Original Article

Superiority of Bayesian Model Averaging to Stepwise Model in Selection of Factors Related to the Incidence of Type II diabetes in Pre-diabetic Women

Yadollah Mehrabi [1], Maryam Mahdavi [2,*], Davood Khalili [3], Ahmad Reza Baghestani [4], Farideh Bagherzadeh-Khiabani[3]

[1]Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran

[2]Department of Biostatistics, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

[3]Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

[4]Department of Biostatistics, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Article Information
Received: 2017-08-30
Revised: 2018-06-26
Accepted: 2018-12-14

Correspondence
Maryam Mahdavi
Email:maryam.mahdavi1368@gmail.com

Cite this article as:
Mehrabi Y, Mahdavi M, Khalili D,Baghestani A R ,Bagherzadeh-Khiabani F
Superiority of Bayesian Model Averaging to Stepwise Model in Selection of Factors Related to the Incidence of Type II diabetes in Pre-diabetic Women. Journal of Paramedical Sciences 2019: 10(1).

Abstract

Introduction: The world prevalence of type 2 diabetes and its related increment mortality rate which needs high controls cost has attracted high scientific attention. Early detection of individuals who face this disease more than the others can prevent getting sick or at least reduce the disease consequences on public health. Regarding the costs and limitations of diagnostic tests, a statistical model is presented that helps predict the time of diabetes incidence and determines its risk factors. Furthermore, this model determines the significant predictor variables on response and considers them as model equation parameters.

Materials and Methods: In this study, 803 pre-diabetic women in the age range of more than 20 years were selected from Tehran lipid and glucose study (TLGS) to examine the predictor variables on time of diabetes incidence. They were entered into the study in the phases 1 and 2 and were followed up to the phase 4. The predictor variables selection was performed using the Stepwise Model (SM) and the Bayesian Model Averaging (BMA). Then, the predictive discrimination was used to compare the results of both models. The Log-rank test was performed and the Kaplan-Meier Curve was plotted. The statistical analyses were performed using R software (version 3.1.3).

Results: The Backward Stepwise Model (BSM), the Forward Stepwise Model (FSM) and the BMA have used 9, 10 and 6 variables, respectively. Although the BMA selected predictor variables number is much lower than the SM, the prediction ability remains nearly constant.

Conclusions: The BMA has averaged on the supported models using dataset. This model has shown nearly constant accuracy despite the selection of lower predictor variables number in comparison to the SM.

Keywords: Bayesian Model Averaging; Stepwise Model; Tehran Lipid and Glucose Study; Women pre-diabetic; Cox regression.

1. Introduction

Diabetes is the most important metabolic human disease and an important factor in ischemic coronary artery disease [1]. Type 2 diabetes risk factors include diabetes family history, obesity, age, high blood pressure and etc. [2]. Due to the personal and social disease burdens, it is important to identify its risk factors [3]. Recent World Health Organization (WHO) estimates shows that the diabetic patients percentage will be increased to 552 million by 2030 [4]. Some reasons for the upward incidence and prevalence trend of type 2 diabetes in recent decades could be changes in lifestyle, prevalent obesity and low physical activity [5]. Before the disease onset, the person suffers from a condition known as the pre-diabetes; it means that a person's blood sugar is higher than normal level but not high enough to be considered as diabetic [6]. Because of the diabetes high risk among the pre-diabetics, it is suggested for these patients to be diagnosed for the diabetes risk factors.

Many variables may be considered as risk factors for the diabetes development. Statistical methods such as logistic and survival regression models are usually employed to find and evaluate the most relevant subset of effective variables. In these regression models, a Stepwise Model (SM) is typically used to select a proper subset of variables [7]. One of the SM disadvantages is that it leads to the selection of a deterministic model, without considering the model uncertainties [8]. Lack of attention to the model uncertainty can lead to bias and inefficiency of the parameters estimation [9]. The Bayesian Model Averaging (BMA) can be used to take into account the model uncertainty. It averages on the possible models which weighted on models posterior probabilities [10, 11]. The basic principle of this model is that it treats models and their parameters as unobservable phenomena and estimates their distribution based on observable data [12]. Although uncertainty in the statistical models is well known, so far, few studies have been considering the uncertainty of the survival analysis models.

The present study was carried out to identify the diabetes risk factors among the pre-diabetics by means of BMA. The BMA was used to select predictor variables that affect diabetes in pre-diabetics. Then, its performance is compared with the SM using Tehran Lipid and Glucose Study (TLGS). Regarding the diagnostic tests' costs and limitations, a statistical model which helps to predict duration of the incidence was presented. Furthermore, this model determines the significant predictor variables on response and considers them as model equation parameters.

2. Materials and Methods

In this cohort study, a dataset associated with 803 pre-diabetic women aged 20 years and over from TLGS participants were selected to investigate the predictor variables on the time of diabetes incidence. Therefore, there are no ethical considerations in this study. TLGS is a cohort study whose design details have been published elsewhere [13]. Women who had fasting blood sugar between 100 and 125 or 2-h blood glucose between 140 and 199 mg/dL were considered as pre-diabetics. These women were entered into the first and second phases of TLGS and they were followed-up to the fourth phase. The follow-up lasted for 10-12 years. If at any phase, a person had a fasting blood sugar of 126 mg/dL or higher and/or 2-h blood glucose of 200 mg/dL or higher or taking blood glucose lowering medicine, she was diagnosed as a diabetic. The number of women who were recruited for the study from 803 persons were 734 who completed the study and their data were considered in data analysis [14]. The participants who left the study before the fourth phase died due to non-diabetes reasons or did not become diabetic by the fourth phase, were considered as a censor.

Initially, based on previous studies, 30 predictor variables were selected as important clinical onset diabetes variables. These variables were examined for the co-linearity presence. The chi-square and the t-tests depending on the type of the variable were used in order to compare risk factors at baseline between participants who become diabetic and those who remained normal by the end of the study. The SM and the BMA were used to determine the factors associated with the time of diabetes incidence in pre-diabetic women using Cox proportional hazard regression. The *Schoenfeld residuals were tested to assess the* proportional hazards assumption of the Cox model. *In addition, Schoenfeld residuals plots for all variables were plotted against log(time) in order to assess* the presence of non-random process [15].

The BMA uses the following equation posterior density for inference and prediction:

$$P(\theta|D) = \sum_{M_k \in A} P(\theta|M_k, D) P(M_k|D), \quad \text{(Eq. 1)}$$

where θ is a vector of parameters of interest and D represents the matrix of investigated data. The K proposed models are selected from the following set.

$$A = \{M_k : \frac{max_L\{P(M_k|D)\}}{P(M_k|D)} \leq c\},$$

where $M = \{M_1, M_2, \ldots, M_k\}$ set is selected by Occam's Window approach. The posterior distributions of $P(\theta/M_k, D)$ are averaged weighted on $\Pr(M_k|D)$ in order to find $p(\theta|D)$. Equation (2) states that the models with posterior probability of less than 1/c of the maximum posterior probability are excluded from the BMA process. The c is a constant value which is selected by the data analyst depending on the field of study. In this study, c was fixed at 20 as proposed by Madigan and Raftery [16]. In the BMA, the variables selection were performed based on the posterior probability from variables with non-zero parameter estimates. For each parameter, $P(\beta_i \neq 0|D)$ was obtained by summing up the posterior probabilities of models which include predictor variable (X_i). Predictor variables with $P(\beta_i \neq 0/D) \geq 0.5$ were kept in the model [17].

The SM and BMA were compared by predictive power after variable selection. Data of 734 participants in this study were randomly divided into two parts: 70 % for training and 30 % for model testing. The 177 women out of 513 participants in training dataset and 80 ones out of 221 in test dataset became diabetic. The Cox proportional hazard models were fitted by each of the BMA and the SM based on the training dataset and then were evaluated for the test dataset.

The predictive discrimination method was used to evaluate the models performance. For this purpose, firstly, all models of $A = \{M_k : \frac{max_L\{P(M_k|D)\}}{P(M_k|D)} \leq c\}$ set were fitted by the BMA in order to estimate the coefficients $(\hat{\theta}_k)$. The hazard scores $(x^T{}_i \hat{\theta}_k)$ were calculated for each person based on each fitted model using train dataset and then their weighted mean was calculated as:

$$\sum_{k=1}^{K} (x_i{}^T \hat{\theta}_k) pr(M_k|D^B) \quad \text{(Eq. 2)}$$

Based on the calculated risk scores from training dataset, cut point of low and high risk groups were considered 50th, 55th, 60th and 65th percentiles. The cut point was used for grouping participants in the test dataset. The predictive discrimination was obtained for the SM through the same way except that the calculations were conducted only for one model. The purpose of the predictive discrimination is that how correctly the model arranges the individuals into discrete low and high risk groups [18]. In addition, the Kaplan-Meier curve and the log-rank test were obtained for the BMA and the SM. The statistical analyses were performed using R software (version 3.1.3) [19].

3. Results

In this study, 257 (25 %) of 734 pre-diabetic women became diabetic by end of the study. The median time of diabetes incidence was 3204 days (8.78 years). Basic characteristics of the participants are shown in Table 1. The S*choenfeld residuals curve and test showed that the proportional hazard assumption is established for the Cox regression model* (P-value = 0.067*).*

In this study, 30 predictor variables were examined to include in the Cox model. To do this in an ordinary manner and considering that each variable can be selected or not, a set of 236 possible models had to be tested. The BMA using Leaps and bound algorithm [20] and Occam's window [21] method reduced the number of models into 68 models. Weighted averaging was done on these models based on the models posterior probabilities. Among them, five models with highest posterior probability were selected which are shown in Table 2.

These top five models calculated 30.11 % of the total posterior probability, with the highest probability of 12.7 %. These results definitely indicated the models uncertainty.

The BMA selected predictor variables that affect the incidence of diabetes among pre-diabetic women are: age, fasting blood sugar, 2-h blood glucose, diabetes family history, hospitalization history during past 3 months and body mass index (Table 3). BSM selected three additional variables of taking aspirin, enzyme drugs and diuretics. Furthermore, FSM selected waist-to-height ratio in addition to both models.

Risk scores were calculated from training dataset, and cut point of low and high risk groups were set as 50^{th}, 55^{th}, 60^{th} and 65^{th} percentiles. In all *cases, the* BMA improved predictive performance but merely the results of 65^{th} percentile are shown in Table 4.

Table 1. Baseline characteristics of pre-diabetic women aged>=20 for newly diagnosed diabetics and non-diabetics

	Diabetic (n=257)	**Non-Diabetic (n=477)**
Quantitative variables		
2-h BG	151.1(25.8)*	137(26.5)†
FBS	103.8(10.3)	97.6(9.1)†
TC	229.4(44)	220.1(42.4)‡
The ratio of TG to HDL	4.8(2.5)	4.2(2.4)‡
The ratio of TC to HDL	5.5(1.4)	5.1(1.4)‡
Cr	.98(.11)	.97(.11)
Pulse pressure	45.3(14)	41.8(13.4)‡
DBP	82.6(9.9)	79.8(9.9) †
WHtR	.62(.06)	.58(0.07) †
BMI	30.8(4.4)	28.6(4.1) †
HDL	42.8(9.5)	44.5(10.6)‡
Age	48.1(10.5)	44.7(11.7) †
Survival time	1725(1029.8)	3315.9(676.4)†
Qualitative variables		
Family history of diabetes	113(44)**	135(28.3) †
History of hyperlipidemia	99(38.5)	119(24.9)†
History of hospitalization	229(89.1)	420(88.1)
History of hospitalization in the last 3 months	4(1.6)	10(2.1)
History of pregnancy hypertension	21(8.2)	40(8.4)
Baby more than 4.5	39(15.2)	42(8.8)‡
History of cardiovascular disease	27(10.5)	52(10.9)
Having goiter	52(20.2)	97(20.3)
The use of diuretics drugs	18(7)	12(2.5)‡
The use of thyroid drugs	21(8.2)	22(4.6)
Use of enzymatic channel blockers drugs	6(2.3)	10(2.1)
Aspirin	37(14.4)	36(7.5) ‡
Smoking in the past	4(1.6)	9(1.9)
Smoking	7(2.7)	11(2.3)
Less than a high school education	147(57.2)	204(42.8)†
Marital status	34(13.2)	57(11.9)
Having a thyroid nodule	22(8.6)	33(6.9)
Intervention to Control Hypertension	31(12.1)	44(9.2)

* Data for quantitative variables as mean (standard deviation) and **qualitative variables as number (percentage)
‡ P less than 0.05 † P less than 0.001 compared with diabetic group
DBP, diastolic blood pressure; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol; TG, triglycerides; FBS, fasting blood sugar; 2-h BG, 2-h blood glucose; Cr, Creatinine; BMI, Body mass index; WHtR, Waist-to-Height

Table 2. Five models with the highest posterior probability selected from 68 models using BMA

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Age	T	T	T	---	---
FBS	T	T	T	T	T
Family history of diabetes	T	T	T	T	T
Aspirin	---	---	T	---	---
Hospitalization history during past 3 months	T	---	T	---	T
2-h BG	T	T	T	T	T
Use of enzymatic channel blockers drugs	---	T	---	---	---
BMI	T	T	T	---	---
Waist-to-Height	---	---	---	T	T
No. of variables	6	6	7	4	5
Posterior model probability	0.127	0.056	0.043	0.039	0.035
BIC	-169.17	-167.53	-167	-166.83	-166.60

T is true (variable selected)

Table 3. Results of BMA and SM for the onset of type II diabetes in pre-diabetic women: TLGS

Variable	Bayesian Model Averaging			BSM	FSM
	Posterior mean	Posterior sd	P($\beta_i \neq 0 \mid D$)	Coef (se)	Coef (se)
2-h BG	.019	.002	1	0.019(0.002)	0.019(0.002)
FBS	.056	.006	1	0.056(0.006)	0.056(0.006)
Family history of diabetes	0.492	.130	1	0.514(0.128)	0.512(0.128)
BMI	0.045	.028	.766	0.061(0.014)	0.058(0.025)
Hospitalization history during past 3 months	-.813	.724	.67	-1.144(0.585)	-1.136(0.587)
Age	.011	.009	0.634	0.017(0.006)	0.016(0.006)
WHtR	0.863	1.626	.243		0.263(1.673)
Use of enzymatic channel blockers drugs	-.163	.403	0.181	-0.891(0.493)	-0.886(0.494)
Education	0.050	.127	0.169		
Aspirin	.050	.146	0.131	0.415(0.187)	0.413(0.188)
Pulse pressure	0.0004	.002	0.059		
Marital status	0.016	.081	0.058		
Use of thyroid drugs	.018	.099	0.047		
The use of diuretic drugs	0.010	.077	0.031	0.392(0.264)	0.392(0.264)
Baby more than 4.5	.001	.023	.011		
Goiter status	.001	.021	.011		
DBP	0.00005	.0008	0.011		
HDL	-.00006	.0009	0.011		
Cr	-.003	.068	.01		
TC	-.00001	.0001	.01		
Former smokers	-.002	.057	.01		
History of hospitalization	-.0009	.021	.009		
History of thyroid nodules	0.001	0.024	0.009		
Cigarette	.001	.038	.009		
History of cardiovascular disease	.0001	.017	0.008		
Prevention for control hypertension	-.0003	.019	.008		
History of hyperlipidemia	.0001	.012	.008		
History of pregnancy hypertension	-.0006	.022	.008		
The ratio of TC to HDL	0.00001	.003	.008		
The ratio of TG to HDL	-.00004	.002	.008		

Table 4. Classification of diabetes in low-risk and high-risk groups by predictive discrimination

Method	Risk group	Non-diabetic	Diabetic
BMA			
	Low risk	106(75.18)	31(38.75)
	High risk	35(24.82)	49(61.25)
BSM			
	Low risk	103(73.05)	37(46.25)
	High risk	38(26.95)	43(53.75)
FSM			
	Low risk	104(73.76)	38(47.5)
	High risk	37(26.24)	42(52.5)

It presented that the BMA indicated higher predictive discrimination in comparison to BSM and FSM. The log-rank test showed a significant difference between low and high risk groups both in the BMA and the SM (P-value<0.001).

This means that pre-diabetic women have been allocated to distinct risk groups. In addition, Kaplan-Meier curve indicated the distinction between low and high risk women (Figure1-3).

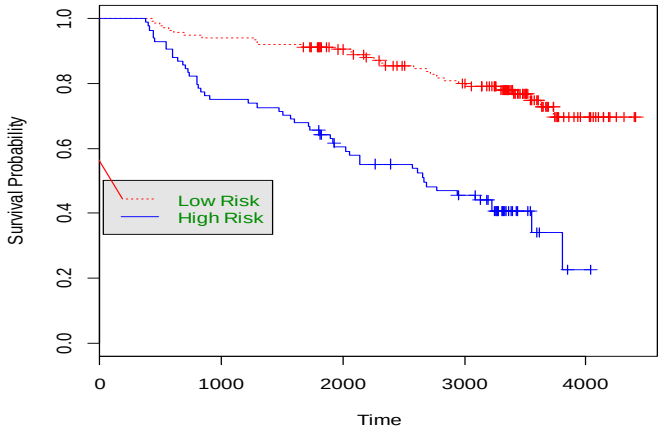

Figure1. Kaplan-Meier plot for the onset of type II diabetes in Bayesian model averaging

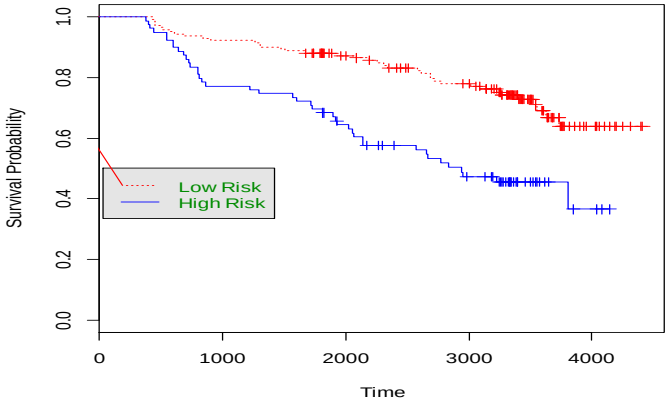

Figure 2. Kaplan-Meier plot for the onset of type II diabetes in the FSM

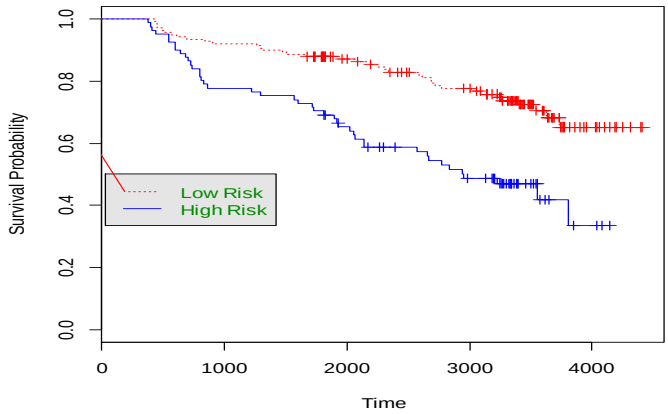

Figure 3. Kaplan-Meier plot for the onset of type II diabetes in the BSM

10 and 6 predictor variables on time of diabetes incidence, respectively.

In this study, the SM variable selection based on the AIC and the BMA were investigated in the Cox regression model. SM procedures involved three basic steps: identifying an initial model; iteratively 'stepping': repeatedly altering the model at the previous step by adding or removing a predictor variable in accordance with the 'stepping criteria' and terminating the search when stepping is no longer possible, given the stepping criteria (AIC). The BMA showed better prediction performance compared to the SM despite the smaller number of variables selection.

4. Discussion

In this study, five models with the highest posterior probabilities were selected by the BMA. They showed that several models with non-negligible posterior probabilities exist but yield different results. It indicates the significance of the model uncertainty. Therefore, the BMA could be useful in this case [22]. In fact, its solution for the model uncertainty is investigating the quantity of interest using the subset of supported models by data and then performing statistical inference using the weighted average of models posterior distributions [23-25]. This model presents the numerical size of the competing models desirability [18] and overcomes the weakness of the individual models. Thus, it must be reliable and accurate [7, 26] . In this paper, the uncertainty between models was considered using the BMA and a small number of variables were selected at the same time, leading to higher accuracy prediction.

The BMA uses Leaps and bound algorithm to reduce the number of models. This algorithm does not check all the possible models. Thus, there is still some uncertainty that has not been calculated; this sort of uncertainty is unavoidable. Nevertheless, because it tries to find all the important models, this additional uncertainty is negligible [10]. The difference between the BMA and the SM is that the BMA results for possible effects are adjusted for all the predictor variables [7]. This model presents a measure of model uncertainty through the posterior probability [27] and it also prevents the dual mode selection for variable effectiveness using continuous scale (0-1) and considers the model uncertainty; hence, it increases the models validity [23]. Doming compared the BMA with the model combination methods and showed that BMA has worse performance [28]. This conclusion is unfitting because the BMA is not an algorithm for the model combination. The model combination works through powering assumption space, and not through estimation of the BMA [29].

In this study, the Kaplan-Meier curves showed higher accuracy prediction of the BMA. A disadvantage of the SM is the instability of variable selection. Altman and Andersen used one hundred Bootstrap samples of a dataset in order to investigate the stability of Cox model selection process. They found that the SM Cox model resulted in different sets of independent variables. They concluded the instability of variable selection using the SM for Cox model [30]. In the SM, the deduction is based on type I and II errors. The SM may lead to the overestimation in coefficients and underestimation of the standard deviation and therefore incorrect P-values [8]. Thus, this method does not consider the models uncertainty [9]. On the other hand, it should be noted that censorship in survival regression models may increase the uncertainty.

Variable selection in the BMA is based on the probability of a variable posterior effect. Thus, it solves the binary problem (variable selection or non-selection). The results of this study showed that in the BMA, the predictor variables of age, fasting blood sugar, 2-h blood glucose, family history of diabetes, BMI, and hospitalization history during past 3 months can effect on time of

diabetes incidence in pre-diabetic women. These variables were selected by SM.
In BMA, BMI was selected as risk factor, while waist-to-height was not. However, the probability of posterior effect for waist-to-height is 0.243. This value showed that the waist-to-height variable in the presence of stronger variable of BMI was not selected. The FSM manifested that in addition to the BMI, the waist-to-height variable was effective on the time of diabetes incidence but the BSM was selected only for the BMI. The adjusted BMA is in the presence of all the variables and it presents the effective intensity of all the variables in the continuous scale. Thus, clinically, it is a flexible model. The Leaps and bounds algorithm in the BMA is not desirable for more than 30 variables, and this is one of the limitations of this method.

5. Conclusion

In the present study, the BMA shows nearly a constant accuracy despite the selection of lower predictor variables number in comparison to the SM. Consequently, it seems that the BMA has presented better performance for the evaluation of the predictor variables on the time of diabetes incidence.

Acknowledgments

The data of this study is taken from Tehran Lipid and Glucose Study that has been performed by the Research Institute for Endocrine Science of Shahid Beheshti University of Medical Sciences. The authors would like to thank all those who were involved in the design and collection of TLGS data as well as the participants. This project was approved by the ethics committee of Shahid Beheshti University of Medical Science with the code IR.SBMU.REC.1396.213.

Conflict of Interest

The authors declare no conflict of interest.

References

1.Keshavarz M, Cheung NW, Babaee GR, Moghadam HK, Ajami ME, Shariati M. Gestational diabetes in Iran: incidence, risk factors and pregnancy outcomes. Diabetes research and clinical practice. 2005;69(3):279-86.
2.Rashidi M, Afkhami-Ardakani M. Risk factors for type 2 diabete. Journal of Shahid Sadoughi University of Medical Sciences 2011(19):266- 80.
3.Azimi-Nezhad M, Ghayour-Mobarhan M, Parizadeh M, Safarian M, Esmaeili H, Parizadeh S, et al. Prevalence of type 2 diabetes mellitus in Iran and its relationship with gender, urbanisation, education, marital status and occupation. Singapore medical journal. 2008;49(7):571.
4.Roche MM, Wang PP. Factors associated with a diabetes diagnosis and late diabetes diagnosis for males and females. Journal of Clinical & Translational Endocrinology. 2014;1(3):77-84.
5.Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes estimates for the year 2000 and projections for 2030. Diabetes care. 2004;27(5):1047-53.
6.Hadaegh F, Shafiee G, Ghasemi A, Sarbakhsh P, Azizi F. Impact of metabolic syndrome, diabetes and prediabetes on cardiovascular events: Tehran lipid and glucose study. Diabetes research and clinical practice. 2010;87(3):342-7.
7.Annest A, Bumgarner RE, Raftery AE, Yeung KY. Iterative bayesian model averaging: A method for the application of survival analysis to high-dimensional microarray data. BMC bioinformatics. 2009;10(1):1.
8.Wieagand RE. Performance of using multiple stepwise algorithms for variable selection. Statistics in medicine. 2010;29(15):1647-59.
9.Amini SM, Parmeter CF. Bayesian model averaging in R. Journal of Economic and Social Measurement. 2011;36(4):253-87.
10.Volinsky CT. Bayesian model averaging for censored survival models: University of Washington; 1997.
11.Genell A, Nemes S, Steineck G, Dickman PW. Model selection in medical research: a simulation study comparing Bayesian model averaging and stepwise regression. BMC medical research methodology. 2010;10(1):108.
12.TAY PL. Iterative Bayesian Model Averaging For Patients Survival Analysis: Universiti Teknologi Malaysia; 2010.
13.Azizi F, Madjid M, Rahmani M, Emami H, Mirmiran P, Hadjipour R. Tehran Lipid and

Glucose Study (TLGS): rationale and design. Iranian journal of endocrinology and metabolism. 2000;2(2):77-86.
14.Bagherzadeh-Khiabani F, Ramezankhani A, Azizi F, Hadaegh F, Steyerberg EW, Khalili D. A tutorial on variable selection for clinical prediction models: feature selection methods in data mining could improve the results. Journal of clinical epidemiology. 2016;71:76-85.
15.Kleinbaum DG, Klein M. Survival analysis: a self-learning text. Springer Science & Business Media; 2006.
16.Madigan D, Raftery AE. Model selection and accounting for model uncertainty in graphical models using Occam's window. Journal of the American Statistical Association. 1994;89(428):1535-46.
17.Raftery AE. Bayesian model selection in social research. Sociological methodology. 1995:111-63.
18.Volinsky CT, Raftery AE. Bayesian information criterion for censored survival models. Biometrics. 2000;56(1):256-62.
19.Raftery AE, Painter IS. BMA: an R package for Bayesian model averaging. R news. 2005;5(2):2-8.
20.Furnival GM, Wilson RW. Regressions by leaps and bounds. Technometrics. 2000;42(1):69-79.
21.Raftery A, Hoeting J, Madigan D. Model selection and accounting for model uncertainty in linear regression models. Citeseer; 1993.
22.Volinsky CT, Madigan D, Raftery AE, Kronmal RA. Bayesian model averaging in proportional hazard models: assessing the risk of a stroke. Journal of the Royal Statistical Society: Series C (Applied Statistics). 1997;46(4):433-48.
23.Rentsch C, Bebu I, Guest JL, Rimland D, Agan BK, Marconi V. Combining epidemiologic and biostatistical tools to enhance variable selection in HIV cohort analyses. PloS one. 2014;9(1):e87352.
24.Lipkovich IA. Bayesian model averaging and variable selection in multivariate ecological models. 2002.
25.Noble Jr RB. Multivariate applications of Bayesian model averaging: Citeseer; 2000.
26.Kim IK, Lee JY, Kwon JK, Park JJ, Cho KS, Ham WS, et al. Prognostic factors for urachal cancer: a bayesian model-averaging approach. Korean journal of urology. 2014;55(9):574-80.
27.Minka TP. Bayesian model averaging is not model combination. Available electronically at http://www stat cmu edu/minka/papers/bma html. 2000:1-2.
28.Domingos P, editor. Bayesian averaging of classifiers and the overfitting problem. ICML; 2000.
29.Domingos PM, editor. Why Does Bagging Work? A Bayesian Account and its Implications. KDD; 1997. Citeseer.
30.Altman DG, Andersen PK. Bootstrap investigation of the stability of a Cox regression model. Statistics in medicine. 1989;8(7):771-83.