her against the need to pay out-front at point of service use in order to obtain access to a defined package of acceptable quality health services (Ministry of Health (MOH), 2004). A number of measures have been put in place to ensure the achievement of this objective however, there are still some challenges threatening the existence and sustainability of the NHIS.

One of the major challenges of the insurance industry is fraud and abuse which causes substantial losses. Gill and Randall (1994), described fraud in the insurance industry as knowingly making a fictitious claim, inflating a claim or adding extra items to a claim, or being in any way dishonest with the intention of gaining more than legitimate entitlement. According to the NHIA (2013), types of fraud by providers include billing for services not rendered, up-coding of services, double billing/duplicate claims, misrepresentation of diagnosis, un-bundling of services, unnecessary services, inappropriate referrals for financial gain and insertion/substitution of medicines. This is supported by the United States Department of Health and Human Services (2014).

Abuse on the other hand describes provider's practices that, either directly or indirectly, result in unnecessary costs to the payer. Abuse includes any practice that is not consistent with the goals of providing patients with services that are medically necessary, meet professionally recognized standards, and are fairly priced (Centers for Disease Control and Prevention, 2012). It is estimated that about 10 percent of health care system expenditure is wasted due to fraud and abuse (Gee, Button, Brooks, & Vincke, 2010). A report on the financial cost of healthcare fraud provides staggering data on the extent of the menace, averaging from 7% to as high as 15% of healthcare expenditure (NHIA, 2013). Therefore, the scale of healthcare fraud and abuse is large enough to make it a priority issue for health systems.

It is difficult to estimate the exact amount of losses caused by fraud, and also to ascertain its true cost in a country as many fraudulent activities go undetected. These acts of fraud and abuses materialize from many sectors under the NHIS, as there are different categories of services provided by the NHIA via the Schemes at Service Provider sites. Some of these fraudulent acts were reported in some of the Health Insurance Scheme in the ten regions. For example, in the Ashanti Region, 19 percent of its total membership accounted for the payment of 30% of fictitious claims whilst in Greater Accra Region, 13% of fictitious claims were detected. The Chief Executive Officer (CEO) of the National Health Insurance Authority in 2011 made these disclosures at a Stakeholders' meeting in Accra (Mensah, 2011).

In order to mitigate fraud, the NHIA created a Claims and Clinical Audit Division (a center to help prevent fraud) in 2009 and commenced full-scale clinical audits in January 2010. The clinical audit process works with multi-disciplinary teams from both the public and the private sector. Provider stakeholder groups such as Ghana Health Service (GHS) and Christian Health Association of Ghana (CHAG) helped develop the methodology for the clinical audits and facilitated implementation of the audits. This made the buy-in more acceptable, relying on peer review to make judgements (NHIA, 2013). However, these traditional methods of detecting health care fraud and abuse are time consuming and inefficient. There is a need for statistically sound legal methods

of measuring fraud and quantifying the effect of reduction in losses due to fraud and tracking their tangible financial benefits. It is generally more preferable to pre-empt fraud rather than just react to it.

The data generated by the health organizations is very vast and complex due to which it is difficult to analyze the data in order to make important decision regarding patient health. This data contains details regarding hospitals, patients, medical claims, treatment cost etc. So, there is a need to generate a powerful tool for analyzing and extracting important information from this complex data. Data Mining is one of the most vital and motivating areas of research with the objective of finding meaningful information from huge data sets. In the present era, Data Mining is becoming popular in healthcare field because there is a need for an efficient analytical methodology for detecting unknown and valuable information in health data. In the health industry, Data Mining provides several benefits such as detection of fraud in health insurance, availability of the medical solution to the patients at lower cost, detection of causes of diseases and identification of medical treatment methods. It also helps the healthcare researchers in making efficient healthcare policies, constructing drug recommendation systems, developing health profiles of individuals etc. (Koh & Tan, 2005 cited in Tomar & Agarwal, 2013).

Putting the foregoing discourse on fraudulent and abusive acts into a closer perspective, one would not be far from right in admitting that fraud and abuse are becoming a canker that needs to be nipped in the bud using appropriate mechanism before they become malignant. The focus of this study is to use data mining to detect fraud and abuse in the health insurance scheme specifically at Awutu-Effutu-Senya District on malaria cases.

**Overview of Malaria cases in Ghana**

Malaria is a common and life-threatening disease in many tropical and subtropical areas. There are currently over 100 countries and territories where there is a risk of malaria transmission, and these are visited by more than 125 million international travellers every year (WHO, 2006). The Corporate Alliance on Malaria in Africa (2011) reported that two hundred and twenty-five (225) million malaria cases were reported in 2009. Out of this, 78% cases were reported in the African region. It further added that malaria can decrease gross domestic product by as much as 1.3% in countries with high disease rates and in some high-burden countries, malaria accounts for 30% to 50% of inpatient hospital admissions.

In Ghana, according to the NHIA, it is known that over 27% of the scheme's medicines cost is attributable to anti-malarial medicines. For example, in 2009 the scheme spent over GH₵51 million on anti-malarial medicines alone (NHIA, 2010). Malaria has been a major cause of poverty and low productivity accounting for about 32.5 percent of all OPD attendances and 48.8 percent of under five years' admissions in the country (National Malaria Control Programme, 2009).

The attempt to control malaria in Ghana began in the 1950s. It was aimed at reducing the malaria disease burden till it is no longer of public health significance. It was also recognized that malaria cannot be controlled by the health sector alone, therefore, multiple strategies were being pursued with other health-related sectors (Ghana Health Service, 2016).

Furthermore, malaria is Ghana's main cause of mortality and morbidity. In 2006, 19% of deaths, 34% of under 5 deaths and 9% of maternal deaths were as a consequence of the disease. In 2006, there were 30,448,596 reported cases of adult malaria and 7,286,356 reported cases of child malaria. Although it is thought that there are many more cases that go unreported and untreated

*Exploring fraud and abuse in National Health Insurance Scheme (NHIS) using data mining technique as a statistical model*

*Kittoe, J. D. & Asiedu-Addo, S. K.*

(Ministry of Health, 2013). As a result of the high cases of malaria reported in Ghanaian hospitals, there is that tendency of fraud and abuse specifically in the billing and provision of drugs, proper treatment conditions, environment and strict compliance to treatment as reported by Hanson et al (2008). Moreover, what government or system failure can cause in the treatment process is a clear depiction that can lead to abuse in the wrong diagnosis resulting in treatment mismatch.

**Data Mining Process**

The objective of data mining is to identify valid novel, potentially useful, and understandable correlations and patterns in existing data (Chung & Gray 1999 cited in Jackson, 2002). With the enormous amount of data stored in files, databases, and other repositories, it is increasingly important, if not necessary, to develop powerful means for analysis and perhaps interpretation of such data and for the extraction of interesting knowledge that could help in decision-making. While data mining and Knowledge Discovery in Databases (or KDD) are frequently treated as synonyms, data mining is actually part of the knowledge discovery process. Figure 1 shows data mining as a step in an iterative knowledge discovery process.
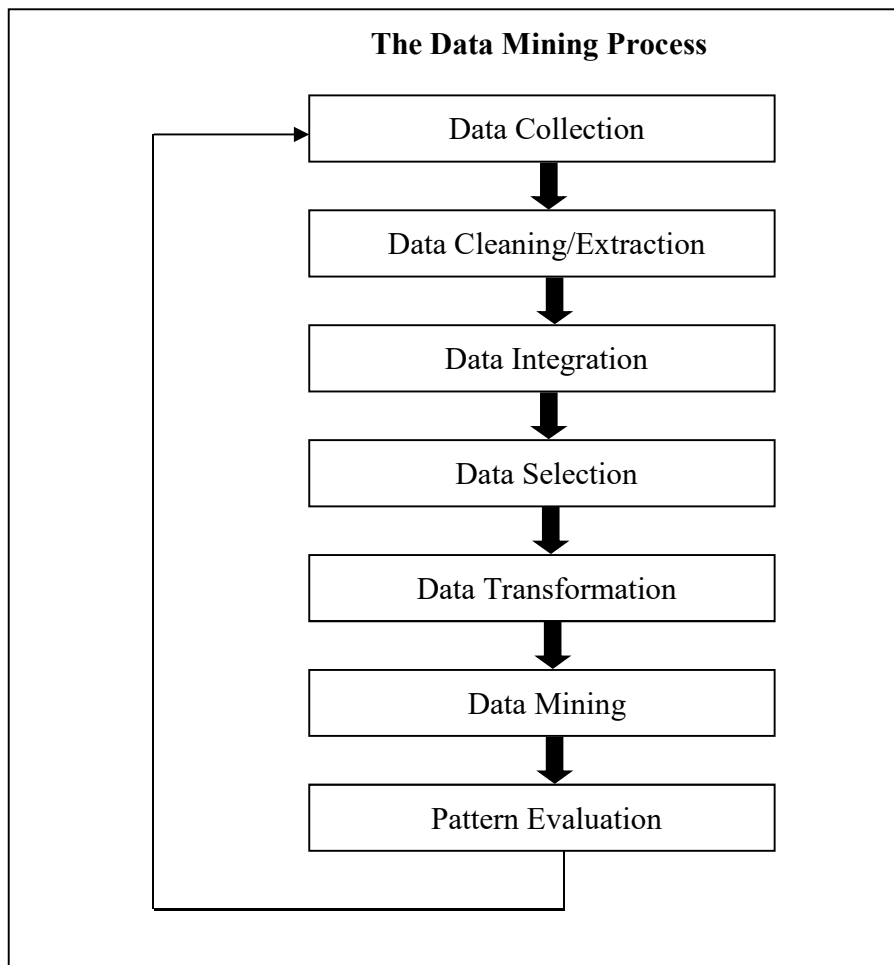
**Figure 1    The data mining process**

As shown in Figure 1, the data mining process is a pipeline containing many phases including data collection, data cleaning/extraction, data integration, data selection, data transformation, data mining and pattern evaluation. These phases can be described as follows;

- Data collection: data collection may require the use of specialized hardware such as a sensor network, manual labour such as the collection of user surveys, or software tools such as a Web document crawling engine to collect documents. While this stage is highly application-specific and often outside the realm of the data mining analyst, it is critically important because good choices at this stage may significantly impact the data mining process. After the collection phase, the data are often stored in a database, or, more generally, a data warehouse for processing (Aggarwal, 2015).

- Data cleaning/extraction: When the data are collected, they are often not in a form that is suitable for processing. For example, the data may be encoded in complex logs or free-form documents. In many cases, different types of data may be arbitrarily mixed together in a free-form document. To make the data suitable for processing, it is essential to transform them into a format that is friendly to data mining algorithms, such as multidimensional, time series, or semi-structured format. It is crucial to extract relevant features for the mining process. The feature extraction phase is often performed in parallel with data cleaning, where missing and erroneous parts of the data are either estimated or corrected. In other words, it is a phase in which noise data and irrelevant data are removed from the collection. In many cases, the data may be extracted from multiple sources and need to be integrated into a unified format for processing. The final result of this procedure is a nicely structured data set, which can be effectively used by a computer program. After the feature extraction phase, the data may again be stored in a database for processing (Zaïane, 1999; Aggarwal, 2015).

- Data integration: at this phase, multiple data sources, often heterogeneous, may be combined in a common source. This involves a data pre-processing technique that merges the data from multiple heterogeneous data sources into a coherent data store.

- Data selection: at this phase, the data relevant to the analysis is decided on and retrieved from the data collection.

- Data transformation: also known as data consolidation, it is a phase in which the selected data is transformed into forms appropriate for the mining procedure.

- Data mining: it is the crucial step in which clever techniques are applied to extract patterns potentially useful.

- Pattern evaluation: in this phase, strictly interesting patterns representing knowledge are identified based on given measures (Zaïane, 1999).

**Problem Statement**

Many were with the belief that the introduction of the NHIS would reduce the burden on household income spent on health care especially with the intervention of the government subsidy. However, this laudable social intervention policy has been bedevilled with a lot of underhand dealings on malaria cases i.e. fraud and abuse from the major stakeholders; the scheme staff, the health care

*Exploring fraud and abuse in National Health Insurance Scheme (NHIS) using data mining technique as a statistical model*

*Kittoe, J. D. & Asiedu-Addo, S. K.*

providers and the clients. These devious practices if not checked will collapse the conceptual framework underpinning the establishment of the scheme under which free health services are rendered to Ghanaian citizens.

In traditional methods of health care fraud and abuse detection, a few auditors handle thousands of paper health care claims. In reality, they have little time for each claim, focusing on certain characteristics of a claim without paying attention to the comprehensive picture of a provider's behaviour (Rashidian, Joudaki, & Vian, 2012). Even though this method is prolonged and unproductive, it is still the dominant picture in many low-income and middle-income countries including Ghana There is therefore the need to invest in dealing with fraud by facilitating stronger national fraud control networks as well as encouraging professional training and accreditation for work on fraud. Approaches to dealing with fraud also need to move away from a 'pay and chase' focus to a proactive and comprehensive approach that identifies and applies the right solutions focused on outcomes, such as reduction in losses.

Data mining has been used in various developed countries to help detect fraud and abuse in the health care services or health insurance setting. It has helped third-party payers such as health insurance organizations to extract useful knowledge from thousands of claims and identify a smaller subset of the claims or claimants for further assessment and scrutiny for fraud and abuse (Rashidian et al., 2012). In this way, the data mining approach is part of a more efficient and effective IT-based auditing system. However, in Ghana, the literature appears to suggest that there has been little investigation involving data mining techniques. Thus, very little studies have applied data mining techniques to detect fraud and abuse in the health insurance setting and also to make relevant recommendations for healthcare services. Meanwhile, there is adequate evidence of fraud and abuse in the National Health Insurance setting (NHIA, 2013). It is against this background that the researchers found it expedient and highly motivated to look for possible application of Data mining techniques on fraud and abuse detection in the Health Insurance Scheme at Awutu-Effutu-Senya District.

## Purpose of the Study and Research Questions

The primary purpose of this study is to explore fraud and abuse of NHIS claims on malaria treatments in Awutu, Effutu and Senya districts in Ghana and the pattern of fraud and abuse in such claims. In pursuance of this purpose the following research questions were formulated to guide the study;

1. What is the mean cost of claims on malaria drugs? How significantly does the cost of claims on malaria drugs differ among the NHIS service providers?
2. What patterns of fraud and abuse exit in the NHIS claims on malaria treatment?

In answering the first research question, the hypothesis below was formulated for the study:

$H_o$: There is no significant difference among the NHIS service providers in the mean cost of claims made on malaria drugs.

$H_a$: There is a significant difference among the NHIS service providers in the mean cost of claims made on malaria drugs.

## Method

*Research Design*

The researcher employed the quantitative research approach with survey design as a strategy of enquiry. The 'survey' is used in a variety of ways but commonly refers to the collection of standardized information from a specific population, or some sample from one, usually but not necessarily by means of questionnaires or interviews (Fraenkel & Wallen, 2003). The survey in this study was used for two major research purposes: descriptive and analytical. In the former, the researcher aimed at getting an accurate description of the Health Insurance Scheme at Awutu-Effutu-Senya District, with a specific focus on malaria claims. For the analytic, the researcher attempted to formulate a hypothesis based on the data collected and checked against further information for differences and relationships.

*Population and Demography*

Awutu-Effutu-Senya, NHIA Office operation covers three political administrations, Effutu Municipal with Winneba as its administrative capital, Awutu Senya District and Awutu Senya East Municipality with Awutu Bereku and Odukpongkpehe Kasoa as their administrative capitals respectively. The Awutu-Effutu-Senya NHIS include Awutu-Beraku, Awutu-Bawjiase, Awutu-Obrachire, Awutu-Bontrase, Awutu-Offankor, Nyarkokwaa and Awutu-Offadaa. The rest are, Awutu-Ahentia, Jei-Krodua, Awutu-Mampong, Kasoa, Effutu-Essuekyir, Effutu-Gyahaadze, Gyangyanadze, Osebonpanyin, Atteitu, Sankor and Winneba. Most of the people living in these catchment areas of the scheme are predominantly peasant farmers and fishermen. The district office of the scheme is located at South Campus of the University of Eduaction, Winneba. The branch became functional in early January, 2006, after its official inauguration on December 19, 2005.

The 2010 population census by Ghana Statistical Service (GSS, 2010) estimated that, the three political areas combined have about 256,000 populations with Awutu Senya East having the highest with 108,422, followed by Awutu Senya District with 86,884 and Effutu Municipality as the least with 68,597. Based on growth rates of various districts projected on 2010 census population census, the current projected figure for 2016 is 314,257. The target population for the study was all health insurance registered members in the Awutu-Effutu-Senya District in the Central Region of Ghana. This comprises active clients totalling 190,640 accessing a total of 42 NHIS service providers. This was influenced by the proximity to the Researchers' place of residence and also the only health Insurance Office available within the researcher's working area.

*Sample and Sampling Technique*

Prior to sampling data and analyzing the NHIS data, some form of central data storage and access facility was created, which is commonly referred to as a data warehouse. Data warehouses enable many groups to access the data, facilitate updating the data, and improve the efficiency of checking the data for reliability and preparing the data for analysis and reporting (Statistical Analysis System (SAS), 2002).

Types of sampling commonly used in data mining projects include simple random sampling, $n-th$ record sampling, first $n$ sampling, cluster sampling and stratified random sampling (SAS, 2002). The researcher employed Stratified Random Sampling in selecting clients. However, convenience sampling was used in selecting clients, of National Health Insurance Scheme of

*Exploring fraud and abuse in National Health Insurance Scheme (NHIS) using data mining technique as a statistical model*

*Kittoe, J. D. & Asiedu-Addo, S. K.*

Awutu-Effutu-Senya District for the study. Convenience sampling was used because of logistic and financial constraints as well as ease of accessibility.

Stratified sampling was used to select a total sample of 720 clients diagnosed of malaria for the year 2013, 2014 and 2015 (240 clients from each year). The clients were selected from 4 NHIS service providers (2 private and 2 public) of National Health Insurance Scheme of Awutu-Effutu-Senya District. In selecting the sample, 10 clients each were selected from the age ranges 0-10, 11-20, 21-30, 31-40, 41-50 and 51 & above for each year. Stratified sampling technique has the advantage of increasing the likelihood of representativeness of the sample and virtually ensures that any key characteristics of individuals in the population are included in the same proportion in the sample (Fraenkel &Wallen, 2003).

*Data Collection*

In line with ethical principles, permission letter was written and sent to the Manager of the National Health Insurance Scheme of Awutu-Effutu-Senya District. Approval and consent were sought from the district scheme manager to enable the researcher to have access to interact with the claims department to obtain appropriate claims data. The data obtained was from a total of 42 NHIS service providers out of which a selection of 4 within the Awutu Effutu Senya Municipal in the Central region of Ghana were made. Data cleaning was performed on the data collated to remove noise and inconsistent data. This made the data suitable for processing, it was also essential to transform them into a format that is friendly to data mining techniques. Secondly, data integration was performed, where multiple data sources, often heterogeneous, were combined in a common source. This involved a data pre-processing technique that merged the data from multiple heterogeneous data sources into a coherent data store. In other words, the data integration was performed in cases where multiple data sources were combined.

Thirdly, data relevant to the analysis task were retrieved and selected for processing. Data transformation was the fourth step, where data were transformed or consolidated into forms appropriate for the mining procedure. Data mining was the fifth and most crucial step carried out, in which clever techniques were applied to extract patterns potentially useful. Pattern evaluation was the sixth step; knowledge-based data mining system was able to identify truly interesting patterns representing knowledge based on given measures. Lastly, knowledge presentation was done where visualization and knowledge representation techniques were used to present the mined knowledge to the user. Data obtained was thus subjected to thorough analysis in mining interesting disease patterns which will ultimately assist for proper decision making in the health sector based on different data of patients with different cases of malaria.

*Data Analysis*

The study sought to explore the application of data mining technique on fraud and abuse detection in the Health Insurance Scheme at Awutu-Effutu-Senya District, with a specific focus on malaria cases. The research generated quantitative data. Emphasis was placed on medical cost for malaria cases and detection of various forms of fraud and abuse in the Health Insurance Scheme with respect to malaria claims. Data mining techniques (data Collection, data cleaning/extraction, data integration, data Selection, data Transformation, data mining, pattern evaluation) were used to explore patterns of fraud and abuse in the malaria claims obtained. In addition, data collected were

analyzed by using descriptive statistics (that is, standard deviation, charts and mean) and inferential statistics (that is, two-way between subjects ANOVA) to interpret the data gathered from the Health Insurance Scheme at Awutu-Effutu-Senya District through data mining techniques. The two-way between subjects ANOVA was used to determine whether there is interaction between NHIS service providers (hospitals) and the cost of drugs.

## Results

*Mean cost of claims on malaria drugs from the NHIS service providers*

The researchers obtained data on the cost of IPD (drugs), drugs prescribed and cost of the drugs recorded in the years 2013, 2014 and 2015 by the NHIS service providers. This provided a way of assembling and analyzing the data obtained into useful information, to help find the mean cost of malaria treatment. Table 1 shows the mean and standard deviation of the cost of malaria drugs recorded in the three years.

**Table 1**      **Mean and standard deviation of the cost of malaria drugs recorded in 2013, 2014 and 2015**

| NHIS Service Providers | 2013 | | 2014 | | 2015 | |
|---|---|---|---|---|---|---|
| | **Mean** | **St. Dev.** | **Mean** | **St. Dev.** | **Mean** | **St. Dev.** |
| Hospital A Private | 5.79 | 1.67 | 7.21 | 3.97 | 4.51 | 2.51 |
| Hospital B Public | 6.36 | 3.71 | 9.05 | 6.04 | 5.74 | 3.21 |
| Hospital A Public | 5.90 | 2.13 | 6.79 | 3.53 | 5.88 | 2.65 |
| Hospital B Private | 7.53 | 2.97 | 16.84 | 6.03 | 16.15 | 8.22 |

In Table 1, data from the year 2013 shows differences in the statistics. It shows Hospital A Private (M=5.79, SD=1.67), Hospital B Public (M=6.36, SD= 3.71), Hospital A Public (M=5.90, SD=2.13) and Hospital B Private (M=7.53, SD=2.97) which reveal that Hospital B Private had the highest mean cost whereas Hospital A Private had the lowest. In the year 2014, Hospital B Private again had the highest mean cost whereas the Hospital A Public had the lowest. In 2015, Hospital B Private again had the highest mean cost whereas the Hospital A Private had the lowest. To ascertain if the differences observed (Table 1) in the costs of malaria drugs recorded by the four NHIS service providers (i.e. two public and two private) are statistically significant the data was subjected to further analysis by a two-way between-subjects Analysis of Variance (ANOVA). The results obtained for the various years are presented in Tables 2, 3 and 4.

*Exploring fraud and abuse in National Health Insurance Scheme (NHIS) using data mining technique as a statistical model*

*Kittoe, J. D. & Asiedu-Addo, S. K.*

**Table 2**     **Two-way ANOVA results for differences in the costs of malaria drugs recorded by the four NHIS service providers in 2013**

| NHIS Service Providers | Mean | SD | Std. Error | 95% Confidence Interval Lower Bound | 95% Confidence Interval Upper Bound | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|---|---|---|---|
| Hospital A Private | 5.79 | 1.67 | 0.33 | 5.14 | 6.43 | | | | | |
| Hospital B Public | 6.36 | 3.71 | 0.33 | 5.71 | 7.00 | | | | | |
| Hospital A Public | 5.90 | 2.13 | 0.33 | 5.26 | 6.55 | 114.513 | 3 | 38.171 | 5.992 | .001 |
| Hospital B Private | 7.53 | 2.97 | 0.33 | 6.89 | 8.18 | | | | | |

In Table 2, there was a significant main effect of the Service Providers, $F(3, 216) = 5.992, p < 0.05$. A post-hoc multiple comparison test using Tukey HSD revealed that the mean cost of drugs at Hospital B Private (M=7.53, SD=2.97) is highly statistically different at 5% level of significance from that at Hospital A Private (M=5.79, SD=1.67) and Hospital B Public (M=6.36, SD=3.71) is also highly statistically different from Hospital A Public (M=5.90, SD=2.13). This indicates that Hospital B Private had the highest mean cost across the hospitals whereas Hospital A Private had the least mean cost in 2013.

**Table 3**     **Two-way ANOVA results for differences in the costs of malaria drugs recorded by the four NHIS service providers in 2014**

| NHIS Service Providers | Mean | SD | Std. Error | 95% Confidence Interval Lower Bound | 95% Confidence Interval Upper Bound | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|---|---|---|---|
| Hospital A Private | 7.21 | 3.97 | 0.58 | 6.07 | 8.34 | | | | | |
| Hospital B Public | 9.05 | 6.04 | 0.58 | 7.91 | 10.19 | | | | | |
| Hospital A Public | 6.79 | 3.53 | 0.58 | 5.66 | 7.93 | 3948.51 | 3 | 1316.17 | 66.06 | 0.00 |
| Hospital B Private | 16.84 | 6.03 | 0.49 | 15.19 | 17.12 | | | | | |

In Table 3, there was a significant main effect of the Service Providers, $F(3, 216) = 66.06, p < 0.05$. A post-hoc multiple comparison test using Tukey HSD revealed that the mean cost of drugs at Hospital B Private (M=16.84, SD=6.03) is highly statistically different at 5% level of significance from that at Hospital A Private (M=7.21, SD=3.97) and Hospital B Public (M=9.05, SD=6.04) is also highly statistically different from Hospital A Public (M=6.79, SD=3.53). This

indicates that Hospital B Private has the highest mean cost across the hospitals whereas Hospital A Public has the least mean cost.

**Table 4**      **Two-way ANOVA results for differences in the costs of malaria drugs recorded by the four NHIS service providers in 2015**

| NHIS Service Providers | Mean | SD | Std. Error | 95% Confidence Interval Lower Bound | 95% Confidence Interval Upper Bound | Type III Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|---|---|---|---|
| Hospital A Private | 4.51 | 2.51 | 0.49 | 3.55 | 5.48 | | | | | |
| Hospital B Public | 5.74 | 3.21 | 0.49 | 4.78 | 6.70 | | | | | |
| Hospital A Public | 5.88 | 2.65 | 0.49 | 4.91 | 6.84 | 5292.68 | 3 | 1764.22 | 122.9 | .000 |
| Hospital B Private | 16.15 | 8.22 | 0.49 | 15.19 | 17.12 | | | | | |

In table 4, there was a significant main effect of the Service Providers, $F(3, 216) = 122.94, p < 0.05$. There was also a significant main effect of the age groups, $F(5, 216) = 5.520, p < 0.05$. A post-hoc multiple comparison test using Tukey HSD revealed that the mean cost of Hospital B Private (M=16.15, SD=8.22) is highly statistically different at 5% level of significance from Hospital A Private (M=4.51, SD=2.51) and Hospital B Public (M=5.74, SD=3.21) is also highly statistically different from Hospital A Public (M=5.88, SD=2.65). This indicates that Hospital B Private has the highest mean cost of drugs across the hospitals whereas Hospital A Private has the least mean cost.

*Patterns of fraud and abuse that exit in the NHIS claims on malaria treatment*

In order to detect possible fraud and abuse in the malaria claims obtained, data mining process was employed as follows;

*Data Collection*

From the analysis of medical cost in the various health facilities for the years 2013, 2014 and 2015, it was revealed that Hospital B private had the highest mean cost of drugs over the three-year period. The mean cost of malaria drug for a client at Hospital B private was GH₵ 7.53, GH₵ 16.84 and GH₵ 16.15 in 2013, 2014 and 2015 respectively. In order to put up a robust classifier using only a small set of data, it was decided to collect the data from Hospital B private to apply data mining strategy on the malaria reported cases since previous analyses proved that their mean cost of malaria drugs was relatively high.

*Data Cleaning/Extraction*

Only the malaria reported cases for 2014 was used for mining patient records to discover categories of fraudulent and abusive cases. The researcher again removed data on cost of service, IPD (drugs), IPD (service). Also, clients with missing data, empty values, nonexistent values or incomplete data were removed from the data collected. After the data cleaning/extraction phase, the data were again stored in another database for further processing.

*Exploring fraud and abuse in National Health Insurance Scheme (NHIS) using data mining technique as a statistical model*

*Kittoe, J. D. & Asiedu-Addo, S. K.*

*Data Integration*

The researcher integrated data by merging data of clients into various age groups namely ages 0-10, 11-20, 21-30, 31-40, 41-50 & 51 and above. Clients were put into these various age groups to allow for homogeneity across the hospitals.

*Data Selection*

Data were thoroughly screened month by month to select fraudulent and abusive cases (of malaria) that appeared in the various months. After data integration phase, the researcher selected three to five clients for each category of fraudulent and abusive acts. These data are presented in tables 5, 6, 7, 8, 9 and 10.

*Data Transformation*

In this phase the selected data were put into appropriate format for data mining. The dates of visit of a client to the hospital, clients' name as well as the e cost of drugs for the visits were included, as shown in tables 5, 6, 7, 8, 9 and 10.

*Data Mining and Pattern Evaluation*

After data selection and transformation, an exhaustive classification/categorization of fraud and abuse was done; the detailed description follows in Tables 5, 6, 7, 8, 9 and 10 below.

**Table 5          Repetition of NHIS Registration number**

| Date | Clients Name | NHIS Reg. No. | Drugs Prescribed | Cost of Drugs |
|------|--------------|---------------|------------------|---------------|
| 27/02/2014 | A1 | NHISAE13423 | Amodiaquine | 12.50 |
| 23/05/2014 | B1 | NHISAE13423 | Artemether | 12.00 |
| 30/07/2014 | C1 | NHISAE13423 | Amodiaquine | 23.50 |
| 12/11/2014 | D1 | NHISAE13423 | Artemether + Lumefantrine | 34.00 |

The NHIS registration number is unique and distinct; it is inherent in the NHIS that each and every registered member of the scheme shall have a unique registration number. This allows for easy identification of persons and makes health record of clients unique. For this reason, not more than one person should have same NHIS registration number. However, thorough search (through data mining) revealed the pattern whereby persons (more than one client) had the same NHIS registration number. Table 5 shows that clients A1, B1, C1 and D1 all shared the same registration number, NHISAE13423. These clients visited Hospital B Private on different dates, given different prescriptions of drugs with different cost.

**Table 6          Over Billing of Medicine**

| Date | Clients Name | NHIS Reg. No. | Drugs Prescribed | Cost of Drugs |
|------|--------------|---------------|------------------|---------------|
| 24/02/2014 | E1 | NHISAE16783 | Amodiaquine + Artesunate (75mg + 135mg) | 9.50 |
| 13/05/2014 | F1 | NHISAE17429 | Artemether + Lumefantrine 20/120mg (6 tablets) | 12.00 |
| 24/06/2014 | G1 | NHISAE24825 | Artemether + Lumefantrine (Powder) | 4.80 |

The NHIA has developed an original list of medicines using a process that involved evaluation of evidence for the management of health problems commonly seen at health facilities in Ghana. A Specialists group comprising medical doctors, pharmacists and a midwife reviewed the evidence for the management of the health problems and selected the appropriate medicines with their respective drugs (NHIA, 2013). When this list was compared to the cost of drugs prescribed for clients at Hospital B Private, another pattern, over-billing of drugs was revealed. From table 6 Amodiaquine + Artesunate (75mg + 135mg) for children was charged GH₵ 9.50 instead of GH₵ 8.50 as in the case of Client E1 (NHISAE16783) who attended hospital on 24/02/2014. Some of the Artemether + Lumefantrine 20/120mg (6 tablets) and Artemether + Lumefantrine (Powder) served have their unit prices wrongly quoted as GH₵ 12.00 and GH₵ 4.80 respectively instead of GH₵ 2.50 and GH₵ 3.50 respectively. These are seen in the cases of Client F1 (NHISAE17429; 13/05/2014) and Client G1 (NHISAE24825; 24/06/2014).

**Table 7          Drug Mismatch**

| Date | Clients Name | NHIS Reg. No. | Drugs Prescribed | Cost of Drugs |
|------|--------------|---------------|------------------|---------------|
| 28/02/2014 | M1 | NHISAE12343 | Amoxicillin + Clavulanic Acid | 12.50 |
| 13/06/2014 | N1 | NHISAE67429 | Carbocisteine | 10.80 |
| 24/06/2014 | R1 | NHISAE26025 | Oxytocin Injection | 8.58 |
| 14/12/2014 | S1 | NHISAE76001 | Ciprofloxacin | 8.60 |

Another pattern revealed was wrong prescription of drugs for malaria cases. Table 7 shows that antibiotics such as Amoxicillin + Clavulanic Acid and Carbocisteine which are supposed to be prescribed for bacterial infections were prescribed for patients diagnosed of malaria as in the case of Client M1 (NHISAE12343; 28/02/2014) and Client N1 (NHISAE67429; 13/06/2014).

*Exploring fraud and abuse in National Health Insurance Scheme (NHIS) using data mining technique as a statistical model*

*Kittoe, J. D. & Asiedu-Addo, S. K.*

**Table 8          Excessive Prescription of Malaria Treatment**

| Date | Clients Name | NHIS Reg. No. | Drugs Prescribed | Cost of Drugs |
|------|--------------|---------------|------------------|---------------|
| 24/02/2014 | Q1 | NHISAE12451 | Amodiaquine + Artesunate (75mg + 135mg), Parafin Tab, Folic acid | 25.70 |
| 03/03/2014 | T1 | NHISAE67451 | Albendazone, Paracetamol, Blood tonic | 13.80 |
| 01/07/2014 | Y1 | NHISAE54313 | Artemether + Lumefantrine, Diclofenac, Diazepam | 46.40 |
| 22/12/2014 | Z1 | NHISAE10313 | Amodiaquine + Artesunate, Blood Tonic, Diclofenac Gel, Folic acid | 38.70 |

The NHIS statute authorizes payment for drugs that are included in each disease's approved plan. The included drugs vary from disease to disease. Only those drugs included in the relevant disease's plan are authorized. Even if a drug is authorized, it is still not covered under NHIS unless it is also medically necessary. From table 8, some clients were unnecessarily given many drugs resulting in unusually high cost of drugs for malaria treatment. For example, Client Y1 was given Artemether + Lumefantrine, Diclofenac and Diazepam at a high cost of GH₵ 46.40, Client Z1 was also given excessive drugs including Amodiaquine + Artesunate, Blood Tonic, Diclofenac Gel and folic acid also at a cost of GH₵ 38.70.\

**Table 9          Duplication of Clients Records (OPD)**

| Date | Clients Name | NHIS Reg. No. | Drugs Prescribed | Cost of Drugs |
|------|--------------|---------------|------------------|---------------|
| 13/02/2014 | A2 | NHISAE10027 | Artemether + Lumefantrine Tablet | 5.80 |
| 13/02/2014 | A2 | NHISAE10027 | Artemether + Lumefantrine Tablet | 5.80 |
| 13/02/2014 | A2 | NHISAE10027 | Artemether + Lumefantrine Tablet | 5.80 |
| 25/05/2014 | B2 | NHISAE34313 | Amodiaquine + Artesunate | 8.70 |
| 25/05/2014 | B2 | NHISAE34313 | Amodiaquine + Artesunate | 8.70 |
| 09/08/2014 | C2 | NHISAE12139 | Artemether with other supplements | 12.5 |
| 09/08/2014 | C2 | NHISAE12139 | Artemether with other supplements | 12.5 |
| 09/08/2014 | C2 | NHISAE12139 | Artemether with other supplements | 12.5 |
| 09/08/2014 | C2 | NHISAE12139 | Artemether with other supplements | 12.5 |

Another form of fraud and abuse discovered was duplication of clients' records. This entails duplication of claims and breach of tariff rules i.e. same person, same NHIS number, same date, same diagnosis, same drug prescription and cost. These are evident in table 9 as seen in the records of clients A2, B2 and C2. As a result of the duplication, the total cost of drugs for a client is higher than normal.

**Table 10        Cost incurred by Fraud and Abuse**

| Category of Fraud and Abuse | Actuals (GHC) | Volume (Gains) | Gains (GHC) | Reimbursement (GHC) |
|---|---|---|---|---|
| Duplication of Clients Records (OPD) | 287.90 | 29 | 302.55 | 590.45 |
| Excessive Prescription of Malaria drugs | 761.56 | 110 | 1,103.98 | 1,865.54 |
| Repetition of NHIS Registration Numbers | 356.85 | 98 | 678.93 | 1,035.78 |
| Overbilling of Medicine | 256.75 | 42 | 504.80 | 761.55 |
| Drug Mismatch | 272.00 | 32 | 752.00 | 1,024 |
| TOTAL | 1,935.06 | 311 | 3,342.26 | 5,277.32 |

The five categories of fraud and abuse identified incurred huge sums of money to be paid by the NHIA. From table 10 the actual amount of money (Actuals) that would have been paid to the hospital if there were no such fraud and abuse was excessively raised (i.e. Gains) due to the fraudulent and abusive practices by the hospital. Table 10 shows 29 cases on Duplication of Clients' records resulting in a sum of GH₵302.55 to be paid by NHIA to the hospital. However, the actual amount that would have been paid to the hospital if there were no Duplications is GH₵287.90. On Excessive prescription of malaria drugs there were 110 cases causing an amount of GH₵1103.98. The actual amount that would have been paid is GH₵761.56. Repetition of NHIS Registration number caused GH₵678.93 while the actual cost, if there were no repetitions would have been GH₵356.85. Moreover, 42 cases were recorded for Overbilling of medicines resulting in a total cost of GH₵504.80 while the actual cost, if there were no overbilling would have been GH₵256.75. Lastly, 32 cases recorded for Drug mismatch resulted in a cost of GH₵752.00. However, the actual cost if there were no mismatch would have been GH₵272.00. In all, the actual total cost that should have been paid to the hospital if there were no such five categories of fraud and abuse is GH₵1,935.06. Instead, a relatively huge amount of GH₵3,342.26 was paid.

Figure 2 gives a visual confirmation of the forms of fraud and abuse identified in the health insurance setting with their respective percentages. Excessive prescription of malaria drugs constituted 35% of the entire fraud and abuse, repetition of NHIS registration number formed 32%, overbilling of medicine and drug mismatch amounted to 14% and 10% respectively and duplication of clients' records formed 9% of the entire fraud and abuse.
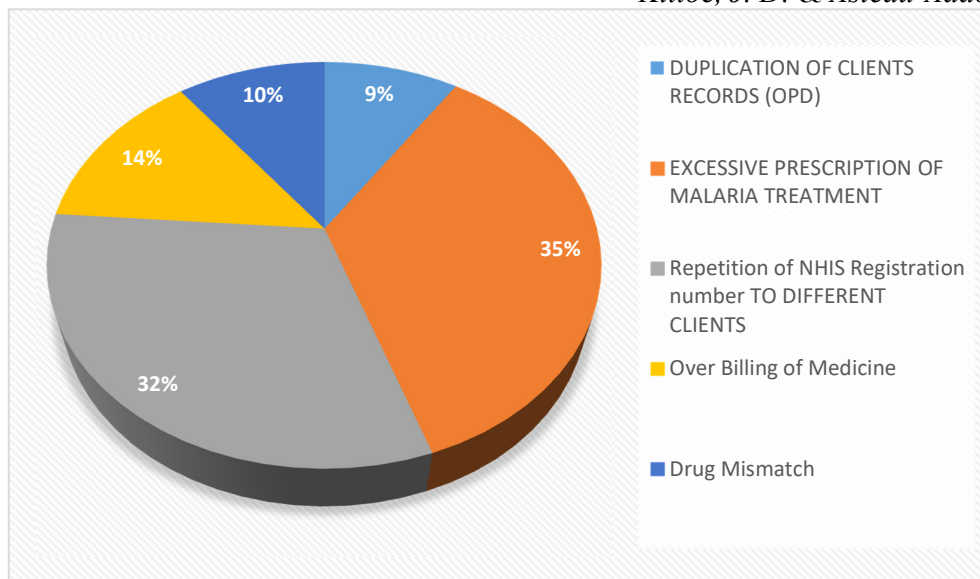
*Exploring fraud and abuse in National Health Insurance Scheme (NHIS) using data mining technique as a statistical model*

*Kittoe, J. D. & Asiedu-Addo, S. K.*

**Figure 2**        **Categories of Fraud and Abuse in the NHIS Setting**

## Discussion

In order to find the mean cost of drugs for the treatment of malaria, a two-way between-subjects Analysis of Variance (ANOVA) was conducted to investigate the costs of drugs administered to patients with malaria related cases across four health facilities (two of which are public and the other two being private) in the Awutu-Effutu-Senya District. The findings from the two-way between-subjects Analysis of Variance (ANOVA) indicated that there was a statistically significant difference in the mean cost of drugs administered to patients with malaria related cases across four health facilities in the Awutu-Effutu-Senya District. The test was conducted for malaria cases recorded in the years 2013, 2014 and 2015 in the four health facilities selected for the study. Hospital B Private had the highest mean cost across the hospitals whereas Hospital A Private had the least mean cost for the three years. The mean cost of malaria drug for a client at Hospital A Private was GH₵5.79, GH₵7.21 and GH₵4.51 in 2013, 2014 and 2015 respectively. That of Hospital B Private was GH₵7.53, GH₵16.84 and GH₵16.15

On fraud and abuse detection, the entire malaria cases in Hospital B Private reported in 2014 was used for mining patient records to discover categories of fraudulent and abusive cases since the Hospital recorded the highest mean cost of drugs for the three-year period. After thoroughly screening the data using data mining approach the following fraudulent and abusive cases (of malaria) were detected; repetition of NHIS registration number, over billing of drugs, drug mismatch, excessive prescription of malaria treatment and duplication of clients records (OPD). These Health care fraud and abuse resulted in substantial loss of money which was very costly to the health insurance system. The actual total cost that should have been paid to Hospital B Private, if there were no such five categories of fraud and abuse is GH₵1,935.06. However, a relatively huge amount of GH₵3,342.26 was paid as a result of the fraud and abuse identified. These findings are also consistent with those of Gee, Button, Brooks and Vincke (2010) and United States

Department of Health and Human Services (2014) who also reported such forms of fraudulent and abusive acts as misusing codes/registration numbers, charging excessively for services or supplies and billing for services that were not medically necessary.

**Conclusion**

Health insurance fraud increases the disorganization and unfairness in our society. Health care fraud leads to substantial losses of money and very costly to health care insurance system. It is horrible because the percentage of health insurance fraud keeps increasing every year in many countries. To address this widespread problem, effective techniques are in need to detect fraudulent claims in health insurance sector. This study provided a comprehensive survey of the statistical data mining methods applied to detect fraud and abuse in health insurance sector and to find the mean cost of the treatment of malaria.

The researchers focused on malaria cases in Out Patient Services (OPD) category due to its sensitive nature on the World Health Organization on the global scale and also to the Government of Ghana on the national front for the reduction of infant mortality rate. Malaria is a common and life-threatening disease in many tropical and subtropical areas. In tackling the issue of malaria in a more cost effective means, patterns were explored in finding the mean cost of drugs for the treatment of malaria. A two-way between-subjects Analysis of Variance (ANOVA) was conducted to investigate the costs of drugs administered to patients with malaria related cases across four health facilities (two of which are public and the other two being private) in the Awutu-Effutu-Senya District. The test was conducted for malaria cases recorded in the years 2013, 2014 and 2015 in the four health facilities selected for the study. Hospital B Private had the highest mean cost across the hospitals whereas Hospital A Private had the least mean cost for the three years. The mean cost of malaria drug for a client at Hospital A Private was GHC 5.79, GHC 7.21 and GHC 4.51 in 2013, 2014 and 2015 respectively.

One of the major objectives of this study was to use data mining technique to detect fraud and abuse in the NHIS with reference to malaria related cases. After thoroughly screening the data using data mining approach the following fraudulent and abusive cases of malaria were detected; repetition of NHIS registration number, over billing of drugs, drug mismatch, excessive prescription of malaria treatment and duplication of clients records (OPD). These Health care fraud and abuse resulted in substantial loss of money which was very costly to the health insurance system. The actual total cost that should have been paid to the Hospital B Private, if there were no such five categories of fraud and abuse is GHC 1,935.06. However, a relatively huge amount of GHC 3,342.26 was paid as a result of the fraud and abuse identified.

**Recommendations**

Based on the findings of the study the following are important considerations that would help improve management of the NHIS to ensure long-term sustainability. The NHIA should:

1. Increase advocacy and sensitization of the impact of fraud and abuse on the health insurance system.
2. Pass specific health insurance fraud laws making it a criminal offence e.g. USA Health Insurance Portability and Accountability (HIPAA) of 1996.
3. Introduce biometric authentications at provider sites for eligibility and membership to generate claims check codes.

*Exploring fraud and abuse in National Health Insurance Scheme (NHIS) using data mining technique as a statistical model*

*Kittoe, J. D. & Asiedu-Addo, S. K.*

4. A disbursement formula for all cost entries should be designed and a check mechanism put in place to ensure all NHIS providers adhere to that.

## References

Aggarwal, C. C. (2015). *Data Mining: The Textbook*. Switzerland: Springer International Publishing Company.

Centers for Disease Control and Prevention (2012). *Center for Global Health Division of Parasitic Diseases and Malaria*. Retrieved April 2, 2016 from: http://www.cdc.gov/malaria/resources/pdf/fsp/cdc_malaria_program.pdf.

Corporate Alliance on Malaria in Africa. (2011). Corporate Alliance on Malaria in Africa. Retrieved May 21, 2016 from http://www.gbchealth.org/wpcontent /uploads/2014/03/CAMA-Best-Practices-Interventions.pdf.

Fraenkel, J. R. & Wallen, N. E. (2003). *How to design and Evaluate Research in Education(5th ed)*. New York: McGraw-Hill,Inc.

Gee, J., Button, M., Brooks, G. & Vincke, P. (2010). *The financial cost of healthcare fraud (Internet)*. Retrieved April 13, 2016 from http://eprints.port.ac .uk/3987/1/The-Financial-Cost-of-Healthcare-Fraud-Final-(2).pdf.

Ghana Health Service (2016). *History of Health System in Ghana*. Retrieved May 7, 2016 from http://www.ghanahealthservice.org.

Ghana Statistical Service (2010). Population and Housing Census. Accra: Ghanaian Times

Gill, A.C. and J.E. Randall, (1994). Xenisthmus balius, a new species of fish from the Persian Gulf (Gobioidei: Xenisthmidae). *Proc. Biol. Soc. Wash*. 107(3):445-450.

Hanson, K., Gilson, L., Goodman, C., Mills, A., Smith, R . & Feachem R. (2008). Is Private Health Care the Answer to the Health Problems of the World's Poor? *PLoS Med.* 5(11): 23-33.

Jackson, J. (2002). Data Mining: A Conceptual Overview. *Communications of the Association for Information Systems.* 8: 267-296.

Mensah, S. A. (2011). NHIA Detects Massive Fraud in two Regions. *The Ghanaian Times*. Retrieved April 14, http://www.peacefmonline.com/pages/local/health/ 201401/186365.php. 2016 from Ministry of Health (2004). *Holistic Assessment of the Health Sector Programme of Work-2004.* Accra, Ghana.

Ministry of Health (2013). *Holistic Assessment of the Health Sector Programme of Work-2012.* Accra, Ghana

National Health Insurance Authority (NHIA) (2010). Annual Report.

National Health Insurance Authority (NHIA) (2013). *Annual report.* Retrieved April 20,  2016 from http://www.nhis.gov.gh/_Uploads/dbsAttachedFiles123%281%29.pdf.

National Malaria Control Programme (2009). *National Malaria Control Programme Annual Report 2009*. Retrieved May 7, 2016 from www.ghanahealthservice. org/ /ghana_malaria_programme review_final.

Rashidian, A., Joudaki, H. & Vian, T. (2012). *No Evidence of the Effect of the Interventions to Combat Health Care Fraud and Abuse: A Systematic Review of Literature.* Retrieved May 23, 2016 from http://journals.plos.org/plosone/article? id= 10.1371/journal.pone.

Statistical Analysis System (SAS) Institute. (2012). *Data Mining and the case for Sampling.* Retrieved June 30, 2016 from http://www.ag.unr.edu/gf/dm/sasdm.pdf.

Tomar, D. & Agarwal, S. (2013). A survey on Data Mining approaches for Healthcare. *International Journal of Bio-Science and Bio-Technology.* 5(5): 241-266.

United States Department of Health and Human Services (2014). *Medicare Fraud & Abuse Prevention, Detection, and Reporting.* Retrieved April 2, 2016 from www.cms.gov/Outreach-and-Education/Medicare-LearningNetworkMLN/MLNProducts/downloads/Fraud_and_Abuse.pdf.

Witter, S., Garshong, B. & Ridde, V. (2013). An exploratory study of the policy process and early implementation of the free NHIS coverage for pregnant women in Ghana. *International Journal for Equity in Health;* 12: 16-23.

World Health Organization (2006). *Malaria: WHO Technical Report Series*, (No. 936). Retrieved May 7, 2016 from www.who.int/ith/ITH_chapter_7.pdf.

Zaïane, O. R. (1999). *Principles of Knowledge Discovery in Databases*. Retrieved April 30, 2016 from http://www.exinfm.com/pdffiles/intro_dm.pdf.