

Receiver Operating Characteristic Curve Analysis Of Diagnostic Tests Results For Gestational Diabetic Mellitus

¹Okeh U.M and ²Oyeka I.C.A

1.Department of Industrial Mathematics and Applied Statistics, Ebonyi State University Abakaliki, Nigeria.

2.Department of Applied Statistics, Nnamdi Azikiwe University Awka, Nigeria.

Abstract: This paper proposes a semi-parametric generalized linear mixed model (GLMM) method of evaluating the impact of covariates on diagnostic tests accuracy of gestational diabetic mellitus (GDM). The proposed method unlike other regression based methods produces smooth Receiver Operating Characteristic curves (ROC) with very good precision and are unbiased. It was however applied in obtaining a given common cut-off value for the screening of GDM using glucose challenge test (GCT) in which a threshold of 177 mg/dl was recommended as the cutoff value of 50 grams GCT for screening of GDM in each trimester in GDM risk women. The proposed methods were seen to compare favorably with other known methods.

Keywords: semi-parametric, ROC curves, AUC, GDM, diagnostic test, parameters

I. Introduction

The constant evolution of medicine over the last two decades has meant that statistics has had to develop methods to solve the new problems that have appeared and has come to play a central part in methods of diagnosis of diseases. A diagnostic method consists of the application of a test with a group of patients in order to obtain a provisional diagnosis regarding the presence or the absence of a particular disease. In this work, a method of obtaining smooth Receiver Operating Characteristics (ROC) curve have been proposed for the purpose of evaluating the impact of covariates on the accuracy of diagnostic tests results. ROC curve is used as a standard statistical tool to assess the accuracy of tests that yield continuous results (Pepe, 2004). These ROC analysis is a technique for visualizing, organizing and selecting classifiers based on their performance and its graphs have long been used in signal detection theory to depict the tradeoff between true positive rates (TPR) and false positive rates (FPR) of classifiers (Swets et al, 2000).

We here propose to develop a semi-parametric generalized linear mixed model (GLMM) method based on ROC analysis. This method will be applied in evaluating the impact of covariates on the accuracy of diagnostic test results by way of obtaining a common cut off value for screening 50g glucose challenge test(GCT) for the three trimesters of pregnancy. Meanwhile, Semi-parametric approach is an intermediate strategy between parametric and non-parametric methods for estimating the ROC curve in the sense that it assumes a parametric bi-normal form for the ROC curve, but does not assume that the test results follow any particular distribution (Pepe, 2004). Unlike the non-parametric method, it yields smooth curve which is important when determining a unique optimal cutoff value (Pepe, 2004). The proposed method will be illustrated using data on gestational diabetic mellitus (GDM) and have been shown to compare favorably with other existing methods in terms of efficiency.

The Proposed Method

Defining ROC curves for Continuous Tests

Let D be a binary indicator of disease status with D=1 for diseased and D=0 for non- diseased subjects. Let Y denote a continuous test result and c be a threshold that any test results greater than c are considered to be positive. For a given threshold c,

$$\left. \begin{aligned} TPR(c) &= \Pr(Y \geq 1 | D = 1) \\ FPR(c) &= \Pr(Y \geq 1 | D = 0). \end{aligned} \right\} \quad (1)$$

Let

S_D and $S_{\bar{D}}$ denote the survivor functions for Y in the diseased and non-diseased populations:

$$\left. \begin{aligned} S_D(y) &= \Pr(Y \geq y | D = 1) \\ S_{\bar{D}}(y) &= \Pr(Y \geq y | D = 0) \end{aligned} \right\} \quad (2)$$

Then, the ROC curve can be expressed as

$$ROC(t) = S_D(S_{\bar{D}}^{-1}(t)), t \in (0, 1). \quad (3)$$

Suppose that Y_D and $Y_{\bar{D}}$ are independent and randomly chosen test results from the diseased and non-diseased population, respectively. By the definition,

$$AUC = \int_0^1 ROC(t) dt = \int_0^1 S_D(S_{\bar{D}}^{-1}(t)) dt. \quad (4)$$

GLM FORMULATION FROM LINEAR REGRESSION MODEL

Generalized linear models (GLMs) are an extension of linear models to model non-normal response variables. Because the diagnostic test results are continuous with normal error, the impact of covariates on the accuracy of a diagnosis test can be modeled through the standard linear regression model given as:

$$Y_i = \mu_i + \varepsilon_i = X_i\beta + \varepsilon_i \quad (5)$$

Where $E(Y_i) = \mu_i$; $\mu_i = X_i\beta$ or $Y_i = X_i\beta + \varepsilon_i$, $\varepsilon_i \sim N(0, \sigma^2)$ and Y follows normal distribution.

GLMs in terms of the (smooth monotonic) link function is given as

$$g(\mu_i) = X_i\beta \quad (6)$$

Glmm For Estimating Probability Under Binary Diagnostic Tests

Generalized linear mixed model (GLMM) extends GLMs by allowing for the incorporation of random effects. The use of GLMM to obtain estimated probability is justified by the fact that it allows for the analysis of binary data and logit models with categorical predictors called risk factors. Under binary diagnostic tests, each subject will have two test results over time or under different tests so that the observations within a given subject will no longer be independent and the intra-subject correlation and variation will be introduced in the analyses. Due to intra-subject variation, the impact of covariates on the accuracy of a diagnostic test will consist of both fixed global effect such as subject's age and family history, etc, as well as individual subject random effect such as hypertension and BMI, etc. In the GLMMs framework, a response Y is to be predicted by covariates (X;Z), where X is a vector-valued covariate associated with fixed effects, and Z is a vector-valued covariate associated with random effects. Because this study is continuous, the random effects from individual subjects will be incorporated in a model by the extension of GLM to GLMM, (Bresloe and Clayton, 1993) given that the linear predictor consists of two effects. Let

$$\mu_i = E(Y_i | b_i) \quad (7)$$

be the conditional mean of an outcome variable Y_i and

let

$$L_i = g(\mu_i) \quad (8)$$

be the link function that connects μ_i with a linear predictor ($X_i\beta$). Then GLMM will be written as:

$$L_i = X_i\beta + Z_i b_i \text{ for } i = 1, \dots, n \quad (9)$$

and the conditional variance is:

$$Var(Y_i | b_i) = R, \quad (10)$$

where Y_i is a $n_i \times 1$ vector of test results for the i th subject, n_i is the number of outcome measures for the i th subject, X_i is $n_i \times p$ containing known covariates/risk factors that are associated with the fixed effects. β , the fixed effect parameter vector is $p \times 1$, and Z_i is $n_i \times k$ representing known covariates that are associated with the random part of the model. b_i , the random effect parameter vector is $k \times 1$, is distributed as $N(0, G_i)$. The conditional distribution of $Y_i | b_i$ will play the same role as the distribution of Y in the fixed effects GLM. Since binary outcome measures results from converting a continuous test by some cutoff value c , let p_{ij} be the probability of being diseased/positive for the i th subject at the j th time point; and L_{ij} be the canonical link function for the i th subject at the j th time point. Modeling the probability of being diseased/positive through GLMM gives:

$$L_{ij} = X_{ij}\beta + Z_{ij}b_i \tag{11}$$

and

$$L_{ij} = g(p_{ij}) = \log(p_{ij}/(1 - p_{ij})) \tag{12}$$

Or

$$\log(p_{ij}/(1 - p_{ij})) = X_{ij}\beta + Z_{ij}b_i \tag{13}$$

Let $\hat{\beta}$ and \hat{b}_i be the estimates of β and b_i while \hat{p}_{ij} is the corresponding estimate of p_{ij} so that we have

$$\log(\hat{p}_{ij}/(1 - \hat{p}_{ij})) = X_{ij}\hat{\beta} + Z_{ij}\hat{b}_i \tag{14}$$

or

$$\hat{p}_{ij} = \frac{\exp(X_{ij}\hat{\beta} + Z_{ij}\hat{b}_i)}{1 + \exp(X_{ij}\hat{\beta} + Z_{ij}\hat{b}_i)} \tag{15}$$

The estimated probability \hat{p}_{ij} where ($i = 1, \dots, n$ and $j = 1, \dots, n_i$) of being diseased/positive will then serve as a bio-marker for constructing the ROC curve for longitudinally discriminating a diseased/positive subject from a non-diseased/negative subject. Let $ROC(t_{X,Z})$ denote the ROC value with FPR at t that is associated with the fixed effect predictors X and random effects predictors Z . By definition, the AUC is:

$$AUC = \int ROC(t_{X,Z}) dt_{X,Z} \tag{16}$$

where the integration limits run from 0 to 1. See Pepe, 2004

ESTIMATION OF AUC FROM PREDICTED PROBABILITIES

In calculating the AUC based on predicted probabilities, we apply the Wilcoxon nonparametric method to compare the magnitude of the predicted probabilities of each discordant (diseased and non-diseased) pair. Since each subject has binary observations and the outcome values could vary from test type to test type at a given time for a given subject (i.e., could be diseased based on a test type at one time and non-diseased based on another test type at another time), the subject cohort can not be simply classified as diseased and non-diseased subject groups. The classification rule needs to be modified and based on the observations rather than individual subjects itself. Let

$$\hat{p}_{ij(D_1)} \text{ where } (i = 1, \dots, n \text{ and } j = 1, \dots, s_i) \tag{17}$$

be the predicted probability of positivity of a disease or condition for the i th subject at the j th time point for the test type that had a diseased/positive observed value, and let $\hat{p}_{kl(D_0)}$ where $(k = 1, \dots, n \text{ and } l = 1, \dots, t_k)$ (18)

be the predicted probability of positivity of a disease or condition for the k th subject at the l th time point that had a non-diseased/ negative observed value. Let

$$N_{D_1} = \sum_{i=1}^n s_i \text{ and } N_{D_0} = \sum_{i=1}^n t_k \tag{19}$$

be the total number of observations with positive or negative observed values. Then the total number of discordant (disagreeing) pairs equals $N = N_{D_1} * N_{D_0}$. Let $I(.)$ be an indicator variable that:

$$I(\hat{p}_{ij(D_1)} > \hat{p}_{kl(D_0)}) = \begin{cases} 1 & \text{if } \hat{p}_{ij(D_1)} > \hat{p}_{ij(D_0)} \\ 0 & \text{if } \hat{p}_{ij(D_1)} < \hat{p}_{ij(D_0)} \end{cases} \tag{20}$$

and

$$I(\hat{p}_{ij(D_1)} = \hat{p}_{kl(D_0)}) = \begin{cases} 1 & \text{if } \hat{p}_{ij(D_1)} = \hat{p}_{ij(D_0)} \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

The AUC is estimated as:

$$\hat{AUC} = \frac{\sum_i \sum_j \sum_k \sum_l \left(I(\hat{p}_{ij(D_1)} > \hat{p}_{kl(D_0)}) + 0.5 * I(\hat{p}_{ij(D_1)} = \hat{p}_{kl(D_0)}) \right)}{N_{D_1} * N_{D_0}} \tag{22}$$

This semi-parametric method proposed and developed above compares favorably in terms of reliability, unbiased ness, consistency, precision of ROC and AUC estimates, accuracy, robustness, smoothness of ROC curve and simplicity with other existing methods such as those from Alonzo and Pepe (2002), Zou and Hall (2000); Cai and Moskowitz (2004).It is however the first contribution of this works to knowledge. Other authors have shown also that semi-parametric method is preferred than other existing methods for evaluating the impact of covariates in diagnostic testing (Colak et al, 2012).

II ILLUSTRATIVE EXAMPLE

The retrospective data for illustration is obtained from five different hospitals in Ebonyi State, Nigeria particularly from the medical records departments based on those that met the inclusion criteria of at least having one of the risk factors for GDM. The data covered for a study period of two years starting from January 2010 to December 2011.The GCT and oral glucose tolerance test (OGTT) results of these women that passed through screening and diagnosis respectively were collected for further analysis.

GCT using 50 grams glucose oral load with plasma glucose measurement after 1 hour was done at the first ANC, less than 24weeks, 24-28 weeks, and 32-34 weeks of gestation. The positive result was defined as plasma glucose of 140 mg/dl or greater than or equal to 7.8 mmol/l. Subsequently, OGTT with 100 grams glucose ingestion was performed for GDM diagnosis using the plasma glucose cutoff values of 105, 190, 165, and 145 mg/dl at the fasting period, 1, 2, and 3 hours, orderly. This is considered to be normal. Thus, adopting the National Diabetes Data Group criteria for the diagnosis of GDM. The OGTT was considered positive (GDM present) when any two of plasma glucose values were equal or greater than the normal criteria. Note to convert from mg/dl to mmol/l, we multiply by 0.0555. Out of the 1114 that met the inclusion criteria and which were subjected to 1h, 50g GCT, in the various hospitals, the following number of pregnant women (422, 391 and 301 for the three trimesters respectively) tested positive for GCT having exceeded the threshold value of 140mg/dl.

The data for the ROC analysis is here presented as follows:

Table 1: Demographic data on GDM for three trimesters in the studied population (n=1114)

	1st Trimester (n=422)	2 nd Trimester (n=391)	3 rd Trimester (n=301)	All Trimesters (n=1114)
Mean Age(yr)	31.0	31.44	30.49	31.02
BMI(kg/m ²)	23.44 ± 4.01	24.97 ± 4.39	27.49 ± 4.74	24.97 ± 4.61

Table 2: The cross classification of screening test results and true disease status for all trimesters

OGTT (Gold standard)

Test result of GCT for GDM	1 st Trimester GDM present =B; GDM absent= \bar{B}			2 nd Trimester GDM present =B; GDM absent= \bar{B}			3 rd Trimester GDM present =B; GDM absent= \bar{B}			All Trimester GDM present =B; GDM absent= \bar{B}		
	B	\bar{B}	Total	B	\bar{B}	Total	B	\bar{B}	Total	B	\bar{B}	Total
Positive (A)	47	13	60	31	20	51	18	18	36	96	51	147
Negative (\bar{A})	124	238	362	85		340	35	230	265	248	719	967
Total	171	251	422	116	275	391	53	248	301	344	770	1114

III. RESULTS OF ROC ANALYSIS

When sensitivity and 1-specificity were used to perform ROC curve, the cutoff value of GCT for screening of GDM in the first trimester was 179 mg/dl(Table 5), which had maximum area under the ROC (AUC=0.7204) (Fig. 1). This cutoff value has respectively the following values for sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) as 27.49% (95% CI 74.4-82.26), 94.82% (95% CI 61.22-70.27), 78.33% (95% CI 23.23-31.74) and 65.75% (95% CI 92.71-96.94). The results were considered statistically significant at p < 0.05 with 95% of confidence interval (CI). In the second trimester, the cutoff value of GCT for screening of GDM was 177 mg/dl(Table 5), which had maximum AUC of 0.6789(Fig.2). From the four-fold table of the cutoff value 177 mg/ dl (Table 3), the sensitivity, specificity, PPV and NPV, were 26.72% (95% CI 55.94-65.62), 92.73% (95% CI 70.71-79.29), 60.78% (95% CI 22.34-31.11) and 75% (95% CI 90.15-95.3) respectively.

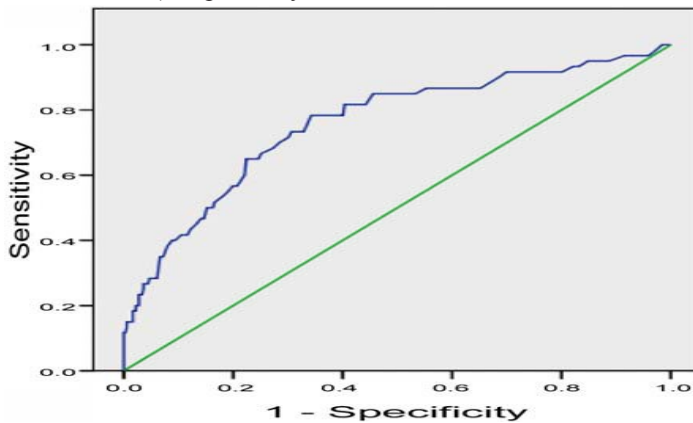


Fig. 1 ROC curve of the 1st trimester

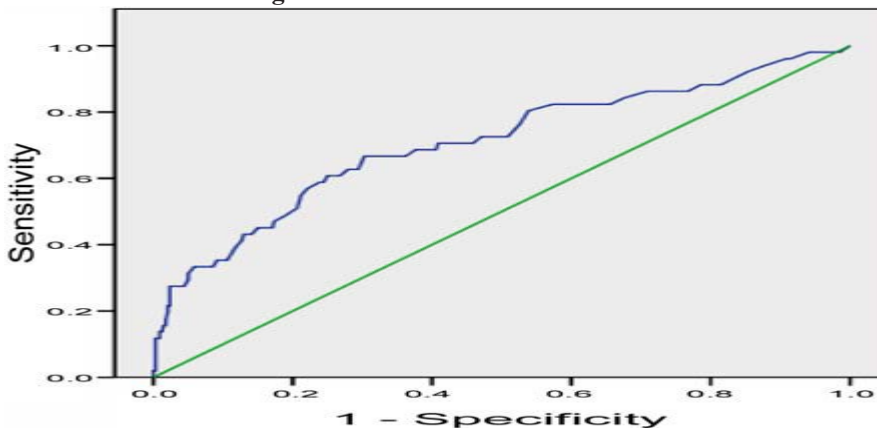


Fig. 2 ROC curve of the 2nd trimester

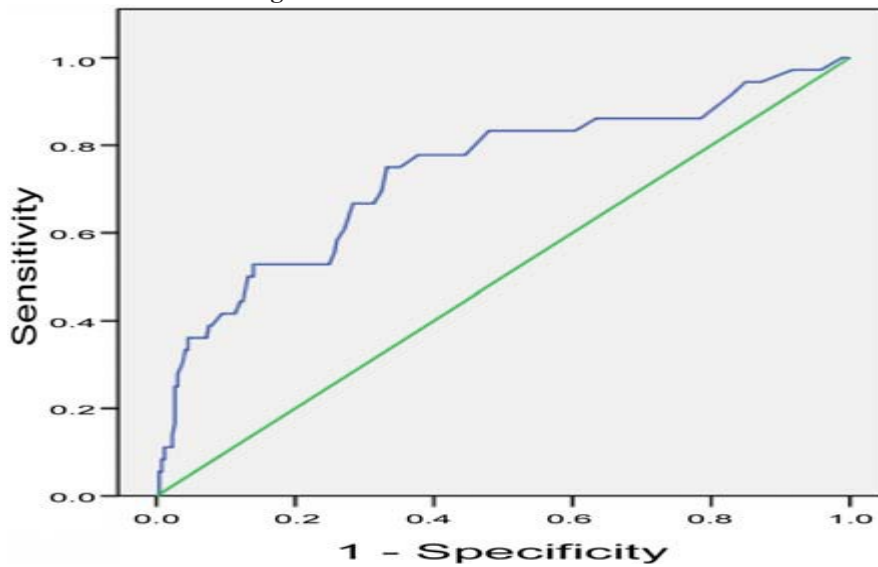


Fig. 3 ROC curve of the 3rd trimester

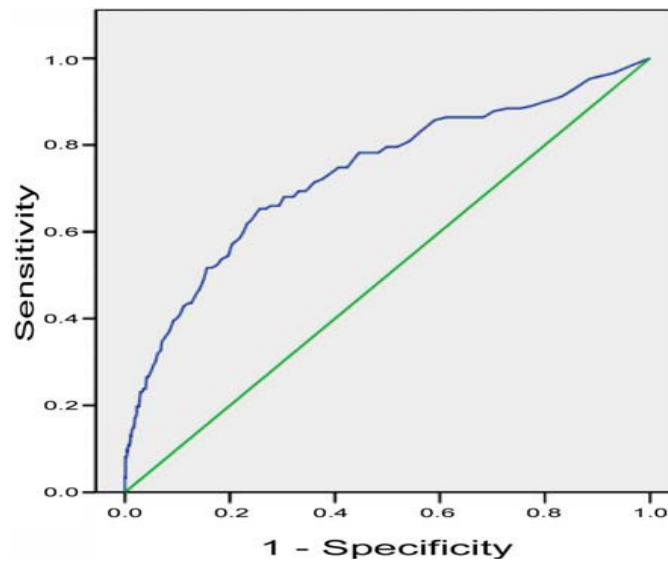


Fig. 4 ROC curve of all trimesters

In the third trimester, the cutoff value of GCT for screening of GDM was 184 mg/dl (Table 5), which had maximum AUC=0.684 (Fig. 3). From the four-fold table of the cutoff value 184 mg/dl, the sensitivity, specificity, PPV and NPV were 33.96% (95% CI 44.35-55.65), 92.74% (95% CI 82.97-90.62), 50% (95% CI 28.61-39.31) and 86.79% (95% CI 89.81-95.67) respectively. In all trimesters, the cutoff value of GCT for screening of GDM was 179 mg/dl (Table 5), which had maximum AUC=0.6983 (Fig. 4). From the four-fold table of the cutoff value 179 mg/dl, the sensitivity, specificity, PPV and NPV were 27.91% (95% CI 62.51-68.1), 93.38% (95% CI 71.79-76.92), 65.31% (95% CI 25.27-30.54) and 74.35% (95% CI 91.92-94.84) respectively.

IV. SUMMARY AND CONCLUSIONS

We therefore conclude that there was still variable cutoff value of GCT for screening of GDM. In the present study, the cutoff values of GCT in 1st, 2nd, 3rd, and all trimesters were 179, 177, 184, and 179 mg/dl respectively. The values had increasing tendency with progressing pregnancy. The authors recommended 177 mg/dl as the cutoff value for screening of GDM in each trimester in potential diabetic pregnancy. The average BMI in the present study was rather low ($24.97 + 4.61 \text{ kg/m}^2$). The BMI was least in the 1st trimester ($23.44 + 4.01 \text{ kg/m}^2$) and most in the 3rd trimester ($27.49 + 4.74 \text{ kg/m}^2$). These findings were due to increasing weight as progressing pregnancy. These cutoff values may be useful for the population that has low BMI or non-obese

pregnancy. Although the recommended cutoff GCT level was 177 mg/dl for screening of GDM in each trimester in gestational diabetic pregnant women, it was rather suitable for low BMI or non-obese pregnancy.

In conclusion, a threshold of 177 mg/dl was recommended as the cutoff value of 50 grams GCT for screening of GDM in each trimester in GDM risk women and suitable for low BMI or non-obese pregnancy.

Reference

- [1]. Alonzo, T.A. and Pepe, M.S. (2002). Distribution-free ROC analysis using binary regression techniques. *Biostatistics* 3, 421-432.
- [2]. Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association* 88, 9-25.
- [3]. Cai, T. and C. S. Moskowitz. (2004). "Semi-parametric estimation of the binormal ROC curve for a continuous diagnostic test," *Biostatistics*, vol. 5, no. 4, pp. 573-586.
- [4]. Colak Ertugrul, FezanMutlu, Cengiz Bal, Setenay Oner, Kazim Ozdamar, Bulent Gok, and Yuksel Cavusoglu. (2012). Comparison of Semi-parametric, Parametric, and Nonparametric ROC Analysis for Continuous Diagnostic Tests Using a Simulation Study and Acute Coronary Syndrome Data. Hindawi Publishing Corporation. *Computational and Mathematical Methods in Medicine* Volume 2012, Article ID 698320, 7 pages doi:10.1155/2012/698320
- [5]. Pepe, M. S.(2004). *The Statistical Evaluation of Medical Tests for Classification and Prediction*, Oxford University Press, New York, NY, USA.
- [6]. Swets, J.A., Dawes, R.M., Monahan, J., 2000. Better decisions through science. *Scientific American* 283, 82-87.
- [7]. Zou, K. H. and Hall, W. J. (2000). Two transformation models for estimating an ROC curve derived from continuous data. *Journal of Applied Statistics* 27:5, 621-631.