British Journal of Applied Science & Technology
16(2): 1-15, 2016, Article no.BJAST.25753
ISSN: 2231-0843, NLM ID: 101664541

SCIENCEDOMAIN *international*
www.sciencedomain.org

Bayesian and Frequentist Comparison: An Application to Low Birth Weight Babies in Ghana

Michael Ofori Fosu[1*], Osborne A. Y. Jackson[2] and Stephen B. Twum[3]

[1]Department of Mathematics and Statistics, P.O.Box 854, Kumasi Polytechnic, Ghana.
[2]Methodist University College Accra, Ghana.
[3]Department of Mathematics, University for Development Studies, Navrongo, Ghana.

Authors' contributions

This work was carried out in collaboration among all authors. Author MOF conceived the idea, secured the data, carried out the coding, did the data analysis and drafted the manuscript. Author OAYJ assisted with the securing of the data and the study design and also did the final proofreading. Author SBT studied the literature, participated in the sequence alignment and also proofread the draft. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJAST/2016/25753
Editor(s):
(1) Qing-Wen Wang, Department of Mathematics, Shanghai University, P.R. China.
Reviewers:
(1) Loc Nguyen, Vietnam.
(2) Hui Li, Icahn Medicine School, USA.
Complete Peer review History: http://sciencedomain.org/review-history/14672

Original Research Article

Received 18th March 2016
Accepted 9th May 2016
Published 17th May 2016

ABSTRACT

The aim of this study is to evaluate the association between maternal factors and birth weight among babies by using and comparing frequentist and Bayesian methods' results from an epidemiologist or public health point of view. Low birth weight babies, defined by WHO as babies born at term who weigh less than 2.5 kg is an important indicator of reproductive health and general health status of any Population. The incidence of low birth weight is quite high in the sub region which has a public health concern.

Our study was based on data from 2011 Multiple Indicator Cluster Survey conducted by Ghana Statistical Service. A total sample size of 10,963 women within the reproductive age were selected throughout the entire country for the survey.

The results from the frequentist and the Bayesian models show that, the two approaches can yield similar results using same data set. However, there are factors that the Bayesian technique can unfold which might not be the case using the frequentist model. We were able to show with our

**Corresponding author: E-mail: mikeoffos@yahoo.com;*

data set that the Bayesian method may have a lot of benefits than the frequentist method. However, in order to narrow the credible intervals, there is the need to bring in informative priors so as to be able to well formulate the null and the alternative hypotheses. However, one can use the Markov Chain Monte Carlo, when using no priors to predict reliable results.
Comparing the two approaches with respect to our data set, we can infer (from Table 4) that using Bayesian model provides better estimates in predicting low birth weight among babies in Ghana. We note however that to better understand the phenomenon under study the two methods could be used together. Our findings further revealed that low birth weight is not only a public health problem but also a socio-cultural issue.

Keywords: Low birth weight; frequentist; Bayesian; informative priors.

1. INTRODUCTION

The debate over Bayesian versus Frequentist statistical inference is largely over in statistics community. Both Bayesian and frequentist ideas have a lot to offer practitioners. Each approach has a great deal to contribute to statistical analysis and each is essential for the full development of the other approach. Both methods often lead to the same solution when no external information (other than the data and the model itself) is introduced into the analysis. But these methods are not the same and do different things. It is very important to understand the assumptions behind the theories and to correctly interpret the mathematical conclusions produced. Using both approaches for an important problem is good in practice [1]. The union of frequentist and Bayesian procedures is discussed extensively by [2], and this study is based partly on their work. Frequentist methods are at times misapplied or used wrongly and sometimes the results are interpreted wrongly. Unlike the frequentist, the Bayesian methods provide many practical benefits, like handling variables that are not observed, small sample sizes as well as errors associated with measurement and using prior information from earlier works [3].

In general, frequentist methods are computationally relatively simple. There is no need for numerical integration. Many of these methods, for sufficiently large data sets, are the locally most powerful tests possible. In many cases the frequentist and Bayesian interpretations are different: Bayesian methods are based on *decision theoretic principles*; actions are dictated by risk management by minimising the expected loss under a chosen 'loss' function. Similar choices are needed in frequentist methodology to determine the optimal procedure (e.g. least squares or maximum likelihood estimation).

Frequentist principle is based on the idea that, with repeated use of a statistical procedure, the actual accuracy of the long run average should not be less than the reported accuracy of the long average. This is really a joint frequency-Bayesian principle.

Low birth weight (LBW) is one of the key reproductive health indicators whose outcome is influenced by consumption of reproductive health care. [4] argue that one of the key measures of child health is that of birth weight. Birth weight is a good gauge of health of the child in the womb because the weight is taken immediately after birth. Consequently, a malnourished fetus will be born at low birth weight. Fetal growth can be affected by factors such as maternal, genetic and environmental. One of the key vulnerable processes in the life cycle of human being is the growth and development of the intrauterine which can lead to a permanent profound influence later in life. According to [5], the intrauterine growth menace has been well studied by birth weight within the developing countries. The birth weight of a child who is less than one year could be used to assess his chances of good health and growth, development and survival [6].

LBW is prevalent in developing countries especially those in the Sub-Saharan region due to the high levels of malnutrition and infectious diseases. Throughout the world, over 20 million children under one year are born with low birth weight. Out of this number about 95% of these babies are located in developing countries. A large proportion of these babies are born in Asia which also have high parity rate and about 22% of them are also born in Africa. Within Asia, about 40% of the low birth weights occur in India with countries in the western world such as Sweden, having between 4% and 6% [7].

The birth weight of a child is a yardstick or a measure of how vulnerable the child may be susceptible to childhood diseases and chances of survival. Sub-Saharan Africa (SSA) has the second highest incidence of low birth weight infants the world over (16%), with South Central Asia being the highest at 27% [8]. The most recent evidence in Ghana shows that approximately 10% of all births are LBW [9]. In particular, the UN envisages a reduction of low birth weight by at least one-third in the proportion of infants. This target is in fact, one of the seven major goals for the current decade of the “A World Fit for Children” programme of the United Nations [10].

The prevalence of LBW situation in Ghana is not so different from what pertains in the sub region. The rate has been hovering around 10% according to the various results contained in the Multiple Indicators Cluster Questionnaire Surveys (MICS) and the Demographic and Health Surveys (DHS) conducted over the years. This includes only a few babies who are weighed at birth or described as being “very small” or “smaller than average” when born. The major challenge is that most babies born in Ghana are not weighed at birth due to the fact that most mothers give birth at home and not at health facilities. For instance, about 79% of babies born in Ghana were not weighed according to the 1998 DHS report (page; 98). Again, in the 2003 DHS, information on birth weight was known for only 28% babies in the five years preceding the survey and for the 2008 DHS, birth weight was reported for only 43% of births in the five years preceding the survey. The 2006 MICS report also indicates that, overall, nearly 2 in 5 babies are not weighed at birth and approximately 9% of infants are estimated to weigh less than 2.5 kg at birth. However, some research findings at various facilities across the country put the prevalence rate above 16% which is higher than the 15% global average threshold making it a public health concern as a country. Again, according to the WHO data published in April 2011 LBW deaths in Ghana reached 6,056 or 3.23% of total deaths in the country. Currently LBW is among the top 20 causes of deaths in Ghana.

The primary objective of our present study is to study and assess maternal factors associated with low birth weight babies by using and comparing both frequentist and Bayesian methods’ results from an epidemiologist or public health point of view.

The justification for the comparison of Bayesian and Frequentist in this study is based on *coherence* and *calibration*. These are two important goals for statistical inference;

i. Bayesian work has tendered to focus on coherence whilst Frequentist work has not been too worried about coherence. The problem with pure coherence is that one can be coherent and completely wrong.
ii. Frequentist work on the other hand tends to focus on calibration whilst Bayesian work has not been too worried about calibration. The problem with pure calibration is that one can be calibrated and completely useless.

Many statisticians therefore make use of both Bayesian perspective and Frequentist perspective, because a blend is often a natural way to achieve both coherent and calibration.

1.1 Conceptual Framework

Reasons for comparing these two approaches’ results are many, namely because frequentist techniques, even though the most universally and widely used, are sometimes considered as misused or misinterpreted [11-13], whilst Bayesian techniques are underused, however, they appear to present several practical advantages, such as accommodating small sample sizes, missing data, covariates measured with error, random effects or a hierarchical structure of variables, unobserved variables along with measurement errors and incorporating information from previous studies [11-14].

Dunson et al. [14] defines Bayesian methods as the explicit quantitative use of external evidence in the design, monitoring, analysis, interpretation and reporting of a study. According to [13], results of epidemiological observational studies provide a likelihood that can be combined with prior information using standard and advanced full Bayesian methods leading to purely probabilistic results. The specific application of Bayesian methods to case-control studies is both feasible and useful, particularly with the advance of advanced computational methods [15,16] coupled with the advent of Markov Chain Monte Carlo (MCMC) methods, Bayesian methods are being implemented with increasing frequency. MCMC methods are computer-intensive technical methods that allow one to simulate draws from the posterior distribution, without having to calculate the posterior distribution [17].

Again, contrary to the classical confidence interval, Bayesian methods offer Bayesian credible interval which has a simple appealing interpretation as the interval containing the true parameter of interest with some probability (e.g., 95%). Most researchers prefer this easy interpretation to that of the classical 95% confidence interval, which is the range of values containing the true parameter 95% of the times in repeated sampling [18]. Furthermore, people sometimes wrongly interpret the confidence interval as if it was a credible interval. This means that Bayesian approach seems more intuitive than the frequentist.

Even though these practical advantages have long been available for epidemiologists, few epidemiological studies have used this powerful tool to assess exposure-disease relations [18]. Controversies are raised by the Bayesian approach, since one is compelled to re-examine fundamental notions about the concept of probability and classical statistical practices [19]. Its usefulness is accepted in specific situations such as in case of sequential data, but its system of inference using priors in other situations is still controversial [19] because of its subjectivity [11]. Therefore, comparing results, interpretations and limitations of both approaches would enrich our discussion and make our conclusions more robust.

2. DATA

Our study was based on data from 2011 Multiple Indicator Cluster Survey (MICS) conducted by Ghana Statistical Service. A total sample size of 10,963 women within the reproductive age were selected throughout the entire country for the survey. All responses were solicited two years preceding the survey. The process for sample selection was the results of representative probability sampling of households conducted nationwide using the 2010 Population and Housing Census Enumeration Areas (EA's). To be able to compare, the MICS used an internationally standardized sampling of two-stratified sample design. The first stage involved selection of a number of EA's from the ten regions which were used as clusters. The second stage involved the selection of households in each region using systematic sampling with probability proportional to size. In all 12,150 households were selected for the sample, however, 11,925 households were duly contacted and interviewed. From the interviewed households, 10,963 women within the reproductive age group were identified for interview.

3. METHODOLOGY

The 2011 MICS was conducted with a sample of 11,925 households from a selected household of 11,970 throughout the ten administrative regions of Ghana translating into about 100% response rate. All the households were selected based on the sizes of the regions. The survey utilized both quantitative and qualitative data collection methods which aimed at providing basic data for measuring the progress of women and children in Ghana. The data set used for analysis in this study is based on information on all births and deaths that occurred two years preceding the survey. The study used SPPS (version 20), R-console and SAS system version 9.4 for extraction and data analysis. Descriptive statistics as well as frequencies of the background characteristics of the mothers and the respective households the children belong to were generated. The association between the dependent and independent variables was ascertained using chi-square analysis procedures. Our dependent variable was based on the outcome of the weight of the baby whether LBW or normal. The independent variables included parity, area of residence, ethnicity region, antenatal care, economics status and mothers' characteristics including; education, religion and age. A critical level of significance of 5 percent ($p<0.05$) was employed to identify the most statistically significant factors of LBW babies. Estimates of LBW were also obtained for the overall administrative regions.

3.1 Model Specification (Logistic Regression)

The following generalized linear logistic model was used

$$\pi = \log\left(\frac{u}{1-u}\right) = \chi\beta + \varepsilon \quad (1)$$

where π links the linear function to $\log\left(\frac{u}{1-u}\right)$. The link is not a linear function, μ is the probability of having a LBW baby and χ is the model matrix including, mothers' educational level, age, religion, ethnicity, antenatal care, economic status of household and parity. The matrix also includes geographical location, such as region of origin and whether the respondent is from rural or urban environment; β is the vector of parameters, and ε is the vector of residuals. We applied the Fisher scoring method (SAS, 2007) to obtain Maximum Likelihood estimates of β. The overall goodness of fit was derived from the Likelihood Ratio Test of the hypothesis $H_{0:}$ c(β) = 0 where a comparison is made between the full model and the reduced model [20]. It is therefore a test for global null hypothesis of the elements of the solution vector.

The odds ratio in this study is the probability that a child will be born of LBW to the probability that the child born has normal birth weight. This means that the outcome variables in the logistic regression should be discrete and dichotomous. Logistic regression was therefore found fit to be used because the outcome variable was in binary form that is a child is born with LBW or otherwise. Again, there were no assumptions to be made about the distributions of the explanatory variables as they did not have to be linear or equal in variance within the group. The model suggests that the likelihood of a woman giving birth to a LBW child varies across all the independent variables to be studied. After fitting the model, the outcomes were used to interpret the existing relationships between LBW babies, household structure and mothers' characteristics.

3.2 Markov Chain Monte Carlo

In drawing an $i.i.d$ sample from a complicated distribution π is difficult, especially in high dimensions. Markov chain Monte Carlo generates a sequence of random variables (X^t) which are dependent and such that the distribution of X^t converges weakly to π as $t \to 1$. Estimation of $\int h(x)\pi(x)dx$ is still based on a law of large numbers, but now for dependent random variables:

$$\int h(x)\pi(x)dx \approx \overline{H}_{N,r} = \frac{1}{N-1}\sum_{t=r+1}^{N} h(X^t). \qquad (2)$$

Here *r* is a "burn-in" period which discards values X^t whose distribution is too far from the target π.

The random variables are constructed recursively: The initial value X^0 is arbitrary, and for each $t \geq 1$, X^t is a deterministic function of X^{t-1} and a uniform random variable U^twhich is independent of $X^0, \ldots, X^{t-1}$

$$X^t = G(X^{t-1}, U^t). \qquad (3)$$

(In practice, often several uniform variables $U^{t,1}, U^{t,2}, \ldots, U^{t,k}$ are used, but this is equivalent).

Because the dependence of X^t on previous random variables is only via X^{t-1}, the sequence (X^t) is called a Markov chain. The conditional distribution of X^t given X^{t-1} is called the transition kernel *P* of the chain

$$\mathbb{P}(X^t \in A|X^0, \ldots, X^{t-1}) = \mathbb{P}(X^t \in A|X^{t-1}) = P(X^{t-1}, A). \qquad (4)$$

It is determined by the function *G* through

$$P(x, A) = \mathbb{P}(G(x, U) \in A) = \mathbb{P}(U \in \{u: G(x, u) \in A\}). \qquad (5)$$

In particular, *P* does not depend on *t* because *G* is the same for all *t*. We therefore call the Markov chain time-homogeneous. In Markov process theory, one usually starts by specifying the transition kernel $P(x, A)$, the conditional probability that the next value of the chain is in *A* given that the current value is equal to *x*. It is always possible to construct a function *G* such that the above equation is satisfied. Because for Markov chain Monte Carlo, we need to draw from $P(x,.)$ for arbitrary values *x*, it is more natural to start with the concrete construction

$$X^t = G(X^{t-1}, U^t). \qquad (6)$$

In order to use Markov chain Monte Carlo to estimate expected values with respect to the target π, we need to find a transition kernel *P* such that we can draw from the distribution $P(x,.)$ for any *x* and such that for $N \to \infty$ 1 the arithmetic mean of the $h(X^t)$ converges to $\int h(x)\pi(x)dx$. The general theory of Markov chains shows that the second requirement holds in a wide range of cases if the chain can reach all sets *A* with $\pi(A) > 0$ and if $X^{t-1} \sim \pi$ implies that $X^t \sim \pi$. If the second condition holds, we call π an invariant or stationary distribution for the transition kernel *P*. Because

$$\mathbb{P}(X^t \in A) = \mathbb{E}(\mathbb{P}(X^t \in A|\, X^{t-1})) = \mathbb{E}(\mathbb{P}(X^{t-1}, A)),$$

π is stationary for *P* if

$$\pi(A) = \int \pi(x)P(x, A)dx \quad \forall A,$$

or, in the case where $P(x,.)$ has the density $p(x, y)$, if

$$\pi(y) = \int \pi(x)p(x, y)dx. \qquad (7)$$

There are two basic recipes for constructing a transition kernel *P* which has a given target distribution π as stationary distribution. The first one is the so-called Gibbs sampler. For this we assume that $x \in \mathbb{R}^{\mathbb{p}}$ and we denote the conditional density of the $i-th$ component of x, x_i given all the other components $x, x_{-i} = (x_j)_{j \neq i}$ by π_i:

$$\pi_i(x_i|x_{-i}) \propto \pi(x)$$

where $\propto$ means up to a term which does not contain x_i. This means that we can identify π_i by inspecting how the target density π depends on the *i-th* component. We don't need any integration. The densities π_i are also called \full conditionals" (because we condition on all other components). The Gibbs sampler depends on a "visiting schedule" $i_t \in \{1,2,\dots,p\}$ and iterates the following steps for $t = 1,2,\dots$

$$X_{i_t}^t \sim \pi_{i_t}\left(x_{i_t} \middle| X_{-i_t}^{t-1}\right) dx_{i_t},\ \ X_{-i_t}^t = X_{-i_t}^{t-1} \tag{8}$$

In words, we leave all components of X^{t-1} except the one that is actually visited unchanged, and we update the visited component according to the conditional distribution of our target. By the definition of the conditional distribution, π is invariant for this transition kernel. The visiting schedule can be either deterministic or it can randomly select one of the components. In order that the chain can reach all sets, we have to visit each possible component infinitely often.

The Gibbs sampler requires that we can sample from the full conditionals. Because these distributions are one-dimensional, this is often possible. If it isn't, we can use the Metropolis-Hastings algorithm instead. It is based on the fact that any reversible distribution is also stationary. Here π is called reversible for the transition kernel *P* if

$$\int_A \pi(x)P(x,B)dx = \int_B \pi(x)P(x,A)dx, \forall A,B, \tag{9}$$

or in words, if $X^t \sim \pi$, then

$$\mathbb{P}(X^t \in A, X^{t+1} \in B) = \mathbb{P}(X^{t+1} \in A, X^t \in B)\ \ \forall A,B. \tag{10}$$

Choosing for *B* the whole space $\mathbb{R}^{\mathbb{p}}$, it follows that a reversible distribution π is stationary.

If $P(x,.)$ has the density $P(x,y)$ for any *x*, then reversibility is equivalent to

$$\pi(x)p(x,y) = \pi(y)p(y,x) \quad \forall x,y. \tag{11}$$

For any pair $x \neq y$, we can therefore choose one of the two values $P(x,y)$ and $P(y,x)$ arbitrarily, whereas the other one is determined by the reversibility equation. However, a solution obtained in this way does in general not satisfy $\int p(x,y)dy = 1$ for any *x* and thus is not the density of a transition kernel. To solve this problem, one can start with an arbitrary transition density *q* and then choose from the two possible solutions

$$p(x,y) = q(x,y), p(y,x) = \frac{\pi(x)q(x,y)}{\pi(y)}$$

and

$$p(x,y) = \frac{\pi(y)q(y,x)}{\pi(x)}, p(y,x) = q(y,x) \tag{12}$$

the one which satisfies $p(x,y) \leq q(x,y)$ and $p(y,x) \leq q(y,x)$ for any $x \neq y$. This solution can be written in the compact form,

$$p(x,y) = q(x,y)\min\left(1, \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}\right). \tag{13}$$

It follows that $\int p(x,y)dy \leq \int q(x,y)dy = 1$ for any *x*, and one can put the missing mass on the diagonal, meaning that the chain does not move. Written in formulae, the transition kernel is given as;

$$p(x,A) = \int_A p(x,y)dy + 1_A(x)\left(1 - \int p(x,y)dy\right) \tag{14}$$

Assuming that we can simulate from the transition density $q(x,.)$ for any *x*, the following algorithm generates a Markov chain with the transition kernel *P*:

- At time *t*, generate $Y^t \sim q(X^{t-1}, x)dx =$ and $U^t \sim uniform\,(0,1)$, independently from each other and independently of previously generated variables.
- Set

$$X^t = \begin{cases} Y^t & if \quad U^t \leq min\left(1, \frac{\pi(Y^t)q(Y^t, X^{t-1})}{\pi(X^t)q(X^{t-1}, Y^t)}\right) \\ X^{t-1} & else \end{cases} \tag{15}$$

This is similar to the accept-or-reject method, but the proposal depends on the most recent value, and in case of a rejection, we do not move. The simplest choice of $q(x,.)$ is a normal density with mean *x* and an arbitrary covariance matrix $\sum$. In this case, $q(x,x') = q(x',x)$ so that the acceptance probability is simply $\min(1, \pi(x')/\pi(x))$. This means moving to value which is more likely than the current value is always accepted whereas the acceptance of a move to a less likely value is given by the likelihood ratio. The algorithm for a symmetric *q* is due to [21] whereas the general case is due to [22].

In general, the posterior distribution, or any of its summary measures, can only be obtained analytically for a restricted set of relatively simple models. Thus, for a long time, researchers could only proceed easily with Bayesian inference when the posterior was available in closed-form or as a (possibly approximate) analytic expression. As a result, practitioners interested in models of realistic complexity did not much use Bayesian inference. This situation changed dramatically with the advent of computer-driven sampling methodology, generally known as Markov chain Monte Carlo (MCMC: e.g., [23,17]) Using MCMC techniques such as Gibbs sampling or the Metropolis–Hastings algorithm, researchers can directly sample sequences of values from the posterior distribution of interest, forgoing the need for closed-form analytic solutions. The current adage is that *Bayesian models are limited only by the user's imagination.*

3.2.1 Gibbs sampling

The formal algorithm can be specified as follows. Let $\boldsymbol{\theta}$ be a vector of model parameters with elements $\theta = \{\theta_1, \ldots, \theta_q\}$. The elements of $\boldsymbol{\theta}$ could be the parameters of a regression model, structural equation model, and so forth. Note that information regarding $\boldsymbol{\theta}$ is contained in the prior distribution *p(*$\boldsymbol{\theta}$*)*. A number of algorithms and software programs are available to conduct MCMC sampling. Following the description given in [17], the Gibbs sampler begins with an initial set of starting values for the parameters, denoted as $\theta^{(0)} = (\theta_1{}^{(0)}, \ldots, \theta_q{}^{(0)})$. Given this starting point, the Gibbs sampler generates $\theta^{(s)}$ from $\theta^{(s+1)}$ as follows:

1. sample
   $\theta_1{}^{(s)} \sim p(\theta_1 | \theta_2{}^{(s-1)}, \theta_3{}^{(s-1)}, \ldots, \theta_q{}^{(s-1)}, y)$

2. sample
   $\theta_2{}^{(s)} \sim p(\theta_2 | \theta_1{}^{(s-1)}, \theta_3{}^{(s-1)}, \ldots, \theta_q{}^{(s-1)}, y)$
   q sample
   $\theta_q{}^{(s)} \sim p(\theta_q | \theta_1{}^{(s-1)}, \theta_3{}^{(s-1)}, \ldots, \theta_{q-1}{}^{(s-1)}, y)$

Then, a sequence of dependent vectors are formed:

$$\theta^{(1)} = (\theta_1{}^{(1)}, \ldots, \theta_q{}^{(1)})$$
$$\theta^{(2)} = (\theta_1{}^{(2)}, \ldots, \theta_q{}^{(2)})$$
$$\theta^{(s)} = (\theta_1{}^{(s)}, \ldots, \theta_q{}^{(s)})$$

This sequence exhibits the so-called *Markov property* insofar as $\theta^{(s)}$ is conditionally independent of $\{\theta_1{}^{(0)}, \ldots, \theta_q{}^{(s-2)}\}$ given $\theta^{(s-1)}$. Under some general conditions, the sampling distribution resulting from this sequence will converge to the target distribution as $s \to \infty$. (*See* [17]) for additional details on the properties of MCMC. In setting up the Gibbs sampler, a decision must be made regarding the number of Markov chains to be generated, as well as the number of iterations of the sampler. With regard to the number of chains to be generated, it is not uncommon to specify multiple chains. Each chain samples from another location of the posterior distribution based on purposefully dispersed starting values. With multiple chains, it may be the case that fewer iterations are required, particularly if there is evidence for the chains converging to the same posterior mean for each parameter. In some cases, the same result can be obtained from one chain, although often requiring a considerably larger number of iterations. Once the chain has stabilized, the iterations prior to the stabilization (referred to as the *burn-in* phase) are discarded. Summary statistics, including the posterior mean, mode, standard deviation, and credibility intervals, are calculated on the post-burn-in iterations. Also, convergence diagnostics are obtained on the entire chain or on post-burn-in iterations.

The Bayesian analysis was performed using models identical to those used in the frequentist analyses. Non informative priors were used in this study. The MCMC method was then applied to derive posterior distributions of the model parameters, using SAS 9.4 and R-Console statistical software packages. Means, 95% credible intervals, and probability of the odds ratios were presented. We also indicatively presented the Deviance Information Criterion, a model goodness of fit measure that usually allows comparing models on the same data set: a lower DIC generally indicates a model that fits better to data [13]. Interpretation of both frequentist and Bayesian analyses were presented, stressing on differences, similarities and complementarities between both methods

4. RESULTS

There were 2873 births registered within the survey period. Out of this figure, 1336 were weighed at birth which is about 46.5%. The LBW incidence found in this study was 13.7% from our sample of no-missing weights. Table 1 gives the description of the various categories in the study. Five regions; Western, Volta, Greater Accra, Brong Ahafo and Eastern all recorded rates lower than the national figure of 13.7%. Women from Central region are more likely to give birth to low birth weight children (about 16%) and those from Western and Volta regions the least likely to give birth to low birth weight children (4.4%). Low birth weight is also predominant among children born in the three northern regions of the country. Women who come from the poorest households, and those who stay in rural households or have up to middle school education are more likely than more advantaged women to give birth to LBW children. For instance, the proportion of LBW among women who have up to middle school education is 76.1%, versus 23.9% for women who have a at least secondary school education. Women in urban households are likely to give birth to children of normal birth weight compared to those in rural households. Women from wealthiest households are more likely to give birth to normal weight children compared to women from poorest households. The probability of giving birth to children of low birth weight among first time mothers and those who have at least four children is higher than second time mothers. Again, women who are at most 24 years or above 35 years have highest proportion of children whose weight is below 2.5 kg. Tables 2 and 3 show the results of our two models (multivariate logistic regression model and Bayesian posteriors model) of maternal factors associated with LBW. The factors observed to be highly significantly associated with LBW included maternal age and children ever born (parity) with with parity^8 being highgly significant. The Bayesian model also show some significance in ANC. The results from (Fig. 1) also show the density plots and the trace plots for the parameters. The density plots portray the credible intervals of the parameters and the trace plots were used to test the convergence of the model (MCMC). From the figure, it is clear that antenatal care, age squared and parity especially parity eight all passed the convergence test, giving credence to the models' validity.

Table 4 shows the results of both frequentist and Bayesian models. We compared the two results based on the odds ratio, standard errors and credible intervals/confidence intervals. From the table, the odds ratio that fall within the Confidence intervals also fall in the Credidible intervals. Generally however, the odds ratio in the Bayesian are higher than that of the frequentist especially whre the parameters are significant. The standard errors in the Bayesian model are also lower compared with the frequentist model. This means that the Bayesian model can be more relied upon than the frequentist model. The credible intervals appear to be more spread than the confidence limits.

5. DISCUSSION

The 13.7% incidence of low birth weight (mean = 2.10) and the normal mean birth weight of 4.012 0.062 kg found in our study can be compared to other studies in the developing world (e.g., Nigeria, which is about 14%). The challenge however is that few mothers in Ghana give birth at health facilities and hence their babies are not weighed at birth. Over 50% of children born in Ghana are not weighed. About 46.5% of the children who were born in the survey period were weighed at birth (1336 of the 2873 births). The descriptive statistics show that locality (residence) has impact on the weight of the new born. Mothers in urban areas tend to give birth to normal weight children than those who live in rural areas.

This is also evidenced by [24]. More so a case control study by [25] to determine the risk factors for LBW in Nagpur city of Maharashtra also found place of residence (rural) to be associated with LBW.

The factors found to be significant using the frequentist models include age and parity. All the other variables like residence, ANC, Region, educational levels and economic status were not significant. This is similar to results obtained by [26,27] whose work revealed educational levels as not being a risk factor in predicting low birth weight. The findings however, is in sharp contrast to [28] who found both age and parity not to be significant to low birth weight in their study of a number of maternal factors including; birth spacing, height, pre-delivery weight and pregnancy weight gain, age, parity, educational level, economic status, ANC, maternal exposure to tobacco and hypertension anaemic.

The Bayesian model on the other hand using same variables as the frequentist model produced age, parity and ANC to be risk factors associated with low birth weight. Again, all the other factors like residence, Region and educational levels considered in the model are not significant. This stands to reason that using Bayesian approach may unveil additional risk factor(s) in predicting low birth weight which might not come out clearly when using the frequentist approach. However the high risk for mothers bearing their eighth child found from both models must be a source of worry.

Comparing the two models we find that the errors associated with the Bayesian model are much lower than the frequentist model; an indication that the Bayesian model may predict better even with non informative priors. This corroborates with the study conducted by [3] who also in comparing the two methods found the Bayesian to perform better.

Table 1. Relationship between socio-demographic characteristics of mothers and LBW

Indicator	LBW (< 2.5 kg)		Normal (≥ 2.5 kg)		Total	
	N	(%)	N	(%)	N	(%)
Baby's weight						
	183	(13.7)	1153	(86.3)	1336	(100)
Mothers age						
≤24	55	(4.1)	232	(17.4)	287	(21.3)
25 – 34	88	(6.5)	609	(45.6)	697	(52.1)
35+	40	(2.8)	312	(23.2)	352	(26.0)
Total	**183**	**13.7**	**1153**	**(86.4)**	**1336**	**(100)**
Antenatal care						
Attended	182	(13.6)	1147	(85.9)	1329	(99.5)
Not attended	1	(0.1)	6	(0.4)	7	(0.5)
Total	**183**	**(13.7)**	**1153**	**(86.3)**	**1336**	**(100)**
Area/location						
Urban	85	(6.4)	511	(38.2)	569	(44.6)
Rural	98	(7.3)	642	(48.1)	740	(55.4)
Total	**183**	**(13.7)**	**1153**	**(86.3)**	**1336**	**(100)**
Children ever born						
1	65	(4.9)	246	(18.4)	311	(23.3)
2	34	(2.5)	199	(14.9)	233	(17.4)
3	29	(2.2)	203	(15.2)	232	(17.4)
≥4	55	(4.1)	504	(37.9)	559	(42.0)
Total	**183**	**(13.7)**	**1153**	**(86.3)**	**1336**	**(100)**
Wealth index quintiles (Economic status)						
Poorest	46	(3.4)	319	(23.9)	365	(27.3)
Second	35	(2.6)	229	(17.1)	264	(19.8)
Middle	40	(3.0)	185	(13.8)	225	(16.8)
Fourth	34	(2.5)	209	(15.6)	243	(18.2)
Richest	28	(2.1)	211	(15.8)	239	(17.9)
Total	**183**	**(13.7)**	**1153**	**(86.3)**	**1336**	**(100)**
Mother's education						
Pre school	0	(0.0)	2	(0.2)	2	(0.2)
Primary	44	(4.9)	227	(25.4)	271	(30.3)
Middle	64	(7.2)	347	(38.8)	411	(46.0)
Secondary+	34	(3.8)	176	(19.6)	210	(23.4)
Total	**142**	**(15.9)**	**752**	**(84.1)**	**894**	**(100)**
Region						
Western	8	(0.6)	81	(6.1)	89	(6.7)
Central	30	(2.2)	161	(12.1)	191	(14.3)
Greater Accra	14	(1.0)	123	(9.2)	137	(10.3)
Volta	8	(0.6)	64	(4.8)	72	(5.4)
Eastern	11	(0.8)	75	(5.6)	86	(6.4)
Ashanti	19	(1.4)	96	(7.2)	115	(8.6)
Brong Ahafo	13	(1.0)	71	(5.3)	84	(6.3)
Northern	22	(1.6)	159	(11.9)	181	(13.5)
Upper East	29	(2.2)	157	(11.8)	186	(13.9)
Upper West	29	(2.2)	166	(12.4)	195	(14.6)
Total	**183**	**(13.7)**	**1153**	**(86.3)**	**1336**	**(100)**

Table 2. Logistic regression model

Parameter	**Estimate (β)**	**Std. error**	**z-value**	**Pr(>|z|)**	**95% confidence intervals**		**Odds ratio**	**MSE**
Intercept	0.085417	1.379637	0.062	0.95063	-2.5697490	2.8503370	1.089171	1.9033700
Age	0.088379	0.097531	0.923	0.35590	-0.1034452	0.2726374	1.092402	0.00951215
Age sqd.	-0.00156	0.001575	-0.992	0.32107	-0.0045774	0.0016117	0.998441	0.00000024
Parity 2	0.342182	0.244936	1.397	0.16241	-0.1334970	0.8290040	1.408017	0.0599927
Parity 3*	0.607238	0.281534	2.157	0.03101	0.0611689	1.1677160	1.835355	0.0792600
Parity 4*	0.783020	0.333863	2.345	0.01901	0.1402631	1.4530570	2.188070	0.1114620
Parity 5*	1.179091	0.402977	2.926	0.00343	0.4119609	2.0000360	3.251417	0.1623880
Parity 6*	1.181578	0.498824	2.369	0.01785	0.2480639	2.2228760	3.259514	0.2488216
Parity 7	0.654110	0.493297	1.326	0.18484	-0.2915003	1.6513880	1.923429	0.2433382
Parity 8**	2.602640	1.086394	2.396	0.01659	0.8607378	5.5521170	13.49933	1.1802343
Parity 9	0.224840	0.703290	0.320	0.74920	-1.0966773	1.7057604	1.252122	0.4940094
Parity 10*	1.820419	1.157665	1.572	0.11584	-0.1206060	4.8441809	6.174445	1.3401682
ANC 2	-0.17949	1.105634	-0.162	0.87104	-2.0207868	2.7847656	0.835696	1.2224080
Locality 2	0.033159	0.202667	0.164	0.87004	-0.3639370	0.4314230	1.033715	0.0410733
Economic status 2	-0.02504	0.248776	-0.101	0.91984	-0.5103671	0.4674914	0.975271	0.0618880
Economic status 3	-0.22517	0.260080	-0.866	0.38661	-0.7333866	0.2883160	0.79838	0.0676405
Economic status 4	0.103000	0.285934	0.360	0.71868	-0.4529114	0.6702375	1.108491	0.08175702
Economic status 5	0.360045	0.326952	1.101	0.27080	-0.2758790	1.0079452	1.433394	0.1068960

Table 3. Bayesian posterior model

Parameter	Mean (β)	Std. dev.	Odds ratio	95% Bayesian credible interval		MSE	Std. error
				Lower	Upper		
Intercept	-4.514712	5.3052	0.010947	-14.49495	6.1892	0.0211000	0.14530
Age	-0.216483	0.1887	0.805346	-0.57182	0.1685	0.0000267	0.00517
Age square*	2.256857	2.0140	9.553017	-1.80848	6.0158	0.0030400	0.05516
Parity 2	0.363909	0.2562	1.438943	-0.13795	0.8831	0.0000492	0.00701
Parity 3	0.619332	0.2946	1.857687	0.06318	1.2127	0.0000651	0.00806
Parity 4	0.827847	0.3352	2.288387	0.17143	1.4874	0.0000842	0.00918
Parity 5	1.249233	0.4146	3.487667	0.44757	2.0626	0.0001280	0.01130
Parity 6	1.259043	0.5216	3.522049	0.26459	2.3289	0.0002040	0.01420
Parity 7	0.711070	0.4961	2.036169	-0.23461	1.6805	0.0001840	0.01350
Parity 8**	3.168589	1.3237	23.77392	1.12369	6.2351	0.0013100	0.03620
Parity 9	0.347425	0.7295	1.415418	-1.03334	1.8242	0.0003990	0.01990
Parity 10	2.461445	1.3712	11.72174	0.14548	6.0874	0.0014100	0.03750
ANC 2	0.255195	1.3712	1.290713	-1.92678	3.4620	0.0014100	0.03750
Locality 2	0.024328	0.2045	1.024626	-0.37678	0.4290	0.0000313	0.00560
Economic status2	-0.007232	0.2531	0.992794	-0.50384	0.5040	0.0000480	0.00693
Economic status3	-0.230446	0.2618	0.794179	-0.75038	0.2933	0.0000514	0.00717
Economic status4	0.096833	0.2888	1.101676	-0.46341	0.6782	0.0000625	0.00791
Economic status5	0.400218	0.3309	1.492150	-0.25450	1.0546	0.0000821	0.00906

Table 4. Comparing results of logistic regression and Bayesian models

	Frequentist analysis (Logistic regression)						Bayesian posterior analysis (Non-informative priors)					
Parameter	**Mean (β)**	**S.E**	**S.D**	**Odds ratio**	**95% confidence intervals**		**Mean (β)**	**S.E**	**S.D**	**Odds ratio**	**95% credible intervals**	
Intercept	0.0854	1.3796	50.37	1.0892	-2.57	2.85	-4.515	0.1453	5.31	0.0109	-14.49	6.19
Age	0.0884	0.0975	3.561	1.0924	-0.10	0.27	-0.216	0.0052	0.19	0.8053	-0.57	0.17
Age sqd.*	-0.002	0.0016	0.058	0.9984	-0.00	0.00	2.2569	0.0552	2.01	9.5530	-1.81	6.02
Parity 2	0.3422	0.2449	8.943	1.4080	-0.13	0.83	0.3639	0.0070	0.26	1.4389	-0.14	0.88
Parity 3	0.6072	0.2815	10.28	1.8354	0.06	1.17	0.6193	0.0081	0.29	1.8577	0.06	1.21
Parity 4*	0.7830	0.3339	12.19	2.1881	0.14	1.45	0.8278	0.0092	0.34	2.2884	0.17	1.49
Parity 5*	1.1791	0.4030	14.71	3.2514	0.41	2.00	1.2492	0.0113	0.41	3.4877	0.45	2.06
Parity 6*	1.1816	0.4988	18.21	3.2595	0.25	2.22	1.2590	0.0142	0.52	3.5220	0.26	2.33
Parity 7*	0.6541	0.3933	18.01	1.9234	-0.29	1.65	0.7111	0.0135	0.50	2.0362	-0.23	1.68
Parity 8**	2.6026	1.0864	39.66	13.499	0.86	5.55	3.1686	0.0362	1.32	23.774	1.12	6.24
Parity 9	0.2248	0.7032	25.68	1.2521	-0.10	1.71	0.3474	0.0199	0.73	1.4154	-1.03	1.82
Parity 10*	1.8204	1.1577	42.27	6.1744	-0.12	4.84	2.4614	0.0375	1.37	11.722	0.15	6.09
ANC 2	-0.179	1.1056	40.37	0.8357	-2.02	2.78	0.2552	0.0375	1.37	1.2907	-1.93	3.46
Locality 2	0.0332	0.2027	7.399	1.0337	-0.36	0.43	0.0243	0.0056	0.20	1.0246	-0.38	0.43
Economic status 2	-0.250	0.2488	9.083	0.9753	-0.51	0.47	-0.007	0.0069	0.25	0.9928	-0.50	0.50
Economic status 3	-0.225	0.2601	9.496	0.7984	-0.73	0.29	-0.230	0.0072	0.26	0.7942	-0.75	0.29
Economic status 4	0.1030	0.2859	10.44	1.1085	-0.45	0.67	0.0968	0.0079	0.29	1.1017	-0.46	0.68
Economic status 5*	0.3600	0.3270	11.94	1.4334	-0.28	1.01	0.4002	0.0091	0.33	1.4922	-0.25	1.05

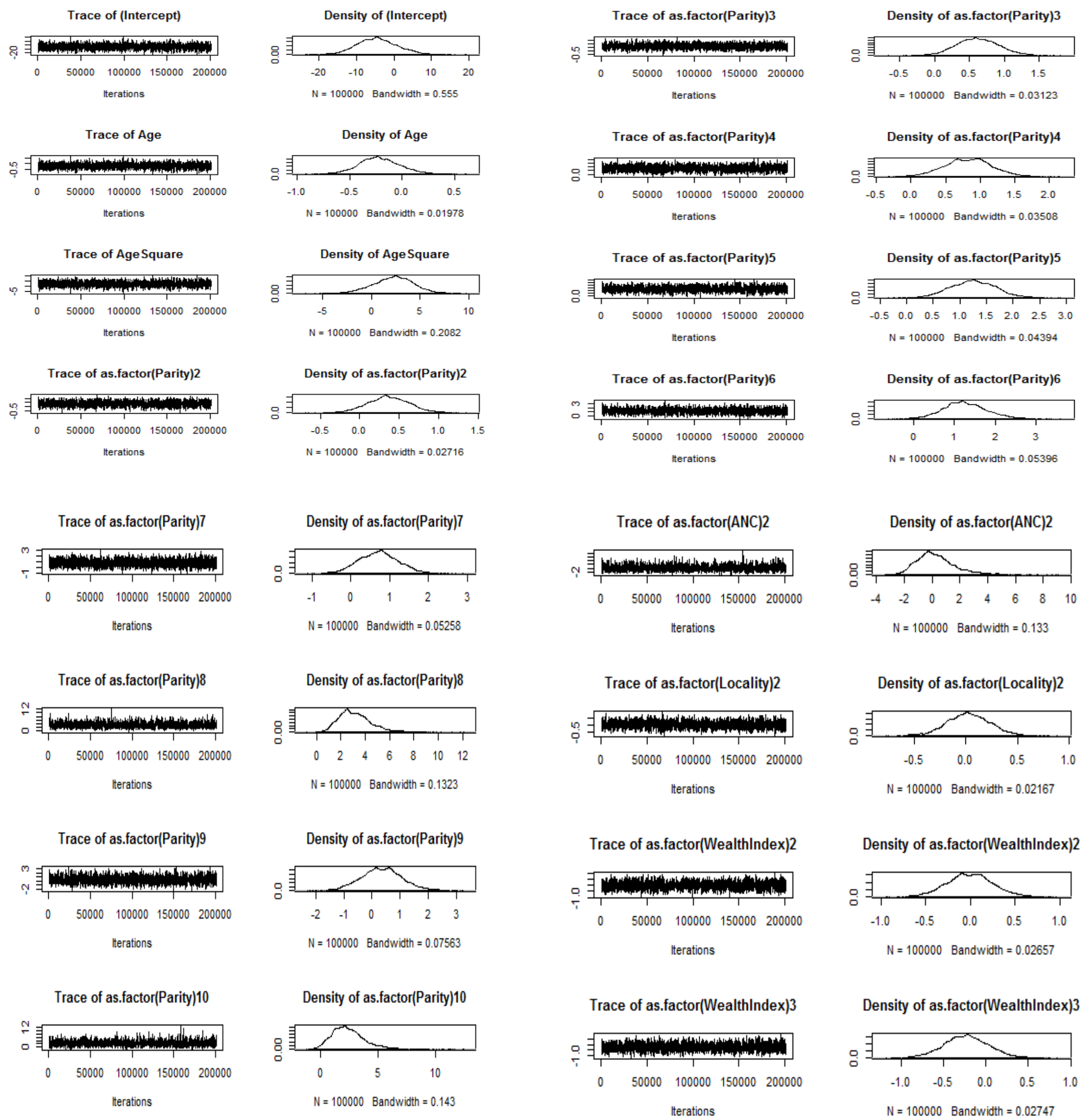

Fig. 1. Density plots and trace plots of the posteriors

6. CONCLUSION

The results from the frequentist and the Bayesian models show that, the two approaches can yield similar results using same data set. However, there are factors that the Bayesian technique can unfold which might not be the case using the frequentist model. With respect to our data set, using the Stepwise approach for the classical methods, antenatal care (ANC) was never found to be significant yet it came out clearly as one of the significant factors in the Bayesian analysis. This stands to reason that the Bayesian models could reveal something that the usual frequentist models could not. We were able to show that the Bayesian method may have several benefits over the frequentist one, particularly with respect to our data. The inclusion of informative priors might however be useful in narrowing the gap of credible interval and provide precise choice between the null and alternative hypothesis. In case of borderline frequentist results however, the MCMC method may be more conservative, particularly without priors.

Comparing the two approaches with respect to our data set, we can infer (from Table 4) that using Bayesian model provides better estimates in predicting LBW among babies in Ghana. We note however that to better understand the phenomenon under study the two methods could be used together. Our findings further revealed that LBW is not only a public health problem but also a socio-cultural issue.

The results of this study indicate that for reducing the incidence of LBW, the measures need to focus attention on nutrition education to facilitate better weight gain during pregnancy focusing more on the girl-child education, regular antenatal care visits and discouraging teenage and old age pregnancy, bearing more children as well as formulating policies that will reduce poverty among rural women. The girl child education policy must also be given all the needed resources it requires in other to achieve the desired set targets.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Berger J. The case for objective Bayesian analysis; 2006.
2. Bayarri MJ, Berger JO. The interplay of Bayesian and frequentist analysis. Statistical Science. 2004;19(1):58-80.
3. Pascale S, Mirna W, Georges K, Michèle D. Bayesian and frequentist comparison for epidemiologists: A non mathematical application on Logistic Regressions; 2014.
4. Rosenzweig, MR, Schultz TP. Estimatinga household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. Journal of Political Economy. 1983;91:723–746.
5. Sachdev HPS. Low birth weight in South Asia. Int J Diab Dev Countries. 2001;21(1):13-33.
6. Park K. Text book of preventive and social medicine. 19th edition. Jabalpur Banarsidas Bhanot Publishers. 2007;426-428.
7. UNICEF, 'Social Monitor 2003'. UNICEF Innocenti Research Centre, Florence; 2003. Available:http://www.uniceficdc.org/publications
8. UNICEF and WHO. Low birth weight; Country, Regional and Global estimates; 2004.
9. Ghana Statistical Service. Multiple Indicator Cluster Survey; GSS; 2011.
10. United Nations. Progress for children. A World Fit for Children Statistical Review; 2004.
11. Greenland S. Bayesian perspectives for epidemiological research: I. Foundations and basic methods. Int J Epidemiol. 2006;35:765-75.
12. Goodman S. Toward evidence-based medical statistics. 1. The P value fallacy. Ann Intern Med. 1999;130:995-1004.
13. Spiegelhalter DJ, Best NG, Carlin BI, Van der Linde A. Bayesian measures of model complexity and fit. J R Statist Soc B. 2002; 64(4):583-639.
14. Dunson D. Practical advantages of Bayesian analysis of epidemiologic data. Am J Epidemiol. 2001;153(12):1222-6.
15. Marshall RJ. Bayesian analysis of case-control studies. Stat Med. 1988;7:1223-30.
16. Zelen M, Parker RA. Case-control studies and Bayesian inference. Stat Med. 1986;5: 261-9.
17. Gilks WR, Richardson S, Spiegelhalter DJ. Eds. Markov chain Monte Carlo in practice. Chapman & Hall: London. 1996;1-19.
18. El–Hakim IE, Uthman MA. Squamous cells carcinoma and Keratoa-canthoma of lower lip associated with 'Goza' and 'Shisha' smoking. Int J Dermatol. 1999;108-10,
19. Etzioni R, Kadane J. Bayesian statistical methods in public health and medicine. Public Health. 1995;16:23-41.
20. Hilbe JM, O'Brien CM. Negative binomial regression. International Statistical Review. 2008;76(2):302–303.
21. Metropolis N, Arianna W, Rosenbluth MN, Rosenbluth AHT, Edward T. Equation of state calculations by fast computing machines; 1953.
22. Hastings WK. Monte Carlo sampling methods using markov chains and their applications. Biometrika. 1970;57(1):97-109.
23. Gamerman D, Lopes HF. Markov chain Monte Carlo: Stochastic simulation for bayesian inference. Boca Raton, FL: Chapman & Hall/CRC; 2006.
24. Deshmukh JS, et al. Low birth weight and associated maternal factors in an urban area. 1996;9.
25. Nagargoje et al. A hospital based case control study of risk factors for low birth

weight in Nagpur city of Maharashtra; 2011.
26. Hogan DP, Park JM. Family factors and social support in the developmental; 2000.
27. Bategeka L, Leah M, Okurut A, Barungi M, Apolot JM. The determinants of birth weight in Uganda. A Final Report Submitted to AERC, Nairobi, Kenya; 2009.
28. Mumbare SS, Maindarkar G, Darade R, Yenge S, Tolani MK, Patole K. Maternal risk factors associated with term low birth weight neonates: A matched-pair case control study; 2011.

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/14672