Cukurova Medical Journal
ÇUKUROVA ÜNİVERSİTESİ TIP FAKÜLTESİ

Cukurova Med J 2018;43(1):41-48
DOI: 10.17826/cumj.339947

ARAŞTIRMA / RESEARCH

Evaluation of diagnostic performance measures with two-stage Bayesian method when chromosomal disorders can not be verified with gold standard test

Kromozomal bozuklukların referans test ile doğrulanamadığı durumda iki aşamalı Bayes yöntemiyle performans ölçülerinin değerlendirilmesi

Didem Derici Yıldırım[1], Bahar Taşdelen[1], Filiz Çayan[2], Özlem İzci Ay[3], Mustafa Ertan Ay[3]

[1]Mersin University, Department of Biostatistics and Medical Informatics, [2]Department of Obstetrics and Gynaecology, [3]Department of Medical Biology and Genetics Mersin, Turkey

Cukurova Medical Journal 2018;43(1):41-48.

Abstract

Purpose: Predicting chromosomal abnormality prevalance and accuracy of diagnostic tests represents a considerable problem when chromosomal disorders could not be verified with gold standard test. Aim of the study was to predict performance of screening tests and chromosomal abnormality prevalance within a bayesian analysis.

Materials and Methods: Retrospective study at the Hospital of Mersin University, including pregnants who were admitted to gynecology and obstetrics clinic between August 2014 and August 2015. Within applicants, 1759 gestational's records were utilized who had ultrasound examination and double screening test result in first trimester period. Two stage Bayesian approach was used. In the first stage, people were classified as patient and healthy. In the second stage, people suspected of having disease were examined by a gold standard test.

Results: In the situation of known prior information, joint sensitivity and joint specificity were estimated as 77% and 99%. In the case of no prior information, joint sensitivity and joint specificity were estimated as 50% and 97%. Considering Deviation Information Criterion and Monte Carlo Error values together, model under prior information was better than the other.

Conclusion: Verification problem will be frequently encountered for screening tests in pregnancy. Performance measures of the tests may be calculated more accurately by utilizing this method.

Key words: Chromosomal abnormalities, two-stage bayesian method, verification problem.

Öz

Amaç: Kromozomal anormallik prevalansını ve diagnostik testlerin doğruluğunu tahmin etmek kromozomal bozukluklar altın Standard test ile doğrulanamadığı durumda ciddi bir problemdir. Bu çalışmanın amacı tanı testlerinin performasını ve kromozomal anomali prevalansını bayes analizi ile tahmin etmektir.

Gereç ve Yöntem: Ağustos 2014 ve Ağustos 2015 tarihleri arasında Mersin Üniversitesi Hastanesi Kadın Hastalıkları ve Doğum polikliniğine başvurmuş gebeler çalışmaya dahil edilmiştir. Başvuran gebeler içerisinden, birinci trimester dönemde ultrasonografi ve ikili tarama test sonucu olan 1759 gebeye ait kayıtlardan yararlanılmıştır. İki Aşamalı Bayes yöntemi kullanılmıştır. İlk aşamada, bireyler hasta ve sağlıklı olarak sınıflandırılmaktadır. İkinci aşamada ise ilk aşamada hasta ya da sağlıklı olduğuna karar verilen bireylere altın standart test uygulanarak durumları doğrulanmaktadır.

Bulgular: Önsel dağılım bilgisi olduğu durumda, birlikte duyarlılık ve seçicilik değerleri %77 ve %99 olarak tahmin edilmiştir. Önsel dağılım bilgisi olmadığı durumda ise, birlikte duyarlılık ve seçicilik değerleri %50 ve %97 olarak tahmin edilmiştir. Sapma Bilgi Ölçütü (DIC) ve Monte Carlo Hata (MC) değerleri birlikte incelendiğinde, önsel bilgi olduğu durumda modele uyum daha iyidir.

Sonuç: Doğrulanamama problemi ile gebelikte uygulanan rutin tarama testleri için sıklıkla karşılaşılmaktadır. Tarama testlerinin doğruluğu bu yöntemden yararlanılarak daha doğru şekilde tahmin edilmektedir.

Anahtar kelimeler: Kromozomal bozukluklar, iki aşamalı bayes yöntemi, doğrulama problemi

Yazışma Adresi/Address for Correspondence: Dr. Didem Derici Yıldırım, Mersin University, Department of Biostatistics and Medical Informatics, Mersin, Turkey. E-mail: didemderici@hotmail.com
Geliş tarihi/Received: 04.01.2017 Kabul tarihi/Accepted: 28.04.2017.

INTRODUCTION

Physical and mental disorders negatively impact people's life such that they encounter many difficulties in chilhood and adolescent period. Many genetic abnormalities can be detected before the birth. The most common chromosomal abnormalities are trisomies caused by nondisjunction during cell division following to meiosis and mitosis. A trisomy is a type of aneuploidy in which there are three instances of a particular chromosome, instead of two and associated with three main syndrome including Trisomy 21 known as Down syndrome, Trisomy 13 (Patau syndrome) and Trisomy 18 (Edward syndrome)[1,2].

Prenatal diagnosis for chromosome abnormalities has been available for over approximately fifty years. It includes all the methods used to obtain information about the embryo or fetus. Also, prenatal diagnosis is one of the best example of the integration of pre-clinical and clinical sciences. It covers both screening and diagnostic methods. First, pregnant women are classified as low or high risky for chromosomal disorders by ultrasonography (USG) or screening tests. Then the most accurate information about the baby's health can be obtained by diagnostic methods such as amniocentesis, cordocentesis or chorionic villus sampling for high risky group according to the results of screening tests. Amniocentesis is the most common invasive method, but it is offered to women that have at least one positive test result after screening because of being an invasive, costly procedure and having miscarriage risk (1/300-1/500)[1-6].

The pregnants having negative results in both tests are not usually verified by a gold standard test. Hence, calculating performance measures become impossible. Based on verification problem for negative results, the primary aim of this study is to predict chromosomal abnormality prevalance and performance of diagnostic tests using two-stage bayesian method. The secondary aim is to evaluate the effect of prior information on predictions.

MATERIALS AND METHODS

In this study, the pregnant women who were admitted to the gynecology and obstetrics clinic for routine follow up between August 1, 2014 and August 1, 2015 were included. From applicants, 18-35 year-old 1759 women's record including ultrasound examination and double screening test results in the first trimester period have been utilized. The study was approved by the clinical research ethics committees of Mersin University on 11/26/2015. (Meeting number/ Decision number: 22/355). All patients provided written informed consent prior to taking part in the study.

Prenatal ultrasonography markers during the late first trimester of gestation were echogenic intracardiac focus, increased fetal nuchal translucency, absent nasal bone, echogenic bowel**,** bilateral renal pyelectasis or shortened fetal femoral length. When, at least two of those markers were detected, a pregnant was evaluated as positive in USG. She was considered as risky if calculated risk value for double screening test was larger than 1/250. All steps of the obsteric evaluations were supervised by the gynecology and obstetric specialist (Co-author) and definite diagnoses of the abnormality were also made by her. Pregnants whose either screening test or USG was positive were directed to amniocentesis. (Co-author) and (Co-author) evaluated amnion fluid who have agreed to amniocentesis.

The evaluation of screening tests have been used since Thomas Bayes first developed the Bayes Theorem in the late 1700's[7]. Then, clinicians have been frequently utilized Bayes Theorem for diagnosing disease or selecting the appropriate treatment according to their experience and previous studies. In generally, two stage approach is used in screening specific disease or condition. In the first stage, people are classified as patient and healthy. In the second stage, people suspected of having disease are examined by a gold standard test for verification. The Bayesian approach is based on not only observed information but also on a prior information[8]. Information from previous studies were prior information for Bayesian analysis. There are many publications applied Bayesian theorem to calculate the sensitivity, specificity of screening tests and prevalance[9-13]. A Bayesian approach was used to evaluate double screening test utilized in 11-14 weeks for detecting chromosomal abnormalities especially trisomy 21 in this study.

The count data of screening test and amniocentesis results indicated in Table I. Where, X_{11} is the number of patients who have positive results for both double screening test and USG, X_{10} is the number of patients who have positive result for

double screening test but negative results for USG, X_{01} is the number of patients who have negative result for double screening test but positive results for USG, X_{00} is the number of patients who have negative results for both of them. Since women that have at least one positive test result are verified using amniocentesis, $[a_{01}]$ and $[a_{00}]$ are unknown. Using sensitivity ($Se_i=P(T_{i+}|D_+)$) and specificity ($Sp_i=P(T_{i-}|D_-)$) for each screening method, covariances and correlations between two tests are calculated for both diseases and nondiseases[14,15]. Where, T_{i+} is true positives, T_{i-} is true negatives, D_+ is the number of all pregnants with chromosomal abnormalities and D_- is the number of all pregnants without chromosomal abnormalities.

Disease prevalance is also shown as $\pi= P(D_+)$. At the same time joint estimation of sensitivity, specificity and positive predictive value of screening methods can be calculated.

Table 1. Data structure for two-stage Bayesian method.

	Ultrasonography				**Amniocentesis**		
Double Screening Test	+	-	Total		D+	D-	Total
+	X11	X10	X1.	At least one test is positive	a11	a10	a1.
-	X01	X00	X0.	Both tests are negative	[a01]	[a00]	a0.
Total	X.1	X.0	n	Total	[a.1]	[a.0]	N=n

Prior distribution information should be known to use Bayeasin approch. Performance measures of screening tests (sensitivity, spesificity, positive predictive value, negative prediactive value) are generally uniformly distributed. Only range value (maximum value-minimum value) information is enough for utilizing uniform prior distribution. Range is [0,1] for performance measures[14-16].

Model evaluation and checking

All estimations in this study can be applied with WinBUGS codes using WinBUGS software[17]. Deviation Information Criterion (DIC), Monte Carlo Error (MC) statistics and Brooks-Gelman-Rubin (BGR) convergence statistics are commonly used model evaluation techniques. DIC, generalisation of Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC), is useful where posterior probabilities are obtained by Monte Carlo Marcov Chain (MCMC) simulation. For model checking, model with the smallest DIC is the best predictor of posterior probabilities. MC error is interpreted as similar to standard error. A rule of thumb is that MC error must be less than 5% of standard deviation of the estimated parameter[18]. Furthermore, Brooks-Gelman-Rubin (BGR) convergence statistics are graphical tools for assessing convergence of multiple chains to the same distribution. There must be at least two chains for utilization of this method and BGR values around 1 indicate convergence. At the same time, Positive predictive value (PPV) is a good performance measure of diagnostic methods because it contains sensitivity, specificity and prevalance. Hence, it is also used for model checking. The difference between the actual PPV value and estimated PPV value must be small and actual PPV value must be included by the estimated 95% Bayesian Credible Interval (BCI) of estimated PPV.

RESULTS

When data archive was investigated, either screening test was positive in 1,13% (20/1759) of enrolled pregnants and underwent to amniocentesis (Table II). Mean age was calculated of all pregnants included the study as 26.42±5,19 years. Three chromosomal abnormalities were detected following amniocentesis and amniocentesis results of pregnants whom at least one test is positive were given in Table 3.

For application of two-stage Bayesian method, two senarios of prior distributions were constructed as below. In the first senario, informative prior distribution was applied. Lower bounds of distributions were arranged such that obtained information from previous studies. According to the literature, sensitivies of double screening test and USG were 0.60 and 0.50 [19] and upper bounds were 1. Because there were no clear information about upper bounds. Information about specificity and

prevalance were not clear and constant. Hence, uniform distribution range for specificity and prevalance were determined between 0 and 1.

Senario 1

Se_1~ uniform (0.60, 1) Se_2 ~ uniform (0.50,1)

Sp_1~ uniform (0,1) , Sp_2 ~ uniform (0,1)

π ~ uniform (0,1)

In the second senario, the uniform distribution in [0,1] was applied to calculation of sensitivity and specificity as a non informative prior distribution in the absence of information from previous studies. Thus, importance of prior information of the sensitivity was demonstrated with real data.

Senario 2

Se_1~ uniform (0,1), Se_2~ uniform (0,1)

Sp_1~ uniform (0,1) , Sp_2 ~ uniform(0,1)

π ~ uniform (0,1)

The diagnostic performance measures were estimated for informative and non-informative prior distribution of sensitivity (Table 4, Table 5).

In the situation of informative prior distribution, joint sensitivity and joint specificity were estimated as 77% and 99%. The positive predictive value was 17% and prevalence of chromosomal disorders was 0.2%. In the case of non informative prior distribution, joint sensitivity and joint specificity were estimated as 50% and 97%. The positive predictive value was 18% and prevalence of chromosomal disorders was 5.6%. The correlation coefficient between two tests under prior information and non information were approximately 0.40.

Table 2. The distribution of pregnants according to screening test and amniocentesis results.

	USG				Amniocentesis		
Double Screening Test	+	-	Total		D+	D-	Total
+	4	13	17	At least one test is positive	3	17	20
-	3	1739	1742	Both tests are negative	[a01]	[a00]	1739
Total	7	1752	1759	Total	[a.1]	[a.0]	1759

Table 3. Amniocentesis results of pregnants whom at least one test is positive

Double Screening Test	USG	Amniocentesis Result
+	+	47,XY+21; 46,XY; 46,XX; 46,XX
+	-	47,XX+21; 46,XY; 46,XX; 46,XY; 46,XX; 46,XX; 46,XX; 46,XY; 46,XX; 46,XX; 46,XY; 46,XY; 46,XX
-	+	47,XY+21; 46,XX; 46,XY

Table 4. Performance measures of USG and double screening test in case of informative prior distribution of sensitivity (Senario 1).

Se_1~ uniform [0.60, 1] Se_2 ~ uniform [0.50,1] Sp_1~ uniform [0,1] , Sp_2 ~ uniform [0,1] π ~ uniform [0,1]			
	Mean±St. Deviation	**Median (95% BCI)**	**MC Error**
Se_1	0.7774±0.1164	0.7676 (0.6064-0.9877)	0.0061
Se_2	0.7595±0,1473	0.7582 (0.5151-0.9944)	0.0081
Se_{je}	0.8982±0.0846	0.9225 (0.6861-0.9980)	0.0053
Sp_1	0.9914±0.0017	0.9915 (0.9880-0.9946)	5.451E-5
Sp_2	0.9974±0.0013	0.9975 (0.9945-0.9956)	6.943E-5
Sp_{je}	0.9899±0.0018	0.9900 (0.9861-0.9931)	6.236E-5
PPV_{je}	0.1762±0.0696	0.1699 (0.0543-0.3290)	0.0022
π	0.0024±0.0010	0.0023 (0,0007-0.0048)	3.321E-5
DIC	25.037		

BCI: Bayesian Credible Interval

Table 5. Performance measures of USG and double screening test in case of non informative prior distribution of sensitivity (Senario 2).

Se_1~ uniform [0,1], Se_2~ uniform [0,1] Sp_1~ uniform [0,1] , Sp_2 ~ uniform[0,1] π ~ uniform [0,1]			
	Mean±St. deviation	**Median (95% BCI)**	**MC Error**
Se_1	0.3637±0.2902	0.3099 (0.0017-0.9608)	0.0109
Se_2	0.3523±0.2930	0.2876 (0.0015-0.9476)	0.0115
Se_{je}	0.5064±0.3171	0.5238 (0.0030-0.9846)	0.0141
Sp_1	0.9723±0.1088	0.9913 (0.6242-0.9946)	0.0071
Sp_2	0.9905±0.0391	0.9970 (0.8878-0.9994)	0.0023
Sp_{je}	0.9693±0.1155	0.9899 (0.5648-0.9932)	0.0076
PPV_{je}	0.1839±0.0780	0.1755 (0.0583-0.3599)	0.0011
π	0.0561±0.1951	0.0043(0,0009-0.9772)	0.0134
DIC	-182,181		

BCI: Bayesian Credible Interval

Considering DIC and MC values together, goodness of model under informative prior distibution of sensitivity was better than non informative prior distribution model. According to BGR statistics, convergency of estimation results was good in case of both informative and non informative prior distribution of sensitivity. When the results of the two scenarios and BGR statistics were examined under informative prior ditribution of sensitivity, the MC error values were smaller than non informative prior distribution and at the same time, mean and median of each parameter were about the same. This shows that estimation results were close to each other and distributed symmetric. (Figure 1.a-1.b) At the same time, 95% BIC of estimated PPV was including actual PPV value. All these results showed that two model fit for BGR statistics, PPV value criteria. Therefore, the model under informative prior distribution was better than non informative prior distribution according to DIC, MC error statistics.

DISCUSSION

In many countries, the usage of screening methods like maternal serum screening and genetic sonogram to detect chromosome anomalies and other birth defects are a routine of prenatal diagnosis in the first and/or second trimesters. However, both of these methods are not recommended because of lower sensitivity and higher false positive rate (2%–7%). Developing countries still use these methods to examine chromosomal anomalies[20]. Because new developed methods like Non Invasive Prenatal genetic testing (NIPT) are too expensive. If one of these tests indicate that a fetus is at risk of aneuploidy, invasive methods like amniocentesis or chorionic villus sampling (CVS) are recommended for diagnosis. Many pregnant hesitate invasive methods because of miscarriage risk or religious causes. But true diagnosis of chromosomal disorders before the birth is very important[21]. It is considered that a large number of pregnant is forced to amniocentesis despite having a healthy baby because of screening tests' results. But there is a bigger problem than this situation that some of pregnants have baby with especially Trisomy 21 despite negative screening results. If at least one of the screening tests is positive, invasive methods can be applied. But the pregnants having negative results in both test are not usually verified by a gold standard test because of ethical reasons.

By now, many methods have been proposed to deal with the verification problem. Firstly performance measures were obtained through latent class models under the assumption of independence of two test errors (Walter,1999)[22,23]. Pepe and Alonzo (2001) suggested to compare relative accuracy of tests by utilizing the ratio of true positive and false negative results[24,25]. Models suggested by Bohning et al. (α model) and Feng Li et al. (θ, RR and matches model) had additional assumptions[16,26,27].

The homogenity assumption of odds ratios under the disease and non disease situation was provided in alpha model. For relative risk (RR) model, homogenity of relative risk values under two condition was assumed. The models suggested before the two-stage Bayesian method use different assumptions as a priori information. However, it is not possible to test this assumptions providing or

not. Bayesian approach is used to resolve unverification problem of negative results for both tests by means of prior information. This method have several advantages. These are consideration of conditional dependency between tests and calculation of joint sensitivity, specificity and positive predictive value. Furthermore, there is no assumption for bayesian approach and it is much easier to understand than other methods[28-31]. Negative results not verified by gold standard test are accepted as missing disease status. The missing disease status mechanism in this study is non-ignorable. Because disease verification depends on both observed risk factors (age, fetal nuchal translucency, double screening test results) and unsaved information connected to the disease (family history, culturel level, socioeconomic level). Bayesian approach is suitable for non-ignorable missing mechanism according to Kosinski and Barnhart study[32-33].

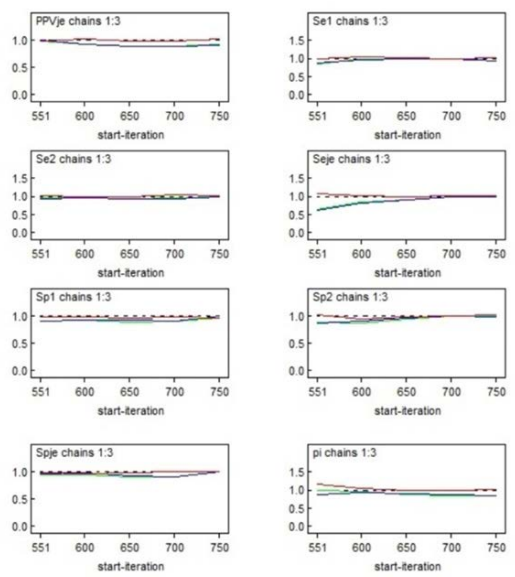

Figure 1.a. Brooks-Gelman-Rubin (BGR) convergence statistics for informative prior distribution of sensitivity

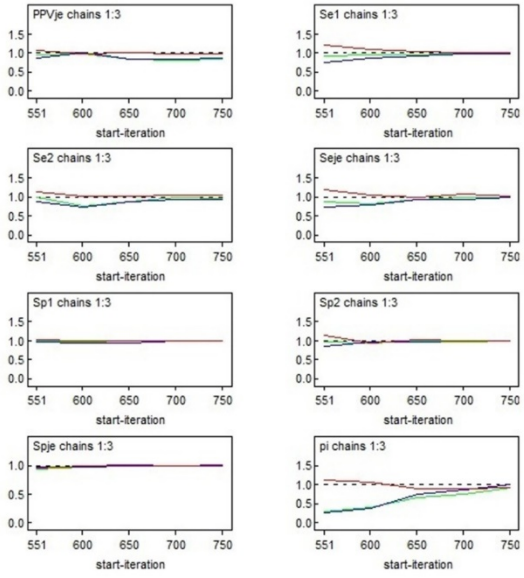

Figure 1.b. Brooks-Gelman-Rubin (BGR) convergence statistics for non informative prior distribution of sensitivity

So, we used two-stage Bayesian method to estimate performance measures and we found good estimations under informative prior distribution of sensitivity. In our study, only pregnants whom 18-35 age range were included to prevent bias. Because age is a risk factor for chromosomal disorder. There are lots of study about performance measures (false positive rate, detection rate, sensitivity or spesificity of methods) of these screening methods. But any literature could not be found about estimation problem of screening tests used in prenatal diagnosis that aneuploidy being verified for positive results only.

In conclusion, ultrasound and double screening test are routine examination in pregnancy, that is done with any doubt, so the number of negative test results for both tests will be considerably more. In this situation, calculation of diagnostic performance measures of screening tests is very important problem. There are issues to be considered by the physician and biostatistician in diagnostic studies. When assessed clinically, previously amniocentesis had suggested all women over the age of 35 because of the high risk of having a chromosomal abnormality. However, clinical evalutions showed that younger women had the babies with Down's syndrome more and the maternal age alone is not an effective factor for prenatal screening.

On the other hand, bias that occured while determining inclusion criteria for selection of cases

and controls in diagnostic studies shoul be avoided. Results that are unconfirmed by gold standard test should be taken into account using advanced methods such as two stage bayesian approach.

The selection of summary statistics that used is an important issue. It is considered that a large number of pregnant is forced to make amniocentesis despite having a healthy baby because of screening tests' results, more precise posteriori estimates of performance measures can be obtained by two stage bayesian method. Study results indicate that sensitivity and specificity values are relatively high, though the low PPV. Indeed, PPV is more important than sensitivity and specificity to pregnants having unverified screening test results because it is the probability that a positive test result indicates a true fetal aneuploidy. Therefore, when the study results and clinical experience are taken into account, more sensitive screening tests for karyotype abnormality should be developed. The addition of Two Stage Bayesian method to the common statistical software will ensure widespread usage.

REFERENCES

1. Antonarakis SE, Lyle R, Chrast R, Scott HS. Differential gene expression studies to explore the molecular pathophysiology of Down syndrome. Brain Res Rev. 2001;36:265-74.
2. Gersen SL, Keagle MB. The Principles of Clinical Cytogenetics. 2nd ed. USA, Humana Press, 2005.
3. Gardner McKinlay RJ, Sutherland GR. Chromosome Abnormalities and Genetic Counseling. 3th ed. England, Oxford University Press. 2004.
4. Kuskucu AC. Fetal kromozom anomalisi tarama testleri. Jjinekoloji Obstetri Pediyatri ve Pediyatrik Cerrahi Dergisi. 2010;2:55-60.
5. Khalil A, Pandya P. Screening for Down syndrome. J Obstet Gynaecol India. 2006;56:205-11.
6. Baer RJ, Currier RJ, Norton ME, Flessel MC, Goldman S, Towner D et al. Outcomes of pregnancies with more than one positive prenatal screening result in the first and second trimester. Prenat Diagn. 2015;35:1223-31.
7. Maxim LD, Niebo R, Utell MJ. Screening test:a review with examples. Inhal Toxicol. 2014;26:811-28.
8. Dunson DB. Commentary: practical advantages of Bayesian analysis of epidemiology data. Am J Epidemiol. 2001;153:1222-6.
9. Berkvens D, Speybroeck N, Praet N, Adel A, Lesaffre E. Estimating disease prevalance in a Bayesian framework using probabilistic constraints. Epidemiology. 2006;17:145-53.
10. Dendukuri N, Joseph L. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. Biometrics. 2001;57:158-67.
11. Geurden T, Claerebout E, Vercruysse J, Berkvens D. Estimation of diagnostic test characteristics and prevalance of Giardia duodenalis in dairy calves in Belgium using a Bayesian approach. Int J Parasitol. 2004;34:1121-7.
12. Geurden T, Berkvens D, Geldhof P, Vercruysse J, Claerebout E. A Bayesian approach for the evaluation of six diagnostic assays and estimation of cryptosporidium prevalance in dairy calves. Vet Res. 2006;37:671-82.
13. Liu P, Yang H, Qiang L, Xiao S, Shi ZX. Estimation of the sensitivty and specificity of assays for screening antibodies to HIV: a comparison between the frequentiest and Bayesian approaches. J Virol Methods. 2012;186: 89-93.
14. Liu J, Chen F, Yu H, Zeng P, Liu L. A two-stage bayesian method for estimating accuracy and disease prevalance for two dependent dichtomous screening tests when the status of individuals who are negative on both tests is unverified. BMC Med Res Methodol. 2014;14:1-11.
15. Li F, Chu H, Nie L. A two-stage estimation for screening studies using two diagnostic tests with binary disease status verified in test positives only. Stat Methods Med Res. 2015;24:635-56.
16. Joseph L, Gyorkos TW, Coupal L. Bayesian estimation of disease prevalance and the parameters of diagnostic tests in the absence of a gold standard. Am J Epidemiol. 1995;141:263-72.
17. Spiegelhalter D, Thromas A, Best N, Lunn D. 2003. WinBUGS User Manual (Version 1.4)
18. Toft N, Innocent GT, Gettinby G, Stuart WJR. Assessing the convergence of Markov Chain Monte Carlo methods: An example from evaluation of diagnostic tests in absence of a gold standard. Prev Vet Med. 2007;79:244-56.
19. Benn PA, Ying J, Beazoglou T, Egan JF. Estimates for the sensitivity and false-positive rates for second trimester serum screening for Down syndrome and trisomy 18 with adjustment for cross-identification and double-positive results. Prenat Diagn. 2001;21:46-51.
20. Abele H, Wagner P, Sonek J, Hoopmann M, Brucker S, Artunc-Ulkumen B et al. First trimester ultrasound screening for Down syndrome based on maternal age, fethal nuchal translucency and different combinations of the additional markers nasal bone, tricuspid and ductus venosus flow. Prenat Diagn. 2015;35:1182-6.
21. Xiao H, Yang YL, Zhang CY, Liao EJ, Zhao HR, Liao SX. Karyotype analysis with amniotic fluid in 12365 pregnant women with indications for genetic amniocentesis and strategies of prenatal diagnosis. J Obstet Gynaecol. 2016;36:293-6.
22. Walter SD. Estimation of test sensitivity and

specificity when disease confirmation is limited to positive results. Epidemiology. 1999;10:67-72.
23. Chu H, Zhou Y, Cole SR, Ibrahim JG. On the estimation of disease prevalance by latent class models for screening studies using two screening tests with categorical disease status verified in test positives only. Stat Med. 2010;29:1206-18.
24. Pepe MS, Alonzo TA. Comparing disease screening tests when true disease status is ascertained only for screen positives. Biostatistics. 2001;2:249-60.
25. Chock C, Irwig L, Berry G, Glasziou P. Comparing dichotomous screening tests when individuals negative on both tests are not verified. J Clin Epidemiol. 1997;50:1211-7.
26. Berry G, Smith CL, Macaskill P, Irwig L. Analytic methods for comparing two dichotomous screening or diagnostic tests applied to two populations of differing disease prevalance when individuals negative on both tests are unverified. Stat Med. 2002;21:853-62.
27. Bohning D, Patılea V. A capture-recapture approach for screening using two diagnostic tests with availability of disease status for the test positivies only. J Am Stat Assoc. 2008;103:212-20.
28. Genc Y, Tuccar E. Effect of verification bias on sensitivity and specificity of diagnostic tests. Journal of Ankara Medical School. 2003;25:107-12.
29. Groot JA, Dendukuri N, Janssen KJM, Reitsma JB, Bossuyt PM, Moons KG. Adjusting for differential-verification bias in diagnostic- accuracy studies: a Bayesian approach. Epidemiology. 2011;22:234-41.
30. Alonzo TA, Brinton JT, Ringham BM, Glueck DH. Bias in estimating accuracy of a binary screening test with differential disease verification. Stat Med. 2011;30:1852-64.
31. Collins J, Huynh M. Estimation of diagnostic test accuracy without full verification: a review of latent class methods. Stat Med. 2014;33:4141-69.
32. Kosinski AS, Barnhart HX. Accounting for nonignorable verification bias in assessment of diagnostic tests. Biometrics. 2003;59:163-71.
33. Buzoianu M, Kadane JB. Adjusting for verification bias in diagnostic test evaluation: a Bayesian approach. Stat Med. 2008;27:2453-73.