行政院國家科學委員會專題研究計畫 期中進度報告

乳癌篩檢貝氏決策分析(2/3)
期中進度報告(精簡版)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 95-2314-B-002-029-
執 行 期 間 ：95 年 08 月 01 日至 96 年 07 月 31 日
執 行 單 位 ：國立臺灣大學公共衛生學院預防醫學研究所

計 畫 主 持 人：陳秀熙
共 同 主 持 人：張淑惠

處 理 方 式 ：期中報告不提供公開查詢

中 華 民 國 96 年 05 月 29 日

中文摘要

前言　利用乳房攝影術進行大規模乳癌篩檢已被西方國家證實可達到 20%-30% 的乳癌死亡率降低，然而大規模篩檢是否適用於乳癌發生率不若西方國家高的亞洲地區則為一個爭議性的議題，是否改採用高危險篩檢？若是，該用怎樣的條件篩選所謂高危險群。本年度我們利用貝氏方法發展適用於一般族群的乳癌危險性計算公式，此外，亦利用貝氏理論發展同時考慮固定效應及隨機效應三階段馬可夫鏈模式以量化乳癌的疾病自然史。

材料與方法　利用羅吉斯迴歸模式可以產生乳癌危險計算公式，再利用病例與對照兩群人的危險分數分佈結合貝氏法則計算不同切點下可以達到的辨識危險比例，並據此決定最佳切點。此外，利用貝氏多層次分析技巧引入不同層次的隨機效應並檢視是否隨機效應的引入可以有效的改善模式表現。

結果　乳癌個案的得病危險分數平均值為-5.05（標準差為 3.14），非病例個案之得病危險分數平均值為-7.76（標準差為 3.75）。切點值與貝氏因子（辨識危險比）呈反比關係。貝氏隨機效應模式結果顯示第一胎足產年齡較晚為統計上具顯著意義的危險因子（OR=1.89，95%信賴區間：1.02-3.66），且隨機效應的存在具統計上顯著意義。

Abstract

Introduction Breast cancer screening with mammography has been demonstrated to reduce 20 to 30% mortality of breast cancer in western country for aged 40-69 (Nystrom et al, 1993; Tabar et al, 2000). Although incidence and mortality of breast cancer have increased over the past decade in many Asian countries, the incidence seems still too low to reach the criteria of mass screening for breast cancer with mammography compared with western country.

This year we developed a Bayesian method in terms of evaluation of different screening strategy in selecting individuals at risk to determine the criteria for the selection of general population at risk of developing breast cancer. In addition, a three-state Markov model including both fixed effect, random effect was developed to capture the heterogeneity caused by individual variation after adjusting for correlated property and measured covariates by Bayesian approach.

Materials and methods The high-risk group was first included the women aged 50-69 years who has the family history of first degree relative of breast cancer, and second defined by a risk score higher than a cutoff point. Risk score based on logistic regression model that incorporates all significant reproductive and menstrual factors obtained from data with mammographic and physical examination of the past community-based out-reaching screening program (1999~2001). We also proposed a Bayesian hierarchical multi-state model to tackle these problems with the incorporation of random-effect parameters. The model was applied to data from a high-risk group (family history of breast cancer) screening for breast cancer. There were 4,867 women with family history derived from 4,464 families attending the screening program. Among them, 130 breast cases were identified by the end of 2002.

For modeling the disease natural history of breast cancer from the data described above, we developed a three-state Markov model including both fixed effect, random effect and to capture the heterogeneity caused by individual variation after adjusting for correlated property and measured covariates by Bayesian approach.

Results For the breast cancer screening with family history of first-degree relative, Of 147 cases, there were only 9 cases with family history. The sensitivity was only 5.8% and the specificity was 97.9%. The average high-risk score of breast cancer cases was –5.05 (SD=3.41) and the average high-risk score of none breast cancer cases was –7.76 (SD=3.75). Apparently, the score was higher in breast cancer rather than in non-breast cancer. The results from Bayesian random-effect Markov model showed that the remarkable effect of age at first full-term pregnancy was seen in the transition from the PCDP to clinical phase. Those who were age at first full-term pregnancy older than 30 years had approximately two-fold risk for the progression from the PCDP to clinical phase (Rate ratio=1.89, 95% credible interval=1.02-3.66). The random effect was still statistically significant.

Introduction

Breast cancer screening with mammography has been demonstrated to reduce 20 to 30% mortality of breast cancer in western country for aged 40-69 (Nystrom et al, 1993; Tabar et al, 2000). Although incidence and mortality of breast cancer have increased over the past decade in many Asia countries, the incidence seems still too low to reach the criteria of mass screening for breast cancer with mammography compared with western country. Even if the efficacy is one of most essential elements to pursue for any screening programme, the efficiency of screening should be also taken into consideration when positive predictive value might be low. However, how to screen women at average risk is no consensus on the optimal modality so far. Providing the genetic testing and counseling for women belonging to family with familial is one of the strategy to reduce the risk for developing breast cancer (Eccles et al, 2000; NICE). According to the number of affected relatives, the age at onset and the cancers associated, three groups of risk have been defined as low, moderate and high for secondary prevention. Women belong to moderate or high risk should undergo annual mammography (Kriege et al, 2004; Kuhl et al, 2005). Similarly, for general population, we should take selective screening into account from the consideration of cost- effectiveness in breast cancer screening programme. Accordingly, in large-scale population screening for breast cancer by high-risk group (i.e., with family history of breast cancer, age at menarch), special attention should be paid to certain number of mammography from the economic aspect. Another viewpoint is the positive predictive value would be improved and save the mammographic cost through selective screening.

No matter what screening modality was considered, the understanding of disease natural history is essential for policy making. However, the estimation of parameters in multi-state process regarding the tumour progression form the

pre-clinical screen-detectable phase (PCDP) to clinical phase is intractable because of multi-level data structure inherent from cluster(family)-based service screening program, which is characterized by correlated property between rounds of screen in the same individual or different subjects but within the same cluster (family), and residual heterogeneity after being explained by measured or known covariates. We proposed a Bayesian hierarchical multi-state model to tackle these problems with the incorporation of random-effect parameters. The model was applied to data from a high-risk group (family history of breast cancer) screening for breast cancer.

This year we developed a Bayesian method in terms of evaluation of different screening strategy in selecting individuals at risk to determine the criteria for the selection of general population at risk of developing breast cancer. In addition, a three-state Markov model including both fixed effect, random effect was developed to capture the heterogeneity caused by individual variation after adjusting for correlated property and measured covariates by Bayesian approach.

Materials and methods

High-risk group predict model

A two-stage breast cancer screening for selecting high-risk group was adopted from July 2002 in Taiwan. In the first stage, we used questionnaire screening to define each individual who was high-risk group or not. Several modalities had been used in communities for risk assessment. The risk assessment tools were provided including reference table and computer software. For reference table, the individual risk score was elucidated by the combination of multi-dimensions with risk indicators on table. The most efficient tool was used by information technology from web-site was

conducted to access risk for all participates on-line. Public health nurse will transfer high-risk group to receive mammography by referred messages in second stage.

The high-risk group was first included the women aged 50-69 years who has the family history of first degree relative of breast cancer, and second defined by a risk score higher than a cutoff point. Risk score based on logistic regression model that incorporates all significant reproductive and menstrual factors obtained from data with mammographic and physical examination of the past community-based out-reaching screening program (1999~2001) (Wu et al, 2006).

According to the cost and the capacity of mammography, the cut-off point of score was defined at –9. It would be getting around 50% high-risk women based on this criterion in general population screening. The score higher than cut-off point was defined as the high-risk cases.

Bayesian analysis for high-risk score

The Bayesian analysis was approached to evaluate the two-stage breast cancer screening. The posterior probability of two-stage breast cancer screening was assess by Bayesian analysis and expressed as follow:

$$\frac{P(BC \mid X = \chi)}{P(\overline{BC} \mid X = \chi)} = \frac{P(BC)}{P(\overline{BC})} \times \frac{P(X = \chi \mid BC)}{P(X = \chi \mid \overline{BC})} \qquad (1)$$

BC is breast cancer and $\overline{BC}$ is none breast cancer. $P(BC)$ is the prior probability of breast cancer and $P(\overline{BC})$ is the prior probability of none breast cancer. The high-risk score is X.

The general form for likelihood as in (1) is expressed by could be written as (2):

$$\frac{P(X=\chi \mid BC)}{P(X=\chi \mid \overline{BC})}=\frac{\frac{1}{\sqrt{2\pi\sigma_1}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right]}{\frac{1}{\sqrt{2\pi\sigma_0}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_0}{\sigma_{\bar{0}}}\right)^2\right]} \quad (2)$$

μ_1 was the average score and σ_1 was the standard deviation for breast cancer cases，μ_0 was the average score and σ_0 was the standard deviation for none breast cancer cases。If the score was follow as Normal distribution. The random variable of risk score of breast cancer was X_1 and assume the distribution of $X_1 \sim N(\mu_1,\sigma_1)$, and the random variable of risk score of none breast cancer was X_0 and assume the distribution of $X_0 \sim N(\mu_0,\sigma_0)$.

The posterior odds can be written as follow:

$$\frac{P(BC \mid X=\chi)}{P(\overline{BC} \mid X=\chi)}=\frac{P(BC)}{P(\overline{BC})}\times\sqrt{\frac{\sigma_{\overline{BC}}}{\sigma_{BC}}}\times\exp\left\{-\frac{1}{2}\left[\left(\frac{x-\mu_{BC}}{\sigma_{\overline{BC}}}\right)^2-\left(\frac{x-\mu_{BC}}{\sigma_{BC}}\right)^2\right]\right\} \quad (3)$$

Finally, Bayesian factor (weight of evidence) can be written as (4):

$$\frac{P(X \mid BC)}{P(X \mid \overline{BC})}=\frac{P(BC \mid X)}{P(\overline{BC} \mid X)}\Big/\frac{P(BC)}{P(\overline{BC})} \quad (4)$$

Bayesian random-effect Markov model for breast cancer screening for women with positive family history

There were 4,867 women with family history derived from 4,464 families attending the screening program. Among them, 130 breast cases were identified by the end of 2002. For modeling the disease natural history of breast cancer from the data described above, we developed a three-state Markov model including both fixed effect, random effect and to capture the heterogeneity caused by individual variation after adjusting for correlated property and measured covariates by Bayesian approach.

Bayesian hierarchical model with random effect

A Bayesian hierarchical model was proposed to estimate the three-state Markov model with the incorporation of random effect and fixed effect.

Figure 2 shows acyclic graph model for estimating the parameters of β's and σ^2's for random-effect of Z's. Three hierarchical levels were formulated due to repeated screening visit level (*v*), subject level (*s*), and family level (*f*).

A typical non-informative but proper prior parameter was specified by a gamma distribution with small but positive values. Hence, Gamma (10^{-3}, 10^{-3}) distribution was used as the prior distribution for transition rates, λ_1 and λ_2, and the precision of random effect parameter, τ, which is equivalent to the inverse of variance, σ^2. For ease of convergence, several boundaries of gamma distribution were set up as 0.001-1 for λ_1, 0.1-1 for λ_2, 0-5 for τ. For modeling the fixed effect, a typical non-informative prior, normal distribution, Normal(0, 10^{-6}), was assigned for regression coefficient, β, with the range between -2 and 2, assuming the odds ratio for the effect of covariate on risk would not be greater then 7.

The Gibbs sampler, an iterative Markov-chain Monte Carol (MCMC) simulation, was used to estimate the posterior distribution. Bayesian analysis software WinBUGS (Spiegelhalter et al., 2004) was used. For each model, 50,000 samples were drawn following a burn-in period of 10,000 iterations. Convergence was assessed by checking the history of iterations and the Kernal density for each parameter in each model.

Model selection

Deviance information criterion (DIC), the summation of posterior mean of the deviance and the effective number of parameters (penalty term for increasing model complexity), was used to compare series of proposed random effect models listed in

Table 2 because only a fraction of degree of freedom as a result of random effect was removed (Spiegelhalter et al., 2002). The model with the smallest DIC was considered as the best one.

Preliminary Results

For the breast cancer screening with family history of first-degree relative, Of 147 cases, there were only 9 cases with family history. The sensitivity was only 5.8% and the specificity was 97.9% (Table 1).

Of all participants, the average high-risk score of breast cancer cases was –5.05 (SD=3.41) and the average high-risk score of none breast cancer cases was –7.76 (SD=3.75). The average high-risk score of breast cancer cases was –4.60 (SD=3.48) and the average high-risk score of none breast cancer cases was –9.364 (SD=3.19) for age 50-59 years. The average high-risk score of breast cancer cases was –5.87 (SD=3.15) and the average high-risk score of none breast cancer cases was –9.36(SD=3.19) for age 60-69 years (Table 2). Apparently, the score was higher in breast cancer rather than in non-breast cancer. The distributions of high-risk score between breast cancer and none breast cancer by age groups were demonstrated as Figure 2. The distributions of high-risk score are closer between breast cancer and none breast cancer for women who belong to age under 60 rather than age over 60.

Table 2 shows the posterior odds conferred by each high-risk score and age group after applying the equation (4) to empirical data with all participants. Bayesian factor decreases with the level of score at decreasing rate. If we take the high risk score between 14 and -13, the highest and lowest, it represented that the ranges of Bayesian factor could reflects the degree of prior odds. Less Bayesian factor presents the less probability to detect breast cancer from the same age group. In contrast with younger women aged under 60, women for aged over 60 has higher Bayesian factor

when score is greater than -9. A posterior odd is the outcome for breast cancer compared with non breast cancer with selecting individuals at risk. For instance, we could detect one breast cancer case for every 1,406 normal cases screened regarding score defined as -9.

Table 3 shows the estimated results by the incorporation of one significant covariate, age at first full-term pregnancy (AP) model, into M_3 models. The remarkable effect of age at first full-term pregnancy was seen in the transition from the PCDP to clinical phase. Those who were age at first full-term pregnancy older than 30 years had approximately two-fold risk for the progression from the PCDP to clinical phase (Rate ratio=1.89, 95% credible interval=1.02-3.66). The random effect was still statistically significant. The model fitting was lacking of statistical significance (P=0.10 for M_3+AP) as the predicted numbers were closer to the observed by taking the heterogeneity from measured covariates and unmeasured covariates.

Table 1 Breast cancer finding by family history of first-degree relative

Family history	Breast Cancer		Total
	Yes	No	
Yes	9(5.8%)	4613(2.1%)	4622(2.1%)
No	146(94.2%)	214054 (97.9%)	214200 (97.9%)
Total	155	218667	218822

Breast cancer finding was included 8 interval cases

Table 2 The posterior odds conferred by each high-risk score and age group

Risk Score	Age					
	50-59*		60-69‡		50-69†	
	Posterior Odds	Bayesian Factor	Posterior Odds	Bayesian Factor	Posterior Odds	Bayesian Factor
14	1: 43.58	18.37	1: 0.56	1689.42	1: 16.79	50.80
13	1: 50.55	15.84	1: 0.79	1197.86	1: 20.36	41.90
12	1: 58.63	13.65	1: 1.11	849.33	1: 24.68	34.57
11	1: 68.01	11.77	1: 1.56	602.21	1: 29.92	28.51
10	1: 78.89	10.15	1: 2.20	426.99	1: 36.27	23.52
9	1: 91.50	8.75	1: 3.11	302.75	1: 43.97	19.40
8	1: 106.14	7.54	1: 4.38	214.66	1: 53.30	16.01
7	1: 123.11	6.50	1: 6.18	152.20	1: 64.62	13.20
6	1: 142.80	5.61	1: 8.72	107.92	1: 78.33	10.89
5	1: 165.64	4.83	1: 12.29	76.52	1: 94.96	8.98
4	1: 192.13	4.17	1: 17.34	54.25	1: 115.12	7.41
3	1: 222.86	3.59	1: 24.46	38.47	1: 139.55	6.11
2	1: 258.50	3.10	1: 34.49	27.27	1: 169.17	5.04
1	1: 299.84	2.67	1: 48.64	19.34	1: 205.09	4.16
0	1: 347.80	2.30	1: 68.61	13.71	1: 248.62	3.43
-1	1: 403.42	1.98	1: 96.76	9.72	1: 301.39	2.83
-2	1: 467.94	1.71	1: 136.47	6.89	1: 365.37	2.33
-3	1: 542.78	1.47	1: 192.47	4.89	1: 442.92	1.93
-4	1: 629.59	1.27	1: 271.45	3.47	1: 536.94	1.59
-5	1: 730.28	1.10	1: 382.84	2.46	1: 650.92	1.31
-6	1: 847.07	0.95	1: 539.94	1.74	1: 789.09	1.08
-7	1: 982.54	0.81	1: 761.52	1.24	1: 956.58	0.89
-8	1: 1139.68	0.70	1: 1074.02	0.88	1: 1159.63	0.74
-9	1: 1321.95	0.61	1: 1514.76	0.62	1: 1405.79	0.61
-10	1: 1533.37	0.52	1: 2136.36	0.44	1: 1704.19	0.50
-11	1: 1778.61	0.45	1: 3013.04	0.31	1: 2065.93	0.41
-12	1: 2063.06	0.39	1: 4249.49	0.22	1: 2504.46	0.34
-13	1: 2393.01	0.33	1: 5993.32	0.16	1: 3036.08	0.28

* Breast Cancer (Mean=-4.60 SD=3.48), None Breast Cancer (Mean=-6.63 SD=3.70)

‡ Breast Cancer (Mean=-5.86 SD=3.15), None Breast Cancer (Mean=-9.36 SD=3.19)

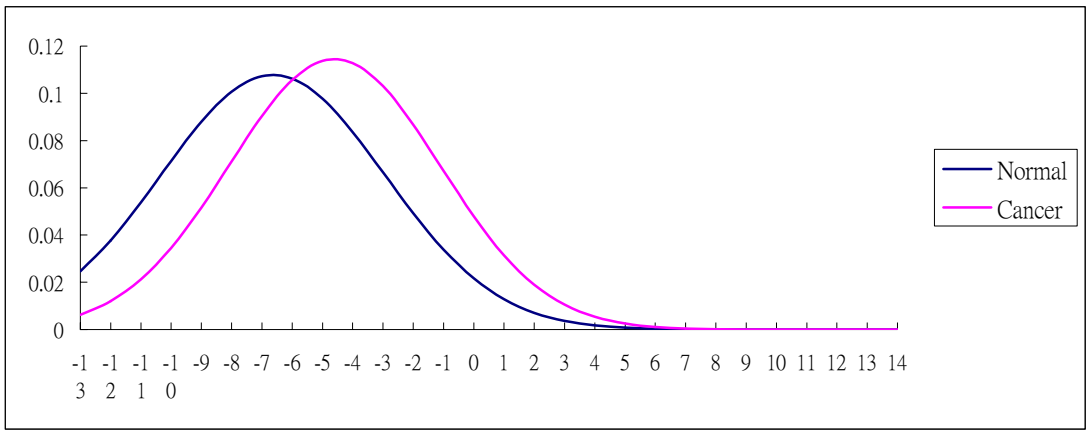
† Breast Cancer (Mean=-5.05 SD=3.41), None Breast Cancer (Mean=-7.76 SD=3.75)

Table 3 Results of random effect model with one covariate, age at first pregnancy (AP).

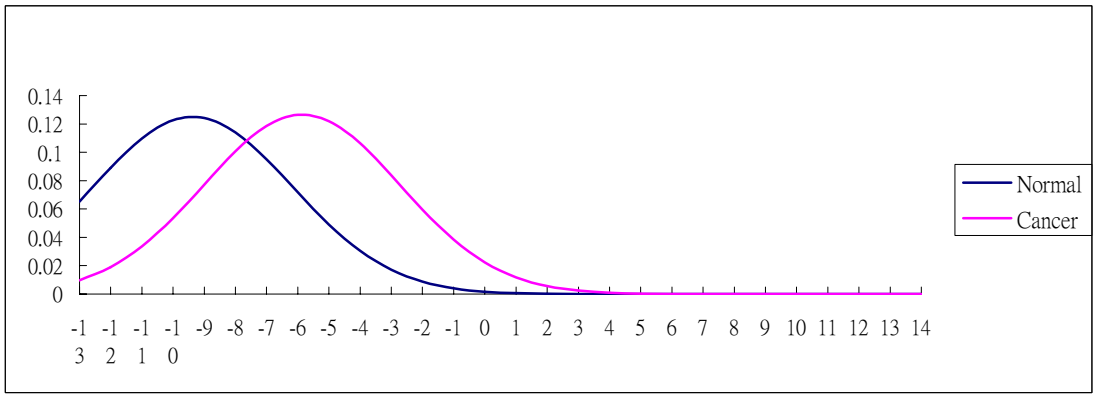
Node	Mean	S.D.	2.50%	Median	97.50%	χ^2	df	P-value
Model: M_3+AP						13.3534	8	0.10025
λ_1	0.0048	0.0006	0.0038	0.0048	0.0060			
λ_2	0.56	0.11	0.38	0.55	0.81			
β_2	0.64	0.33	0.02	0.63	1.30			
exp(β_2)	1.89	1.38	1.02	1.88	3.66			
σ^2_{2f}	0.56	0.35	0.38	0.58	0.79			

Figure 1 Distribution of Risk Score between Case and Non-case

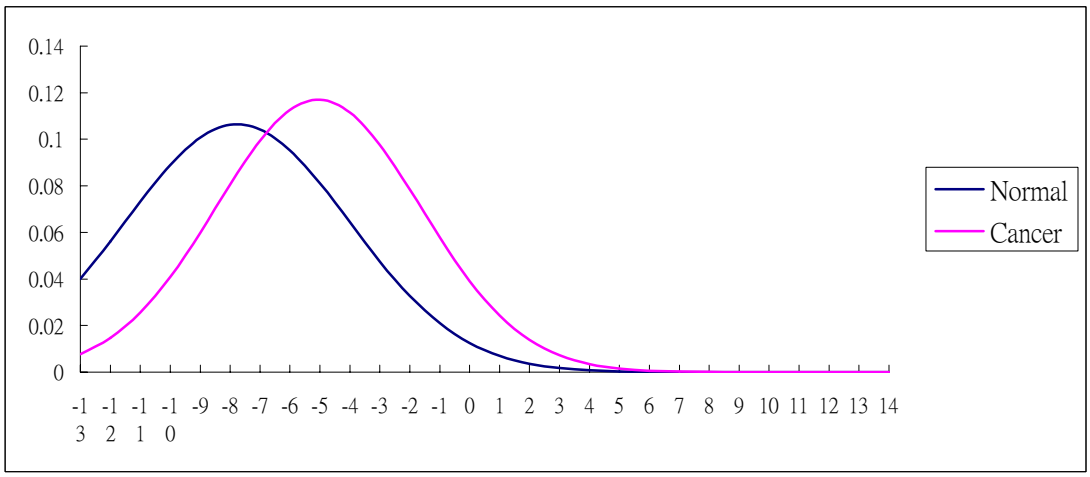
(a)



(b)



(c)



(a) Risk Score Distribution for Aged 50-59

(b) Risk Score Distribution for Aged 60-69

(c) Risk Score Distribution for Aged 50-69

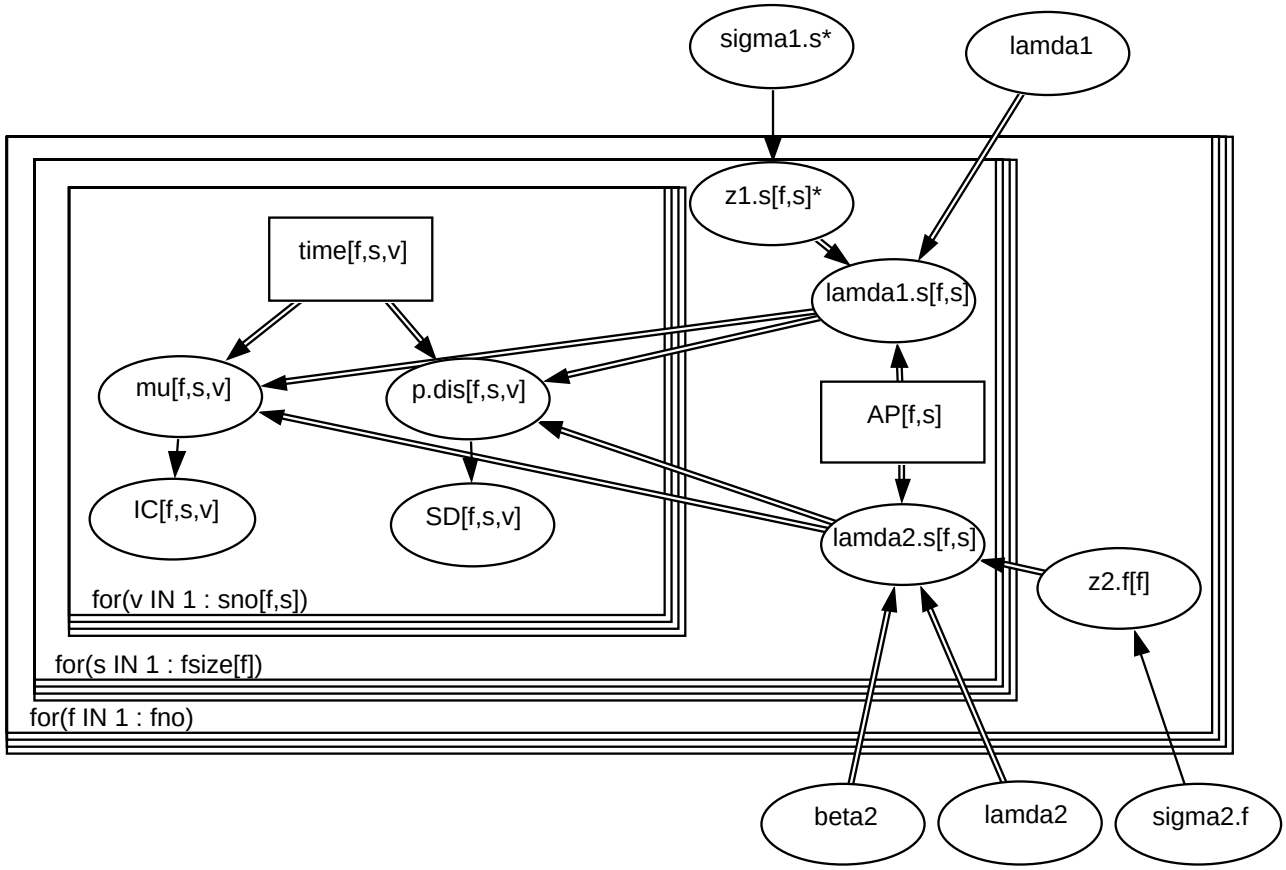


Figure 2. The acyclic graphic model for estimating the random effect and fixed effect of multi-state process.**

*The nodes were excluded in the final model with random effects and fixed effects, M_3+AP.

參考文獻

Eccles DM, Evans DG, Mackay J: Guidelines for a genetic risk based approach to advising women with a family history of breast cancer. UK Cancer Family Study Group (UKCFSG). J Med Genet 2000, 37:203-209.

Kriege M, Brekelmans CT, Boetes C et al. Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition. N Engl J Med 2004, 351:427-437.

Kuhl CK, Schrading S, Leutner CC et al. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. J Clin Oncol 2005, 23:8469-8476.

NICE: Guidelines for familial breast cancer. [http://www.nice.org.uk].

Nystro¨mL, Rutqvist LE, Wall S, et al. Breast cancer screening with mammography: overview of Swedish randomized trials. Lancet. 1993;341:973–978.

Spiegelhalter, D., Thomas, A., Best, N., Lunn, D. Win BUGS User Manual Version 2.0. Cambridge: MRC Biostatistics Unit, 2004.

Spiegelhalter, D. J. (2002). Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society, Series B 2002; 64, Part 4, 583-639.

Tabar L, Vitak B, Chen HH, et al. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. Radiol Clin North Am. 2000;38:625–651.

Wu GH, Chen LS, Chang KJ et al. Evolution of breast cancer screening in countries with intermediate and increasing incidence of breast cancer. J Med Screen. 2006 Dec;13 Suppl 1:23-7.

計畫結果自評

In this study, a two-stage screening modality was made the contribution of finding the most of cases. If the risk prediction could be take into account, the posterior odds could increase by the function of Bayesian factor to 3.43 fold for early detection of breast cancer when risk score is greater than 0. When risk score is greater than -9, the Bayesian factors are higher for women aged between 60 and 69 in comparison with women aged between 50 and 59. Compared with disusing, more information could be gained after conducting this screening tool in elder rather than younger.

Moreover, the quantitative score is convenience to define the optimal cut-off point using operator receiver curve (ROC) method. For all ages 50–69 years, the best cut-off was –6 with a sensitivity of 63% and specificity of 68%. The cut-off score actually used to refer to the second stage had been fixed at –9. This gave a sensitivity of 87% and specificity of 37% (Kriege et al, 2004). Given a lower cut-off point, higher sensitivity is accompanied by lower specificity. It need to be concerned for that, at the lower cut-off value, unnecessary mammographic resulting from higher false positive cases with risk assessment by questionnaire but high utility as a result of lower false negative cases would be expected. By contrast, at higher cut-off value, false positive rate could be reduced but false negative cases were increased. Cost-effectiveness analysis therefore should be provided and applied to determining the best cut-off value between sensitivity and specificity in the future.

We presented a Bayesian hierarchical multi-state model to estimate transition parameters of the disease natural history based on multi-level data of breast cancer screening for women with relatives suffering from breast cancer. The major of our approach is to introduce random-effect parameters corresponding to different

hierarchical levels to capture the heterogeneity resulting from correlations, measured covariates, and residual variation that cannot be explained by measures covariates. It can be found that the consideration of random-effect not only affects the mean value but also the variance. In the three-state model, it can be seen that making allowance for random-effects led to higher transition rate and wider confidence of lambda2, but lower annual incidence rate and narrower confidence of lambda1 (see Table 3). This suggests that a shorter screening interval may be needed if such heterogeneity is considered.