Journal of Data Science 16(2017), 575-588

USING INFORMATIVE PRIOR FROM META-ANALYSIS IN BAYESIAN APPROACH

Esin AVCI

Faculty of Art and Science, Department of Statistics, Giresun University, Giresun, Turkey

*Abstract:*In a Bayesian approach, uncertainty explained by a prior distribution that contains information about an uncertain parameter. Determination of the prior distribution is important in because it impacts the posterior inference. The objective of this study is to use meta-analysis for proportion to obtain prior information about patients with breast cancer stage I who undergoing modified radical mastectomy treatment and applied Bayesian approach. R and WinBUGS programs are performed for meta-analysis and Bayesian approach respectively.

Keywords: Bayesian approach, meta -analysis, informative prior, binomial, pooled proportion.

1. Introduction

Until the late 1980s, Bayesian approach was taken into account as an alternative to the frequentists approach. At the beginning of the 21st century, Bayesian approaches become popular in science. The main tool of Bayesian approach is probability theory. The probability considered as a measure of the frequency of repeated events by frequentists, while Bayesians consider as a measure of the degree of

certainty. Since the frequentists consider parameters to be fixed and data to be random, while Bayesians consider parameters to be random and data to be fixed. The parameters are characterized by a prior distribution. Frequentists throwing prior information away. In Bayesian approach, statistical inference is based on the posterior distribution of the parameters which obtained by combining the prior distribution and likelihood. Because of the difficulty of the calculation of the posterior distribution, the Bayesian approach was not handled as a well-accepted approach for data analysis. [1] In the past few years, to solutions problems asymptotic methods have been developed by using computer algorithms (e.g. Gibbs sampler, Metropolis-Hasting algorithm, and MCMC) to draw a random sample from the posterior distribution, without having to completely evaluate it.

Determination of the prior distribution is important because it impacts the posterior inference. If prior information is available, it should be appropriately summarized by the prior distribution. Such distributions are called informative distribution. In the case of no prior information is available, we define prior as a will not affect the posterior distribution. Such distributions are called non-informative distribution [1]. The importance of prior distribution reveals when the sample size is small, or the data supply only indirect information about the parameters [3].

A suggested approach is to obtain informative prior and produce quantitative results from published studies, meta-analysis can be conducted. Meta-analysis refers to the statistical synthesis of results from a series of studies. If the studies have been collected systematically, the synthesis will be meaningful. Provide more powerful test, summarize numerous and inconsistent findings and investigate the consistency of effect across different samples are the reasons of using meta-analysis [2].

Miller et al. (2009) used Bayesian adaptation of the summary roc curve method for meta-analysis of diagnostic test performance [4].

The objective of this study is to use meta-analysis for proportion to obtain prior information about patients with breast cancer stage-I who undergoing modified radical mastectomy treatment and applied Bayesian approach.

2. Bayesian Approach

Bayesian approach is derived from the application of Bayes' theorem, which was developed by Thomas Bayes in the 1700s. The principle of Bayesian approach is using probabilities that are conditional on data to explain beliefs about parameters. The Bayesian approach also combines past knowledge into the analysis, and so it can be viewed as the updating of prior beliefs with current data.

For θ parameters and data Bayes' theorem is expressed as

$$g(\theta|data) = \frac{g(\theta)\times f(data|\theta)}{\int g(\theta)\times f(data|\theta)d\theta} \quad (1)$$

A closed form of (1) existed only in a few simple cases, such as for a normal sample with a normal prior. In other cases, the numerical solution required had to be done for integration. In the past few years, computer algorithms (e.g. Gibbs sampler, Metropolis-Hasting algorithm, and MCMC) have been developed to draw a random sample from the posterior distribution. Depending on choosing the prior, the generated posterior distribution can impact either strongly (subjective or informative prior) or minimally (objective or non-informative prior).

Non-informative priors have a minimal impact on the posterior distribution. Most used non-informative priors are Uniform and Jeffreys' priors. Uniform prior distribution is frequently used in Bayesian approach because they yield non-informative priors and proper posterior distributions. Uniform prior that gives equal weight to all possible values. Jeffreys' prior (Jeffreys 1961) based on the observed Fisher information matrix. Because it is locally uniform, it is a non-informative prior. It is a useful prior because it doesn't change much over the region in which the likelihood is significant and doesn't have large values outside that range the local uniformity property. If the prior and posterior distributions are from the same family, such prior called conjugate prior. Conjugate priors result in closed-form solutions for the posterior distribution, enabling either direct inference or the construction of efficient Markov chain Monte Carlo sampling algorithms [5].

An informative prior dominates the likelihood, and thus it has a visible impact on the posterior distribution. The informative prior can be obtained by using the knowledge about parameters based on other data or expert opinion. This information should be appropriately summarized by the prior distribution [5].

3. Meta-Analysis for Breast Cancer

Meta-analysis collect information from multiple independent studies in order to obtain an average estimate. Depending on the statistic to be reported Different meta-analysis method exist. Examples of statistics of interest include association measures such as one-dimensional binomial or continuous measures such as proportions or means risk difference, the difference in means, odds ratio, risk ratio. There are three main aspects in meta-analysis: a) create the analysis framework, b) decide the model (fixed or random) and c) the choice of the method to estimate the heterogeneity parameter [2].

Fixed-effects model assume that the parameter of interest is identical across studies and the difference between the observed proportion and the mean is only due to sampling error. In the random-effects model, the observed difference between the proportions and the mean consist of sampling error and other factors such as differences in study population, study designs, etc. Each study estimates a different parameter, and the pooled estimate describes the mean of the distribution of the estimated parameters. The variance parameter describes the heterogeneity among the studies and in the case where the variance is zero, this model simply reduces to the fixed-effects model [2].

There are three approaches to the modeling of binomial data. First, using transformations (logit, arcsine) to an approximation to the normal distribution and is known as the approximate likelihood approach [6]-[8]. The second approach identifies the true nature of the data and is known as the exact likelihood approach. In this approach, the special relationship between the mean and the variance as characterized by binomial data is captured by the binomial distribution [9]. The beta-binomial distribution [10] can be used to fit

a random-effects model such that the beta distribution describes the distribution of the varying binomial parameters. The third approach is a compensation between approximate and exact likelihood.

4. Data Analysis and Discussion

In this section, an empirical application to determine the proportion of patients with breast cancer stage-I who undergoing modified radical mastectomy treatment in Bayesian approach was considered.

To obtain the posterior distribution of the proportion of the patients with breast cancer stage-I who undergoing modified radical mastectomy treatment, first, non-informative priors (flat and Jeffreys' priors) and informative priors used, respectively. In any case, it is most convenient to represent prior to a beta distribution for a single proportion.

Second, the likelihood of the data x needs to be constructed. Since the observed events (patients with breast cancer stage-I who undergoing modified radical mastectomy treatment) ri are assumed to have a binomial distribution with parameters pi and sample size ni,

$$ri \sim binomial(pi, ni) \quad (2)$$

The non-informative priors was selected as Beta(1,1) (flat) and Beta(0.5,0.5) (Jeffreys').

Informative prior was determined by conducting a meta-analysis. Three electronic databases EMBASE, PubMed, and Cochrane Controlled Trials Register Databases were searched for studies on patients with stage-1 breast cancer who were treated with modified radical mastectomy. The search for studies was restricted to studies published in English-Turkish language. The following search terms were used: (early stage breast cancer OR early stage breast carcinoma) AND (mastectomy OR modified radical mastectomy OR radical mastectomy). Studies with MRM were considered for inclusion if they reported raw data available for stage knowledge and a clearly defined type of surgery. Studies published in other language were excluded. Women with early (stage-I) breast cancer undergoing

MRM who did not have a history of prior cancer and evidence of metastatic disease were included.

By database search strategy yielded a total of 932 candidate abstracts. 19 satisfying inclusion criteria. In all, 19 studies including a total of 22733 patients. The analysis was carried out using R software.

In the systematic review, the random effects model (because of studies combined from published literature) with the exact likelihood approach (p) measure of treatment effect was used throughout. The meta-analysis contained 19 studies having large sample sizes ranging between 30 and 14249. For calculating the weights inverse variance method and to estimate the between-study variance (τ^2), DerSimonian-Laird method was used. The result was summarized in Forest plot which was given in Figure 1.

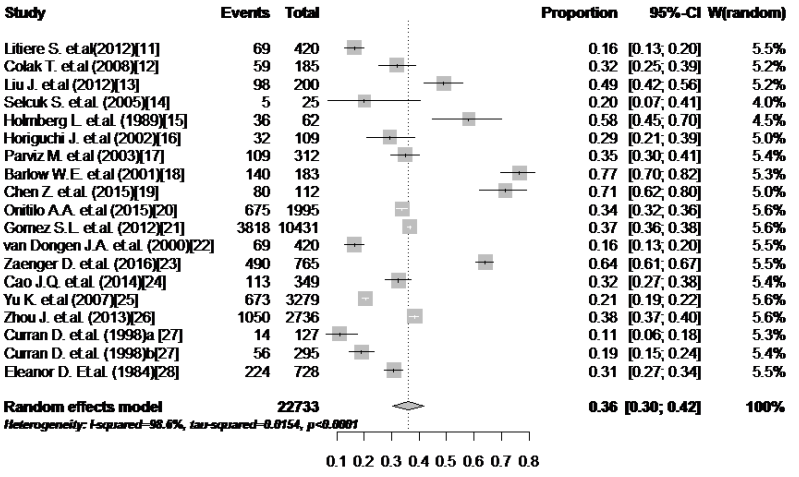

Figure 1: Forest plot for stage-I

Forest plot in a meta-analysis can principally be divided into six columns. Individual study results are displayed in rows. The first column ("study") lists the individual study ID's included in the meta-analysis. Second and third columns represent the number of outcomes and a total number of participants. The fourth column visually displays the study results. The size of the box is directly related to the

"weighting" of the study in the meta-analysis. The weighting ranging between 4 and 5.6.The horizontal line (Whiskers) through the boxes depict the length of the confidence interval. The longer the lines, the wider the CI, the less precise the study results. The diamond in the last row of the graph illustrates the pooled result of the meta-analysis. The width of the diamond depicts the width of the pooled CI. The pooled proportion was 0.36 (95% CI, 0.30 to 0.42), At the bottom of the graph on the left hand the most reliable test for heterogeneity. It ranges between 0 and 100% measures the variability between studies in the meta-analysis. The results indicated significant heterogeneity between study (Q=1309.77, p<0.0001) with $I^2 = 98.6\%$, which means that 99% of observed variance comes from real differences between studies.

To assess the publishing bias, Egger's bias coefficients for proportion was selected. The Egger's test was 1.25 and p-value =0.69 greater than 0.05 so the test did not suggest the presence of bias.

A flexible choice of a prior distribution for a Binom probability is Beta(α,β) The mean of a Beta distribution is $\alpha/(\alpha+\beta)$ and variance $\alpha\beta/[(\alpha+\beta)^2(\alpha+\beta+1)]$. In random effect model the pooled estimate describes the mean of the distribution of the estimated parameters. The pooled proportion was 0.36 and variance 0.0164 and was solved for α and β. The results was showed that Beta(4.73,8.35) was sensible.

To assessment of the influence of priors on posterior distribution, randomly observed events for different sample size (10, 50, 100 and 1000) was simulated. The analysis was done by using the commercially available WinBugs software employing the Markov chain Monte Carlo (MCMC) methodology. The Bugs language allows a concise expression of the parametric model to denote stochastic relationships and to denote deterministic relationships. Thus, using an MCMC method of parameter estimation with informative and non-informative priors, one is able to obtain the posterior estimates and credible regions of estimates of these effects. Graphical displays of convergence history and posterior densities affirm the stability of the results.

Table 1: Posterior estimates, credible regions and DIC

Prior Distribution	*Mean*	*SD*	*MC error*	*2.5%*	*97.5%*	*DIC*
For x=2 and n=10						
Beta(1,1)	0.25	0.12	0.0008	0.06	0.52	4.026
Beta(0.5,0.5)	0.23	0.12	0.0008	0.04	0.51	4.258
Beta(4.73,8.35)	0.29	0.09	0.0006	0.13	0.49	3.562
For x=16 and n=50						
Beta(1,1)	0.33	0.06	0.0004	0.21	0.46	6.277
Beta(0.5,0.5)	0.32	0.07	0.0004	0.20	0.46	6.336
Beta(4.73,8.35)	0.33	0.06	0.0004	0.22	0.45	5.888
For x=26 and n=100						
Beta(1,1)	0.26	0.04	0.0003	0.18	0.35	6.777
Beta(0.5,0.5)	0.26	0.04	0.0003	0.18	0.35	6.811
Beta(4.73,8.35)	0.27	0.04	0.0003	0.19	0.36	6.590
For x=173 and n=1000						
Beta(1,1)	0.17	0.01	0.00009	0.15	0.20	8.745
Beta(0.5,0.5)	0.17	0.01	0.00009	0.15	0.20	8.747
Beta(4.73,8.35)	0.18	0.01	0.00009	0.15	0.20	8.734

Table 1 presented posterior estimation and credible regions with Beta priors. To determine the number of iterations to obtain samples that can be used for posterior inference. By saving the more samples, the more accurate will be obtained in posterior estimation. One way to assess the accuracy of the posterior estimates is by calculating the MC error for each parameter. As a rule of thumb, the simulation should be run until the MC error for each parameter of interest is less than about 5% of the sample standard deviation (SD). The iterating process was carried beyond 25,000 that the estimates proved to be very stable. To

reduce the potential bias, the first 2,500 samples as burn-in was discarded. Table 1 showed that, as sample size increases the effect of prior on posterior distribution played a much smaller role. For all condition the informative prior had smaller Deviance Information Criteria (DIC), that means better supported by data.

Figure 2 displayed posterior densities and convergence history for all three priors. Note that in the case of sample size was small, only informative prior of the proportion had symmetry in the posterior densities. The history plots looked like nice oscillograms around a horizontal line without any trend. The Markov chain was most likely to be sampling from the stationary distribution and is mixing well.

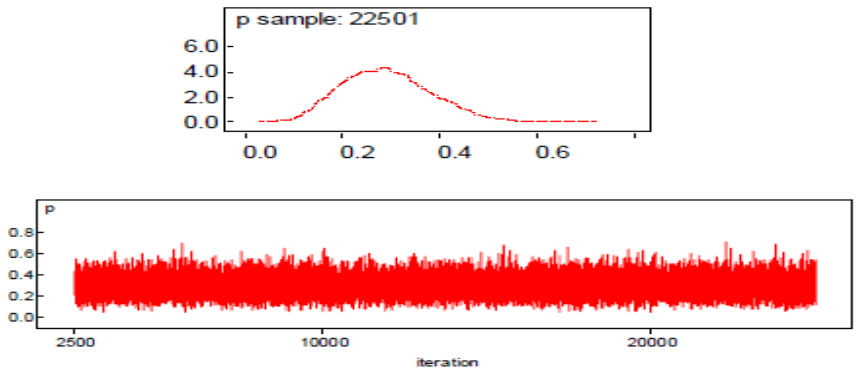

Figure 2 Marginal posterior densities and history plots of proportion

4. Conclusions

One of the advantages of Bayesian approach on frequentist approach is to use both source of information: the prior information and the information about the process included in the data. In the case of no prior information is available, the prior should be defined as a minimal impact on the posterior distribution. If prior information is available, it should be appropriately summarized by the prior distribution.

A suggested approach is to obtain informative prior and produce quantitative results from published studies. To achieve this aim meta-analysis was conducted. By conducting meta-analysis more powerful and consistency of prior information was obtained.

The probability of patients with breast cancer stage-I who undergoing modified radical mastectomy treatment was considered. Informative prior was obtained via meta-analysis for proportion. To reveal the effectiveness of informative prior, Bayesian approach was applied to different sample size (10, 50, 100 and 1000) of breast cancer patients who undergoing modified radical mastectomy treatment with randomly observed stage-I patients and was compared with non-informative prior results. The results showed that, especially for small sample size, informative prior had smaller Deviance Information Criteria (DIC), that means better supported by data.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

References

[1] I. Ntzoufras, Bayesian Modeling Using WinBUGS, Hoboken-New Jersey-USA, Wily&Sons, 2009.

[2] M. Borenstein, L.V. Hedges, J. P.T. Higgins, H.R. Rothstein, Introduction to Meta-Analysis, West Sussex-UK Wily&Sons, 2009.

[3] A. Gelman, Prior Distribution, Encyclopedia of Environmetrics, 3, Chichester, (2002), 1634-1637.

[4] S. W. Miller, D. Sinha, E. H. Slate, D. Garrow, J. Romagnuolo, Bayesian Adaptation of the Summary ROC Curve Method for Meta-analysis of Diagnostic Test Performance, Journal of Data Science, 7(3), (2009), 349-364.

[5] W. M. Bolstad, Introduction to Bayesian Statistics, 2nd Edition, Wiley-Interscience, A John Wiley&Sons, Inc., Publication, 2007.

[6] M. E. Glickman, D. A. van Dyk, Basic Bayesian Methods, Methods in Molecular Biology, vol. 404: Topics in Biostatistics, W. T. Ambrosius © Humana Press Inc., Totowa, NJ, (2007), 319-388.

[7] A. Agresti, B. A. Coull, Approximate is better than 'exact' for interval estimation of binomial proportions, Am Stat, 52(2), (1998),119–126.

[8] N. E. Breslow, D. G. Clayton, Approximate inference in generalized linear mixed models, J AmStat Assoc, 88, (1993), 9–25.

[9] J. J. Miller, The inverse of the Freeman-Tukey double arcsine transformation, Am Stat, 32(4), (1978),138.

[10] T. H. Hamza, H. C. van Houwelingen, T. Stijnen, The binomial distribution of meta-analysis was preferred to model within-study variability, J Clin Epidemiol, 61, (2008), 41–51.

[11] G. Molenberghs, G. Verbeke, S. Iddib, C. G. B. Demétrio, A combined beta and normal random effects model for repeated, over-dispersed binary and binomial data, JMultivar Anal, 111, (2012), 94–109.

[12] S. Litière, G. Werutsky, I.S. Fentiman, E. Rutgers, M.R. Christiaens, E.V. Limbergen, M.H.A. Baaijens, J. Bogaerts, H. Bartelink. Breast conserving therapy versus mastectomy for stage I–II breast cancer: 20-year follow-up of the EORTC 10801 phase 3 randomized trial, Lancet Oncol, 13, (2012), 412-419.

[13] T. Colak, E. Alimoglu, A. Mesci, E. Pestereli, A. Kabaalioglu, S. Karaveli, M. Akaydın. Meme kanserli kadınlarda cerrahi yöntemin seçimini etkileyen faktörler, Meme Sağlığı Dergisi, 4(1), (2008), 29-33.

[14] J.J. Liu, S. Zhang, X. Hao, J. Xie, J. Zhao, J. Wang, L. Liu, P.P. Wang, J. Zhang. Breast-conserving therapy versus modified radical mastectomy: Socioeconomic status determines who receives what—Results from case–control study in Tianjin, China, Cancer Epidemiology, 36, (2012), 89–93.

[15] S. Selcuk , N. Zalluhoglu, A. Gurkan, S. Kacar, S. Kilic, C. Karaca, C. Varilsuha, Effect of breast preserving surgery in the treatment of early breast cancer (retrospective analysis), Turkish Journal of Surgery, 21(3), (2005), 135-140.

[16] L. Holmberg, M. Omne-Ponten, T. Burns, H. 0. Adami, R. Bergstrom, Psychosocial Adjustment After Mastectomy and Breast-Conserving Treatment, Cancer, 64, (1989), 969-974.

[17] J. Horiguchi, Y. lino, Y. Koibuchi, T. Yokoe, H. Takei, M. Yamakawa, T. Nakajima, T. Oyama, T. Ando, T. Ishida, K. Endo, Y. Takai, H. Suzuki, T. Fujii, T. Yokomori, Y. Morishita. Breast-conserving therapy versus modified radical mastectomy in the treatment of early breast cancer in japan, Breast Cancer,9(2), (2002),160-165.

[18] M. Parviz, J.B. Cassel, B.J. Kaplan, S.E. Karp, J.P. Neifeld, L.T. Penberthy, H.D. Bear, Breast conservation therapy rates are no different in medically ındigent versus ınsured patients with early stage breast cancer. Journal Of Surgical Oncology, 84, (2003), 57–62.

[19] W.E. Barlow, S.H. Taplin, C.K. Yoshida, D.S. Buist, D. Seger, M. Brown. Cost comparison of mastectomy versus breast-conserving therapy for early-stage breast cancer, Journal of the National Cancer Institute, 93(6), (2001), 447-455.

[20] Z. Chen, Y. Xu, J. Shu, N. Xu, Breast-conserving surgery versus modified radical mastectomy in treatment of early stage breast

cancer: A retrospective study of 107 cases, Journal of Cancer Research and Therapeutics, 11, (2015), 29-31.

[21] A.A. Onitilo, J.M. Engel, R.V. Stankowski, S.A.R. Doi, Survival Comparisons for Breast Conserving Surgery and Mastectomy Revisited: Community Experience and the Role of Radiation Therapy, Clinical Medicine & Research, 13(2), (2015), 65-73.

[22] S.L. Gomez, D.J. Press, D. Lichtensztajn, T.H.M. Keegan, S.J. Shema, G.M. Le, A.W. Kurian, Patient, Hospital, and Neighborhood Factors Associated with Treatment of Early-Stage Breast Cancer among Asian American Women in California, Cancer Epidemiol Biomarkers & Prevention, 21(5), (2012), 821–834.

[23] J.A. Van Dongen, A.C. Voogd, S.I. Fentiman, et al. Long-term results of a randomized trial comparing breast-conserving therapy with mastectomy: European Organization for Research and Treatment of Cancer 10801 Trial, J Natl Cancer Inst, 92, (2000), 1143–1150.

[24] D. Zaenger, B.M. Rabatic, B. Dasher, W.F. Mourad, Is Breast Conserving Therapy a Safe Modality for Early-Stage Male Breast Cancer?, Clinical Breast Cancer, 16(2), (2016), 101-104.

[25] J.Q. Cao, P.T. Truong, I.A. Olivotto, R. Olson, G. Coulombe, M. Keyes, L. Weir, K. Gelmon, V. Bernstein, R. Woods, C. Speers, S. Tyldesley, Should women younger than 40 years of age with invasive breast cancer have a mastectomy?:15-year outcomes in a population-based cohort. Int J Radiation Oncol Biol Phys, 90(3), (2014), 509-517.

[26] K. Yu, G. Di, J. Wu, J. Lu, K. Shen, Z. Shen, Z. Shao, Development and Trends of Surgical Modalities for Breast Cancer in China: A Review of 16-Year Data, Annals of Surgical Oncology, 14(9), (2007), 2502–2509.

[27] J. Zhou, L. Enewold, S.H. Zahm, I. Jatoi, C. Shriver, W.F. Anderson, D.D. Jeffery, A. Andaya, J.F. Potter, K.A. McGlynn, K. Zhu, Breast conserving surgery versus mastectomy: the, influence of comorbidities on choice of surgical operation in the Department of Defense health care system, The American Journal of Surgery, 206, (2013), 393-399.

[28] D. Curran, J.P. Van Dongen, N.K. Aaronson, G. Kiebert, I.S. Fentiman, F. Mignolet, H. Bartelin, Quality of life of early-stage breast cancer patients treated with radical mastectomy or breast-

conserving procedures: results of EORTC trial 10801, Eur J Cancer, 34(3), (1998), 307-314.

[29] D.M. Eleanor, C.A. Frederick, R.S. Sylvia, M.R. Marvin, Conservation surgery and irradiation as an alternative to mastectomy in the treatment of clinically favorable breast cancer, Cancer, 54, (1984), 2668-2672.