Bayesian and Frequentist Comparison for Epidemiologists: A Non Mathematical Application on Logistic Regressions

Pascale Salameh[*,1], Mirna Waked[2], Georges Khayat[3] and Michèle Dramaix[4]

[1]*Lebanese University - Faculties of Pharmacy & of Public Health, Beirut, Lebanon*

[2]*Saint George Hospital - Beirut & Balamand University - Faculty of Medicine, Beirut, Lebanon*

[3]*Hôtel Dieu de France Hospital - Beirut & Saint Joseph University- Faculty of Medicine, Beirut, Lebanon*

[4]*Université Libre de Bruxelles, School of Public Health, Brussels, Belgium*

Abstract: *Background*: Statistical frequentist techniques are sometimes misinterpreted or misused, while Bayesian techniques seem to present several practical advantages, such as accommodating small sample sizes, unobserved variables along with measurement errors and incorporating information from previous studies. The primary objective of this study was to evaluate the association between waterpipe dependence and chronic obstructive pulmonary disease (COPD), by comparing frequentist and Bayesian methods' results.

Methods: It is a multicenter case-control study, comparing a group of COPD patients with a control group. COPD diagnosis was held after clinical and paraclinical testing, while a standardized questionnaire was used to evaluate smoking history. Both frequentist and Bayesian analyses were performed.

Results: Although carried out on the same dataset, the results quantitatively differed between the frequentist and Bayesian analysis. Whenever the frequentist results were clear cut such as in case of cigarette smoking association with COPD, performing the MCMC method helped to increase the accuracy of the results, but did not change the direction of hypothesis acceptance, except in doubtful cases. When the frequentist p-value was ≤ 0.100, such as in case of smoking more than 15 waterpipe-years, the MCMC method improved deciding between the null and alternative hypothesis.

Conclusion: The Bayesian approach may have advantages over the frequentist one, particularly in case of a low power of the frequentist analysis, due to low sample size or sparse data; the use of informative priors might be particularly useful in narrowing credible interval and precising the choice between the null and alternative hypothesis. In case of borderline frequentist results, the MCMC method may be more conservative, particularly without priors. However, in case of large sample sizes, using frequentist methods is preferred.

Keywords: Bayesian, case-control, comparison, COPD, frequentist, waterpipe.

INTRODUCTION

Waterpipe (WP) or narguileh is a form of tobacco smoking, quickly spreading in Middle Eastern countries and around the world [1, 2]. Details regarding waterpipe smoking can be found in selected studies [3-6]. Waterpipe smokers are potentially exposed to the same substances that are harmful in cigarettes and studies concerning waterpipe tobacco smoking health effects suggest that it leads to cancers, genetic damages, diseases of the lungs and other conditions in active and passive smokers [7-13]. In a cross-sectional study, we had found relative risks of 2.5 of physician diagnosed respiratory diseases and chronic bronchitis symptoms in waterpipe smokers *versus* never smokers [14]. However, since waterpipe smoking is a habit that has dimensions of physiological, psychological and

social nature [15], it may manifest in several patterns in real life: a substantial percentage of waterpipe smokers are not dependent on it (around 80%) [14], and may thus smoke irregularly. This is why we would consider waterpipe dependence as an exposure variable of interest, in addition to waterpipe smoking, to be both linked with chronic obstructive pulmonary disease. The latter is a chronic disease of relatively high prevalence in Lebanon (10%), mostly related to cigarette smoking, but also suspected to be associated with waterpipe smoking [16].

Objective of the Study

The primary objective of this study is thus to evaluate the association between waterpipe dependence and chronic obstructive pulmonary disease (COPD), by using and comparing frequentist (both Fisherian and Neyman-Pearson approaches) and Bayesian methods' results from an epidemiologist point of view.

Conceptual Framework

Reasons for choosing to compare these approaches' results are numerous, namely because frequentist techniques,

*Address correspondence to this author at the Jdeidet El Meten, Chalet Suisse Street, Ramza Azzam bldg, 5th Floor, Beirut, Lebanon;
Tel: 009613385542; Fax: 009611696600;
E-mails: psalameh@ul.edu.lb; pascalesalameh1@hotmail.com

although the most universally used, are sometimes considered as misinterpreted or misused [17-19], while Bayesian techniques are underused while they seem 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 [17-20].

Bayesian methods are defined as the explicit quantitative use of external evidence in the design, monitoring, analysis, interpretation and reporting of a study [20]. Results of epidemiological observational studies provide a likelihood that can be combined with prior information using standard and advanced full Bayesian methods [19], 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 [21-23]: 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 [24, 25].

Moreover, in parallel to the classical confidence interval, Bayesian methods offer what is called the Bayesian credible interval; it 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 [9]; furthermore, people sometimes wrongly interpret the confidence interval as if it was a credible interval. In this sense, Bayesian approach seems more intuitive.

Although these practical advantages have long been presented for epidemiologists, few epidemiological studies have used this powerful tool to assess exposure-disease relations [9]. Controversies are raised by the Bayesian approach, since it forces one to reexamine fundamental notions about the concept of probability and classical statistical practices [26]: 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 [26] because it is considered subjective [27]. Thus, comparing results, interpretations and limitations of both approaches would enrich our discussion and make our conclusions more robust.

METHODS

Study Design

It is a multicenter case-control study, comparing a group of patients having COPD with a control group. WP exposure, WP dependence and socio-demographic characteristics were compared between both groups. Since the study was observational and there was no traceability of patients (data was collected anonymously), the Internal Review Board (IRB) of the Lebanese University waived the need for an official approval to perform the study, provided the study respected patients' autonomy and confidentiality. Participants gave an oral informed consent before data

collection starting; a response rate of more than 90% was recorded.

Study Population

The COPD group was composed of incident outpatient cases of mild, moderate or severe COPD. Inclusion criteria encompassed being 40 years of age or more and free of other diagnosed respiratory diseases such as asthma, tuberculosis, lung cancer or fibrosis. Exclusion criterion was a previous diagnosis of COPD.

The control group was composed of outpatients aged 40 years of age or more and consulting for diverse problems: they were taken from outpatient clinics of cardiology, endocrinology, dermatology, nephrology, hematology, urology, gastroenterology, gynecology, ophthalmology, pre-surgery consultation, and oto-rhino-laryngology. Moreover, people accompanying cases (family members or friends) could also be taken as controls. Exclusion criteria included: having previous or current diagnosis or any other diagnosed respiratory diseases such as asthma, tuberculosis, lung cancer or fibrosis, or suffering from any chronic respiratory symptom.

Data Collection

COPD diagnosis was held after clinical and paraclinical testing: patients were evaluated for chronic respiratory symptoms such as dyspnea, chronic cough, and chronic expectorations. A spirometry was performed to classify COPD patients according to GOLD guidelines [28].

A standardized questionnaire was used to evaluate the following: detailed cigarette and waterpipe smoking history, passive smoking, exposure to smokes and fumes at work and home, polluted areas residency history, cigarette and waterpipe dependence scores according to Fagerström Test for Nicotine Dependence (FTND) [29] and LWDS-11 [15], respectively. Other potential confounding variables included age, gender, height, weight, education, work status, and marital status. Additional methodological details are available in a separate publication [30].

Sample Size Calculation

Sample size calculation was performed with an alpha risk of 5% and a beta risk of 20%; since exposure probability to waterpipe was equal to 50%, representing the mean percentage of Lebanese individuals who smoked and could theoretically be exposed to waterpipe [31], and since 20% of these are known to be dependent to waterpipe [14], 10% of the general population was thus considered waterpipe dependent. The minimal sample size necessary to show a twofold increase in risk consisted of 721 subjects, divided as 2.5 controls (n=515) for 1 case (n=206). Cases and controls were not paired.

Statistical Analysis

Data entry was performed by independent laypersons that were unaware of the objectives of the study. Quality control of data entry and data cleaning were carried out by researchers. Statistical frequentist analysis was performed using SPSS software, version 13.0. A p-value of 0.05 was considered significant. For continuous variables between groups' comparison, we used ANOVA or Students' tests for

variables with adequate normal distribution. For non normally distributed continuous variables and for non continuous quantitative and ordinal variables, Kruskal-Wallis test was used. For categorical variables, the Chi2 and Fisher exact tests were used when applicable.

The main dependent variable was being diagnosed with COPD or not. Independent variables were waterpipe smoking and waterpipe dependence. Multivariate analyses were carried out to compare measures between groups of comparison, taking into account potential confounding variables: gender, residency, age, height, weight, body mass index, education, work status, marital status and the number of smokers to which a person was potentially exposed (passive smoking at home and work). Since the dependent variable was dichotomous, a logistic regression model was performed, using a stepwise descendent based on likelihood ratio method, and the final model was retained. Point estimates of adjusted Odds ratios (ORa), 95% confidence intervals and p-values were presented. A p-value is the long-term probability of obtaining a test statistic, at least as large as the one observed, if data were to be repeatedly generated under identical conditions from the larger population in which the null hypothesis is true. For a small p-value (<0.05), we rejected the null hypothesis. Traditionally, most statistical analyses in epidemiology have been carried out from this frequentist perspective.

On the other hand, a Bayesian analysis was performed using models identical to those used in the frequentist analyses. Informative priors and non informative priors were used; informative priors were taken from a previous cross-sectional national study performed by our team to evaluate the prevalence of COPD in Lebanon; we note that a similar questionnaire was used in that study on a representative sample of the Lebanese population, and we thus could obtain priors for the majority of the variables that were used in the present analysis [16]; logarithm of the prior ORa we used followed a normal distribution.

The Markov Chain Monte Carlo method was then applied to derive posterior distributions of the model parameters, using the WinBUGS14 (Bayesian inference Using Gibbs Sampling) statistical software package. Medians, 95% credible intervals, and probability of the odds ratios being more than 1 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 [32].

The Bayes factor, defined as how well a hypothesis predicts the data *versus* another, was also calculated [33]. We considered that an OR>1 represented the alternative hypothesis ($P_{H1|data}$), while an OR≤1 represented the null hypothesis ($P_{H0|data}$). The abovementioned posterior distribution served us to calculate the Bayes factor ($BF_{H0|H1}$):

$$BF01 = \frac{(PH0|data)}{(PH1|data)} * \frac{P(H1)}{P(H0)}$$

According to Goodman (1999) [33], the BF01 gives the following evidence strength against the null hypothesis:

weak if ≥0.20, moderate if included in [0.1- 0.2[, moderate to strong if included in [0.05-0.1[, strong if included in [0.05-0.01[, and very strong if ≤0.01.

A sensitivity analysis was performed using the different priors specified above, and posterior distributions were graphically presented. Interpretation of both frequentist and Bayesian analyses were presented, stressing on differences, similarities and complementarities between both methods.

RESULTS

Models Choice

Using one dataset, we performed three stepwise descending logistic regressions on SPSS, where the dependent variable was the COPD, while the major independent variables were respectively: dichotomous ever smoking of more than 15 waterpipes-years, while the second model encompassed dichotomous waterpipe dependence and previous waterpipe smoking in the subgroup of waterpipe and never smokers.

The results of these logistic regressions were reported in Tables **1** and **2**: From the frequentist point of view, ever smoking more than 15 waterpipe-years might increase the risk of COPD, but the association did not reach statistical significance (ORa=1.74[0.90-3.36]; p=0.100) (Table **1**). Moreover, waterpipe dependence in current waterpipe smokers and previous waterpipe smoking were both associated with COPD: ORa=4.73[1.14-19.64]; p=0.032 and ORa=8.96[2.03-39.60]; p=0.004, respectively (Table **2**).

Bayesian Modeling

To be able to compare, we used the variables kept by the previously generated models.

Let p_i denote the probability of COPD for the *i*th patient. Let $x_{1i}, ..., x_{ki}$ denote the values of the k predictor variables measured on the *i*th patient. We fitted the model:

$$\log(p_i/(1-p_i)) = alpha_0 + a_1 x_{1i} + ... + a_k x_{ki}$$

In the BUGS language, these models were respectively:

For Ever smoking more than 15 waterpipes-years:

```
model {for (i in 1:N) {

COPD[i] ~ dbin(p[i], N)

logit(p[i)) <- alpha0 + alpha2*education[i]
+ alpha3*work[i] + alpha4*closeroad[i] +
alpha5*heatelec[i] + alpha6*heatair[i] +
alpha7*familyhistory[i] +
alpha8*cigarette[i] + alpha9*everarg15[i] +
alpha10*ageclass[i] + alpha11*smokwork[i]
```

For waterpipe dependence in waterpipe and never smokers:

```
model {for (i in 1:N) {

COPD[i] ~ dbin(p[i], N)

logit(p[i]) <- alpha0 + alpha2*education[i]
+ alpha3*closeroad[i] + alpha4*heatelec[i] +
alpha5*heatdiesel[i] + alpha6*previousarg[i]
+ alpha7*evercigarette[i] +
alpha8*dependencearg[i] + alpha9*ageclass[i]
```

Table 1. COPD association with dichotomous waterpipe smoking (n=735).

Analysis Method	Frequentist Analysis: Logistic Regression			Bayesian Analysis: Logit Model; Non Informative Priors				Bayesian analysis: Logit Model; Informative Priors*		
Variable	ORa	95% Confidence Interval	p-value	ORa	95% Credible Interval	Pr (OR>1)	Bayes Factor (H0/H1)	ORa	95% Credible Interval	Pr (OR>1)
Smoking > 15 WP-years	1.74	0.90-3.36	0.100	3.32	0.74-8.67	0.940	0.063	3.25	1.52-5.99	1.000
Older age class**	1.96	1.67-2.30	<0.001	1.46	1.31-1.65	1.000	0.000	1.48	1.33-1.66	1.000
Living close to a busy road	1.83	1.13-2.97	0.014	1.25	1.31-1.65	0.920	0.087	1.25	0.92-1.65	0.920
Lower education	1.45	1.14-1.85	0.003	1.15	1.00-1.33	0.975	0.026	1.15	1.00-1.32	0.972
Ever smoking cigarettes	16.61	8.61-32.02	<0.001	8.06	4.10-15.57	1.000	0.000	7.34	4.31-12.17	1.000
Heating home by hot air	2.31	1.08-4.94	0.030	1.34	0.83-2.01	0.880	0.136	1.34	0.84-2.00	0.890
Heating home not electrically	1.85	1.15-3.03	0.012	1.23	0.93-1.69	0.930	0.075	1.23	0.93-1.69	0.930
Lung disease family history	1.80	1.04-3.12	0.037	1.17	0.84-1.56	0.820	0.220	1.17	0.85-1.56	0.820
Having smokers at work	2.51	1.24-5.09	0.011	1.89	1.10-3.04	0.990	0.010	1.92	1.12-3.06	0.990
Being exposed to toxics and fumes at work	1.75	0.95-3.23	0.091	1.32	0.88-2.08	0.910	0.099	1.30	0.88-2.04	0.910
Model adequacy	Hosmer-Lemeshow p-value=0.074	Nagelkerke R²=0.622		Deviance Information Criterion =700.745; Percentage of p-values <0.05 is 4%‡				Deviance Information Criterion =698.825 Percentage of p-values <0.05 is 3.5% ‡		

*Priors available for: Smoking >15 WP-years (ORa=5.84[3.58-9.53]), Age class (ORa=1.73[1.55-1.94]) and Ever smoking cigarettes (ORa=3.33[2.19-5.05]).
**Age classes increase by 5 years increments; ‡ Low percentages show adequate modeling to data.

Table 2. COPD association with waterpipe dependence in waterpipe smokers (n=414).

Analysis Method	Frequentist Analysis: Logistic Regression			Bayesian Analysis: Logit Model; Non Informative Priors				Bayesian Analysis: Logit Model; Informative Priors*		
Variable	ORa	95% Confidence Interval	p-Value	ORa	95% Credible Interval	Pr (OR>1)	Bayes Factor	ORa	95% Credible Interval	Pr (OR>1)
Waterpipe dependence	4.73	1.14-19.64	0.032	3.08	0.91-7.60	0.963	0.018	3.08	1.06-6.93	0.981
Previous WP smoking	8.96	2.03-39.60	0.004	2.85	0.84-7.04	0.947	0.056	2.87	0.99-6.38	0.974
Older age class**	2.02	1.50-2.72	<0.001	1.57	1.24-1.97	1.000	0.000	1.57	1.24-1.98	0.999
Lower education	2.27	1.35-3.85	0.002	1.33	0.98-1.89	0.967	0.034	1.33	0.97-1.85	0.964
Ever cigarette smoking	13.06	3.15-54.18	<0.001	3.99	1.31-10.05	0.994	0.006	3.72	1.41-8.60	0.996
Heating home by diesel	4.14	1.05-16.29	0.042	1.45	0.61-2.95	0.768	0.302	1.52	0.66-3.02	0.817
Heating home not electrically	3.13	1.04-10.00	0.043	1.72	0.92-3.70	0.955	0.047	1.64	0.88-3.45	0.948
Living close to a busy road	5.93	1.91-18.45	0.002	2.48	1.13-4.89	0.990	0.010	2.50	1.14-4.92	0.990
Model adequacy	Hosmer-Lemeshow p-value=0.904	Nagelkerke R²=0.707		Deviance Information Criterion=210.532 Percentage of p-values <0.05 is 1%‡				Deviance Information Criterion =210.574 Percentage of p-values <0.05 is 0.8%‡		

*Priors available for: Waterpipe dependence (ORa=3.34[1.13-9.89]); Previous WP smoking (ORa=3.68[1.36-9.97]); Older Age class (ORa=1.87[1.52-2.30]); Ever smoking cigarettes (ORa=5.00[1.94-12.96]); Heating home by diesel (ORa=3.22[1.60-6.48]); Heating home not electrically (ORa=2.35[1.16-4.75]);
** Age classes increase by 5 years increments;
‡ Low percentages show adequate modeling to data.

Two chains were used for model compilation; initial values for MCMC procedures were defined at 0 (chain 1) and 0.3 (chain 2) for all parameters. By convention, one then assumes diffuse or non-informative priors on the parameters to be estimated. This implies that our prior beliefs about the parameters' probability distribution were vague and

imprecise, and that the parameter could assume values over a large range, with approximately equal likelihood. The previous probability distribution for the regression coefficients was assumed to be a normal distribution, with a mean of zero and a variance of 1000000.

Afterwards, we used informative priors according to a previous study we carried out in Lebanon, as stated above. For model 1, the available priors were for smoking >15 WP-years (ORa=5.84[3.58-9.53]), age class (ORa=1.73[1.55-1.94]) and ever smoking cigarettes (ORa=3.33[2.19-5.05]).

For model 2, the available priors were respectively: waterpipe dependence (ORa=3.34[1.13-9.89]), previous WP smoking (ORa=3.68[1.36-9.97]), older age class (ORa=1.87[1.52-2.30]), ever smoking cigarettes (ORa=5.00[1.94-12.96]), heating home by diesel (ORa=3.22[1.60-6.48])]; and heating home not electrically (ORa=2.35[1.16-4.75]).

In the two models, the same process was adopted to generate unknown parameters. First of all, the model was specified along with data loading; two chains were compiled to increase the assertiveness of convergence. Five thousands iterations were burned out; afterwards, the program was allowed to run 10000 iterations, making a total of 20,000 samples used for statistics generation. This number of iterations allowed obtaining a Monte Carlo error for parameters of interest of less than 5% of the sample standard deviation [24, 34]. Convergence of the results was checked by looking at the iterations history of major independent variables parameters (Fig. **1**) and using the Brooks Gelman Rubin (BGR) diagnosis tool (Fig. **2**).

The WinBUGS program was designed to estimate the posterior distribution of the unknown parameters exponentials, which represent the posterior distributions of the Adjusted Odds Ratios (ORa); we could thus generate the posterior distributions along with the quintiles of the ORa (Figs. **3**, **4**). Moreover, we looked at the distribution of the probability function of having an OR that is higher than 1 by using the "step" function of WinBUGS. We considered that P(H1)=P(H0)=0.5 to evaluate the Bayes Factor of the data without prior information, a situation that is generally considered the best for coherence between the frequentist and Bayesian approaches [33]. We also looked at the percentage of p-values that were below 0.05; the lower this percentage was, the better the model adequacy to the data [32]. Respective results were presented in Tables **1** and **2**.

COPD Association with Dichotomous Waterpipe Smoking

In Table **1**, the Markov Chain Monte Carlo drawing of 20,000 samples from the data reached results that were different from the frequentist logistic regression performed by SPSS. For COPD association with waterpipe smoking, when using non informative priors, there was an increase from a non significant (p=0.100) frequentist ORa of 1.74[0.9-3.36] to a strong association of 3.32; however, the credible interval still contained 1 [0.74-8.67], showing a low (0.06) but possible probability for the ORa being ≤1 (moderate evidence against H0 from the Bayes factor). When using the informative prior, we obtained a close ORa=3.25,

a narrower credible interval [1.52-5.99], with a null probability for the ORa being lower or equal to 1. In this case, the results go very strongly against the null hypothesis for the association between COPD and dichotomous waterpipe smoking.

We note that for other variables where priors were available (cigarette smoking and age), Bayesian results were closer to these priors, as expected. For other variables where priors were not available and where p-values were >0.01, a decrease in point estimate was noted, with credible intervals containing 1 (driving us away from alternative hypothesis, such as for heating home by hot air, not electrically, and having a lung disease family history), or not containing 1 (confirming alternative hypothesis, such as for having smokers at work).

Moreover, as expected, the Deviance Information Criteria were similar with or without informative priors, because we were using the same model and including the same number of parameters. The model adequacy to the data was also adequate in both cases (percentage lower than 5%). Figs. (**3a**, **b**, **4a**, **b**) clearly show the improvement in parameters estimation robustness. We note that analogous results were obtained for other variables.

COPD Association with Waterpipe Dependence in Waterpipe and Never Smokers

The association of waterpipe dependence with COPD was strong but had a borderline statistical significance (ORa=4.73[1.14-19.64]; p=0.032) in the frequentist analysis. In the Bayesian analysis, non informative priors gave an ORa=3.08[0.91-7.60]: although the credible interval included 1, it was narrower than the confidence interval and the Bayes factor was in favor of the alternative hypothesis (Table **2**). A further narrowing of the credible interval was further witnessed when using informative priors, removing 1 from it, and definitely shifting the results towards the alternative hypothesis. For waterpipe previous smoking, a change in the point estimate of the ORa was found; however, the evidence against the null hypothesis remained moderate to strong. For other variables, results similar to model 1 were noticed: when p-values were between 0.01 and 0.05 (heating home by diesel and not electrically), the credible intervals given by the MCMC method did not contain 1.

The Deviance Information Criteria were similar with or without informative priors; the model adequacy to the data was also adequate in both cases (percentage lower than 5%). Figs. (**3c**, **d**, **4c**, **d**) showed that parameters estimation robustness remained similar with informative priors.

DISCUSSION

In this analysis, we were able to show that waterpipe smoking of more than 15 WP-years was associated with COPD. We refer the readers to a related publication for a full epidemiological discussion [30]. The results we showed in this work focused on the comparative use of different methods of analysis (frequentist and Bayesian). Most importantly, when comparing the results of both methods, several points warranted our attention and were discussed below.
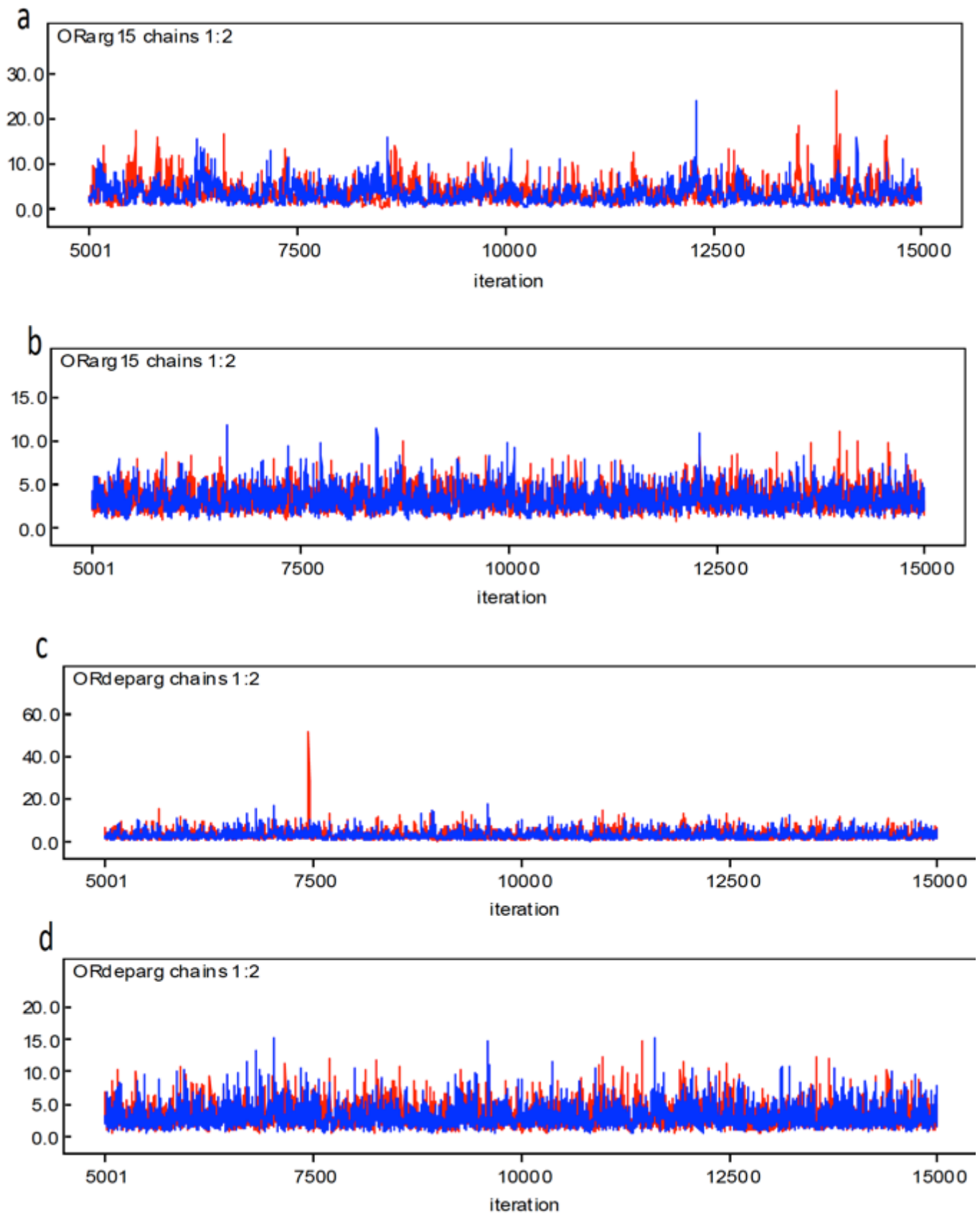
Fig. (1). Convergence history for major independent variables. a=Convergence history for OR of COPD & Ever smoking more than 15 waterpipe-years; non informative priors; b= Convergence history for OR of COPD & Ever smoking more than 15 waterpipe-years; informative priors: c= Convergence history for OR of COPD & waterpipe dependence; non informative priors: d= Convergence history for OR of COPD & waterpipe dependence; informative priors. Models convergence is clear in all situations, because of chains overlapping.
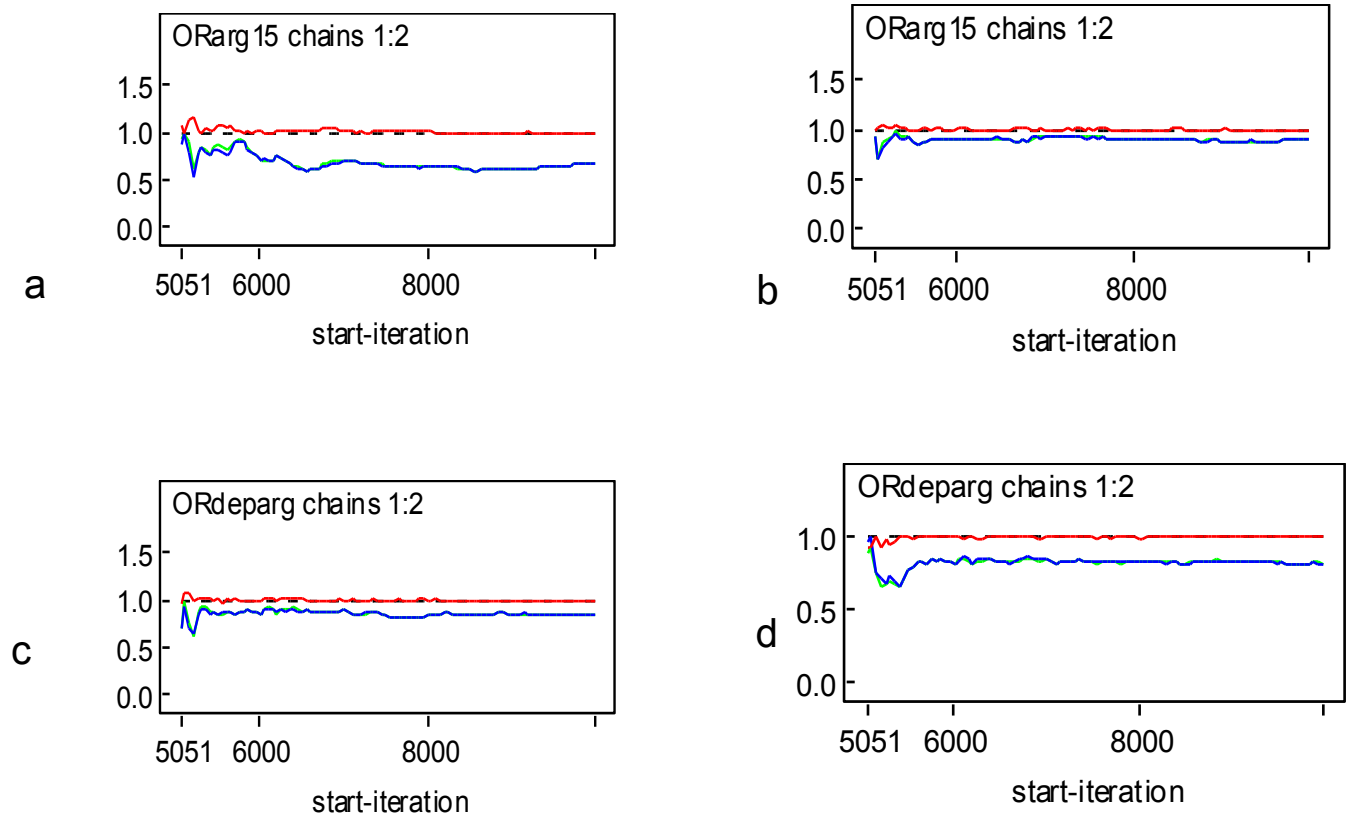
Fig. (2). Brooks Gelman Rubin convergence diagnosis for Major independent variables. a= OR of COPD & Ever smoking more than 15 waterpipe-years; non informative priors: b= OR of COPD & Ever smoking more than 15 waterpipe-years; informative priors: c= OR of COPD & waterpipe dependence; non informative priors; d= OR of COPD & waterpipe dependence; informative priors. The models are adequate and results were convergent according to this diagnostic procedure in all situations. Blue: Width of the 80% interval of the pooled runs; Green: Average width of the 80% intervals within the individual runs; Red: R = pooled / within; Interval widths are normalized to have an overall maximum of 1; R >1 if starting values are over-dispersed; convergence is adequate in case of convergence of R to 1, and convergence of pooled and within interval.

Although carried out on the same dataset, the results quantitatively differed between the frequentist and Bayesian analysis: we obtained ORa that were different between both types of analysis; in fact, the Gibbs sampling method usually gives a median of the ORa generated from the 20,000 samples, while the frequentist analysis works on the available dataset. Thus, the WinBUGS program resulted in a more logical estimation of parameters [34, 35]; this can drive the results in both directions of confirming or infirming the frequentist results. The most likely explanation of this change could be the data sparsity (low level of exposure in the sample) that might have overinflated parameters, a problem that could be corrected by the "shrinkage" used in Bayesian analysis. This process usually improves overall accuracy of estimation and prediction [36]. The prior information, even when vague, improves the validity of the fitted model [37]; whenever the sample size is large, however, the priors' generated information is dominated by that of the dataset, and its added value is rather questionable.

Whenever the frequentist results were clear cut (due to a large sample size or a strong association), such as in the case of cigarette smoking association with COPD, performing the MCMC method helped to increase the accuracy of the results by narrowing the credible interval, but did not change the direction of hypothesis acceptance; as stated above, the

Gibbs sampling might give closer results to the truth and using informative priors might further help to improve credible intervals. In these situations, the frequentist likelihood ratio calculations are generally known to give results similar to Bayesian analysis [9], and Bayesian credible intervals may be asymptotically similar to optimal frequentist confidence intervals [35]. Thus, the interest of Bayesian analysis decreased whenever sample size increased.

When the frequentist p-value is higher than 0.05 but ≤0.100 and the confidence interval of the OR includes 1, epidemiologists usually renounce to conclude and recommend studies with a higher sample size to be able to choose between the null and alternative hypothesis, because of a weak evidence against the null hypothesis [33]. There are chances however that the association exists but the low sample size could have precluded showing it (under power of the analysis). In other words, there is a non negligible probability that the alternative hypothesis is true although the p-value is higher than 0.05. In our analysis, we found that the Bayesian approach improved deciding between the null and alternative hypothesis, even with non informative priors, in case of smoking more than 15 waterpipe-years. The best result appeared to occur when we used informative priors: in this case, we witnessed an increase in the decision possibility
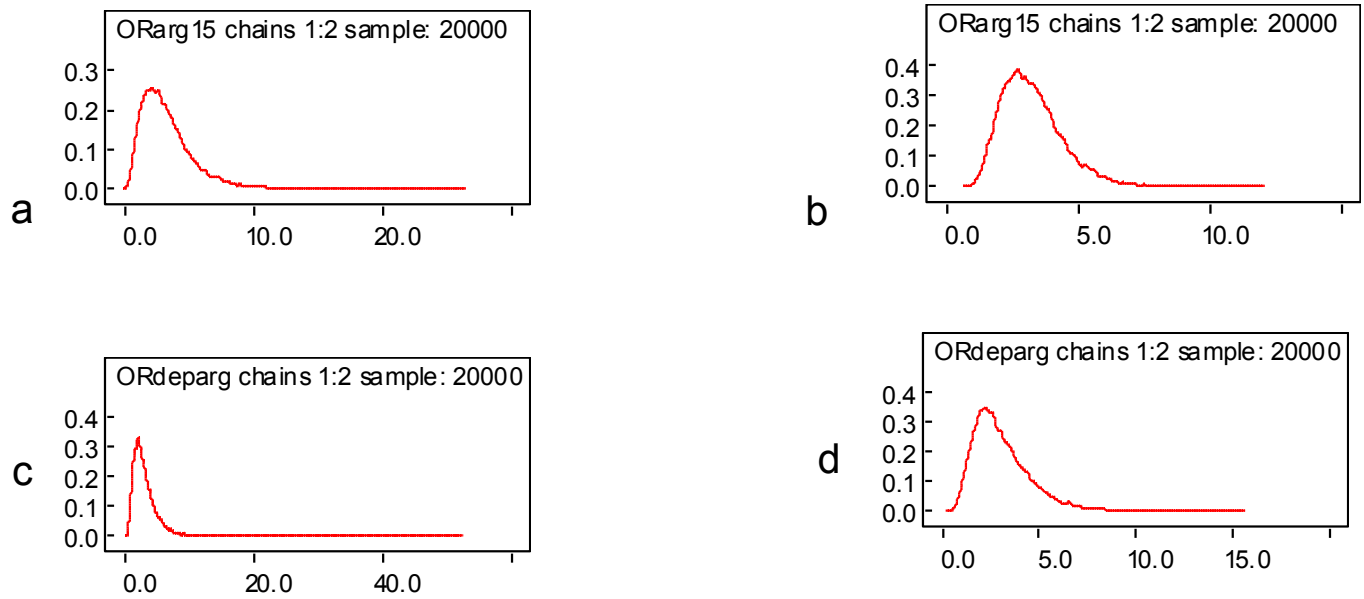
Fig. (3). Posterior distribution for major independent variable. a=Posterior distribution for OR of COPD & Ever smoking more than 15 waterpipe-years; non informative priors: b= Posterior distribution for OR of COPD & Ever smoking more than 15 waterpipe-years; informative priors: c= Posterior distribution for OR of COPD & waterpipe dependence; non informative priors: d= Posterior distribution for OR of COPD & waterpipe dependence; informative priors. Note the narrowing of the posterior distribution after informative priors (b & d in particular).
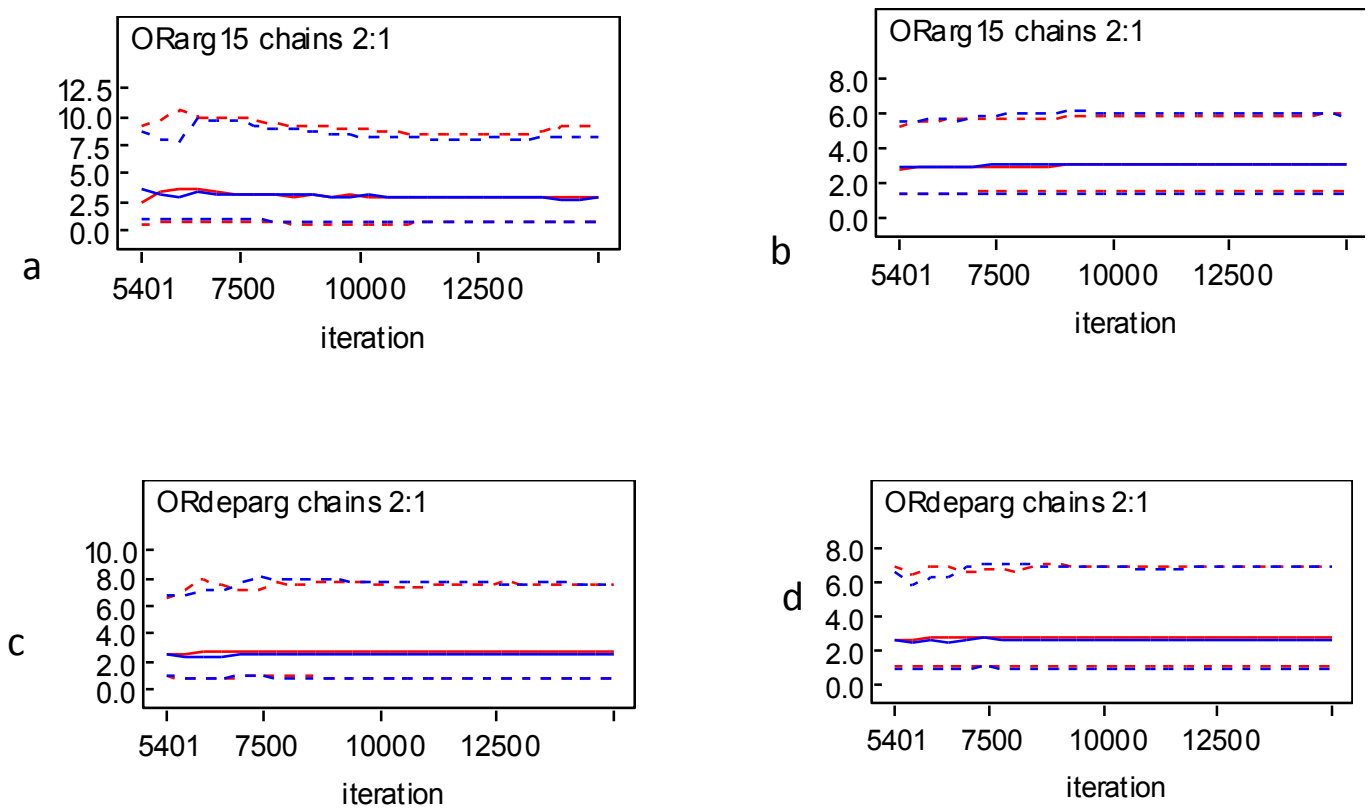

Fig. (4). Quintiles for major independent variable. a=Quintiles for OR of COPD & Ever smoking more than 15 waterpipe-years; non informative priors: b= a=Quintiles for OR of COPD & Ever smoking more than 15 waterpipe-years; informative priors: c= Quintiles for OR of COPD & waterpipe dependence; non informative priors: d= Quintiles for OR of COPD & waterpipe dependence; informative priors. Note the narrowing of quintiles with informative priors and stabilization with increasing number of iterations, particularly with informative priors.

and a narrowing of credible intervals, based on which the null hypothesis could be rejected with higher confidence.

A better decision between null and alternative hypotheses could also be made whenever in frequentist analysis the p-value is of borderline significance, i.e. between 0.03 and

0.05; in this situation, working epidemiologists carefully reject the null hypothesis and accept the alternative one. However, it is now known that this result is not robust; it only gives moderate strength of evidence against the null hypothesis [33], and it has been shown that the probability of repetition of a significant p-value is low [38, 39]. In fact, lower point estimates and a shift towards the null hypothesis were noticed with Bayesian models, with or without priors, such as for the variables "heating home by hot air", "by diesel", or "not electrically". In these cases, Bayesian credible interval estimates obtained from MCMC procedures were narrower and maybe more appropriate than confidence intervals, which are most of the times calculated by assuming large sample approximations [9]. Thus, there is a "natural protection" in Bayesian testing against the tendency to wrongly reject the null hypothesis in frequentist testing [40].

On another hand, the advantages of the Bayes factor over the p-value have been widely discussed in the literature [32, 41]. Furthermore, in line with our results on an observational study, many Bayesian re-analyses of clinical trials concluded that the observed differences were not likely to be true [42, 43]; the other way round, the clinical community seems to decline results of clinical trials that counteract strong prior beliefs [44, 45]. Despite this, health professionals are still reluctant to adopt the Bayesian approach [20], at least as a complementary method to further validate their studies. In fact, several authors have also come to conclusions that the frequentist and Bayesian approaches are complementary [46-50]. We hope that this analysis will help epidemiologists moving forward towards this goal; we recommend doing both types of analyses in epidemiological studies, particularly in case of low sample size or sparsity of data. In case of contradiction between both analyses, extreme caution should be applied before coming to conclusions.

Our study may have several limitations. We did not try to mathematically explain the advantages and disadvantages of every approach; we only compared the final results useful for working analysts (epidemiologists and biostatisticians). We did not take into account the possible biases that are associated with the design of the study during computations; the choice of variables to include in the regression was also made with the frequentist approach prior to the comparison process. All these were already discussed in the previous related publication [30]. We also used priors taken from a study that was performed on a different setting than the current one; this may induce some errors, but this is the closest prior to the truth that is available to us, since no previous study was performed on the same topic in a similar setting. Moreover, the results we found may not be applicable for all models that could be used in health, and our validation approach could gain with an application on other types of models, particularly multiple regression and survival analyses. Furthermore, we did not go into all details of convergence diagnosis [51, 52] and models goodness of fit measures [32, 53]. Additional studies, showing other advantages of Bayesian approach, are necessary to convince the frequentist school experts about its added value in the health field; we also recommend carrying out additional comparative analyses on other types of models, while repeating similar studies to ours to improve decision making,

particularly in case of contradictory results between both analyses types.

CONCLUSION

In conclusion, we were able to show that the Bayesian approach may have advantages over the frequentist one, particularly in case of a low power of the frequentist analysis, due to low sample size or scarce data; the use of informative priors might be particularly useful in narrowing credible interval and precising the choice between the null and alternative hypothesis. In case of borderline frequentist results, the MCMC method may be more conservative, particularly without priors. However, in case of large sample sizes, using frequentist methods is preferred.

ABBREVIATIONS

ANOVA = Analysis of Variance

COPD = Chronic Obstructive Pulmonary Disease

FTND = Fagerström Test for Nicotine Dependence

GOLD = Global initiative for Obstructive Lung Disease

MCMC = Markov Chain Monte Carlo

WP = Waterpipe

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1] Serota JA. A pack in 30 minutes. J Ped Health Care 2007; 21: 180-1 & 209-10.

[2] Chouachi K. The medical consequences of Narguile use in the world. Rev Epidemiol Sante Publique 2007; 55: 165-70.

[3] World Health Organization. World Health Organization study group on tobacco product regulation (TobReg) (2005). Waterpipe tobacco smoking: health effects, research needs and recommended actions by regulators. www.who.org [February 2007].

[4] Chaouachi K. Tabacologie du narguilé. Rev Alcool 1999; 21(1/83): 88-9.

[5] Knishkowy B, Amitai Y. Water-pipe (Narghile) smoking: an emerging health risk behavior. Pediatrics 2005; 116: e113-9.

[6] Chaouachi K. A critique of the WHO TobReg's "Advisory. Note" report entitled:"Waterpipe tobacco smoking: health effects, research needs and recommended actions by regulators". J Negat Results BioMed 2006; 5: 17-26.

[7] Tamim H, Terro A, Kassem H, *et al*. Tobacco use by university students, Lebanon 2001. Addiction 2003; 98(7): 933-9.

[8] Aydin A, Kiter G, Durak H, can ES, Kaya GC, Ceylan E. Water-pipe smoking effects on pulmonary permeability using technetium-99m DTPA inhalation scintigraphy. Ann Nucl Med 2004; 18(4): 285.

[9] El-Hakim IE, Uthman MA. Squamous cell carcinoma and keratoa-canthoma of the lower lip associated with "Goza" and "Shisha" smoking. Int J Dermatol 1999; 108-10.

[10] Al-Fayez SF, Salleh M, Ardaawi M, *et al*. Effects of sheesha and cigarette smoking on pulmonary function of Saudi males and females. Trop Geogr Med 1988; 40: 115-23.

[11] Nuwayhid IA, Yamout B, Azar G, Kambris MA. Waterpipe (Hubble-bubble) smoking, low birth weight, and other pregnancy outcomes. Am J Epidemiol 1998; 148(4): 375-83.

[12] Mohammad Y, Kakah M, Mohammad Y. Chronic respiratory effect of narguileh smoking compared with cigarette smoking in women from the East Mediterranean region. Int J Chron Obstruct Pulmon Dis 2008; 3(3): 405-14.

[13] Radwan GN, Mohamed MK, El-Setouhy M, *et al*. Review on water-pipe smoking. J Egypt Soc Parasitol (Egypt) 2003; 33 (Suppl 3): 1051-71.

[14] Salameh P, Aoun Z, Waked M. Saliva cotinine and exhaled carbon monoxide in real life narghile (waterpipe) smokers: a post hoc analysis. Tob Use Insights 2009; (2): 1-10.

[15] Salameh P, Waked M, Aoun Z. Narguileh smoking: Construction and validation of the LWDS-11 dependence scale. Nicotine Tob Res 2008; 10(1): 148-59.

[16] Waked M, Khayat G, Salameh P. Chronic Obstructive Pulmonary Disease in Lebanon: a cross-sectional study. Clin Epidemiol 2011; 3: 315-23.

[17] Greenland S. Bayesian perspectives for epidemiological research: I. Foundations and basic methods. Int J Epidemiol 2006; 35: 765-75.

[18] Goodman S. Toward evidence-based medical statistics. 1. The P value fallacy. Ann Intern Med 1999; 130: 995-1004.

[19] Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. Bayesian methods in health technology assessment: a review. Health Technol Assess 2000; 4(38): 1-130.

[20] Dunson D. Practical advantages of Bayesian analysis of epidemiologic data. Am J Epidemiol 2001; 153(12): 1222-6.

[21] Marshall RJ. Bayesian analysis of case-control studies. Stat Med 1988; 7: 1223-30.

[22] Nurminen M, Mutanen P. Bayesian analysis of case-control studies. Stat Med 1989; 8: 1023-4.

[23] Zelen M, Parker RA. Case-control studies and Bayesian inference. Stat Med 1986; 5: 261-9.

[24] Gilks WR, Richardson S, Spiegelhalter DJ. Introducing Markov chain Monte Carlo. In: Gilks WR, Richardson S, Spiegelhalter DJ, Eds. Markov Chain Monte Carlo in Practice. Chapman & Hall: London 1996; pp. 1-19.

[25] Gilks WR, Thomas A, Spiegelhalter DJ. A language and program for complex Bayesian modelling. Statistician 1994; 43: 169-78.

[26] Etzioni R, Kadane J. Bayesian statistical methods in public health and medicine. Public Health 1995; 16: 23-41.

[27] Greenland S. Bayesian perspectives for epidemiological research: I. Foundations and basic methods. Int J Epidemiol 2006; 35: 765-75.

[28] Buist SA, Anzueta A, Calverly P, *et al*. 2006. Global Initiative for Chronic Obstructive Lung Disease: Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease. Available at: www.goldcopd.org [Accessed: December 2010].

[29] Moolchan ET, Radzius A, Epstein DH, *et al*. The fagerström test for nicotine dependence and the diagnostic interview schedule: do they diagnose the same smokers? Addict Behav 2002; 27: 101-13.

[30] Salameh P, Waked M, Khayat G, Dramaix M. Waterpipe smoking and dependence are associated with chronic obstructive pulmonary disease: a case-control study. Open Epidemiol J 2012; 5: 36-44.

[31] Maziak W, Ward KD, Afifi Soweid RA, Eissenberg T. Tobacco smoking using a waterpipe: a re-emerging strain in a global epidemic. Tob Control 2004; 13(4): 327-33.

[32] 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.

[33] Goodman S. Toward evidence-based medical statistics. 2: The Bayes Factor. Ann Intern Med 1999; 130(12): 1005-13.

[34] Lunn DJ, Thomas A, Best N, *et al*. WinBUGS—a Bayesian modeling framework: concepts, structure, and extensibility. Stat Comput 2000; 10: 325-37.

[35] Berger JO. Could Fisher, Jeffreys and Neyman have agreed on testing? Stat Sci 2003; 18(1): 1-32.

[36] Greenland S. Bayesian perspectives for epidemiological research. II. Regression analysis. Int J Epidemiol 2007; 36: 195-202.

[37] Greenland S. Methods for epidemiologic analyses of multiple exposures: A review and a comparative study of maximum-likelihood, preliminary testing, and empirical-Bayes regression. Stat Med 1993; 12: 717-36.

[38] Goodman SN. A comment on replication, P-values and evidence. Stat Med 1992; 11(7): 875-9.

[39] Killeen P. An alternative to null-hypothesis significance tests. Psychol Sci 2005; 16(5): 345-53.

[40] Johnson W. Comment on: Berger JO. Could Fisher, Jeffreys and Neyman have agreed on testing? Stat Sci 2003; 18(1): 13-16.

[41] Kass R, Raftery A. Bayes factors and model uncertainty. J Am Stat Assoc 1995; 90: 773-95.

[42] Lilford R, Braunholtz D. The statistical basis of public policy: a paradigm shift is overdue. BMJ 1996; 313: 603-7.

[43] Brophy JM, Joseph L. Placing trials in context using Bayesian analysis. GUSTO revisited by Reverend Bayes. JAMA 1995; 273: 871-5.

[44] Belanger D, Moore M, Tannock I. How American oncologists treat breast cancer: an assessment of the influence of clinical trials. J Clin Oncol 1991; 9: 7-16.

[45] Omoigui NA, Silver MJ, Rybicki LA, *et al*. Influence of a randomized clinical trial on practice by participating investigators: lessons from the Coronary Angioplasty *Versus* Excisional Atherectomy Trial (CAVEAT). CAVEAT I and II Investigators. J Am Coll Cardiol 1998; 31: 265-72.

[46] Rubin D. Bayesianly justifiable and relevant frequency calculations for the applied statistician. Ann Stat 1984; 12: 1151-72.

[47] Shafer G. Savage revisited. Stat Sci 1986; 1: 463-501.

[48] Walley P. Statistical reasoning with imprecise probabilities. London: Chapman and Hall 1991.

[49] Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. In: Slovic P, Tversky A, Kahneman D, Eds. Cambridge: Cambridge University Press 1982; pp. 1-20.

[50] Samaniego FJ, Reneau DM. Toward a reconciliation of the Bayesian and frequentist approaches to point estimation. J Am Statist Assoc 1994; 89: 947-57.

[51] Brooks S, Gelman A. General methods for monitoring convergence of iterative simulations. J Comput Graph Stat 1998; 7(4): 434-55.

[52] Plummer M, Best N, Cowles K, Vines K. CODA: convergence diagnosis and output analysis for MCMC. R News 2006; 6(1): 7-11.

[53] Gelman A, Meng XL, Stern H. Posterior predictive assessment of model fitness *via* realized discrepancies. Stat Sinica 1996; 6: 733-807.