

January 2012

# Evaluation of Repeated Biomarkers: Non-parametric Comparison of Areas under the Receiver Operating Curve Between Correlated Groups Using an Optimal Weighting Scheme

Ping Xu

University of South Florida, nkxp@yahoo.com

Follow this and additional works at: <http://scholarcommons.usf.edu/etd>

 Part of the [American Studies Commons](#), and the [Biostatistics Commons](#)

## Scholar Commons Citation

Xu, Ping, "Evaluation of Repeated Biomarkers: Non-parametric Comparison of Areas under the Receiver Operating Curve Between Correlated Groups Using an Optimal Weighting Scheme" (2012). *Graduate Theses and Dissertations*.  
<http://scholarcommons.usf.edu/etd/4261>

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact [scholarcommons@usf.edu](mailto:scholarcommons@usf.edu).

Evaluation of Repeated Biomarkers:  
Non-parametric Comparison of Areas under the Receiver Operating Curve  
Between Correlated Groups Using an Optimal Weighting Scheme

By

Ping Xu

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
Department of Epidemiology and Biostatistics  
College of Public Health  
University of South Florida

Major Professor: Yougui Wu, Ph.D.  
Yiliang Zhu, Ph.D.  
Getachew Dagne, Ph.D.  
Jeffrey Krischer, Ph.D.

Date of Approval:  
July 13, 2012

Keywords: Prediction Accuracy, Receiver Operation Curve, AUC, Repeated  
Biomarker, Optimal Weight, Correlated Groups

Copyright © 2012, Ping Xu

## ACKNOWLEDGEMENTS

I am grateful for the encouragement from my committee members, support from my family and faith from friends during last six years of my life's journey. Without them, I would never have been able to finish my doctoral dissertation.

I would like to express my sincerest gratitude to my major advisor, Dr. Yougui Wu, for his guidance, patience, and providing me with the critical thinking for doing research.

I would like to thank Dr. Jeffrey Krischer, who provided me the opportunity to gain experience in the field and master practical statistical skills beyond textbook examples, and for patiently reviewing my writing.

I would also like to thank Dr. Yiliang Zhu and Dr. Getachew Dagne for guiding my research for the past several years and helping me to develop my background in statistics research. Thank you for your valuable comments. My research would not have been possible without your helps.

Finally, the work is dedicated to my husband Giffe for emotional support, and dearest Aeon and Ashton for being my inspiration. You were always there to cheer me up and have accompanied me through the good times and the bad.

## Tables of Contents

List of Tables.....	iv
List of Figures.....	v
Abstract.....	vi
Chapter 1: Introduction and Literature Review.....	1
1.1 Overview of biomarker evaluation in health research.....	1
1.1.1 What is a biomarker?.....	1
1.1.2 What is the utility of evaluating biomarkers?.....	2
1.1.3 How are biomarkers measured?.....	3
1.2 Measures of accuracy for continuous biomarker.....	4
1.2.1 The history of ROC curves.....	5
1.2.2 Definition of ROC.....	5
1.2.3 Basic principal of the ROC curve.....	6
1.2.4 Area under the ROC curve.....	9
1.3 Non-parametric estimation of AUC.....	10
1.3.1 Non-parametric methods.....	11
1.3.2 Comparison of two AUCs.....	13
1.4 Biomarker evaluation in longitudinal settings.....	15
1.4.1 Statistical methods for repeated biomarkers.....	16
1.4.2 Area under curve for repeated biomarkers.....	17

1.5 Aims and Objectives.....	19
Chapter 2: Weight Optimization for Comparing Areas under the ROC Curve	
between Correlated Groups for a Repeated Biomarker.....	20
2.1 Motivation and background.....	20
2.2 Non-parametric comparison of two correlated AUC.....	24
2.2.1 Notations.....	24
2.2.2 Non-parametric estimate of area under the ROC.....	25
2.3 The optimal weighting scheme.....	26
2.4 Asymptotic variance and the relative efficiency.....	31
2.5 Simulation results.....	35
2.6 Application on the thyroidism study data.....	37
2.7 Concluding remarks.....	38
2.8 Acknowledgements.....	39
2.9 Appendix.....	40
2.10 Bibliography.....	46
Chapter 3: Prognostic Performance of Metabolic Indexes in Predicting	
Onset of Type 1 Diabetes.....	52
3.1 Introduction.....	52
3.2 Subjects and methods.....	54
3.3 Laboratory measurements.....	55
3.4 Data analysis.....	56
3.5 Results.....	57
3.6 Discussions and conclusions.....	59

3.7 Acknowledgements.....	63
3.8 Bibliography.....	64
Chapter 4: Concluding Remarks and Future Research.....	72
4.1 Summary and conclusions .....	72
4.2 Future research.....	74
Bibliography.....	76
Appendices .....	81
Appendix A: R/S-PLUS Codes for Theoretical Manuscript.....	82
Appendix B: SAS Codes for Application Manuscript .....	97
Appendix C: Permission for Use of Published Article .....	122
About the Author.....	End Page

## LIST OF TABLES

Table 1.1 The Definition and Types of Biomarkers (NIH, 2001).....	2
Table 2.1 Results for Comparing Accuracy of TPOAb in Predicting Thyroid Function.....	37
Table 2.2 Simulation Results for AUCs Comparison (n=50).....	48
Table 2.2 Simulation Results for AUCs Comparison (n=50) (Continued).....	49
Table 2.3 Simulation Results for AUCs Comparison (n=100).....	50
Table 2.3 Simulation Results for AUCs Comparison (n=100) (Continued).....	51
Table 3.1 Clinical and Metabolic Characteristics of the Study Subjects.....	68
Table 3.2 AUC, Specificity, Sensitivity, PPV and NPV at the Optimal Cut-off Value.....	69

## LIST OF FIGURES

Figure 1.1	ROC Curve.....	7
Figure 1.2	The Trade-off by Varying the Cut-off Point.....	8
Figure 2.1	The Effect of Within-subject Correlation Coefficient ( $\rho_{w}$ ), Between-group Coefficient ( $\rho_{b}$ ), and the Proportion of Disease ( $\tau$ ) on the Asymptotic Relative Efficiencies.....	47
Figure 3.1	ROC Curves for Various Metabolic Indexes. ....	70
Figure 3.2	Kaplan-Meier Curves by the Level of OGTT 2-hour Glucose.....	71
Figure 3.3	Kaplan-Meier Curves by the Level of IVGTT FPIR/HOAM-IR.....	71



## ABSTRACT

Receiver Operating Characteristic (ROC) curves are often used to evaluate the prognostic performance of a continuous biomarker. In a previous research, a non-parametric ROC approach was introduced to compare two biomarkers with repeated measurements. An asymptotically normal statistic, which contains the subject-specific weights, was developed to estimate the areas under the ROC curve of biomarkers. Although two weighting schemes were suggested to be optimal when the within subject correlation is 1 or 0 by the previous study, the universal optimal weight was not determined. We modify this asymptotical statistic to compare AUCs between two correlated groups and propose a solution to weight optimization in non-parametric AUCs comparison to improve the efficiency of the estimator. It is demonstrated how the Lagrange multiplier can be used as a strategy for finding the weights which minimize the variance function subject to constraints. We show substantial gains of efficiency by using the novel weighting scheme when the correlation within group is high, the correlation between groups is high, and/or the disease incidence is small, which is the case for many longitudinal matched case-control studies. An illustrative example is presented to apply the proposed methodology to a thyroid function dataset. Simulation results suggest that the optimal weight performs well with a sample size as small as 50 per group.

# Chapter 1

## Introduction and Literature Review

### 1.1 Overview of biomarker evaluation in health research

#### 1.1.1 What is a biomarker?

The term biomarker, i.e. biological marker, was first introduced in 1989 as a medical subject heading term: measurable and quantifiable biological parameters (e.g. specific enzyme concentration, specific hormone concentration, specific gene phenotype distribution in a population, presence of biological substances) which serve as indices for health-related and physiology-related assessments, such as disease risk, psychiatric disorders, environmental exposure and its effects, disease diagnosis, metabolic processes, substance abuse, pregnancy, cell line development, epidemiologic studies, etc (Vasan, 2006). Biomarkers are anatomic, physiologic, biochemical, or molecular parameters associated with the presence and severity of specific disease states. In clinical or public health studies, a biomarker is defined as an objective measurement that acts as an indicator of normal biological processes, pathogenic processes or pharmacologic responses to therapeutic intervention (e.g. cholesterol, c-peptide secretion, or PSA) and qualification should be taken to mean demonstrating the utility of the marker(s) for use in clinical or public health studies (NIH-FDA

Conference, 1998). In 2001, an NIH working group standardized the definition of a biomarker as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention" and defined types of biomarkers (Table 1.1) (Biomarkers Definitions Working Group, 2001).

Table 1.1: The Definition and Types of Biomarkers (NIH, 2001).

Biological marker (biomarker):	A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to a therapeutic intervention.
Type 0 biomarker:	A marker of the natural history of a disease and correlates longitudinally with known clinical indices.
Type I biomarker:	A marker that captures the effects of a therapeutic intervention in accordance with its mechanism of action.
Surrogate end point:	A marker that is intended to substitute for a clinical end point; a surrogate end point is expected to predict clinical benefit (or harm or lack of benefit or harm) on the basis of epidemiological, therapeutic, pathophysiological, or other scientific evidence.
Risk factor:	A risk factor is associated with a disease because it is in the causal pathway leading to the disease.
Risk marker:	A risk marker is associated with the disease (statistically) but need not be causally linked; it may be a measure of the disease process itself.
Clinical end point:	A characteristic or variable that reflects how a patient feels, functions, or survives.
Intermediate end point:	A true clinical end point (a symptom or measure of function, such as symptoms of angina frequency or exercise tolerance) but not the ultimate end point of the disease, such as survival or the rate of other serious and irreversible morbid events.

### 1.1.2 What is the utility of evaluating biomarkers?

The use of biomarkers is of growing importance in clinical prognosis or diagnosis of disease. Biomarkers are seen as valuable tools for: diagnosing disease; risk stratification and depicting the mechanism of action of therapeutic agents; and, investigating the safety profile of an intervention

drug and the response of patients to treatment. For example, in a clinical study for drug development an appropriate biomarker can help reduce the number of participants needed for the study and enable critical decisions on the efficacy and safety of a new drug to be made more quickly, thus reducing patient exposure to minimum and reducing cost. Identification and improved use of disease biomarkers are also important for prevention and early diagnosis, which are essential to reduce mortality and complications from disease. Depending on the type of information biomarkers provide, biomarkers can be classified as antecedent biomarkers (identifying the risk of developing an illness), screening biomarkers (screening for subclinical disease), diagnostic biomarkers (recognizing overt disease), staging biomarkers (categorizing disease severity), or prognostic biomarkers (predicting future disease course, including recurrence and response to therapy, and monitoring efficacy of therapy). Biomarkers may also serve as surrogate end points (Mayeux, 2004).

### **1.1.3 How are biomarkers measured?**

Biomarkers are detectable and measureable by a variety of methods, including physical examination, laboratory assays and medical imaging. A biomarker may be measured in a biological sample (such as a blood, urine, or glucose test), it may be a recording obtained from a person (such as blood pressure, weight, or temperature), or it may be an imaging test (such as echocardiogram or CT scan). New medical technologies promise a vast array of tools for diagnosis and screening. In recent years, there has been much focus on studying biomarkers that could potentially provide non-invasive and accurate ways of detecting disease, predicting disease progression, and evaluating patients' response to treatment. Tools such as gene expression profiling and protein mass spectrometry provide a plethora of possible biomarkers that must be studied further to assess their accuracy.

It has been recognized that the qualification of biomarkers is under-resourced, in part because it is seen as a major undertaking and may be high-risk. However, the risk in exploring a putative biomarker for clinical studies is reduced if the biomarker is biologically plausible, that is, it can be linked closely to the pathology of the disease or is in a pathway that is closely linked to the effect of a treatment. Biomarker measurements can yield the qualitative results, with simply positive or negative. However, many of these biomarkers that are used in a qualitative manner are measured on a continuous scale or an ordinal scale. Statistical techniques to evaluate biomarkers are developed mainly to assess the strength of association of biomarker with the disease. One of most important aspects for the evaluation of biomarkers to be used in clinical research is their sensitivity, specificity, and overall diagnostic/prognostic ability. In this dissertation focus will be placed on evaluating the the diagnostic/prognostic ability of tests measured by the continuous biomarker.

## **1.2 Measures of accuracy for continuous biomarker**

Often, biomarkers are measured on a continuous scale, such as an amount or concentration. The receiver Operating Characteristics (ROC) Curve is one of the best statistical methods for evaluating the diagnostic/prognostic performance of qualitative biomarkers. It is a technique for visualizing, organizing and selecting classifiers based on their predictive accuracy.

ROC has been used in many areas of research since it has been developed. ROC methodology proves to be particularly useful not only in measuring the predictive accuracy of a test or decision-making strategy, but also in dealing with sensitivity/specificity tradeoffs when a classifier variable is actually used in final classification decision-making.

### 1.2.1 The history of ROC curves

The ROC methodology was first developed by electrical engineers and radar engineers during World War II for detecting enemy objects in battle fields, also known as the signal detection theory. In the 1950s, ROC curves were employed in psychophysics to assess human (and occasionally non-human animal) detection of weak signals (Green et al., 1966). In the late 1960s, it was first used in medicine for the assessment of imaging devices (reviewed by Zweig et al., 1993). Since then, it has been increasingly used for the evaluation of clinical laboratory tests (Henderson, 1993; Smith et al., 1995; Greiner et al., 1997) in epidemiology and medical research. More recently, ROC curves also proved useful for the evaluation of engineering techniques (Nockemann, 1991). For example, Spackman demonstrated the value of ROC curves in comparing and evaluating different classification algorithms of machine learning (Spackman, 1989).

### 1.2.2 Definition of ROC

An ROC curve is a plot of TPR (or sensitivity) (plotted on the y axis) versus its FPR (or 100%- specificity) (plotted on the x axis) (Figure 1.1). The scales of the curve, sensitivity and specificity are the basic measures of accuracy and are easily read from the plot. Two parameters (sensitivity and specificity) can fully describe the probabilities of the four possible test outcomes (TPR=True Positive Rate, TNR=True Negative Rate, FPR=False Positive Rate and FNR=False Negative Rate).

Mathematically, using a threshold  $c$  to define a binary test from a continuous biomarker  $Y$  ( $y \in (-\infty, +\infty)$ ) as

Positive if  $Y \geq c$ ,

Negative if  $Y \leq c$ ,

let the corresponding true and false positive rate at the threshold  $c$  be  $TPR(c)$  and  $FPR(c)$  respectively, we define:

$$TPR(c) = P[Y \geq c | D = 1],$$

$$FPR(c) = P[Y \geq c | D = 0],$$

and the ROC curve is the entire set of possible true and false positive fractions obtained by dichotomizing  $Y$  with different thresholds. That is:

$$ROC(\cdot) = \{(FPR(c), TPR(c)), c \in (-\infty, +\infty)\}.$$

As the threshold increases, both  $FPR(c)$  and  $TPR(c)$  decrease. As  $c$  approaches positive infinity,  $\lim_{c \rightarrow \infty} TPR(c) = 0\%$  and  $\lim_{c \rightarrow \infty} FPR(c) = 0\%$ . At the other end, as  $c$  approaches negative infinity,  $\lim_{c \rightarrow -\infty} TPR(c) = 100\%$  and  $\lim_{c \rightarrow -\infty} FPR(c) = 100\%$ .

### 1.2.3 Basic principal of the ROC curve

Area under the ROC curve can take on values between 0% and 100%. It is a monotonic increasing function. A biomarker with an area under the ROC curve of 100% is perfectly accurate because the true positive rate (i.e sensitivity) is 100% and the false positive rate (i.e 1-specificity) is 0%. In contrast, a biomarker with an area of 0% is perfectly inaccurate. ROC plots for diagnostic tests with perfect discrimination between disease and disease free reference samples (no overlap of values of the two groups) pass through the co-ordinates (0%, 100%) which represent 100% sensitivity (TPR) and specificity (1-FPR). In this case, the AUC would be 100%. The diagonal divides the ROC space. Points above the diagonal represent good classification results, points

below the line poor results. Biomarkers that better discriminate disease from disease-free have ROC curves that are higher and to the upper left corner.

The ROC plot has many advantages over single measurement of sensitivity and specificity. The ROC curve displays all possible thresholds and illustrates how sensitivity and specificity vary together.

ROC curves are invariant with respect to monotonic transformations of the original test data such as the linear (with positive slope), logarithmic and square root (Campbell, 1994). The curve does not depend on the scale of the test results, that is, we can alter the test results by adding or subtracting a constant or taking the logarithm or square root without any change to the ROC curve.

Because sensitivity and specificity are independent of disease prevalence, so too is the ROC curve.

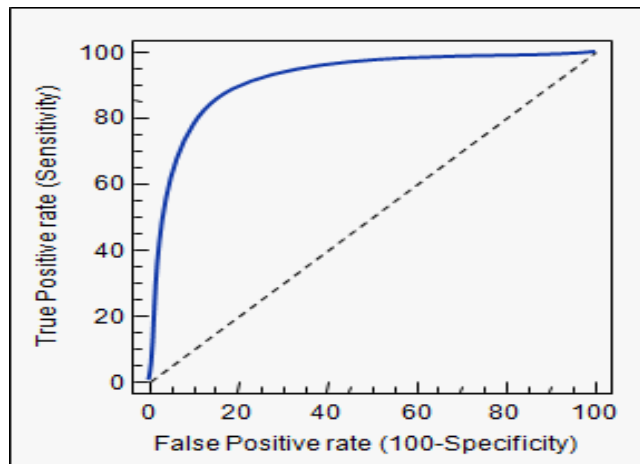


Figure 1.1: ROC Curve.

A theoretical optimal prediction can achieve 100% sensitivity (i.e. predict all people from



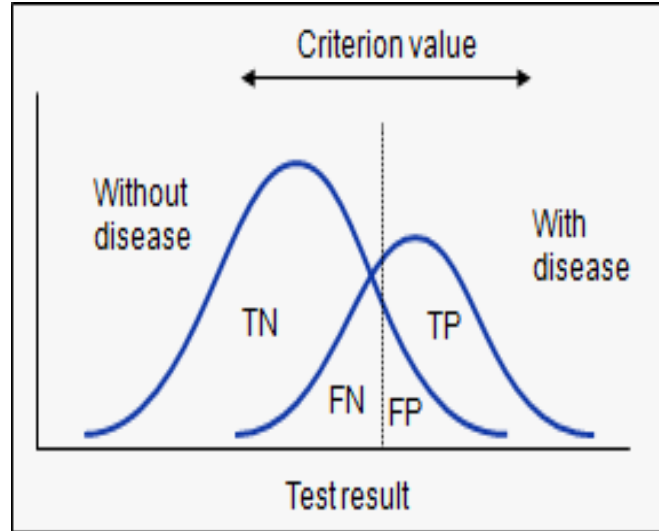


Figure 1.2: The Trade-off by Varying the Cut-off Point.

the sick group as sick) and 100% specificity (i.e. not predict anyone from the healthy group as sick). However, there is usually a trade-off between the measures for the test. The choice of a suitable cut-off value will vary with circumstances. Three criteria are commonly used to find optimal cut-off value from ROC curve, including point on curve that is closest to the (0, 1), Youden index, and minimize average cost criterion (Metz, 1978; Perkins et al., 2005). For example, we can make a test have very high sensitivity, but this sometimes results in low specificity. As described by Schfer (Schafer, 1989), the cut-off value and the resulting sensitivity (or specificity) can be obtained for a pre-selected specificity (or sensitivity). It may also depend on the availability of health care resources and perceived gravity of the disease condition. That is, the choice of criterion value depends on the trade-off that is acceptable between failing to detect disease and falsely identifying disease with the test as depicted in Figure 1.2. Optimally, the cut-off selection procedure is an informed decision that takes into account the epidemiologic situation (e.g. prevalence in the target population) and the relative consequences of FN and FP test results (which may differ

for every different decision-making situation). Given a disease of low prevalence and high cost of false-positive diagnoses, it may be advisable to choose a cut-off at the lower part of the curve to maximize specificity. On the other hand, if the disease occurs at high prevalence and missing any diseased animal has serious consequences, a cut-off value towards the upper part of the curve would be selected to maximize sensitivity. For example, for evaluating women with a personal history of breast cancer, we need a cut point with good sensitivity even if the specificity is low.

### 1.2.4 Area under the ROC curve

The area under the ROC curve (AUC) is a global (that is based on all possible threshold values) summary statistic of diagnostic accuracy. AUC is a measure of how well a biomarker can distinguish between two diagnostic groups (diseased/diseased-free). It is widely used as a summary index for the ROC. Mathematically, AUC is defined as:

$$AUC = \int_0^1 ROC(t)dt \quad (1.1)$$

AUC under the ROC curve area has several interpretations: (a) the average value of sensitivity for all possible values of specificity,

$$\begin{aligned} AUC &= \int_0^1 Sen(Spe^{-1}(t))dt \\ &= \int_{-\infty}^{\infty} Sen(c)dSpe(c) \\ &= Average(Sensitivity) \end{aligned}$$

or (b) the average value of specificity for all possible values of sensitivity (Metz, 1978; Hanley et

al, 1982), and

$$\begin{aligned}AUC &= \int_0^1 Sen(Spe^{-1}(t))dt \\ &= \int_{-\infty}^{\infty} Spe(c)dSen(c) \\ &= Average(Specificity)\end{aligned}$$

or (c) the probability that a randomly selected patient with disease has a test result that indicates greater suspicion than a randomly chosen patient without disease (Swets, 1988).

$$\begin{aligned}AUC &= \int_{-\infty}^{\infty} Sen(c)dSpe(c) \\ &= \int_{-\infty}^{\infty} (1 - G_Y(c))dF_X(c) \\ &= \int_{-\infty}^{\infty} F_X(c)dG_Y(c) \\ &= Pr(Y > X)\end{aligned}$$

The area measures discrimination, that is, the ability of the test to correctly classify those with and without the disease. Consider the situation in which patients are already correctly classified into two groups. You randomly pick one from the disease group and one from the disease-free group and do the test on both. The patient with the more abnormal test result should be the one from the disease group. The area under the curve is the percentage of randomly drawn pairs for which this is true, that is, the test correctly classifies the two patients in the random pair.

### 1.3 Non-parametric estimation of AUC

A subject is assessed as diseased (positive) or non-diseased (negative) depending on whether the corresponding biomarker value is greater than or less than or equal to a given threshold value.

Associated with any threshold value  $c$  is the probability of a true positive (sensitivity) and the probability of a true negative (specificity). The theoretical ROC curve is a plot of  $q = \text{sensitivity}$  versus  $p = 1 - \text{specificity}$  for all possible threshold values. Suppose that biomarker results  $x_1, x_2, \dots, x_m$  and  $y_1, y_2, \dots, y_n$  are available from the non-diseased and diseased population having cumulative distribution functions  $F(x)$  and  $G(y)$ , respectively. Then at threshold  $c$ ,  $q = 1 - F(c)$  and  $p = 1 - G(c)$ . The theoretical ROC curve is a plot of  $(1 - G(c), 1 - F(c))$  for all possible values of  $c$  or, equivalently, a plot of  $(p, q)$  where  $p$  ranges from 0 to 1 and  $q = 1 - F(G^{-1}(1 - p))$ . AUC is the area under this curve. AUC can be estimated parametrically, with the assumption that either the biomarkers themselves or some unknown monotonic transformation of the biomarkers follows a bi-normal distribution. Nevertheless, if the normal assumption is not met or tenuous, the non-parametric methods using the empirical ROC curve are preferred. In this section, we review two non-parametric approaches for estimating the AUC. One is to use the trapezoidal rule to approximate AUC by integration and the second is to use the Mann-Whitney statistic.

### 1.3.1 Non-parametric methods

If sensitivity and specificity are denoted by  $q$  and  $p$ , respectively, the trapezoidal rule calculates the area by joining the points  $(q, p)$  at each interval value of the continuous biomarker and draws a straight line joining the x-axis. This forms several trapezoids and their area can be easily calculated and summed.

Another interesting parallel meaning for AUC is that AUC is the probability that a randomly observed test result in diseased subjects is higher than a randomly observed test result in non-diseased subjects. It arises if we think just about comparing the distributions of biomarker between

non-diseased and diseased subjects. At the most elementary level, one might want to know whether the distributions of two groups had means that were statistically different. If the distributions were normal, a two-sample t-test would be a logical starting point to answer this question. The non-parametric analog to the t-test is the Wilcoxon rank-sum test, or synonymously the Mann-Whitney U test, which is a non-parametric test. That is, AUC is the Mann-Whitney U statistic in the version of the two-sample rank-sum test (Shapiro, 1999). It has been shown that the area under an empirical ROC curve, when calculated by the trapezoidal rule, is equal to the Mann-Whitney two sample statistics applied to the two samples (Bamber, 1975; Hanley et al., 1982). The Mann-Whitney statistics estimated the probability,  $\theta$ , that a randomly selected observation from the population represented by A will be less than or equal to a randomly selected observation represented by B. The estimate,  $\hat{\theta}$ , can be computed as the average over a kernel,  $\psi$ , as

$$\hat{\theta} = \frac{1}{mn} \sum_{j=1}^m \sum_{i=1}^n \Psi(X_i, Y_j) \quad (1.2)$$

$$\text{Where } \Psi(X_i, Y_j) = \begin{cases} 1 & \text{if } Y < X; \\ 1/2 & \text{if } Y = X; \\ 0 & \text{if } Y > X; \end{cases}$$

In terms of probabilities  $E(\hat{\theta}) = \theta = Pr(Y > X) + 1/2Pr(Y = X)$ . For continuous distributions,  $Pr(Y = X) = 0$ .

This non-parametric interpretation requires no assumptions regarding the distribution of the disease and healthy reference samples.

### 1.3.2 Comparison of two AUCs

We will review the methods that can be used to compare two AUCs, which could arise when it is necessary to compare AUCs from two different biomarkers; or compare AUCs from the same biomarker by multi-readers. An important problem concerns the comparison of two correlated AUCs, such as two AUCs derived from two biomarkers measured on the same set of patients. Consider a comparison for two continuous biomarkers BM1 and BM2 that were measured on the same  $m$  non-diseased subjects and  $n$  diseased subjects. The subjects are assumed to be mutually independent. We denote the bivariate outcomes as  $(X_{1j}, X_{2j})(j = 1, 2, \dots, m)$  and  $(Y_{1k}, X_{2k})(k = 1, 2, \dots, n)$ , respectively. The respective bivariate distributions are  $F(x_1, x_2)$  and  $G(y_1, y_2)$ ; the marginal distributions are  $F_1(x_1), F_2(x_2), G_1(y_1)$  and  $G_2(x_2)$ . For  $i$ th biomarker, it is called positive if this biomarker value exceeds a cut-off value  $c$ . The corresponding specificity for  $i$ th biomarker is  $F_i(c)$  and the corresponding sensitivity  $1 - G_i(c)$ . As  $c$  takes on all possible values, we can generate the ROC curve plot  $(1 - F_i(c), 1 - G_i(c))$ . To compare two biomarkers, we can test if the difference of two AUCs is equal to zero. The null hypothesis is:  $\Delta = \theta_1 - \theta_2 = 0$ . Denoting the sensitivity at specificity  $p$  by  $S_i(p) = 1 - G_i(F_i^{-1}(p))$ , this comparison can be expressed as:

$$\Delta = \int_0^1 (S_1(p) - S_2(p))dW(p) \quad (1.3)$$

$$\text{where } W(p) = \begin{cases} 0 & \text{if } 0 < p < p_1 \\ \frac{p-p_1}{p_2-p_1} & \text{if } p_1 \leq p \leq p_2 \\ 1 & \text{if } p_2 < p < 1 \end{cases}$$

A nonparametric procedure for comparing diagnostic biomarkers with paired or unpaired

data based on a weighted average of sensitivities was developed by Wiend, Gail, and Hanley (Wiend et al., 1983). Pearson correlation coefficients were used to estimate the correlation of the two AUCs. Hanley and McNeil (Hanley et al., 1983) used the properties of U-statistics to compare two AUCs from paired data. The approach suggested using the binormal assumption for estimation of the covariance between two area estimators. DeLong et al. (DeLong et al., 1988) proposed a more completely nonparametric approach which was based on the properties of the Mann-Whitney statistics to obtain an estimated covariance matrix. Under this framework, the AUCs of any two biomarkers can be compared by evaluating the difference of the AUCs which is asymptotically normal.

The Mann-Whitney statistics estimates the probability,  $\theta$ , that a randomly selected observation from the population represented by A will be less than or equal to a randomly selected observation from the population represented by B. It can be computed as  $\hat{\theta} = \frac{1}{mn} \sum_{j=1}^m \sum_{i=1}^n \Psi(X_i, Y_j)$ . Asymptotic normality and an expression for the variance of the Mann-Whitney statistic can be derived from theory developed for generalized U-statistics by Hoeffding (Hoeffding, 1948). Define:

$$\xi_{10} = E[\Psi(X_i, Y_j)\Psi(X_i, Y_k)] - \theta^2, j \neq k;$$

$$\xi_{01} = E[\Psi(X_i, Y_j)\Psi(X_k, Y_j)] - \theta^2, j \neq k;$$

$$\xi_{11} = E[\Psi(X_i, Y_j)\Psi(X_i, Y_j)] - \theta^2.$$

Then, we obtain the variance of  $\hat{\theta}$ :

$$var(\hat{\theta}) = \frac{(n-1)\xi_{10} + (m-1)\xi_{01}}{mn} + \frac{\xi_{11}}{mn} \quad (1.4)$$

Hoeffding's theory extends to a vector of U-statistics. Let  $\hat{\theta} = (\hat{\theta}^1, \hat{\theta}^2, \dots, \hat{\theta}^k)$  be a vector

of statistics, representing the areas under the ROC curves derived from the values of  $k$  different biomarkers. We define,

$$\xi_{10}^{rs} = E[\Psi(X_i^r, Y_j^r)\Psi(X_i^s, Y_k^s)] - \theta^r \theta^s, j \neq k;$$

$$\xi_{01}^{rs} = E[\Psi(X_i^r, Y_j^r)\Psi(X_k^s, Y_j^s)] - \theta^r \theta^s, j \neq k;$$

$$\xi_{11}^{rs} = E[\Psi(X_i^r, Y_j^r)\Psi(X_i^s, Y_j^s)] - \theta^r \theta^s.$$

The covariance of the and statistics can then be written as:

$$cov(\hat{\theta}^r, \hat{\theta}^s) = \frac{(n-1)\xi_{10}^{rs} + (m-1)\xi_{01}^{rs}}{mn} + \frac{\xi_{11}^{rs}}{mn} \quad (1.5)$$

A method of structural components by Sen (Sen, 1960) is used to provide consistent estimates of the elements of the variance-covariance matrix of U-statistics, which is equivalent to jackknifing. The resulting test statistic for comparing two or more correlated AUCs has been proved to be asymptotically chi-square distributed.

## 1.4 Biomarker evaluation in longitudinal settings

Some large-scale cross-sectional screening programs have come under heavy criticism for being ineffective and wasteful of public health resources, largely because the biomarkers used have poor specificity, that is, high false positive rate and/or low sensitivity, that is, high false negative rate. Using a biomarker by comparing the current observed level with a given critical value can lead to poor specificity and sensitivity because of between-subject variability. Because intra-subject variability is substantially less than inter-subject variability, a series of repeated biomarker can result in a more sensitive or statistically powerful test. Moreover, a series of measurements taken



periodically enable the entire response pattern to be used to assess the subject's disease status. Longitudinal biomarker analysis seeks to evaluate whether changes in a biomarker process are correlated with the clinical outcome such as disease onset or death. Repeated measures design has now grown tremendously in biological, medical and health services research because of its advantages, but statistical methods for evaluating repeated biomarkers, have been challenging for such designs due to the dependency of data points.

### **1.4.1 Statistical methods for repeated biomarkers**

One useful way to analyze such repeated biomarker is to describe the marginal mean of the outcome (disease status) as a function of the covariates (biomarkers) while taking into account the correlation between the measurements. This is the generalized estimating equations (GEE) (Liang et al., 1986; Zeger et al., 1986) approach for analyzing correlated categorical data using a variety of different working correlation matrices, such as exchangeable, banded, unstructured. Another popular way for longitudinal biomarker analysis is to employ the Cox proportional hazards regression model with a time dependent covariate (Cox, 1972; Breslow, 1975). Prentice et al. (Prentice et al., 1978) used this approach to investigate the electiveness of the grade of graft versus host disease as a risk indicator for the recurrence of Leukemia in bone marrow transplant patients.

More recently, a number of researchers have concentrated on the joint distributional modeling of the longitudinal biomarker data and the onset of a clinical event using a semi-parametric longitudinal model and a proportional hazard model (Tsiatis et al., 1995). A biomarker is assumed to follow a semi-parametric mixed model where covariate effects are modeled parametrically and subject-specific time measurements are modeled non-parametrically using a population smoothing

spline and subject-specific random stochastic processes. The Cox model is assumed for survival data by including both the current measure and the rate of change of the underlying longitudinal trajectories as covariates. DeGruttol et al. applied this method to model progression of CD-4 lymphocyte count and its relationship to survival time (DeGruttol et al., 1993; DeGruttol et al., 1994). All the methods above evaluate the statistical significance of associations between the marker process and occurrence of the disease outcome of interest and none of them directly address the issue of the sensitivity of the biomarker as a diagnostic instrument. However, it does help identify high risk patients for recruitment into prevention trials.

#### **1.4.2 Area under curve for repeated biomarkers**

The use of AUC for repeated biomarker simplifies the statistical analyses by transforming the multivariate data into univariate space, especially when the numbers of repeated measurements are high and there is a need to summarize the information. The computation of the area under the curve (AUC) is a useful method particularly in endocrinological research and other chronic disease research to comprise information that is contained in repeated measurements over time. Parker and DeLong extended the semi-parametric ROC method for cross-sectional studies (DeLong et al., 1985), whereby the biomarker test is applied once to each subject in a population, to accommodate a framework in which the biomarker test is repeated over time to monitor for occurrence of an event. Within this frame work, sensitivities are calculated directly from the data. Specificities are estimated via a discrete logistic regression model. Calculation of variances and covariance for estimates of the areas under ROC curves is simplified by relying on a one-step approximation to parameter estimates from the discrete logistic regression models with individual observations deleted (Parker et al., 2003). However, their method requires that each subject could have at most

one event.

Emir et al. (Emir et al., 1998) proposed a non-parametric approach to estimate both sensitivity and specificity of a biomarker when each subject could have more than one event. The approach involves estimating subject-specific sensitivities and specificities and taking weighted averages of these estimates over all subjects. Emir et al. (Emir et al, 2000) further extended it to the estimation of the AUC for a repeated biomarker and to comparison of two biomarkers with repeated measurements. Using new definitions of specificity and sensitivity, an asymptotically normal statistic is used to compare the average of sensitivities across all specificities or a range of specificities.

When the subjects are limited and cost is high, the efficiency of studies is crucial. In many studies, the correlated-groups designs are preferred by researchers due to increased statistical power. These designs do not include random assignment but, nevertheless, provide equivalent groups at the start of the study and allow other controls to be applied. There are two ways of introducing the correlation among participants in correlated-groups designs: (1) by having a single group of participants exposed to all of the conditions (repeated measures designs); and (2) by matching participants on some important variable(s)(matched-subjects designs). In a matched-subject study with the repeated measurements, it may be of clinical interest to compare the prognostic performance of the biomarker between correlated groups to determine if the prognostic accuracy in one of the groups is superior. The statistics for comparing the diagnostic/prognostic accuracy of a biomarker between the correlated populations is challenging because it needs to deal with not only within-subject but also between-group correlations. In the current literature, no method is available to address this situation.

## 1.5 Aims and objectives

In this dissertation, an extensive analysis of area under the ROC curve in longitudinal setting will be presented. The aim of this research was to develop novel and robust statistical methods for the evaluation of a repeated biomarker particularly where the diagnosis/prognosis accuracy of biomarker needs to be compared between the correlated populations, such as in a matched design study. Specifically, the objectives are to modify Emir's asymptotic normal statistics to compare two AUCs between correlated groups for the repeated biomarker; to research novel methods for weight optimization in comparing non-parametric AUCs of a repeated biomarker between correlated groups; to demonstrate the asymptotic relative efficiency of this novel weighting scheme; and to show how the correlations and the incidence rate of disease impact the efficiency through the simulation studies. As well, I will present a published work to demonstrate how ROC analysis can be applied in type 1 diabetes intervention research.

## Chapter 2

# Weight Optimization for Comparing Areas under the ROC Curve between Correlated Groups for a Repeated Biomarker

### 2.1 Motivation and Background

The association between thyroid disorders and risk of breast cancer has long been a subject of debate. SunCoast CCOP Research Base at the University of South Florida is conducting a matched-subject design trial to study thyroid function and breast cancer: A pilot study to estimate the prevalence of thyroid dysfunction in women diagnosed with breast cancer and the magnitude of change in thyroid function post-chemotherapy. Two hundred and fifty breast cancer patients between ages of 25 and 75, diagnosed with primary, operable, stage I-III B breast cancer with planned chemotherapy regimen Adriamycin/Cytosan (AC) plus a taxane were the trial candidates. Two hundred and fifty healthy volunteers, from the same general demographic area, without prior history of cancer and within 5 years of the patient's age (+/- 5 years) were matched for each cancer patient. All subjects were followed for two years or until disease onset. All biomarkers were collected at baseline, yearly, and at the end of study. Free T4 and TSH have been identified as reliable

and the most effective tool for diagnosing thyroid function. They are used to define the status of hypothyroidism as the gold standard. The antibody TPO is a useful biomarker for establishing the presence of thyroid autoimmunity as the cause of hypothyroidism. Previous studies have shown that the higher the TPOAb concentration, the more rapid the development of hypothyroidism. If the women with breast cancer are at higher risk of having thyroid disorders or changes in antibody level that might result in worsened hypothyroidism, it is important to understand the mechanism of these changes. The characterization of prediction accuracy is essential to the success of future prevention trials for thyroidism disorder among the women with breast cancer. Of particular interests, clinicians hypothesized that the breast cancer may alter the performance of TPO antibody titers in predicting thyroidism disorder. In this study data, there are two types of correlation among participants: (1) within-subject correlation due to repeated measurements; and (2) between-group correlation introduced by matching participants on some important variable(s) (matched-subject design). Under these circumstances, the statistics for evaluating the biomarker need to deal with not only within-subject but also between-group correlations.

The diagnostic/prognostic accuracy of a binary biomarker is most commonly measured by its sensitivity and specificity. Sensitivity is the proportion of subjects who have positive test result among those who truly have the disease. Similarly, specificity is the proportion of subjects who have negative test result among those who truly do not have the disease. When the biomarker is measured on a continuous scale, the sensitivity and specificity of the test depends on the specific threshold selected. The receiver operating characteristic (ROC) curve is a plot of sensitivity and 1-specificity for all possible thresholds in the study dataset. To evaluate the discriminatory ability of a continuous biomarker, we commonly summarize the information of the ROC curve into a single

global index. The most used index of the overall performance of a continuous biomarker is area under the ROC curve (AUC).

For evaluating a repeated biomarker in a single group, Emir et al. (1998) proposed a non-parametric estimate of sensitivity and specificity. In 2000, Emir et al. further extended the earlier research and derived the estimation of AUC under the ROC curve for a repeated biomarker. An asymptotic normal statistic was developed to test area under the ROC curve of a repeated biomarker in a single group and the bootstrap method was used to obtain the variance of this estimate. The approach allows for estimating AUC in the presence of within-subject (repeated) correlation. It involves estimating subject-specific sensitivities and specificities and taking weighted averages of these estimates over all subjects.

The nonparametric estimate of AUC proposed by Emir et al. (2000) involves assigning a weight to each subject. Two weighting schemes were suggested: (1) assigning equal weights to all biomarker measurements, and (2) assigning equal weights to all subjects. It was suggested that weighting scheme 1 would be optimal when the correlation within subject is 1 and that weighting scheme 2 would be optimal when the correlation within subject is 0. However, the universal optimal weights were not determined. More recently, a novel solution to the weight optimization problem was introduced by Wu et al. (2011), who proposed the Lagrange Multiplier Method to find the optimal weights, which minimize the variance of AUC estimate for a single repeated biomarker or comparing AUCs from two biomarkers in a single group. The optimal weights for AUC comparison proposed by Wu et al. (2011) dealt with the difference of paired biomarkers from the same subject. To compare AUCs of a repeated biomarker between two correlated groups, we may apply Wu et al.'s method (2011) to each group and obtain two sets of weights, one for each

group. However, these weights would not be optimal because the method cannot take the between-group correlation into consideration as the disease progression may be different, and the number of biomarker measurements may be different between groups. A generalization of interest for us is to find the universal optimal weights for comparing AUCs of a repeated biomarker between two correlated groups, which is often of interest to clinicians or researchers to select the best environment or condition for the biomarker and/or better understand the mechanism of the disease. The goals of this article are to modify the asymptotic statistics by Emir et al. (2000) to compare AUCs between two correlated groups and obtain the variance of this estimate; to determine the optimal weight sets which can minimize the variance of this estimate; to demonstrate the asymptotic relative efficiency of the optimal weighting scheme; and to show how the between-group correlation, the within-subject correlation, and the incidence rate of disease impact the efficiency.

The remainder of this article is organized as follows. In Section 2, the non-parametric approach for comparison of AUCs between two correlated groups is introduced and two simple weight schemes are discussed in detail. The optimal weights are derived using the Lagrange Multiplier Method in Section 3. In Section 4, the asymptotic relative efficiencies comparing different weight schemes are presented; here we also show how the correlation between groups, the correlation within subject, and/or the incidence of disease may impact the efficiency. In Section 5, we illustrate how we can apply the method on the motivation data example. In Section 6, the finite sample performance of the estimators using different weighting schemes are revealed through simulation studies. The article ends with concluding remarks in Section 7.



## 2.2 Non-parametric comparison of two correlated AUCs

### 2.2.1 Notations

Suppose that we have a random sample of  $n$  subjects who were exposed to some existing conditions (the exposed group). Each subject in the exposed group was matched to a subject who was not exposed to the conditions, based on some confounding variables, such as age or gender. The  $n$  subjects resulting from the matching form the non-exposed group. All subjects are being followed for the outcome of disease progression. A biomarker was repeatedly measured over time until the end of study or the disease progression. Let  $X_{ijl}$  be the continuous random variable whose observations are the biomarker values obtained from the  $i$ th group ( $i = 1$  if exposed group and  $i = 2$  if non-exposed group), for the  $j$ th subject ( $j = 1, 2, \dots, n$ ), at the  $l$ th non-progression visit ( $l = 1, 2, \dots, m_{ij}$ ), where  $m_{ij}$  is the number of non-progression visits for subject  $j$  in group  $i$ . Let  $Y_{ij}$  be the continuous random variable associated with values for the same biomarker from the  $j$ th subject in the  $i$ th group at the progression visit. Also let  $\delta_{ij} = 1$  if subject  $j$  in the  $i$ th group became a progressor and  $= 0$  otherwise. Define  $D_i = \sum_{j=1}^n \delta_{ij}$  as the total number of progressors in the  $i$ th group. Let  $F_i$  and  $G_i$  be the distribution functions of  $X_{ijl}$  and  $Y_{ij}$ , respectively. Further, let  $\theta_i$  represent area under the respective ROC curves of biomarker in the  $i$ th group. In this article, we discuss test for the null hypothesis of  $H_0 : \theta_1 - \theta_2 = 0$  versus the alternative hypothesis  $H_1 : \theta_1 - \theta_2 \neq 0$ .

## 2.2.2 Non-parametric estimate of area under the ROC

Assume that the biomarker will be considered to be positive if the value of  $X_{ijl}$  or  $Y_{ij}$  exceeds a predetermined threshold  $c$ . Non-parametric estimate of area under the ROC curve for the biomarker can be obtained using standard method based on the empirical cumulative distribution functions (CDFs) corresponding to  $F_i$  and  $G_i$ . Area under the ROC curve of a repeated biomarker for the  $i$ th group is  $\theta_i = \int_{-\infty}^{\infty} F_i(c)dG_i(c)$ . The corresponding area estimate,  $\hat{\theta}_i$ , is given by

$$\hat{\theta}_i = \int_{-\infty}^{\infty} \hat{F}_i(c)d\hat{G}_i(c), \quad (2.1)$$

where

$$\begin{aligned} \hat{F}_i(c) &= \sum_{j=1}^n w_{ij} \left\{ \frac{1}{m_{ij}} \sum_{l=1}^{m_{ij}} \mathbf{I}(x_{ijl} \leq c) \right\}, \\ \hat{G}_i(c) &= \frac{1}{D_i} \sum_{j=1}^n \delta_{ij} \mathbf{I}(y_{ij} \leq c), \end{aligned}$$

and  $(w_{i1}, \dots, w_{in})$  is a set of weights assigned to subjects in the  $i$ th group, satisfying  $w_{ij} > 0, j = 1, \dots, n$  and  $\sum_{j=1}^n w_{ij} = 1$ .

Let  $\Delta = \theta_1 - \theta_2$  be the true AUC difference between two correlated groups. The non-parametric estimate of  $\Delta$  is  $\hat{\Delta} = \hat{\theta}_1 - \hat{\theta}_2$ , given by

$$\hat{\Delta} = \int_{-\infty}^{\infty} \hat{F}_1(c)d\hat{G}_1(c) - \int_{-\infty}^{\infty} \hat{F}_2(c)d\hat{G}_2(c). \quad (2.2)$$

In next section, we will consider the optimal weights to minimize the variance of  $\hat{\Delta}$ .

## 2.3 The optimal weighting scheme

The estimate of  $\hat{\Delta}$  involves a set of weights in each group. We can use two simple weighting schemes that were provided in Emir et al. (2000): (1) assigning equal weights to all biomarker observations in each group, i.e.,  $w_{ij} = m_{ij} / \sum_{j'=1}^n m_{ij'}$ , when the within-subject correlation is low; or (2) assigning equal weights to all subjects in each group, i.e.,  $w_{ij} = 1/n$  when the within-subject correlation is high. Alternatively, we may use the Lagrange Multiplier Method to derive the optimal weights which can minimize the variance of  $\hat{\Delta}$ .

To derive the optimal weights, we utilize the following fact

$$\hat{\Delta} - \Delta = \sum_{j=1}^n (\epsilon_{1j} + \xi_{1j} - \epsilon_{2j} - \xi_{2j}) + o(n^{-1/2}), \quad (2.3)$$

where

$$\begin{aligned} \epsilon_{ij} &= \frac{\delta_{ij}}{D_i} \int_{-\infty}^{\infty} F_i(c) d \{I(y_{ij} \leq c) - G_i(c)\}, \\ \xi_{ij} &= \frac{w_{ij}}{m_{ij}} \sum_{l=1}^{m_{ij}} \int_{-\infty}^{\infty} \{I(x_{ijl} \leq c) - F_i(c)\} dG_i(c). \end{aligned}$$

The proof of (2.3) above can be found in Appendix of publication by Emir et. al (2000).

Defining the transformations

$$U_{ijl} = G_i(x_{ijl}), \quad V_{ij} = F_i(y_{ij}), \quad (2.4)$$

we can express the variances of  $\hat{\theta}_1$  and  $\hat{\theta}_2$ , and the covariance between  $\hat{\theta}_1$  and  $\hat{\theta}_2$  in terms of  $U_{ijl}$

and  $V_{ij}$ . They are

$$\begin{aligned}
\text{Var} \sum_{j=1}^n (\epsilon_{1j} + \xi_{1j}) &= \sigma_{u1}^2 \sum_{j=1}^n \frac{\sum_{l \neq l'} \text{corr}(U_{1jl}, U_{1jl'})}{m_{1j}^2} w_{1j}^2 \\
&\quad - \frac{2\sigma_{u1}\sigma_{v1}}{D_1} \sum_{j=1}^n \frac{w_{1j}\delta_{1j}}{m_{1j}} \sum_{l=1}^{m_{1j}} \text{corr}(U_{1jl}, V_{1j}) + \frac{\sigma_{v1}^2}{D_1}, \\
\text{Var} \sum_{j=1}^n (\epsilon_{2j} + \xi_{2j}) &= \sigma_{u2}^2 \sum_{j=1}^n \frac{\sum_{l \neq l'} \text{corr}(U_{2jl}, U_{2jl'})}{m_{2j}^2} w_{2j}^2 \\
&\quad - \frac{2\sigma_{u2}\sigma_{v2}}{D_2} \sum_{j=1}^n \frac{w_{2j}\delta_{2j}}{m_{2j}} \sum_{l=1}^{m_{2j}} \text{corr}(U_{2jl}, V_{2j}) + \frac{\sigma_{v2}^2}{D_2}, \\
\text{Cov}(\sum_{j=1}^n (\epsilon_{1j} + \xi_{1j}), \sum_{j=1}^n (\epsilon_{2j} + \xi_{2j})) &= \sigma_{u1}\sigma_{u2} \sum_{j=1}^n \sum_{j=1}^n \frac{w_{1j}w_{2j}\text{corr}(U_{1jl}, U_{2jl})}{m_{1j}m_{2j}} \\
&\quad + \frac{\sigma_{u2}\sigma_{v1}}{D_1} \sum_{j=1}^n \frac{w_{2j}\delta_{1j}}{m_{2j}} \sum_{l=1}^{m_{2j}} \text{corr}(U_{2jl}, V_{1j}) \\
&\quad + \frac{\sigma_{u1}\sigma_{v2}}{D_2} \sum_{j=1}^n \frac{w_{1j}\delta_{2j}}{m_{1j}} \sum_{l=1}^{m_{1j}} \text{corr}(U_{1jl}, V_{2j}) \\
&\quad + \frac{\sigma_{v1}\sigma_{v2} \sum_{j=1}^n \delta_{1j} \sum_{j=1}^n \delta_{2j} \text{corr}(V_{1j}, V_{2j})}{D_1 D_2},
\end{aligned}$$

where  $\sigma_{u1}^2 = \text{Var}(U_{1jl})$ ,  $\sigma_{v1}^2 = \text{Var}(V_{1j})$ ,  $\sigma_{u2}^2 = \text{Var}(U_{2jl})$ , and  $\sigma_{v2}^2 = \text{Var}(V_{2j})$ .

Thus, the variance of  $\hat{\Delta}$  given  $m_{ij}$  can be expressed as

$$\begin{aligned}
\text{Var}(\hat{\Delta}|m_{ij}) &= \text{Var} \sum_{j=1}^n (\epsilon_{1j} - \epsilon_{2j} + \xi_{1j} - \xi_{2j}) \\
&= \text{Var} \sum_{j=1}^n (\epsilon_{1j} + \xi_{1j}) + \text{Var} \sum_{j=1}^n (\epsilon_{2j} + \xi_{2j}) - 2\text{Cov}(\sum_{j=1}^n (\epsilon_{1j} + \xi_{1j}), \sum_{j=1}^n (\epsilon_{2j} + \xi_{2j})) \\
&= \sum_{j=1}^n a_{1j}w_{1j}^2 - 2 \sum_{j=1}^n b_{1j}w_{1j} + \sum_{j=1}^n a_{2j}w_{2j}^2 - 2 \sum_{j=1}^n b_{2j}w_{2j} - 2 \sum_{j=1}^n c_j w_{1j}w_{2j} + d, \quad (2.5)
\end{aligned}$$

where

$$\begin{aligned}
a_{1j} &= \frac{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1 u_1}}{m_{1j}^2}, \\
a_{2j} &= \frac{\sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2 u_2}}{m_{2j}^2}, \\
b_{1j} &= \frac{\delta_{1j} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_1}}{D_1 m_{1j}} - \frac{\delta_{2j} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_2}}{D_2 m_{1j}}, \\
b_{2j} &= \frac{\delta_{2j} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_2}}{D_2 m_{2j}} - \frac{\delta_{1j} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_1}}{D_1 m_{2j}}, \\
c_j &= \frac{\sum_k \sum_l \sigma_{kl}^{u_1 u_2}}{m_{1j} m_{2j}}, \\
d &= \frac{\sigma^{v_1 v_1}}{D_1} + \frac{\sigma^{v_2 v_2}}{D_2} - 2 \frac{\sigma^{v_1 v_2}}{D_1 D_2},
\end{aligned}$$

with  $\sigma_{ll'}^{u_i u_i} = Cov(U_{ijl}, U_{ijl'})$ ,  $\sigma_l^{u_i v_i} = Cov(U_{ijl}, V_{ij})$  and  $\sigma^{v_i v_i} = Cov(V_{ij}, V_{ij})$ .

The method of Lagrange Multiplier provides a strategy for finding the maxima or minima of a function subject to constraints. By applying the Lagrange Multiplier Method in this case, the optimal weights can be obtained to minimize the variance function (2.5) with constraints  $w_{1j} > 0$ ,  $w_{2j} > 0$ , ( $j = 1, \dots, n$ ),  $\sum_{j=1}^n w_{1j} = 1$  and  $\sum_{j=1}^n w_{2j} = 1$ . The Lagrange function is defined as

$$L(w_{11}, \dots, w_{1n}; w_{21}, \dots, w_{2n}; \lambda; \mu) = Var(\hat{\Delta} | m_{ij}) + \lambda(1 - \sum_{j=1}^n w_{1j}) + \mu(1 - \sum_{j=1}^n w_{2j}),$$

where  $\lambda$  and  $\mu$  are the two Lagrange multipliers.

The partial derivatives of the Lagrange function with respect to  $w_{1j}$  and  $w_{2j}$  are

$$\begin{cases} \frac{dL(w_{11}, \dots, w_{1n}; w_{21}, \dots, w_{2n}; \lambda; \mu)}{dw_{1j}} = 2a_{1j}w_{1j} - 2b_{1j} - 2c_j w_{2j} - \lambda = 0; \\ \frac{dL(w_{11}, \dots, w_{1n}; w_{21}, \dots, w_{2n}; \lambda; \mu)}{dw_{2j}} = 2a_{2j}w_{2j} - 2b_{2j} - 2c_j w_{1j} - \mu = 0. \end{cases} \quad (2.6)$$

Let

$$H_j = \begin{pmatrix} a_{1j} & -c_j \\ -c_j & a_{2j} \end{pmatrix}, \quad \tilde{w}_j = \begin{pmatrix} w_{1j} \\ w_{2j} \end{pmatrix}, \quad \tilde{b}_j = \begin{pmatrix} b_{1j} \\ b_{2j} \end{pmatrix},$$

and

$$\tilde{\chi} = \begin{pmatrix} \lambda \\ \mu \end{pmatrix},$$

the system equations (2.6) can be expressed in a matrix format. That is

$$2H_j\tilde{w}_j - \tilde{\chi} - 2\tilde{b}_j = 0.$$

We obtain:

$$\tilde{w}_j = \frac{1}{2}H_j^{-1}(\tilde{\chi} + 2\tilde{b}_j). \quad (2.7)$$

Let

$$\tilde{I} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

the constraint equations can be written in the matrix format. That is

$$\sum_{j=1}^n \tilde{w}_j = \sum_{j=1}^n \frac{1}{2}H_j^{-1}(\tilde{\chi} + 2\tilde{b}_j) = \tilde{I},$$

we can solve for  $\tilde{\chi}$ , which is given by

$$\tilde{\chi} = 2\left(\sum_{j=1}^n H_j^{-1}\right)^{-1} \left[ \tilde{I} - \left(\sum_{j=1}^n H_j^{-1}\tilde{b}_j\right) \right]. \quad (2.8)$$

Next, plugging (2.8) into (2.7), we obtain the optimal weights as

$$\tilde{w}_j = H_j^{-1} \left[ \left(\sum_{j=1}^n H_j^{-1}\right)^{-1} \tilde{I} - \left(\sum_{j=1}^n H_j^{-1}\right)^{-1} \left(\sum_{j=1}^n H_j^{-1}\tilde{b}_j\right) + \tilde{b}_j \right]. \quad (2.9)$$

The optimal weights involve the unknown parameters, which can be consistently estimated from estimated transformed data  $(\hat{U}_{ijl}, \hat{V}_{ij})$

$$\hat{U}_{ijl} = \hat{G}_i(x_{ijl}), \hat{V}_{ij} = \hat{F}_i(y_{ij}),$$

The estimates of unknown parameters calculated from the estimated transformed data are given as

$$\begin{aligned} \hat{\sigma}_{ll'}^{u_1 u_1} &= \frac{\sum_{j=1}^n \delta_{1jl} \delta_{1j'l'} (\hat{U}_{1jl} - \bar{U}_{1l}) (\hat{U}_{1j'l'} - \bar{U}_{1l'})}{\sum_{j=1}^n \delta_{1jl} \delta_{1j'l'}}, \\ \hat{\sigma}_{ll'}^{u_2 u_2} &= \frac{\sum_{j=1}^n \delta_{2jl} \delta_{2j'l'} (\hat{U}_{2jl} - \bar{U}_{2l}) (\hat{U}_{2j'l'} - \bar{U}_{2l'})}{\sum_{j=1}^n \delta_{2jl} \delta_{2j'l'}}, \\ \hat{\sigma}_{kl}^{u_1 u_2} &= \frac{\sum_{j=1}^n \delta_{1jk} \delta_{2jl} (\hat{U}_{1jk} - \bar{U}_{1k}) (\hat{U}_{2jl} - \bar{U}_{2l})}{\sum_{j=1}^n \delta_{1jk} \delta_{2jl}}, \\ \hat{\sigma}_l^{u_1 v_1} &= \frac{\sum_{j=1}^n \delta_{1jl} \delta_{1j} (\hat{U}_{1jl} - \bar{U}_{1l}) (\hat{V}_{1j} - \bar{V}_1)}{\sum_{j=1}^n \delta_{1jl} \delta_{1j}}, \\ \hat{\sigma}_l^{u_1 v_2} &= \frac{\sum_{j=1}^n \delta_{1jl} \delta_{2j} (\hat{U}_{1jl} - \bar{U}_{1l}) (\hat{V}_{2j} - \bar{V}_2)}{\sum_{j=1}^n \delta_{1jl} \delta_{2j}}, \\ \hat{\sigma}_l^{u_2 v_2} &= \frac{\sum_{j=1}^n \delta_{2jl} \delta_{2j} (\hat{U}_{2jl} - \bar{U}_{2l}) (\hat{V}_{2j} - \bar{V}_2)}{\sum_{j=1}^n \delta_{2jl} \delta_{2j}}, \\ \hat{\sigma}_l^{u_2 v_1} &= \frac{\sum_{j=1}^n \delta_{2jl} \delta_{1j} (\hat{U}_{2jl} - \bar{U}_{2l}) (\hat{V}_{1j} - \bar{V}_1)}{\sum_{j=1}^n \delta_{2jl} \delta_{1j}}, \\ \hat{\sigma}^{v_1 v_1} &= \frac{\sum_{j=1}^n \delta_{1j} (\hat{V}_{1j} - \bar{V}_1) (\hat{V}_{1j} - \bar{V}_1)}{\sum_{j=1}^n \delta_{1j}}, \\ \hat{\sigma}^{v_2 v_2} &= \frac{\sum_{j=1}^n \delta_{2j} (\hat{V}_{2j} - \bar{V}_2) (\hat{V}_{2j} - \bar{V}_2)}{\sum_{j=1}^n \delta_{2j}}, \\ \hat{\sigma}^{v_1 v_2} &= \frac{\sum_{j=1}^n \delta_{1j} \delta_{2j} (\hat{V}_{1j} - \bar{V}_1) (\hat{V}_{2j} - \bar{V}_2)}{\sum_{j=1}^n \delta_{1j} \delta_{2j}}, \end{aligned}$$

where  $\delta_{ijl} = 1$  if the  $j$ th subject in the  $i$ th group has the  $l$ th non-progression visit and =0 otherwise;

$\delta_{ij} = 1$  if the  $j$ th subject in the  $i$ th group has the disease progression visit and =0 otherwise;

$\bar{U}_{ijl} = \sum_{j=1}^n \delta_{ijl} \hat{U}_{ijl} / \sum_{j=1}^n \delta_{ijl}$  and  $\bar{V}_{ij} = \sum_{j=1}^n \delta_{ij} \hat{V}_{ij} / \sum_{j=1}^n \delta_{ij}$  and so on. The estimated optimal

weights are then obtained by plugging those estimates into (2.9). Since  $\hat{G}_i(x_{ijl})$  includes  $w_{ij}$ , we

may iteratively estimate  $w_{ij}$  and  $G_i(\cdot)$  until it converges, or, to allow for a closed-form solution, we may replace  $\hat{w}_{ij}$  in  $\hat{G}_i(x_{ijl})$  with any simple weight such as  $w_{ij} = m_{ij} / \sum_{j'=1}^n m_{ij'}$  or  $w_{ij} = 1/n$ .

Now consider three special cases: (1) there is no within-subject correlation between any two time points,  $\rho_{u1u1} = \rho_{u1v1} = 0$  and  $\rho_{u2u2} = \rho_{u2v2} = 0$ . In this case, the optimal weight becomes  $w_{ij} = m_{ij} / \sum_{j'=1}^n m_{ij'}$ , which means that the simple weighting scheme 1 suggested by Emir et al. (2000) is optimal; (2) there is perfect within-subject correlation between any two time points,  $\rho_{u1u1} = \rho_{u1v1} = 1$  and  $\rho_{u2u2} = \rho_{u2v2} = 1$ . In this case, the optimal weight becomes  $w_{ij} = 1/D_i$  for subject  $j$  in group  $i$  who became a progressor at some time point, and  $w_{ij} = 0$  for subject  $j$  in group  $i$  who remained a non-progressor until the end of study, which means that the simple weighting scheme 2 suggested by Emir et al. (2000) is not optimal unless  $D_i = n$ , that is, all subjects in both groups became progressors at some time point during the study period; (3) there is no between-group correlation at any time point,  $\rho_{u1u2} = \rho_{u1v2} = \rho_{u2v1} = 0$ . In this case, the weights by Wu et al. (2011) are not optimal unless the number of visits and disease incidence rates are the same in two groups.

## 2.4 Asymptotic variance and the relative efficiency

Let  $\tau_1 = \lim_{n \rightarrow \infty} D_1/n$ , and  $\tau_2 = \lim_{n \rightarrow \infty} D_2/n$ , so  $\tau_1$  and  $\tau_2$  are the disease incidence rates in group 1 and 2 respectively. Let  $\hat{\Delta}_1$  be the estimator of  $\Delta$  using simple weighting scheme  $w_{1j} = m_{1j} / \sum_{j'=1}^n m_{1j'}$  and  $w_{2j} = m_{2j} / \sum_{j'=1}^n m_{2j'}$ ; and  $\hat{\Delta}_2$  be the estimator of  $\Delta$  using simple weighting scheme  $w_{1j} = 1/n$  and  $w_{2j} = 1/n$ . To compare our optimal estimator with  $\hat{\Delta}_1$  and  $\hat{\Delta}_2$ , we consider the asymptotic variances of  $\hat{\Delta}_1$  and  $\hat{\Delta}_2$ .



Emir et al. (2000) proved that  $\hat{\Delta}/\sqrt{Var(\hat{\Delta})}$  is approximately normal with  $N(0, 1)$ . We can

show that  $\sqrt{n}(\hat{\Delta}_1 - \Delta)$  is approximately normal with  $N(0, \sigma_1^2)$

$$\begin{aligned} \sigma_1^2 = & \frac{E \sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1 u_1}}{(Em_1)^2} - 2 \frac{E \sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1 u_2}}{(Em_1)(Em_2)} + \frac{E \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2 u_2}}{(Em_2)^2} \\ & - 2 \frac{E \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_1}}{Em_1} + 2 \frac{E \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_2}}{Em_1} - 2 \frac{E \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_2}}{Em_2} + 2 \frac{E \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_1}}{Em_2} \\ & + \tau_1^{-1} \sigma^{v_1 v_1} + \tau_2^{-1} \sigma^{v_2 v_2}, \end{aligned} \quad (2.10)$$

where  $m_i = \frac{1}{n} \sum_{j=1}^n m_{ij}$  and  $E$  is the expected value taken with respect to the random variable of  $m_{ij}$ .

Similarly, we can also show that  $\sqrt{n}(\hat{\Delta}_2 - \Delta)$  is approximately normal with  $N(0, \sigma_2^2)$

$$\begin{aligned} \sigma_2^2 = & E \frac{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1 u_1}}{m_1^2} - 2E \frac{\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1 u_2}}{m_1 m_2} + E \frac{\sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2 u_2}}{m_2^2} \\ & - 2E \frac{\sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_1}}{m_1} + 2E \frac{\sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_2}}{m_1} - 2E \frac{\sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_2}}{m_2} + 2E \frac{\sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_1}}{m_2} \\ & + \tau_1^{-1} \sigma^{v_1 v_1} + \tau_2^{-1} \sigma^{v_2 v_2}. \end{aligned} \quad (2.11)$$

Let  $\hat{\Delta}_{op}$  be the estimate of  $\Delta$  obtained by using the estimated optimal weights. We can show that  $\sqrt{n}(\hat{\Delta}_{op} - \Delta)$  is approximately normal with mean 0 and variance

$$\sigma_{op}^2 = \left( \tilde{I}^t - \tilde{\phi}^t \right) \Sigma^{-1} \left( \tilde{I} - \tilde{\phi} \right) - v + \tau_1^{-1} \sigma^{v_1 v_1} + \tau_2^{-1} \sigma^{v_2 v_2}, \quad (2.12)$$

where

$$\tilde{\phi} = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix},$$

$$\begin{aligned}
e_1 &= \lim_{n \rightarrow \infty} \sum_{j=1}^n \frac{a_{2j}b_{1j} + b_{2j}c_j}{a_{1j}a_{2j} - c_jc_j} \\
&= \tau_1^{-1} E \frac{m_1 \left( \sum_{l=1}^{m_{1j}} \sigma_l^{u_1v_1} - \sum_{l=1}^{m_{1j}} \sigma_l^{u_1v_2} \right) \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2}}{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} - \left( \sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2} \right)^2} \\
&\quad + \tau_2^{-1} E \frac{m_1 \left( \sum_{l=1}^{m_{2j}} \sigma_l^{u_2v_2} - \sum_{l=1}^{m_{2j}} \sigma_l^{u_2v_1} \right) \sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2}}{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} - \left( \sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2} \right)^2},
\end{aligned}$$

$$\begin{aligned}
e_2 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \frac{a_{1j}b_{2j} + b_{1j}c_j}{a_{1j}a_{2j} - c_jc_j} \\
&= \tau_2^{-1} E \frac{m_2 \left( \sum_{l=1}^{m_{2j}} \sigma_l^{u_2v_2} - \sum_{l=1}^{m_{2j}} \sigma_l^{u_2v_1} \right) \sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1}}{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} - \left( \sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2} \right)^2} \\
&\quad + \tau_1^{-1} E \frac{m_2 \left( \sum_{l=1}^{m_{1j}} \sigma_l^{u_1v_1} - \sum_{l=1}^{m_{1j}} \sigma_l^{u_1v_2} \right) \sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2}}{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} - \left( \sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2} \right)^2},
\end{aligned}$$

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix},$$

$$\begin{aligned}
\sigma_{11} &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \frac{a_{2j}}{a_{1j}a_{2j} - c_jc_j} \\
&= E \frac{m_1^2 \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2}}{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} - \left( \sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2} \right)^2},
\end{aligned}$$

$$\begin{aligned}
\sigma_{12} &= \sigma_{21} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \frac{c_j}{a_{1j}a_{2j} - c_jc_j} \\
&= E \frac{m_1 m_2 \sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2}}{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} - \left( \sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2} \right)^2},
\end{aligned}$$

$$\begin{aligned}
\sigma_{22} &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \frac{a_{1j}}{a_{1j}a_{2j} - c_jc_j}, \\
&= E \frac{m_2^2 \sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1}}{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} - \left( \sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2} \right)^2},
\end{aligned}$$

and

$$\begin{aligned}
v &= \lim_{n \rightarrow \infty} n \sum_{j=1}^n \frac{a_{2j}b_{1j}^2 + 2b_{1j}b_{2j}c_j + a_{1j}b_{2j}^2}{a_{1j}a_{2j} - c_jc_j} \\
&= \frac{1}{\tau_1^2} E \frac{\sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} \left(\sum_{l=1}^{m_{1j}} \sigma_l^{u_1v_1}\right)^2 + \sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \left(\sum_{l=1}^{m_{2j}} \sigma_l^{u_2v_1}\right)^2}{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} - \left(\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2}\right)^2} \\
&\quad - 2 \frac{1}{\tau_1^2} E \frac{\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1v_1} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2v_1}}{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} - \left(\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2}\right)^2} \\
&\quad + \frac{1}{\tau_2^2} E \frac{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \left(\sum_{l=1}^{m_{2j}} \sigma_l^{u_2v_2}\right)^2 + \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} \left(\sum_{l=1}^{m_{1j}} \sigma_l^{u_1v_2}\right)^2}{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} - \left(\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2}\right)^2} \\
&\quad - 2 \frac{1}{\tau_2^2} E \frac{\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1v_2} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2v_2}}{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} - \left(\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2}\right)^2} \\
&\quad + 2 \frac{1}{\tau_1\tau_2} E \frac{\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2} \left(\sum_{l=1}^{m_{1j}} \sigma_l^{u_1v_1} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2v_2} + \sum_{l=1}^{m_{1j}} \sigma_l^{u_1v_2} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2v_1}\right)}{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} - \left(\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2}\right)^2} \\
&\quad - 2 \frac{1}{\tau_1\tau_2} E \frac{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2v_2} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2v_1} + \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1v_1} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1v_2}}{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1u_1} \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2u_2} - \left(\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1u_2}\right)^2}.
\end{aligned}$$

The proofs of (2.10), (2.11) and (2.12) are included in Appendix.

Let  $ARE_1 = \sigma_{op}^2/\sigma_1^2$  be the asymptotic relative efficiency for comparing  $\hat{\Delta}_1$  with  $\hat{\Delta}_{op}$ , and  $ARE_2 = \sigma_{op}^2/\sigma_2^2$  be the asymptotic relative efficiency for comparing  $\hat{\Delta}_2$  with  $\hat{\Delta}_{op}$ . To demonstrate how the within-subject correlation, the between-group correlation, and the incidence of disease affect  $ARE_1$  and  $ARE_2$ , we consider a special case where variances of  $U_{ijl}$  and  $V_{ij}$  are homogeneous across visits and groups, within-subject non-progression and non-progression correlation coefficients are homogeneous across groups, within-subject progression and non-progression correlation coefficients are homogeneous across groups, and between-group correlation coefficients

are homogeneous across visits. Specifically,

$$\begin{aligned} \text{Var}(U_{ijl}) &= \text{Var}(V_{ij}) = \sigma^2, \\ \text{Corr}(U_{ijl}, U_{ijl'}) &= \text{Corr}(U_{ijl}, V_{ij}) = \rho_w, \\ \text{Corr}(U_{ijl}, U_{i'jl}) &= \text{Corr}(U_{ijl}, V_{i'j}) = \rho_b, i \neq i'. \end{aligned}$$

We graph  $ARE_1$  and  $ARE_2$  against  $\rho_w$  for three different values of  $\rho_b$ ,  $\rho_b = 0.3, 0.5, 0.7$  and three different sets of  $(\tau_1, \tau_2)$ ,  $(\tau_1, \tau_2) = (0.5, 0.5), (0.3, 0.3), (0.1, 0.1)$ . The uniform distribution of  $m_1$ ,  $P(m_1 = k) = 1/6, k = 1, \dots, 6$ , and the distribution of  $m_2$ ,  $P(m_2 = k) = 1/3, 1/3, 1/3, 0, 0, 0$ , were used to calculate the expectations involved in  $ARE_1$  and  $ARE_2$ . Figure 2.1 shows that the efficiency gain of our optimal estimator increases dramatically as  $\rho_w$  increases and  $(\tau_1, \tau_2)$  decreases, and increases as  $\rho_b$  increases. The loss of efficiency by using the two simple weighting schemes instead of our optimal weights can be severe when the within-subject correlation is large and the incidence of disease is small.

## 2.5 Simulation results

Simulations were conducted to examine the finite sample properties of the proposed weighting scheme relative to the simple weight scheme 1 and the simple weight scheme 2 suggested by Emir et al. (2000) with small sample size. In the absence of disease progression, the subjects will have the biomarker observed every month for a total of at most six monthly visits per subject. For each pair of subjects, we generate an independent multivariate normal random vector  $X_i = (X_{11}, X_{12}, \dots, X_{16}, X_{21}, X_{22}, \dots, X_{26})$ , of size  $12 \times n$  with mean vector 0 and variance-covariance matrix  $\Sigma = (\sigma_{ll'})_{12 \times 12}$  with

$$\sigma_{ll'} = \begin{cases} 1 & \text{if } l = l' \\ \gamma & \text{if } l, l' = 1, \dots, 6 \text{ or } l, l' = 7, \dots, 12 \\ \lambda & \text{if } l = 1, \dots, 6; l' = 7, \dots, 12 \text{ or } l = 7, \dots, 12; l' = 1, \dots, 6. \end{cases}$$

Note that  $(X_{11}, X_{12}, \dots, X_{16})$  are the observed biomarker values for the subjects in the exposed group and  $(X_{21}, X_{22}, \dots, X_{26})$  for the matched subjects in the non-exposed group, and that  $\gamma$  and  $\lambda$  are the parameters which introduce correlations within-subject and between-group respectively. The values of  $\lambda$  and  $\gamma$  were chosen so that the variance-covariance matrix  $\Sigma$  is positive definite.

We generate disease onset times for the subjects using an exponential distribution such that the expected disease rates at six months in two groups are  $\psi_1$  and  $\psi_2$ . If a simulated disease onset time is greater than six months, we define the subject to be a non-progressor at all six visits and use all six values of  $X_i$  for the biomarker. If the disease onset time occurs before six months, we assume that the disease progression is detected clinically at the next visit. For example, if a disease onset in the exposed group occurs between the third and fourth visit, the simulated biomarker values for the first three visits are  $(x_{11}, x_{12}, x_{13})$ . We assume the expected value of the biomarker is increased by 1 at the time of disease onset, so we define the biomarker value at this fourth visit to be  $Y_1 = x_{14} + 1$ . With this set-up for both groups, the true AUC is equal to 0.75 in both groups and the true AUC difference of the biomarker between groups is  $\Delta = 0$ . In our simulation, we set the disease rates to be either equal or unequal between groups using  $(\psi_1, \psi_2) = (0.5, 0.4), (0.4, 0.4), (0.2, 0.4)$ . The disease rates were chosen to show how the change of disease rate in any group would impact the efficiency. The correlations were either low, moderate or high, using  $\gamma = 0.3, 0.5, 0.7, 0.9$  and  $\lambda = 0.0, 0.3, 0.5, 0.7$ . The sample size of 50 and the sample size of 100 per group were considered

for each combination of  $(\lambda, \gamma, \psi_1, \psi_2)$ .

A total of 2000 simulations were replicated. The efficiency of the estimators was estimated using the mean square error (MSE), which was computed using the variance of the parameter estimates and the estimated bias. Table 2.2 (n=50) and Table 2.3 (n=100) summarize the results using the relative efficiency. The results are consistent between both sample sizes. These simulations show that the estimator using optimal weights outperforms both  $\hat{\Delta}_1$  and  $\hat{\Delta}_2$  especially when the within-group correlation is large, which is consistent with the conclusion from the asymptotic comparisons. As well, the performance is increasing as the between-group correlation increases and the disease incidence rate decreases. Additionally, we also compared the optimal weights to the weights using Wu et al.'s method (2011). The results ( $RE_3$ ) indicate up to 20% of efficiency gain when the between-group correlation is being taken account in the optimal weights.

## 2.6 Application on the thyroidism study data

The study is ongoing. The use of this data is for illustration of statistical methods only. It does NOT represent the final outcome reporting of the ongoing study. In this illustrative example, only 213 pairs of subjects who completed the study as 05/01/2012 and had at least one measurement were included. Hypothyroidism is defined using arbitrary thresholds, with TSH greater than 2.2 mIU/L and Free T4 less than 1.76 pmol/L. The incidence rate of hypothyroidism was only 12% in the breast cancer group and approximate 9% in health control group during the study period.

The raw values of TPOAb titers were treated as the original data of  $(x_{1jl}, y_{1j})$  for the breast cancer group and  $(x_{2jl}, y_{2j})$  for the healthy control group. The optimal weights were obtained by

Table 2.1: Results for Comparing Accuracy of TPOAb in Predicting Thyroid Function.

	Optimal Weight	Simple Weight 1	Simple Weight 2
Estimate	-0.014	-0.031	-0.030
Variance	0.0016	0.0025	0.0023
Relative efficiency	1	0.64	0.70

following these procedures: First, obtain the estimates of  $F_i$  and  $G_i$  using simple weight 1; second, generate the transformed data  $(\hat{u}_{ijl}, \hat{v}_{ij})$ ; next, obtain the estimates of  $\sigma^{u_i u_i}$ ,  $\sigma^{u_i v_i}$  and  $\sigma^{v_i v_i}$  from the transformed data; finally, plug these estimated variance-covariance parameters into (2.9) to get the estimated optimal weights. The estimated optimal weights were then plugged into equation (2.1) to calculate the point estimate of the two AUCs difference and plugged into (2.5) to calculate the variance of  $\hat{\Delta}_{op}$ . Table 2.1 summarizes both the point estimates and their variance estimates using all three weight schemes. The optimal weight scheme outperformed the other two simple weight schemes with a substantial gain of efficiency because the disease incidence rates were low in both groups and the within-subject correlation was moderate to high in this study.

## 2.7 Concluding remarks

In previous studies, the Cox proportional hazard regression model with a time-dependent covariant was traditionally utilized for evaluating the biomarker with repeated measurements (Cox, 1972; Kalbfleisch, 1980). However, when the samples are correlated, the Cox proportional hazard model cannot be used. There has been an increased use of the ROC curve to examine the prognostic performance of some continuous biomarkers with repeated measurements. A non-parametric ROC approach, due to Emir et al. (2000) was introduced to estimate area under the ROC curve of a repeated biomarker. We modified Emir et al.'s approach (2000) to compare the prediction accuracy

of a biomarker between two correlated groups. When the study subjects are limited and cost is high, the efficiency of the statistical methods is a crucial consideration when designing clinical studies. We extended Wu et al.'s idea (2011) to provide a solution to weight optimization which minimize the variance of AUCs comparison estimate. Our simulation results show substantial gains of efficiency by using an optimal weighting scheme when the correlation within subject is high, the correlation between groups is high, and/or the disease incidence rate is low. The proposed optimal weighting scheme is generally preferred to the two weighting schemes suggested by Emir et al. (2000). Thus, it is recommended.

This research has several limitations. First, the optimal weights could be out of boundary when the sample size is too small. Secondly, the homogenous variance-covariance structure was used in the simulation studies. In the future study, we may consider other more complicated structure. The methodology presented in this paper focused on comparing the predictive accuracies of a repeated biomarker between two correlated groups in the one to one matched-subject design study. This methodology may be adapted to compare AUCs in matched-subject design studies when there are more than one matches per subject. As well, the methodology can be applied to other correlated-group design study, such as cross-over design study. Futhermore, the extension of the proposed methodology can be used to accommodate the clustered data, such as there are recurrent outcome events. Future research of these, based on the weighting optimization methods described here, may produce better solutions, particularly where the efficiency is concerned.



## 2.8 Acknowledgments

A pilot study to estimate the prevalence of thyroid dysfunction in women diagnosed with breast cancer and the magnitude of change in thyroid function post-chemotherapy was sponsored by the National Cancer Institute (grant number: NCI 5U10CA081920). The study is lead by Dr. Jeffrey Krischer (Principal Investigator) in affiliation with the Suncoast CCOP Research Base and Dr. Nagi Kumar (Study Chair) in affiliation with Moffit Cancer Center, Tampa, FL.

## 2.9 Appendix

Using the fact (2.3) to prove (2.10), (2.11) and (2.12), it suffices to show that

$$\sigma^2 = \lim_{n \rightarrow \infty} n \text{Var}(\hat{\Delta}).$$

**Proof of (2.10)**

$$\begin{aligned}
nVar(\hat{\Delta}_1) &= n \left[ \sum_{j=1}^n a_{1j} w_{1j}^2 - 2 \sum_{j=1}^n c_j w_{1j} w_{2j} + \sum_{j=1}^n a_{2j} w_{2j}^2 \right] \\
&\quad - 2n \sum_{j=1}^n b_{1j} w_{1j} - 2n \sum_{j=1}^n b_{2j} w_{2j} + nd \\
&= n \sum_{j=1}^n \frac{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1 u_1}}{m_{1j}^2} \left( \frac{m_{1j}}{\sum_{i=1}^n m_{1j}} \right)^2 + n \sum_{j=1}^n \frac{\sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2 u_2}}{m_{2j}^2} \left( \frac{m_{2j}}{\sum_{i=1}^n m_{2j}} \right)^2 \\
&\quad - 2n \sum_{j=1}^n \frac{\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1 u_2}}{m_{1j} m_{2j}} \left( \frac{m_{1j}}{\sum_{i=1}^n m_{1j}} \right) \left( \frac{m_{2j}}{\sum_{i=1}^n m_{2j}} \right) \\
&\quad - 2n \sum_{j=1}^n \frac{\delta_{1j} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_1}}{D_1 m_{1j}} \left( \frac{m_{1j}}{\sum_{i=1}^n m_{1j}} \right) + 2n \sum_{j=1}^n \frac{\delta_{2j} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_2}}{D_2 m_{1j}} \left( \frac{m_{1j}}{\sum_{i=1}^n m_{1j}} \right) \\
&\quad - 2n \sum_{j=1}^n \frac{\delta_{2j} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_2}}{D_2 m_{2j}} \left( \frac{m_{2j}}{\sum_{i=1}^n m_{2j}} \right) + 2n \sum_{j=1}^n \frac{\delta_{1j} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_1}}{D_1 m_{2j}} \left( \frac{m_{2j}}{\sum_{i=1}^n m_{2j}} \right) \\
&\quad + nd \\
&= \frac{\frac{1}{n} \sum_{j=1}^n \sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1 u_1}}{\left( \frac{1}{n} \sum_{j=1}^n m_{1j} \right)^2} - 2 \frac{\frac{1}{n} \sum_{j=1}^n \sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1 u_2}}{\left( \frac{1}{n} \sum_{j=1}^n m_{1j} \right) \left( \frac{1}{n} \sum_{j=1}^n m_{2j} \right)} \\
&\quad + \frac{\frac{1}{n} \sum_{j=1}^n \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2 u_2}}{\left( \frac{1}{n} \sum_{j=1}^n m_{2j} \right)^2} \\
&\quad - 2 \frac{\frac{1}{D_1} \sum_{j=1}^n \delta_{1j} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_1}}{\left( \frac{1}{n} \sum_{j=1}^n m_{1j} \right)} + 2 \frac{\frac{1}{D_2} \sum_{j=1}^n \delta_{2j} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_2}}{\left( \frac{1}{n} \sum_{j=1}^n m_{1j} \right)} \\
&\quad - 2 \frac{\frac{1}{D_2} \sum_{j=1}^n \delta_{2j} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_2}}{\left( \frac{1}{n} \sum_{j=1}^n m_{2j} \right)} + 2 \frac{\frac{1}{D_1} \sum_{j=1}^n \delta_{1j} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_1}}{\left( \frac{1}{n} \sum_{j=1}^n m_{2j} \right)} + nd.
\end{aligned}$$

Hence,

$$\begin{aligned}
\lim_{n \rightarrow \infty} nVar(\hat{\Delta}_1) &= \frac{E \sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1 u_1}}{(Em_1)^2} - 2 \frac{E \sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1 u_2}}{Em_1 Em_2} + \frac{E \sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2 u_2}}{(Em_2)^2} \\
&\quad - 2 \frac{E \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_1}}{Em_1} + 2 \frac{E \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_2}}{Em_1} - 2 \frac{E \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_2}}{Em_2} \\
&\quad + 2 \frac{E \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_1}}{Em_2} + \tau_1^{-1} \sigma^{v_1 v_1} + \tau_2^{-1} \sigma^{v_2 v_2}.
\end{aligned}$$

**Proof of (2.11)**

$$\begin{aligned}
nVar(\hat{\Delta}_2) &= n \left[ \sum_{j=1}^n a_{1j} w_{1j}^2 - 2 \sum_{j=1}^n c_j w_{1j} w_{2j} + \sum_{j=1}^n a_{2j} w_{2j}^2 \right] \\
&\quad - 2n \sum_{j=1}^n b_{1j} w_{1j} - 2n \sum_{j=1}^n b_{2j} w_{2j} + nd \\
&= n \sum_{j=1}^n \frac{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1 u_1}}{m_{1j}^2} \frac{1}{n^2} - 2n \sum_{j=1}^n \frac{\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1 u_2}}{m_{1j} m_{2j}} \frac{1}{n^2} \\
&\quad + n \sum_{j=1}^n \frac{\sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2 u_2}}{m_{2j}^2} \frac{1}{n^2} - 2n \sum_{j=1}^n \frac{\delta_{1j} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_1}}{D_1 m_{1j}} \frac{1}{n} \\
&\quad + 2n \sum_{j=1}^n \frac{\delta_{2j} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_2}}{D_2 m_{1j}} \frac{1}{n} - 2n \sum_{j=1}^n \frac{\delta_{2j} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_2}}{D_2 m_{2j}} \frac{1}{n} \\
&\quad + 2n \sum_{j=1}^n \frac{\delta_{1j} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_1}}{D_1 m_{2j}} \frac{1}{n} + nd \\
&= \frac{1}{n} \sum_{j=1}^n \frac{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1 u_1}}{m_{1j}^2} - \frac{2}{n} \sum_{j=1}^n \frac{\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1 u_2}}{m_{1j} m_{2j}} + \frac{1}{n} \sum_{j=1}^n \frac{\sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2 u_2}}{m_{2j}^2} \\
&\quad - \frac{2}{D_1} \sum_{j=1}^n \frac{\delta_{1j} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_1}}{m_{1j}} + \frac{2}{D_2} \sum_{j=1}^n \frac{\delta_{2j} \sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_2}}{m_{1j}} - \frac{2}{D_2} \sum_{j=1}^n \frac{\delta_{2j} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_2}}{m_{2j}} \\
&\quad + \frac{2}{D_1} \sum_{j=1}^n \frac{\delta_{1j} \sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_1}}{m_{2j}} + nd.
\end{aligned}$$

Hence,

$$\begin{aligned}
\lim_{n \rightarrow \infty} nVar(\hat{\Delta}_2) &= E \frac{\sum_{l,l'=1}^{m_{1j}} \sigma_{ll'}^{u_1 u_1}}{m_1^2} - 2E \frac{\sum_{k=1}^{m_{1j}} \sum_{l=1}^{m_{2j}} \sigma_{kl}^{u_1 u_2}}{m_1 m_2} + E \frac{\sum_{l,l'=1}^{m_{2j}} \sigma_{ll'}^{u_2 u_2}}{m_2^2} \\
&\quad - 2E \frac{\sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_1}}{m_1} + 2E \frac{\sum_{l=1}^{m_{1j}} \sigma_l^{u_1 v_2}}{m_1} - 2E \frac{\sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_2}}{m_2} + 2E \frac{\sum_{l=1}^{m_{2j}} \sigma_l^{u_2 v_1}}{m_2} \\
&\quad + \tau_1^{-1} \sigma^{v_1 v_1} + \tau_2^{-1} \sigma^{v_2 v_2}.
\end{aligned}$$

**Proof of (2.12)** First, let  $\tilde{w}_j = (w_{1j}, w_{2j})^t, j = 1, \dots, n$  and rewrite (2.5) as

$$\begin{aligned} nVar(\hat{\Delta}) &= n \left[ \sum_{j=1}^n a_{1j} w_{1j}^2 - 2 \sum_{j=1}^n c_j w_{1j} w_{2j} + \sum_{j=1}^n a_{2j} w_{2j}^2 \right] \\ &\quad - 2n \sum_{j=1}^n b_{1j} w_{1j} - 2n \sum_{j=1}^n b_{2j} w_{2j} + nd \\ &= n \sum_{j=1}^n \tilde{w}_j^t H_j \tilde{w}_j - 2n \sum_{j=1}^n \tilde{b}_j^t \tilde{w}_j + nd. \end{aligned}$$

Using the fact:

$$\tilde{w}_j = H_j^{-1} \left[ \left( \sum_{k=1}^n H_k^{-1} \right)^{-1} \tilde{I} - \left( \sum_{k=1}^n H_k^{-1} \right)^{-1} \left( \sum_{j=k}^n H_k^{-1} \tilde{b}_k \right) + \tilde{b}_j \right],$$

let

$$\tilde{o} = \left( \sum_{k=1}^n H_k^{-1} \right)^{-1} \tilde{I} - \left( \sum_{k=1}^n H_k^{-1} \right)^{-1} \left( \sum_{j=k}^n H_k^{-1} \tilde{b}_k \right),$$

so we have

$$\tilde{w}_j = H_j^{-1} (\tilde{o} + \tilde{b}_j),$$

then plug it in the variance function,

$$\begin{aligned}
nVar(\hat{\Delta}_{op}) &= n \sum_{j=1}^n \left[ H_j^{-1}(\tilde{o} + \tilde{b}_j) \right]^t H_j \left[ H_j^{-1}(\tilde{o} + \tilde{b}_j) \right] \\
&\quad - 2n \sum_{j=1}^{n_1} \tilde{b}_j^t \left[ H_j^{-1}(\tilde{o} + \tilde{b}_j) \right] + nd \\
&= n \sum_{j=1}^n (\tilde{o}^t + \tilde{b}_j^t) H_j^{-1}(\tilde{o} + \tilde{b}_j) - 2n \sum_{j=1}^n \tilde{b}_j^t \left[ H_j^{-1}(\tilde{o} + \tilde{b}_j) \right] + nd \\
&= n\tilde{o}^t \sum_{j=1}^n H_j^{-1}\tilde{o} - n \sum_{j=1}^n (\tilde{b}_j^t H_j^{-1}\tilde{b}_j) + nd \\
&= \left( \tilde{I}^t - \sum_{j=k}^n \tilde{b}_k^t H_k^{-1} \right) \sum_{j=1}^n \left( \frac{1}{n} H_j^{-1} \right)^{-1} \left( \tilde{I} - \sum_{j=k}^n H_k^{-1} \tilde{b}_k \right) \\
&\quad - n \sum_{j=1}^n (\tilde{b}_j^t H_j^{-1} \tilde{b}_j) + nd.
\end{aligned}$$

Note that:

$$\begin{aligned}
\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n H_j^{-1} &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \begin{pmatrix} \frac{a_{2j}}{a_{1j}a_{2j}-c_j^2} & \frac{c_j}{a_{1j}a_{2j}-c_j^2} \\ \frac{c_j}{a_{1j}a_{2j}-c_j^2} & \frac{a_{1j}}{a_{1j}a_{2j}-c_j^2} \end{pmatrix} \\
&= \lim_{n \rightarrow \infty} \begin{pmatrix} \frac{1}{n} \sum_{j=1}^n \frac{a_{2j}}{a_{1j}a_{2j}-c_j^2} & \frac{1}{n} \sum_{j=1}^n \frac{c_j}{a_{1j}a_{2j}-c_j^2} \\ \frac{1}{n} \sum_{j=1}^n \frac{c_j}{a_{1j}a_{2j}-c_j^2} & \frac{1}{n} \sum_{j=1}^n \frac{a_{1j}}{a_{1j}a_{2j}-c_j^2} \end{pmatrix} \\
&= \Sigma,
\end{aligned}$$

$$\begin{aligned}
\lim_{n \rightarrow \infty} \sum_{j=1}^n H_j^{-1} \tilde{b}_j &= \lim_{n \rightarrow \infty} \sum_{j=1}^n \begin{pmatrix} \frac{a_{2j}}{a_{1j}a_{2j}-c_j^2} & \frac{c_j}{a_{1j}a_{2j}-c_j^2} \\ \frac{c_j}{a_{1j}a_{2j}-c_j^2} & \frac{a_{1j}}{a_{1j}a_{2j}-c_j^2} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \end{pmatrix} \\
&= \lim_{n \rightarrow \infty} \begin{pmatrix} \sum_{j=1}^n \frac{b_{1j}a_{2j}+b_{2j}c_j}{a_{1j}a_{2j}-c_j^2} \\ \sum_{j=1}^n \frac{b_{2j}a_{1j}+b_{1j}c_j}{a_{1j}a_{2j}-c_j^2} \end{pmatrix} \\
&= \tilde{\phi},
\end{aligned}$$

$$\begin{aligned}\lim_{n \rightarrow \infty} \sum_{j=1}^n \tilde{b}_j^t H_j^{-1} &= \lim_{n \rightarrow \infty} \left( \sum_{j=1}^n \frac{b_{1j} a_{2j} + b_{2j} c_j}{a_{1j} a_{2j} - c_j^2} \quad \sum_{j=1}^n \frac{b_{2j} a_{1j} + b_{1j} c_j}{a_{1j} a_{2j} - c_j^2} \right) \\ &= \tilde{\phi}^t,\end{aligned}$$

and

$$\begin{aligned}\lim_{n \rightarrow \infty} n \sum_{j=1}^n \tilde{b}_j^t H_j^{-1} \tilde{b}_j &= \lim_{n \rightarrow \infty} n \sum_{j=1}^n \begin{pmatrix} b_{1j} & b_{2j} \end{pmatrix} \begin{pmatrix} \frac{a_{2j}}{a_{1j} a_{2j} - c_j^2} & \frac{c_j}{a_{1j} a_{2j} - c_j^2} \\ \frac{c_j}{a_{1j} a_{2j} - c_j^2} & \frac{a_{1j}}{a_{1j} a_{2j} - c_j^2} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \end{pmatrix} \\ &= \lim_{n \rightarrow \infty} \sum_{j=1}^n n \begin{pmatrix} b_{1j} & b_{2j} \end{pmatrix} \begin{pmatrix} \frac{b_{1j} a_{2j} + b_{2j} c_j}{a_{1j} a_{2j} - c_j^2} \\ \frac{b_{2j} a_{1j} + b_{1j} c_j}{a_{1j} a_{2j} - c_j^2} \end{pmatrix} \\ &= \lim_{n \rightarrow \infty} \left( n \sum_{j=1}^n \frac{b_{1j}^2 a_{2j} + 2b_{1j} b_{2j} c_j + b_{2j}^2 a_{1j}}{a_{1j} a_{2j} - c_j^2} \right) \\ &= v.\end{aligned}$$

Consequently, we arrive at

$$\lim_{n \rightarrow \infty} n \text{Var}(\hat{\Delta}_{op}) = \left( \tilde{I}^t - \tilde{\phi}^t \right) \Sigma^{-1} \left( \tilde{I} - \tilde{\phi} \right) - v + \tau_1^{-1} \sigma^{v_1 v_1} + \tau_2^{-1} \sigma^{v_2 v_2},$$

which completes the proof.

## 2.10 Bibliography

- [1] Cox DR. Regression models and life tables(with discussion). *Journal of the Royal Statistical Society, Series B* **34**:187-220 (1972).
- [2] Kalbfleisch JD Prentice RL . *The Statistical Analysis of Failure Time Data*, Wiley, New York **34**:187-220 (1980).
- [3] Emir B, Wieand S, Su J, Cha S. Analysis of repeated markers used to predict progression of cancer. *Statistics in Medicine* **17**:2563–2578 (1998).
- [4] Emir B, Wieand S, Jung S, Ying Z. Comparison of diagnostic markers with repeated measurements: a non-parametric ROC curve approach. *Statistics in Medicine* **19**:511–523 (2000).
- [5] Wu YG, Wang XF. Optimal weight in estimating and comparing areas under the receiver operating characteristic curve using longitudinal data. *Biometrical Journal* **53**:764–778 (2011).

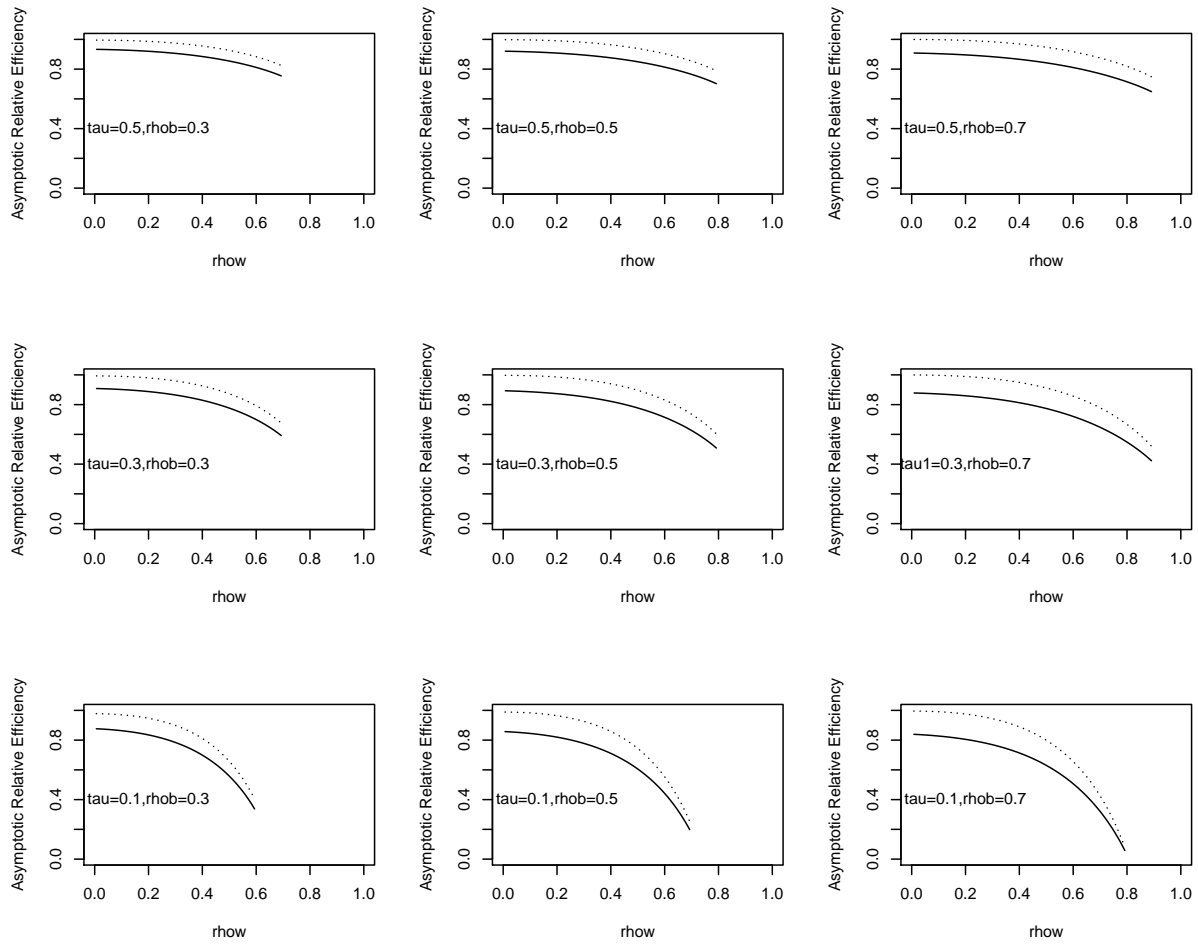


Figure 2.1: The Effect of Within-subject Correlation Coefficient( $\rho_w$ ), Between-group Correlation Coefficient( $\rho_b$ ), and the Incidence Rate of Disease( $\tau$ ) on the Asymptotic Relative Efficiencies,  $ARE_1$  (Solid Line) and  $ARE_2$  (Broken Line).



Table 2.2: Simulation Results for AUCs Comparison (n=50)

$n$	$\psi_1$	$\psi_2$	$\lambda$	$\gamma$	$RE_1$	$RE_2$	$RE_3$
50	0.5	0.4	0.0	0.3	0.92	0.95	0.99
50	0.5	0.4	0.0	0.5	0.89	0.90	0.99
50	0.5	0.4	0.0	0.7	0.85	0.88	0.97
50	0.5	0.4	0.0	0.9	0.76	0.79	0.97
50	0.5	0.4	0.3	0.3	0.90	0.91	0.96
50	0.5	0.4	0.3	0.5	0.88	0.89	0.97
50	0.5	0.4	0.3	0.7	0.83	0.85	0.97
50	0.5	0.4	0.3	0.9	0.77	0.80	0.97
50	0.5	0.4	0.5	0.3	0.84	0.85	0.87
50	0.5	0.4	0.5	0.5	0.86	0.87	0.93
50	0.5	0.4	0.5	0.7	0.84	0.86	0.98
50	0.5	0.4	0.5	0.9	0.78	0.80	0.97
50	0.5	0.4	0.7	0.3	0.82	0.83	0.88
50	0.5	0.4	0.7	0.5	0.81	0.82	0.90
50	0.5	0.4	0.7	0.7	0.80	0.82	0.89
50	0.5	0.4	0.7	0.9	0.79	0.81	0.96
50	0.4	0.4	0.0	0.3	0.90	0.92	0.99
50	0.4	0.4	0.0	0.5	0.84	0.86	0.98
50	0.4	0.4	0.0	0.7	0.80	0.85	0.98
50	0.4	0.4	0.0	0.9	0.72	0.73	0.97
50	0.4	0.4	0.3	0.3	0.88	0.89	0.94
50	0.4	0.4	0.3	0.5	0.86	0.87	0.95
50	0.4	0.4	0.3	0.7	0.78	0.80	0.97
50	0.4	0.4	0.3	0.9	0.71	0.72	0.97
50	0.4	0.4	0.5	0.3	0.79	0.80	0.86
50	0.4	0.4	0.5	0.5	0.79	0.81	0.90
50	0.4	0.4	0.5	0.7	0.80	0.82	0.93
50	0.4	0.4	0.5	0.9	0.73	0.75	0.96
50	0.4	0.4	0.7	0.3	0.75	0.76	0.82
50	0.4	0.4	0.7	0.5	0.77	0.78	0.87
50	0.4	0.4	0.7	0.7	0.78	0.79	0.91
50	0.4	0.4	0.7	0.9	0.72	0.74	0.94

Continued on next page

Table 2.2: Simulation Results for AUCs Comparison (n=50) (Continued)

$n$	$\psi_1$	$\psi_2$	$\lambda$	$\gamma$	$RE_1$	$RE_2$	$RE_3$
50	0.2	0.4	0.0	0.3	0.79	0.80	0.97
50	0.2	0.4	0.0	0.5	0.72	0.77	0.93
50	0.2	0.4	0.0	0.7	0.61	0.64	0.92
50	0.2	0.4	0.0	0.9	0.49	0.51	0.92
50	0.2	0.4	0.3	0.3	0.78	0.79	0.88
50	0.2	0.4	0.3	0.5	0.72	0.74	0.92
50	0.2	0.4	0.3	0.7	0.54	0.55	0.91
50	0.2	0.4	0.3	0.9	0.47	0.48	0.89
50	0.2	0.4	0.5	0.3	0.63	0.63	0.74
50	0.2	0.4	0.5	0.5	0.67	0.68	0.83
50	0.2	0.4	0.5	0.7	0.56	0.58	0.84
50	0.2	0.4	0.5	0.9	0.51	0.52	0.90
50	0.2	0.4	0.7	0.3	0.59	0.59	0.68
50	0.2	0.4	0.7	0.5	0.62	0.64	0.75
50	0.2	0.4	0.7	0.7	0.61	0.63	0.84
50	0.2	0.4	0.7	0.9	0.43	0.44	0.82

Table 2.3: Simulation Results for AUCs Comparison (n=100)

$n$	$\psi_1$	$\psi_2$	$\lambda$	$\gamma$	$RE_1$	$RE_2$	$RE_3$
100	0.5	0.4	0.0	0.3	0.83	0.86	0.99
100	0.5	0.4	0.0	0.5	0.78	0.79	0.98
100	0.5	0.4	0.0	0.7	0.74	0.76	0.98
100	0.5	0.4	0.0	0.9	0.71	0.74	0.97
100	0.5	0.4	0.3	0.3	0.82	0.84	0.95
100	0.5	0.4	0.3	0.5	0.77	0.79	0.96
100	0.5	0.4	0.3	0.7	0.71	0.75	0.95
100	0.5	0.4	0.3	0.9	0.68	0.70	0.92
100	0.5	0.4	0.5	0.3	0.78	0.82	0.84
100	0.5	0.4	0.5	0.5	0.76	0.78	0.86
100	0.5	0.4	0.5	0.7	0.69	0.73	0.82
100	0.5	0.4	0.5	0.9	0.65	0.68	0.81
100	0.5	0.4	0.7	0.3	0.75	0.81	0.83
100	0.5	0.4	0.7	0.5	0.71	0.75	0.85
100	0.5	0.4	0.7	0.7	0.67	0.70	0.80
100	0.5	0.4	0.7	0.9	0.62	0.69	0.82
100	0.4	0.4	0.0	0.3	0.82	0.84	0.99
100	0.4	0.4	0.0	0.5	0.78	0.82	0.99
100	0.4	0.4	0.0	0.7	0.70	0.80	0.98
100	0.4	0.4	0.0	0.9	0.68	0.73	0.96
100	0.4	0.4	0.3	0.3	0.80	0.82	0.95
100	0.4	0.4	0.3	0.5	0.76	0.79	0.92
100	0.4	0.4	0.3	0.7	0.69	0.74	0.91
100	0.4	0.4	0.3	0.9	0.67	0.71	0.89
100	0.4	0.4	0.5	0.3	0.73	0.78	0.88
100	0.4	0.4	0.5	0.5	0.70	0.76	0.85
100	0.4	0.4	0.5	0.7	0.65	0.69	0.83
100	0.4	0.4	0.5	0.9	0.61	0.64	0.84
100	0.4	0.4	0.7	0.3	0.65	0.75	0.85
100	0.4	0.4	0.7	0.5	0.61	0.70	0.82
100	0.4	0.4	0.7	0.7	0.59	0.66	0.81
100	0.4	0.4	0.7	0.9	0.52	0.59	0.80

Continued on next page

Table 2.3: Simulation Results for AUCs Comparison (n=100) (Continued)

$n$	$\psi_1$	$\psi_2$	$\lambda$	$\gamma$	$RE_1$	$RE_2$	$RE_3$
100	0.2	0.4	0.0	0.3	0.62	0.72	0.97
100	0.2	0.4	0.0	0.5	0.59	0.65	0.97
100	0.2	0.4	0.0	0.7	0.51	0.56	0.96
100	0.2	0.4	0.0	0.9	0.44	0.52	0.95
100	0.2	0.4	0.3	0.3	0.61	0.70	0.88
100	0.2	0.4	0.3	0.5	0.55	0.64	0.85
100	0.2	0.4	0.3	0.7	0.50	0.54	0.81
100	0.2	0.4	0.3	0.9	0.44	0.48	0.79
100	0.2	0.4	0.5	0.3	0.61	0.65	0.82
100	0.2	0.4	0.5	0.5	0.60	0.62	0.84
100	0.2	0.4	0.5	0.7	0.55	0.51	0.81
100	0.2	0.4	0.5	0.9	0.42	0.49	0.80
100	0.2	0.4	0.7	0.3	0.58	0.59	0.72
100	0.2	0.4	0.7	0.5	0.53	0.54	0.76
100	0.2	0.4	0.7	0.7	0.49	0.50	0.74
100	0.2	0.4	0.7	0.9	0.41	0.44	0.79

## Chapter 3

# Prognostic Performance of Metabolic Indexes in Predicting Onset of Type 1 Diabetes

The non-parametric area estimation and comparisons are most commonly used in practice due to its simplicity. Here I will apply Delong's method on a Type 1 Diabetes biomarker research. The application in this section is published by the peer review journal Diabetes Care (Xu et al., 2010).

Copyright 2010 American Diabetes Association From Diabetes Care, Vol. 33, 2010; 2508 – 2513. Reprinted by permission of the American Diabetes Association.

### 3.1 Introduction

Early disease prediction and prevention is one of the most important strategies in health care. Preventative care can substantially decrease mortality and morbidity and significantly reduce public health costs (1-2). As genetic/familial factors and auto-immune factors have become available to screen subjects for the risk of developing type 1 diabetes, early intervention trials for this

disease have become a reality (3-7). The characterization of this risk may be refined by additional factors to more precisely target individuals who would benefit from preventative treatment. To most accurately select individuals who are at risk for developing disease, beyond screening for antibodies and genetic factors, metabolic risk indicators are being investigated for the development of a more effective clinical prognostic index (8-12). The principal metabolic indexes currently being evaluated as prognostic indicators for type 1 diabetes have been focused on measurements from oral glucose tolerance tests (OGTTs) and intravenous glucose tolerance tests (IVGTTs). Previous research from DPT-1 has indicated that some metabolic indexes derived from OGTT provide substantial predictive value in receiver operating characteristic area under the curve (ROCAUC) analysis (13). IVGTT derived indexes, such as FPIR, HOMA-IR, and FPIR/HOMA-IR have also demonstrated prognostic value (14-16). However, indexes from both methods have not been compared for predictive accuracy in moderate risk subjects who are antibody positive and have genetic risk factors, but do not have impaired glucose tolerance. As subjects in this population who would progress to disease are in an early stage of disease progression, they are an important subgroup to target for preventative intervention. If OGTT or IVGTT measurements produce superior predictive indexes compared to one another, future trials can reduce costs by relying on a single method of measurement that produces the greatest predictive accuracy.

In addition to determining the superior testing method of producing predictive indexes (OGTT versus IVGTT) there also remains a need to produce effective prognostic thresholds to select between individuals who will progress to disease and who will not, as screening for familial, genetic, and immunoglobulin risk factors are not precise enough to accurately select subjects, particularly for those at an early stage of disease progression who do not exhibit impaired glucose

tolerance. Future intervention trials will depend on a refined selection tool to choose subjects for early intervention to ensure an accurate characterization of treatment effects. Optimal cut-off values derived from ROCAUC analysis from metabolic indexes would provide valuable guidance for clinicians and researchers in evaluating patient risk for progressing to type 1 diabetes by providing a threshold, above which the risk is characterized with greater precision than is provided by their underlying risk factors.

In this investigation, we assessed the prognostic accuracy of nine metabolic indexes for predicting the progression to clinical onset of type 1 diabetes over a five year period using the data from Diabetes Prevention Trial-Type 1 (DPT-1). The optimal cut-off values of metabolic indexes were determined to provide previously unavailable guidance to clinicians and researchers in selecting patients likely to progress to disease, who are therefore candidates for early preventative intervention.

## **3.2 Subjects and methods**

The Diabetes Prevention Trial-Type 1 (DPT-1) was a longitudinal study in North America, which was designed to determine if Type 1 Diabetes can be prevented or delayed by preclinical intervention of oral insulin supplement. DPT-1 screened for ICA-positive subjects in 103, 390 first- and second-degree nondiabetic relatives of individuals in whom type 1 diabetes had been diagnosed before the age of 45 years. The 3, 483 relatives positive for islet-cell antibodies (ICA) were staged to quantify the projected 5-year risk of diabetes (17). Staging consisted of ICA confirmation, HLA-DQ typing, determination of IAA autoantibody, IVGTT and OGTT.

A total of 372 subjects whose 5-year risk was considered to be 25 to 50 without metabolic abnormality or loss of first phase insulin response (defined as moderate risk), were entered into the DPT-1 oral insulin trial. To study the natural history of the disease, the current investigation evaluated 186 subjects who were randomized to the placebo arm of the study.

All subjects were examined every six months from enrollment until diabetes onset or study end after randomization. All subjects (and/or their parent) signed a written consent form approved by the participating study center's human subjects committee.

### **3.3 Laboratory measurements**

An intravenous glucose tolerance test was performed after an overnight fast. The IV-GTT was done as recommended by the ICARUS study group. This includes instructions for a diet containing at least 150 g of carbohydrate per day for the 3 days before the test. Blood samples for determination of glucose and insulin levels were drawn at -10 and -4 minutes. A solution of 25 dextrose (0.5 g/kg body weight up to a maximum of 35 g) was then infused over 3 minutes (15 seconds). Repeat blood samples for determination of glucose and insulin levels were drawn at 1, 3, 5, 7, and 10 minutes after the glucose infusion. FPIR was calculated as the sum of the serum insulin concentrations at 1 and 3 minutes after intravenous injection of glucose. FPIR was above threshold if  $\geq$  10th percentile for siblings, offspring, and second-degree relatives ( $\geq$  100 U/ml if age  $\geq$  8 years;  $\geq$  60 U/ml if age  $<$  8 years) and  $\geq$  1st percentile for parents ( $\geq$  60 U/ml). These thresholds were determined from the Gainesville, Florida Family Study and a general school population study (18). FPIR above threshold was required for eligibility. HOMA-IR was calculated



as the fasting insulin (mU/L) x fasting glucose (mmol/L)/22.5 from the mean of fasting insulin at 4 and 10 minutes and fasting glucose at 4 minutes before each IVGTT performed.

Oral glucose tolerance tests (OGTT) were administered to assess glycemic status. The dose of oral glucose was 1.75 g per kilogram (maximum, 75 g of carbohydrate). Blood samples were obtained for C-peptide measurements in the fasting state, and then 30, 60, 90 and 120 minutes after. Peak C-peptide was the maximum point of all measurements. The AUC C-peptide was calculated using Trapezoid's rule. A normal OGTT during staging was required for eligibility.

Diabetes was diagnosed according to the American Diabetes Association (ADA) criteria: fasting glucose  $\geq$  126 mg/dl; two-hour glucose  $\geq$  200 mg/dl with confirmation by either an elevated fasting or two-hour glucose level at a follow-up visit; or random plasma glucose  $\geq$  200 mg/dl accompanied by symptoms of polyuria, polydipsia, and /or weight loss (19).

### **3.4 Data analysis**

For baseline characteristics, categorical variables were compared by Pearson's chi-squared test, and continuous variables were evaluated by t-test for the differences in means or by Wilcoxon rank sum test for differences in order between progressors and non-progressors. Receiver operating characteristic (ROC) curves were used to compare the discriminative power of nine different metabolic indexes to predict progression to type 1 diabetes. The global performance of each measurement in predicting progression to type 1 diabetes was summarized by the area under the curve (AUC) and results are presented as the mean and 95% confidence interval. Confidence intervals that exclude 0.5 were considered to indicate significant results (20). Areas under the ROC curves

were compared using the algorithm suggested by DeLong et al. 1988, which is a nonparametric approach to the analysis of areas under correlated ROC curves by using the theory on generalized U-statistics to generate an estimated covariance matrix (21). The ROC curve is constructed by varying the cut point used to determine which values of the observed variable will be considered abnormal and then plotting the resulting sensitivities against the corresponding false positive rates (1-specificity). The optimal cut-off points are the values yielding maximum sums of sensitivity and specificity from the ROC curves. The Cox proportional hazard model was utilized to calculate the hazard ratio. The log-rank test was used for survival curves comparison. Analyses were performed by using SAS version 9.2 (SAS Institute Inc., Cary, NC) software.

### **3.5 Results**

A summary of clinical and metabolic characteristics comparing those who progressed to clinical disease onset and those who did not are shown in Table 3.1. By design, subjects were ICA and IAA positive with a normal FPIR and normal glucose tolerance (n=186) giving them a projected risk of 25-50 for progression to clinical diabetes over five years. The actual risk was 35 over 5 years. The subjects were followed for a median of 1,213 days (3.3 years; interquartile range 726-1718). Annual rate of loss to follow-up was 0.2. The subjects who were lost to follow-up before the end of the study were considered to be non-progressors. The progressors and non-progressors did not significantly differ in age, sex, race, or relationship to the proband.

The areas under the ROC curves (AUCs) of various metabolic indexes for the prediction of progression to type 1 diabetes are summarized in Table 3.2. Fasting glucose from both IVGTT and

OGTT performed poorly and did not demonstrate prognostic ability with the same AUC value of 0.49 (0.40-0.59). A analysis of IVGTT fasting insulin demonstrated some prognostic value with an AUC value of 0.59 (0.5 - 0.68) though the estimate had borderline significance. The AUC estimate of HOMA-IR and FPIR exceeded 0.5, though the lower confidence limit for both variables fell slightly below 0.5 at 0.49 and 0.48 respectively, rendering them non-significant predictors. However, when the ratio of FPIR/HOMA-IR was analyzed, it resulted in AUC value of 0.66 (0.57-0.74), representing the best index among the indexes derived from IVGTT. The only statistically significant AUC among the standard indexes derived from OGTT testing was two-hour glucose, which yielded the greatest AUC value of all examined metabolic indexes at 0.67 (0.59-0.76). A composite index that included AUC glucose and peak c peptide was developed, using the proportional hazard model [index =  $3.54 \times 10^4 \times (\text{AUC glucose}) - 0.15 \times (\text{peak c-peptide})$ ]. The ROCAUC result for the OGTT composite index was 0.71 (0.63-0.79). Although higher than the FPIR/HOMA-IR and the two-hour glucose, the differences were not significant ( $p=0.68$ ;  $p=0.64$ ). The prediction performance of antibody titers was evaluated for comparison. AUCs for ICA titer and IAA titer were 0.69 (0.61-0.77), 0.67 (0.58-0.76) respectively. They did not provide better prediction than two-hour glucose or FPIR/HOMA-IR ( $P \leq 0.05$ ) in this population. The cut-off values for optimal prediction of progression to type 1 diabetes using ROC analysis for all metabolic indexes are summarized in Table 3.2 Although FPIR demonstrated high sensitivity (0.76) at the optimal cut-off value, the low specificity (0.40) diminishes FPIR as a useful prognostic index. Optimal cut-off values for fasting insulin and AUC C-peptide also had high positive predictive ability with sensitivities greater than 0.80, but were lacking in specificity similar to FPIR, which restricts them from being useful prognostic indexes in this moderate risk group. FPIR/HOMA-IR and two-hour glucose OGTT produced cut-off values with both sensitivity and specificity levels above 0.6, demonstrating

the greatest potential as prognostic indexes compared to all other metabolic indexes. The positive predictive value (PPV) was 0.46 and 0.41 for two-hour glucose and FPIR/HOMA-IR, respectively. The negative predictive value (NPV) was 0.83 for both indexes. The optimal cut-off value of 114 mg/dL for OGTT two-hour glucose was notably less than the value currently used to define the lower range of impaired glucose tolerance (140 mg/dL). The hazard ratio for those with two-hour glucose equal to or in excess of the optimal cut-off value (114) compared to those under the optimal cut-off was 2.96 (1.67-5.22) after adjusting for age, gender and BMI. The 5-year risk for those with baseline values equal to or over the optimal two-hour glucose cut-off level was 46 compared to 17 for those with baseline values under the optimal cut-off. Kaplan-Meier curves (Figure 3.2 3.3) illustrate the risk of diabetes over the study period by the level of OGTT two-hour glucose ( $P < 0.001$ ) and the level of IVGTT FRIR/HOMA-IR ( $P < 0.001$ ).

### **3.6 Discussions and conclusions**

The comprehensive baseline data from DPT-1 has provided the opportunity to assess and validate the accuracy of metabolic risk indicators in predicting the future development of type 1 diabetes. Elevated fasting glucose and substantially impaired glucose tolerance are the primary metabolic indicators currently used to identify those with significant risk for progressing to type 1 diabetes; however, these indicators cannot be used to effectively predict disease progression among individuals who had not demonstrated gross metabolic abnormality. In order to generate accurate prognostic indexes for those with such moderate risk factors, the current investigation evaluated nine metabolic indicators for their prediction accuracy among those with the following underlying

risk factors: ICA/IAA-positive relatives of those diagnosed with type 1 diabetes who were not classified as having impaired glucose tolerance. The results of this investigation identified two metabolic indicators as having a significant predictive accuracy for identifying individuals likely to progress to clinical onset of type 1 diabetes within 5 years: FPIR/HOMA-IR and two-hour postprandial glucose.

Investigations into the progression of type 2 diabetes have shown that insulin resistance is demonstrated long before overt diabetes is diagnosed, and that this can be a powerful predictor of disease progression (20-23). Abdul-Ghani et al. 2007 investigated the prognostic performance of insulin secretion/insulin resistance indexes using ROC analysis to determine their predictive accuracy for progression to type 2 diabetes and reported that insulin secretion/insulin resistance is the best predictor of type 2 diabetes demonstrated by substantial sensitivity and specificity (20). In the current investigation, the role of insulin resistance is also implied from the analysis of FPIR/HOMA-IR, which demonstrated significant predictive accuracy in the study population, with an AUC cut-off point that maintained 0.68 sensitivity and 0.61 specificity for detecting progression to clinical onset. The etiological significance of this observation is the indication that the early stages of type 1 diabetes may be demarcated by a disturbance in the balance between insulin response and insulin activity, though the effects of an insulin secretion/insulin resistance disturbance may be more subtle in type 1 diabetes progression (24). The counter-regulatory hormones associated with puberty may play a role in the modulation of insulin secretion and insulin resistance during this developmental period (24-25).

The clinical utility of FPIR/HOMA-IR must be considered in the context of the predictive accuracy found in a standard index derived from OGTT, a test that is both less burdensome to the

patient, and in some instances, may be administered at lower cost. FPIR/HOMA-IR is actually a composite of three measures: two from IVGTT to determine FPIR and a fasting status measurement to determine HOMA-IR. In contrast, the relatively simplistic two-hour postprandial glucose test from OGTT maintained the largest area under the ROC curve, and the greatest degree of both sensitivity and specificity at the optimal cut-off value compared to all other metabolic indicators evaluated in this investigation. As well, when compared at the same negative predictive value, two-hour glucose retained a higher estimated positive predictive value than FPIR/HOMA-IR, though this difference was not statistically significant. However, since HOMA-IR does not require IVGTT, future research may indicate that insulin resistance can be incorporated as a part of a composite to increase predictive value.

OGTT is the gold standard method to diagnose type 1 diabetes diagnosis and impaired glucose tolerance. A standard two-hour glucose tolerance test contributes substantial predictive accuracy beyond screening for the underlying risk factors. This result is consistent with the findings by Sosenko et al., who determined the accuracy of two-hour glucose was 0.66 for the combined study population of subjects with and without impaired glucose tolerance at baseline (13). The prognostic accuracy of this index is likely due to two-hour glucose being influenced by insulin production and insulin resistance, two important factors that modulate the progression of this disease. This is evidenced by an observed level of prognostic accuracy similar to FPIR/HOMA-IR. As a principal objective of this investigation was to examine metabolic indexes to provide clinical guidance for selecting individuals at high risk for progression to disease despite the absence of clinical metabolic abnormality, we have determined the optimal cut-off point for the metabolic index with the largest area under the ROC curve, two-hour glucose at 114 mg/dL. This threshold

derived from standard OGTT diagnostics may provide cost effective guidance for clinicians and researchers for selecting patients likely to progress to disease if not provided preventative care, even among those presenting with normal glucose tolerance, in future prevention trials.

This investigation contributes to a growing body of evidence that metabolic indexes derived from OGTT testing are the most efficient and effective analytical method of determining the risk for progression to type 1 diabetes in those with known genetic and autoimmune risk factors(9, 13). Measurements from OGTT are the clinical standard for diagnosing impaired glucose tolerance and clinical diabetes, and are therefore a necessary component in evaluating subjects for type 1 diabetes. In the absence of superior predictive value from indexes produced by IVGTT measurements, IVGTT essentially replicates the results from OGTT measurements. Future intervention trials may consider eliminating IVGTT measurements as an effective cost reduction strategy.

The success of preventative medicine is dependent on accurately identifying patients with risk for disease development at an early stage of disease progression. In this investigation, we analyzed the prediction accuracy of nine common and novel metabolic indicators for identifying patients with moderate risk factors, yet display no clinical metabolic abnormality, who progress to type 1 diabetes within 5 years. Both FPIR/HOMA-IR from IVGTT and two-hour glucose from OGTT provided significant prognostic value. The standard OGTT index of two-hour glucose is preferred because it achieved the largest prognostic accuracy in predicting disease onset, making FPIR/HOMA-IR a redundant, and unnecessary, predictive index. As well, our findings confirm prior DPT-1 findings that accuracy for selecting at risk patients could be improved by utilizing two-hour glucose values below the current cutoff for impaired glucose tolerance.

Future analysis should focus on the combination of metabolic biomarkers, immunological

biomarkers and/or genetic biomarkers in order to improve the modest positive predictive values derived from solitary predictors. Our results indicated that autoantibody titers may not provide a cost/benefit improvement over OGTTs as a sole predictor, but they may be an important component of a composite modeled score. This is a secondary analysis of DPT-1 data with the sample size limited to the placebo arm of the oral insulin protocol. Nonetheless, it provides much needed guidance for clinicians and researchers for selecting subjects in future prevention trials in populations with underlying risk factors and clinically normal glucose tolerance by providing both the evidence for a preferred index and a threshold for selection.

### **3.7 Acknowledgements**

This work was sponsored by cooperative agreements with the Division of Diabetes, Endocrinology, and Metabolic Diseases of the National Institute of Diabetes and Digestive and Kidney Diseases, the National Institute of Allergy and Infectious Diseases, the National Institute of Child Health and Human Development, the National Center for Research Resources, the American Diabetes Association, and the Juvenile Diabetes Research Foundation.

I would also like to thank all the co-authors, Dr. Jeffrey Krischer, Mr. David Cuthbertson, Dr. Jay Skyler, Dr. Jay Sosenko, Dr. Giffe Johnson, Dr. Yougui Wu, Dr. Yiliang Zhu and Dr. Getachew Dagne for reviewing, editing the manuscript and contributions to the discussion.



## 3.8 Bibliography

- [1] Hensrud, D. D. "Clinical Preventive Medicine in Primary Care: Background and Practice: 1. Rationale and Current Preventive Practices." Mayo Clinic Proceedings 75(2):165-172, 2000.
- [2] Hensrud, D. D. "Clinical Preventive Medicine in Primary Care: Background and Practice: 2. Delivering Primary Preventive Services." Mayo Clinic Proceedings 75(3):255-264, 2000.
- [3] Skyler JS. Prediction and prevention of type 1 diabetes: progress, problems, and prospects. Clin Pharmacol Ther. 81(5):768-771, 2007.
- [4] Harrison LC: Risk assessment, prediction and prevention of type1diabetes. PediatrDiabetes 2: 71-83, 2001.
- [5] Riley WJ, Krischer J, Spillar RP, Silverstein JH, Schatz DA, Schwartz S, Malone J, Shah S, Vadheim C, Rotter JI: A prospective study of the development of diabetes in relatives of patients with insulin-dependent diabetes. N Engl J Med 323:1167-1172, 1990.
- [6] Eisenbarth G: Prediction of type1 diabetes: the natural history of the prediabetic period. In Immunology of Type 1 Diabetes, 2nd ed. Eisenbarth GS, Ed. NewYork: Springer Publishing, p.268-281, 2004.
- [7] Greenbaum CJ, Sears KL, Kahn SE, Palmer JP: Relationship of  $\beta$ -cell function and autoantibodies to progression and nonprogression of subclinical type 1 diabetes. Diabetes 48:170 -175, 1999.
- [8] Sosenko JM, Palmer JP, Greenbaum CJ, Mahon J, Cowie C, Krischer JP, Chase HP, White NH, Buckingham B, Herold KC, Cuthbertson D, Skyler JS. Patterns of metabolic progression to type 1

diabetes in the Diabetes Prevention Trial-Type 1. *Diabetes Care* 29(3): 643-649, 2006.

[9] Sosenko JM, Krischer JP, Palmer JP, Mahon J, Cowie C, Greenbaum CJ, Cuthbertson D, Lachin J, Skyler JS, and Diabetes Prevention Trial-Type 1 Study Group: A Risk Score for Type 1 Diabetes Derived From Autoantibody-Positive Participants in the Diabetes Prevention Trial-Type 1. *Diabetes Care* 31: 528-533, 2008.

[10] Sosenko JS, Palmer JP, Rafkin-Mervis L, Krischer JP, Cuthbertson D, Matheson D and Skyler JS: Glucose and C-Peptide Changes in the Perionset Period of Type 1 Diabetes in the Diabetes Prevention Trial-Type 1 *Diabetes Care* 31:2188-2192, 2008.

[11] Ferrannini E, Mari A, Nofrate V, Sosenko JM, Skyler JS, the DPT-1 Study Group. Progression to diabetes in relatives of type 1 diabetic patients: Mechanisms and mode of onset. *Diabetes* 59 (3): 679-685, 2009.

[12] Palmer JP, Fleming GA, Greenbaum CJ, Herold KC, Jansa LD, Kolb H, Lachin JM, Polonsky KS, Pozzilli P, Skyler JS, Steffes MW: C-peptide is the appropriate outcome measure for type 1 diabetes clinical trials to preserve  $\beta$ -cell function. *Diabetes* 53:250-264, 2004.

[13] Sosenko JM, Palmer JP, Greenbaum CJ, Mahon J, Cowie C, Krischer JP, Chase HP, White NH, Buckingham B, Herold KC, Cuthbertson D, Skyler JS: Increasing the accuracy of oral glucose tolerance testing and extending its application to individuals with normal glucose tolerance for the prediction of type 1 diabetes: the Diabetes Prevention Trial-Type 1. *Diabetes Care* 30:38-42, 2007.

[14] Mrena S, Virtanen SM, Laippala P, Kulmala P, Hannila ML, kerblom HK, Knip M: Models for predicting type 1 diabetes in siblings of affected children. *Diabetes Care* 29:662- 667, 2006.

[15] Chase HP, Cuthbertson DD, Dolan LM, Kaufman F, Krischer JP, Schatz DA, White NH,

Wilson DM, Wolfsdorf J: First-phase insulin release during the in- Travenous glucose tolerance test as a risk factorfortype1diabetes. J Pediatr 138: 244-249, 2001.

[16] Xu P, Cuthbertson D, Greenbaum C, Palmer JP, Krischer JP; Diabetes Prevention Trial-Type 1 Study Group. Role of insulin resistance in predicting progression to type 1 diabetes. Diabetes Care 30(9):2314-2320, 2007.

[17] Krischer JP, Cuthbertson DD, Yu L, Orban T, Maclaren N, Jackson R, Winter WE, Schatz DA, Palmer JP, Eisenbarth GS: Screening strategies for the identification of multiple antibody positive relatives of individuals with type 1 diabetes. J Clin Endocrinol Metab 88:103-108, 2003.

[18] Schatz D, Krischer J, Horne G, Riley W, Spillar R, Silverstein J: Islet cell antibodies predict insulin dependent diabetes in U.S. school age children as powerfully as in unaffected relatives. J Clin Invest 93:2403-7,1994.

[19] American Diabetes Association Expert Committee on the Diagnosis and lassification of Diabetes Mellitus: Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care 20:1183-1197, 1997.

[20] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143: 29-36, 1982.

[21] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves; a nonparametric approach. Biometrics 44: 837-845, 1988.

[22] Abdul-Ghani M, DeFronzo RA, Williams K, Stern M : What is the best predictor of future type 2 diabetes? Diabetes care 30: 1544-1548, 2007.

[23] Greenbaum CJ, Kahn SE, Palmer JP: Nicotinamide's effects on glucose metabolism in subjects at risk for type 1 diabetes. *Diabetes* 45:1631-1634,1996.

[24] Greenbaum CJ: Insulin resistance in type 1 diabetes. *Diabetes Metab Rev* 18:192-200, 2002.

[25] Acerini CI, Cheetham TD, Edge JA, Dunger DB: Both insulin sensitivity and insulin clearance in children and young adults with type 1 (insulin-dependent diabetes) vary with growth hormone concentrations and with age. *Diabetologia* 43:61-68, 2000.

Table 3.1: Clinical and Metabolic Characteristics of the Study Subjects.

Characteristic:	Progressor	Non-progressor	P value
n(%)	53(28)	133(72)	
Age-yr (median)	9.78(6.35)	12.79(9.23)	0.67
BMI Z-score (median) *	-0.67(-1.56-0.53)	-1.10(-2.61- 0.33)	0.07
Race n(%)			0.64
White	47(88.68)	116(89.23)	
African American	0(0.00)	2(1.54)	
Hispanic	4(7.55)	10(7.69)	
Other	2(3.77)	2(1.54)	
Gender n(%)			0.53
Male	28(52.83)	77(57.89)	
Female	25(47.17)	56(42.11)	
Relationship to proband n (%)			0.09
Sibling	36(67.92)	72(54.14)	
Offspring	15(28.30)	38(28.57)	
Parent	1(1.89)	6(4.51)	
Second Degree	1(1.89)	17(12.78)	
Immunological factors:			
ICA titer (JDF Units**) (median)	160.00(80.00-320.00)	80.00(20.00-160.00)	0.01
IAA titer (nU/ml) (median)	385.10(125.40-672.00)	156.70(73.70-343.00)	0.01
ICA512 antibodies n (%)		0.01	
Positive	32(64.00)	50(42.37)	
Negative	18(36.00)	68(57.63)	
GAD antibodies n (%)		0.67	
Positive	38(76.00)	86(72.88)	
Negative	12(24.00)	32(27.12 )	
Metabolic factors:			
Fasting Glucose (mmol/L)-IV	87.66(9.54)	87.12(9.18)	0.67
Fasting Insulin (mU/L)-IV	17.02(10.07)	14.84(9.50)	0.17
FPIR (l/ml)-IV	145.22(80.49)	163.90(105.84)	0.20
HOMA-R- IV	3.79(2.69)	3.25(2.20)	0.16
FPIR/HOMA-R- IV	44.28(19.20)	59.93(36.58)	0.01
Fasting Glucose (mg/dL) – OG	86.26(7.69)	86.18(7.85)	0.95
Two-hour Glucose (mg/dL) – OG	113.25(18.71)	102.75(19.17)	0.01
Peak C-Peptide (nmol/L) – OG	5.09(1.98)	5.59(2.26)	0.16
AUC C-Peptide(nmol/L) – OG	472.13(172.51)	523.96(21.698)	0.12

Note: Data are mean ( SD), n (%), or median (Inter-quartile range).

\*BMI Z-score from 2000 CDC Growth chart.

\*\*JDF denotes Juvenile Diabetes Foundation.

Table 3.2: AUC, Specificity, Sensitivity, PPV and NPV at the Optimal Cut-off Value.

Index/testing	AUC(95% CI)	Optimal cutoff	Sen	Spe	PPV	NPV
Fasting Glucose (IV)	0.49 (0.40 - 0.59)	96	0.23	0.88	0.46	0.74
Fasting Insulin (IV)	0.59 (0.50 - 0.68)	10.23	0.82	0.29	0.32	0.80
FPIR (IV)	0.57 (0.48 - 0.66)	156	0.76	0.4	0.34	0.81
HOMA-IR (IV)	0.58 (0.49 - 0.67)	2.64	0.66	0.48	0.33	0.77
FPIR/HOMA-IR (IV)	0.66 (0.57 - 0.74)	49.22	0.68	0.61	0.41	0.83
Fasting Glucose (OG)	0.49 (0.40 - 0.59)	88	0.47	0.56	0.27	0.7
Two-hour Glucose (OG)	0.67 (0.59 - 0.76)	114	0.62	0.71	0.46	0.83
Peak C-Peptide (OG)	0.56 (0.47 - 0.66)	5.3	0.7	0.46	0.34	0.79
AUC C peptide (OG)	0.56 (0.47 - 0.65)	595	0.81	0.31	0.33	0.8

Note: AUCs of Two-hour Glucose and FPIR/HOMA-IR are significant better than AUC's fasting glucose derived from IVGTT(IV) or OGTT(OG) ( $p \leq 0.01$ ).

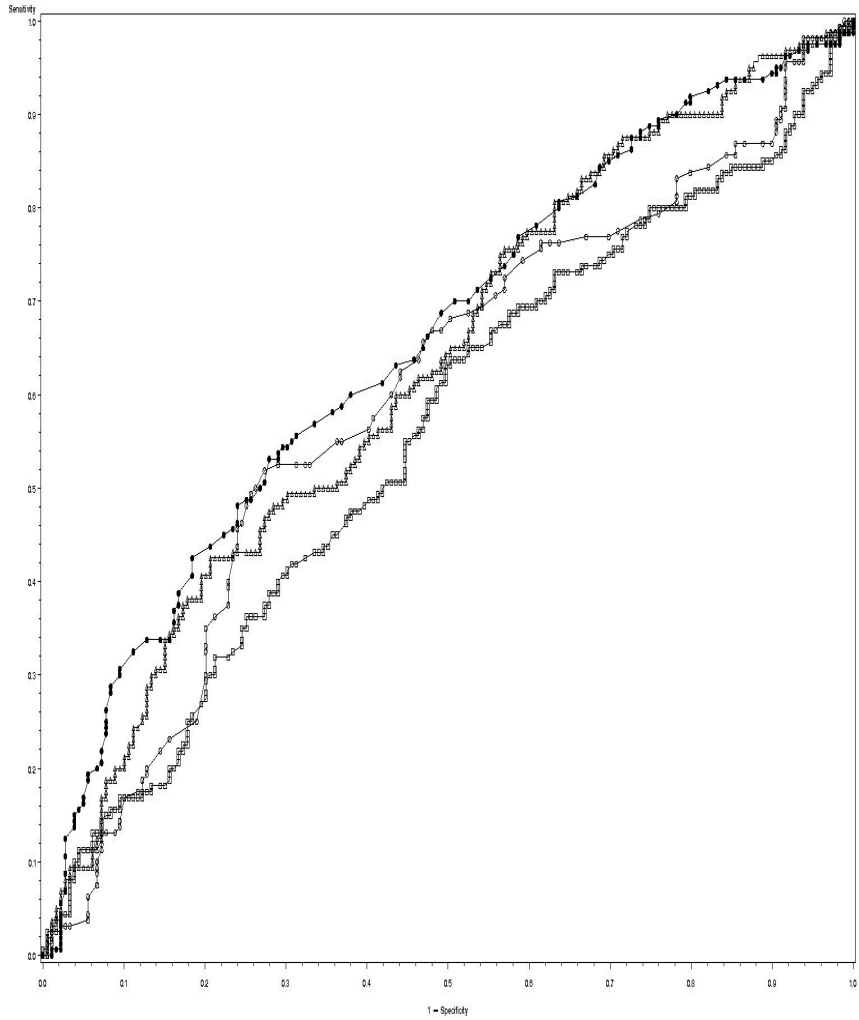


Figure 3.1: ROC Curves for Various Metabolic Indexes.  
 (black circle=2-hour glucose--OGTT; white circle=FPIR;  
 white triangle= FPIR/HOMA-IR; white square=HOMA-IR)

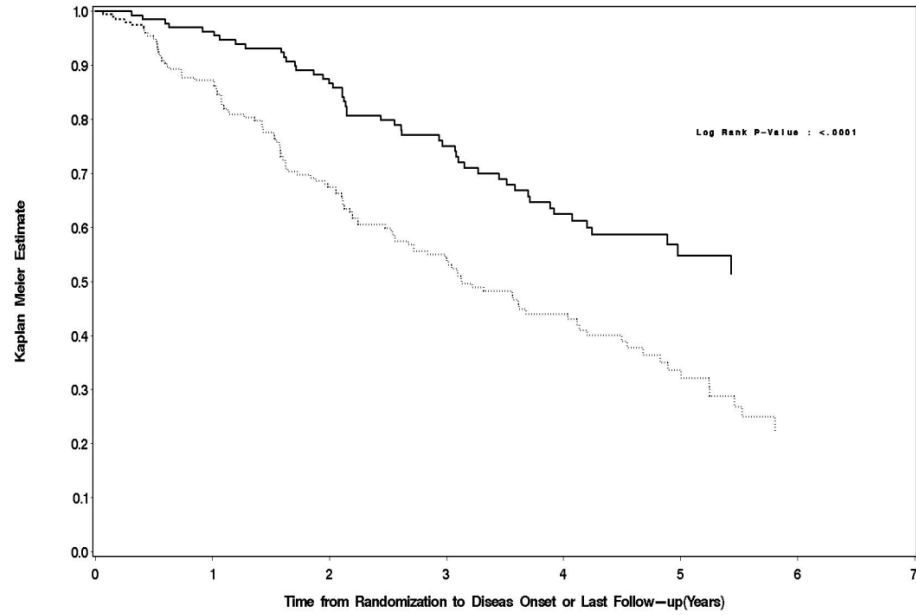


Figure 3.2: Kaplan-Meier Curves by the Level of OGTT 2-hour Glucose.  
 (solid line= 2-hour glucose < 114; broken line=2-hour glucose  $\geq 114$ );

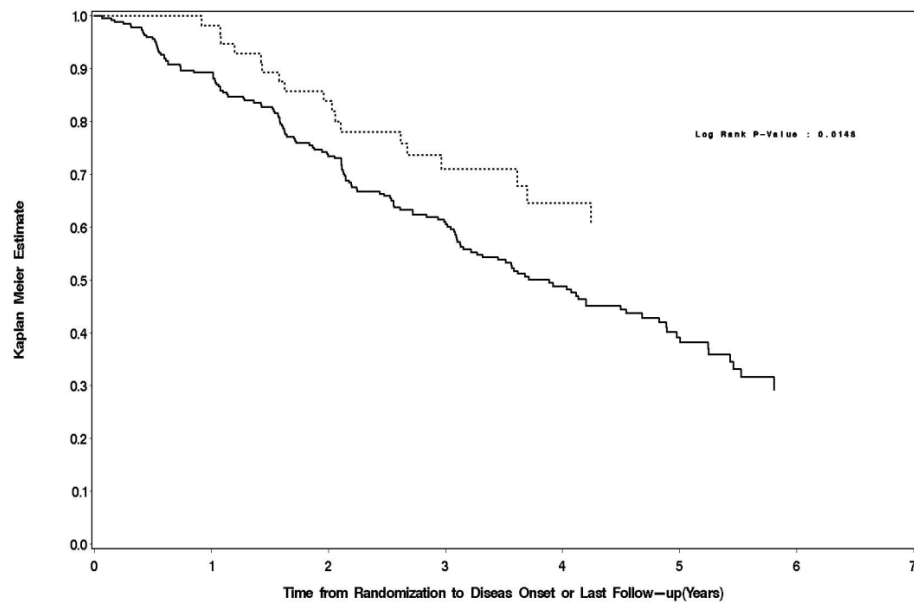


Figure 3.3: Kaplan-Meier Curves by the Level of IVGTT FPIR/HOAM-IR.  
 (solid line= FPIR/HOAM-IR < 49.22; broken line=FPIR/HOM-IR  $\geq 49.22$ );



# Chapter 4

## Concluding Remarks and Future Research

### 4.1 Summary and conclusions

In the first section, I have reviewed the definition of biomarker in health research field, explained some of the biological, pathological and physiological reasons why thorough evaluation is necessary, and discussed the common methods of evaluating the qualitative and/or quantitative biobiomarkers. Next, the receiver operating characteristics (ROC) curve, a particular statistical technique for evaluating the overall diagnosis/prognosis accuracy of continuous biomarkers was introduced. The basic probability theory that would be used to explain existing ROC statistics and underpinning the novel methods was presented in this section. Lastly, the literature on evaluation of repeated biobiomarker was reviewed. The non-parametric AUC statistics as a measure of diagnostic/prognostic accuracy for the repeated biobiomarker was described and how it was an improvement as it comparing to traditional statistics, which only measures association.

In Section 2, a proposed manuscript was presented regarding methodological development. In this manuscript, the non-parametric approach for comparison of AUCs proposed by Emir et al.

(2000) was modified to compare AUCs between the correlated groups. The optimization of weights using the Lagrange multiplier method was demonstrated. The asymptotic relative efficiencies were obtained. The optimal weight scheme achieved the substantial gain of efficiency, comparing to the two simple weighting schemes. The finite sample performance of the estimators was compared through simulation studies. The results were consistent with the theoretical results. The results also revealed the loss of the efficiency if we ignore the correlated group design and use the calculation procedure by Wu et al. (2011). It has been shown how this methodology can be applied on a matched case-control study. In conclusion, the efficiency using the new optimal weight scheme for comparing the non-parametric AUCs in the correlated studies increases as the between-group correlation increases, and/or within-subject correlation increases, and/or the incidence rate of disease decreases. Thus, it is recommended.

In Section 3, an application manuscript was presented. It demonstrated application of the non-parametric methods for AUC comparisons on a type 1 diabetes research. The research concluded that the combination of metabolic biomarkers, binary immunological biomarkers and/or genetic biomarkers can improve the modest predictive accuracies derived from solitary predictors. The results also indicated that autoantibody titers may not provide a cost/benefit improvement over OGTTs as a sole predictor, but they may be an important component of a composite modeled score among the subjects who have had the detectable autoimmunity but have not yet have the impaired glucose and/or first phase insulin secretion. It provided much needed guidance for clinicians and researchers for selecting subjects in future prevention trials in populations with underlying risk factors and clinically normal glucose tolerance by providing both the evidence for a preferred biobiomarker and a threshold for selection.

## 4.2 Future research

In the recent years, more and more correlated study designs have been used in the clinical studies due to the increased power. So developing the methodology to evaluate biomarker in the correlated group studies is important. Benefits include, but not limited to: (i) More elective diagnostic/prognostic measures for early detection of disease in a screening program; especially important when the disease are curable at only early stages. (ii) Derivation of optimal testing intervals for screening strategy plan. (iii) Earlier detection of disease occurrence while monitoring patients in a treatment trial. (iv) Improvement of the efficiency of designs of epidemiologic and clinical studies, by using the methodology to aid the identification of high risk patients for inclusion in the study population. (v) Improvement of the efficiency of comparative treatment trials by use of potential biomarker as surrogate response variables or intermediate end point variables.

The methodology presented in this dissertation focused on comparing the predictive accuracies of biomarker between two correlated groups in the one to one matched case-control study design. There are three main areas for future work. First, the methodology can be extended to other correlated study design, such as cross-over study design, N-of -1 study design. As well, the methodology can be further modified to the matched case-control studies where there is more than one matched control per case;

Secondly, an alternative summary statistic that is often used is the partial area under the ROC curve. In theory, the approach could be applied to this summary statistic although several issues need to be addressed. The partial AUC is rather sensitive to heterogeneity. Comparisons between groups are more difficult, especially if an empirical truncation process is used. Moreover,

the partial area lacks a useful symmetry property that the full AUC has. Although some work has been done addressing the comparison of partial ROC curves (Jiang et al., 1996; Dodd et al., 2003; Walter, 2005), the statistical methodology has been much less developed than for full AUCs under ROC curves.

Other directions of future research include extension of the proposed methodology to accommodate the multiple-comparison setting, a commonly used design in which the conditions of biomarker need to be evaluated using different modalities. Additionally, the methodology can be adapted to analyze the clustered data, such as when there are recurrent outcome events. Future research of these, based on the novel weighting optimization methods described here, may produce better solutions, particularly where small sample sizes are concerned.

# Bibliography

- [1] Vasan RS. Biobiomarkers of cardiovascular disease: molecular basis and practical considerations. *Circulation*. 2006, 113: 2335-2362.
- [2] NIH-FDA Conference: Biobiomarkers and Surrogate Endpoints: Advancing Clinical Research and Applications. Abstracts. *Disease Biomarkers*; 1998,14: 187-334.
- [3] Biobiomarkers Definitions Working Group. Biobiomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001; 69: 89-95.
- [4] Mayeux R. Biobiomarkers: Potential Uses and Limitations. *NeuroRx*. 2004; 1(2): 182-188.
- [5] Green DM, Swets JA. Signal detection theory and psychophysics. New York, NY, 1996; John Wiley and Sons Inc.
- [6] Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry*.1993; (8): 561-577.
- [7] Henderson AR. Assessing test accuracy and its clinical consequences: a primer for receiver operating characteristic curve analysis. *Ann Clin Biochem*. 1993; 30:521-539.
- [8] Smith RD. Evaluation of diagnostic tests. In: *Veterinary Clinical Epidemiology. A Problem-Oriented Approach*. CRC Press, Boca Raton, FL, 1995; 31-43.

- [9] Greiner M, Kumar S, Kyeswa C. Evaluation and comparison of antibody ELISAs for serodiagnosis of bovine trypanosomosis. *Vet. Parasitol.* 1997; 73, 197-205.
- [10] Spackman KA. Signal detection theory: Valuable tools for evaluating inductive learning. *Proceedings of the Sixth International Workshop on Machine Learning.* San Mateo, CA: 1989; 160-163.
- [11] Nockemann C, Heidt H, Thomsen N. Reliability in NDT: ROC study of radiographic weld inspections. *Nondestructive Testing and Evaluation International* 1991; 24:235-245.
- [12] Campbell G. Advances in statistical methodology for the evaluation of diagnostic and laboratory tests. *Stat. Med.* 1994; 13, 499-508.
- [13] Schafer H. Constructing a cut-off point for a quantitative diagnostic test. *Stat. Med.* 1989; 8: 1381-1391.
- [14] Perkins NJ, Schisterman EF. The Youden Index and Optimal Cut-point Corrected for Measurement Error. *Biometrical Journal.* 2005; 47:428441.
- [15] Metz CE. Basic principles of ROC analysis. *Semin Nucl Med.* 1978;8(4):283-98.
- [16] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic curve. *Radiology.* 1982; 143, 29-36.
- [17] Swets JA. Measuring the accuracy of diagnostic systems. *Science.* 1988; 240, 1285-1293.
- [18] Shapiro DE. The interpretation of diagnostic tests. *Statistical Methods in Medical Research.* 1999; 8:113-134.

- [19] Bamber DC. The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *Journal of Mathematical Psychology* 1975; 12:387-415.
- [20] Wieand HS, Gail MM, Hanley JA. A nonparametric procedure for comparing diagnostic tests with paired or unpaired data. *I.M.S. Bulletin* 1983; 12: 213-214.
- [21] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Area under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* 1988; 44(3): 837-845. 20.
- [22] Wieand HS, Gail M, James B, James K. A family of nonparametric statistics for comparing diagnostic biomarkers with paired or unpaired data. *Biometrika* 1989; 76: 585- 592. 21.
- [23] Metz CE, Wang P-L, Kronman HB. A new approach for testing the significance of differences between ROC curves measured from correlated data. *Information Processing in Medical Imaging VIII*, F. Deconick (ed.) 1984; 432-445.
- [24] Hoeffding W. A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics* 1948; 19(3): 293-325.
- [25] Sen PK. On some convergence properties of U-statistics. *Cal. Statistic Association Bulletin*.1960; 10: 1-18.
- [26] Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73:13-22.
- [27] Zeger SL, Liang KY. The analysis of discrete and continuous longitudinal data. *Biometrics*. 1986;42:121-30.

- [28] Cox DR. Regression Models and Life-Tables. Journal of the Royal Statistical Society. 1972; 34 (2): 187-220.
- [29] Breslow NE. Analysis of Survival Data under the Proportional Hazards Model. International Statistical Review / Revue Internationale de Statistique. 1975; 43 (1): 45-57.
- [30] DeGruttola V, TV XM. Modelling progression of CD-4 lymphocyte count and its relationship to survival time. Biometrics. 1994; 50: 1003-1014.
- [31] DeGruttola V, Wulfsohn M, Fischl M, and Tsiatis AA.. Modeling the relationship between survival and CD4 lymphocytes in patients with AIDS and AIDS-related complex. Journal of AIDS. 1993; 6: 359-365.
- [32] Tsiatis AA., DeGruttola V, Wulfsohn M S. Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS. Journal of the American Statistical Association. 1995; 90: 27-37.
- [33] Ghizzoni L, Bernasconi S, R. Viridis, A. Vottero, M. Ziveri, C. Volta, L. Lughetti, G. Giovannelli Dynamics of 24-hour pulsatile cortisol, 17-hydroxyprogesterone, and androstenedione release in prepubertal patients with nonclassic 21-hydroxylase deficiency and normal prepubertal children. Metabolism, 43 (1994), pp. 372-377.
- [34] O'Brien JT, Ames D, Schweitzer I, Mastwyk M, Colman P. Enhanced adrenal sensitivity to adrenocorticotrophic hormone (ACTH) is evidence of HPA axis hyperactivity in Alzheimer's disease. Psychol. 1996; Med, 26: 7-14.
- [35] Tucci JR, Tonino RP, Emkey RD, Peverly CA, Kher U, Stantora AC. Effect of three years of oral alendronate treatment in postmenopausal women with osteoporosis. Am. J. 1996; Med.,



101: 488-501.

- [36] DeLong ER, Vernon WB, Bollinger R R. Sensitivity and specificity of a monitoring test. *Biometrics*. 1985; 41: 947 - 958 .
- [37] Parker CB, DeLong ER. ROC methodology within a monitoring framework. *Statistics in Medicine*. 2003; 22: 3473-3488.
- [38] Emir B, Wieand S, Jung S, Ying Z. Comparison of diagnostic markers with repeated measurements: a non-parametric ROC curve approach. *Statistics in Medicine*. 2000; 19: 511-523.
- [39] Emir B, Wieand S, Su J, Cha S. Analysis of repeated markers used to predict progression of cancer. *Statistics in Medicine*. 1998; 17: 2563-2578.
- [40] Xu P, Wu YG, Zhu YL, Dagne G, Johnson G, Cuthbertson D, Krischer JP, Sosenko JM, Skyler JS. Prognostic performance of metabolic indexes in predicting onset of type 1 diabetes. *Diabetes Care*. 2010; 33(12):2508-13.
- [41] Jiang Y, Metz CE, Nishikawa RM. A receiver operating characteristic partial area index for highly sensitive diagnostic tests. *Radiology*. 1996; 201:745-750.
- [42] Dodd LE, Pepe MS. Partial AUC estimation and regression. *Biometrics* 2003; 59:614 -623.
- [43] Walter SD. The partial area under the summary ROC curve. *Stat Med*. 2005 Jul 15;24(13):2025-40.

## Appendices

## Appendix A: R/S-PLUS codes for comparing AUCs between correlated groups.

Following are the R/S-PLUS codes that were needed for comparing the prognostic accuracy of repeated biomarker between two correlated groups for the manuscript in Section Two.

```
##compute the optimal weights##
weight.fun<-function(x1mat,x2mat,y1vec,y2vec, x1deltamat,
x2deltamat,y1deltavec,y2deltavec)
{
n <- nrow(x1mat)
r1 <- ncol(x1mat)
r2 <- ncol(x2mat)
D1 <- sum(y1deltavec)
D2 <- sum(y2deltavec)
y1vec1 <- y1vec[y1deltavec == 1]
y2vec1 <- y2vec[y2deltavec == 1]
x1long <- rep(x1mat, D1)
x2long <- rep(x2mat, D2)
y1long <- rep(y1vec1, rep(n * r1, D1))
y2long <- rep(y2vec1, rep(n * r2, D2))
xyarray_1 <- array(ifelse(x1long - y1long<= 0,1, 0), dim = c(n, r1, D1))
xyarray_2 <- array(ifelse(x2long - y2long<= 0,1, 0), dim = c(n, r2, D2))
x1deltalong <- rep(x1deltamat, D1)
x2deltalong <- rep(x2deltamat, D2)
y1deltalong <- rep(rep(1, D1), rep(n * r1, D1))
y2deltalong <- rep(rep(1, D2), rep(n * r2, D2))
xydeltaarray_1 <- array(x1deltalong * y1deltalong, dim = c(n, r1, D1))
xydeltaarray_2<- array(x2deltalong * y2deltalong, dim = c(n, r2, D2))
xyarray1_1 <- array(ifelse(x1long - y1long> 0,1, 0), dim = c(n, r1, D1))
xyarray1_2 <- array(ifelse(x2long - y2long> 0,1, 0), dim = c(n, r2, D2))
m1.vec <- apply(x1deltamat, 1, sum)
m2.vec <- apply(x2deltamat, 1, sum)
w1.vec<-m1.vec/sum(m1.vec)
w2.vec<-m2.vec/sum(m2.vec)
wmarray_1 <- array(rep(rep(w1.vec/m1.vec, r1), D1), dim = c(n, r1, D1))
wmarray_2 <- array(rep(rep(w2.vec/m2.vec, r2), D2), dim = c(n, r2, D2))
u1mat <- apply(xyarray1_1, c(1, 2), sum)/sum(y1deltavec)
u2mat <- apply(xyarray1_2, c(1, 2), sum)/sum(y2deltavec)
v1vec <- apply(xyarray_1 * wmarray_1 * xydeltaarray_1, c(3), sum)
v2vec <- apply(xyarray_2 * wmarray_2 * xydeltaarray_2, c(3), sum)
v1vec1<-replace(y1deltavec,y1deltavec>0,v1vec)
v2vec1<-replace(y2deltavec,y2deltavec>0,v2vec)
u1deltamat<- x1deltamat
u2deltamat<- x2deltamat
v1deltavec<- y1deltavec
```

```

v2deltavec<- y2deltavec
temp<-cov(cbind(v1vec1*v1deltavec,v2vec1*v2deltavec))
rv1hat<-temp[1,1]
rv2hat<-temp[2,2]
rhov1v2hat<-temp[1,2]
p<-ncol(u1mat)
ru1u1hat<-matrix(rep(0,p^2),ncol=p)
rhoulu1hat<-matrix(rep(0,p^2),ncol=p)
  for(i in 1:p){
    for(j in 1:p){
      temp<-matrix(c(u1mat[u1deltamat[,i]*u1deltamat[,j]==1,i],
u1mat[u1deltamat[,i]*u1deltamat[,j]==1,j]),ncol=2)
      ru1u1hat[i,j]<-var(temp)[1,2]
      rhoulu1hat[i,j]<-var(temp)[1,2]/sqrt(var(temp)[1,1]*var(temp)[2,2])
    }
  }
p<-ncol(u2mat)
ru2u2hat<-matrix(rep(0,p^2),ncol=p)
rhoulu2hat<-matrix(rep(0,p^2),ncol=p)
for(i in 1:p){
  for(j in 1:p){
    temp<-matrix(c(u2mat[u2deltamat[,i]*u2deltamat[,j]==1,i],
u2mat[u2deltamat[,i]*u2deltamat[,j]==1,j]),ncol=2)
    ru2u2hat[i,j]<-var(temp)[1,2]
    rhoulu2hat[i,j]<-var(temp)[1,2]/sqrt(var(temp)[1,1]*var(temp)[2,2])
  }
}
p<-ncol(u1mat)
ru1u2hat<-matrix(rep(0,p^2),ncol=p)
rhoulu2hat<-matrix(rep(0,p^2),ncol=p)
for(i in 1:p){
  for(j in 1:p){
    temp<-matrix(c(u1mat[u1deltamat[,i]*u2deltamat[,j]==1,i],
u2mat[u1deltamat[,i]*u2deltamat[,j]==1,j]),ncol=2)
    ru1u2hat[i,j]<-var(temp)[1,2]
    rhoulu2hat[i,j]<-var(temp)[1,2]/sqrt(var(temp)[1,1]*var(temp)[2,2])
  }
}
ru1v1hat<-c(rep(0,p-1),0)
rhoulv1hat<-c(rep(0,p-1),0)
for(i in 1:(p-1))
{
temp<-
matrix(c(u1mat[u1deltamat[,i]*v1deltavec==1,i],v1vec1[u1deltamat[,i]*v1deltavec==1]),
ncol=2)
ru1v1hat[i]<-var(temp)[1,2]

```

```

rhou1v1hat[i]<-var(temp)[1,2]/sqrt(var(temp)[1,1]*var(temp)[2,2])
}
ru2v2hat<-c(rep(0,p-1),0)
rhou2v2hat<-c(rep(0,p-1),0)
for(i in 1:(p-1))
{
temp<-
matrix(c(u2mat[u2deltamat[,i]*v2deltavec==1,i],v2vec1[u2deltamat[,i]*v2deltavec==1]),
ncol=2)
ru2v2hat[i]<-var(temp)[1,2]
rhou2v2hat[i]<-var(temp)[1,2]/sqrt(var(temp)[1,1]*var(temp)[2,2])
}
ru1v2hat<-c(rep(0,p-1),0)
rhou1v2hat<-c(rep(0,p-1),0)
for(i in 1:(p-1))
{
temp<-
matrix(c(u1mat[u1deltamat[,i]*v2deltavec==1,i],v2vec1[u1deltamat[,i]*v2deltavec==1]),
ncol=2)
ru1v2hat[i]<-var(temp)[1,2]
rhou1v2hat[i]<-var(temp)[1,2]/sqrt(var(temp)[1,1]*var(temp)[2,2])
}
ru2v1hat<-c(rep(0,p-1),0)
rhou2v1hat<-c(rep(0,p-1),0)
for(i in 1:(p-1))
{
temp<-
matrix(c(u2mat[u2deltamat[,i]*v1deltavec==1,i],v1vec1[u2deltamat[,i]*v1deltavec==1]),
ncol=2)
ru2v1hat[i]<-var(temp)[1,2]
rhou2v1hat[i]<-var(temp)[1,2]/sqrt(var(temp)[1,1]*var(temp)[2,2])
}
a1.vec<-rep(0,length(m1.vec))
b1.vec<-rep(0,length(m1.vec))
for(i in 1:length(m1.vec))
{
a1.vec[i]<-sum(ru1u1hat[1:m1.vec[i],1:m1.vec[i]])
a1.vec[i]<-a1.vec[i]/m1.vec[i]^2
b1.vec[i] <-(v1deltavec[i]*sum(ru1v1hat[1:m1.vec[i]]))/(sum(v1deltavec)*m1.vec[i])-
(v2deltavec[i]*sum(ru1v2hat[1:m1.vec[i]]))/(sum(v2deltavec)*m1.vec[i])
}

a2.vec<-rep(0,length(m2.vec))
b2.vec<-rep(0,length(m2.vec))
for(i in 1:length(m2.vec)){
a2.vec[i]<-sum(ru2u2hat[1:m2.vec[i],1:m2.vec[i]])

```

```

a2.vec[i]<-a2.vec[i]/m2.vec[i]^2
b2.vec[i] <- (v2deltavec[i]*sum(ru2v2hat[1:m2.vec[i]]))/(sum(v2deltavec)*m2.vec[i])-
(v1deltavec[i]*sum(ru2v1hat[1:m2.vec[i]]))/(sum(v1deltavec)*m2.vec[i])
}
c1.vec<-rep(0,length(m1.vec))
a.vec<-rep(0,length(m1.vec))
  for(i in 1:length(m1.vec))
{
a.vec[i]<-sum(ru1u2hat[1:m1.vec[i],1:m2.vec[i]])
c1.vec[i]<-a.vec[i]/(m1.vec[i]*m2.vec[i])
}
c2.vec<-rep(0,length(m1.vec))
  for(i in 1:length(m1.vec))
{
a.vec[i]<-sum(ru1u2hat[1:m1.vec[i],1:m2.vec[i]])
c2.vec[i]<-a.vec[i]/(m1.vec[i]*m2.vec[i])
}
d1.vector<-(c1.vec)/(2*((a1.vec*a2.vec)-(c1.vec*c2.vec)))
e1.vector<-(a2.vec)/(2*((a1.vec*a2.vec)-(c1.vec*c2.vec)))
f1.vector<-((2*b2.vec*c1.vec+2*a2.vec*b1.vec)/(2*((a1.vec*a2.vec)-(c1.vec*c2.vec))))
d2.vector<-(a1.vec)/(2*((a1.vec*a2.vec)-(c1.vec*c2.vec)))
e2.vector<-(c2.vec)/(2*((a1.vec*a2.vec)-(c1.vec*c2.vec)))
f2.vector<-((2*b1.vec*c2.vec+2*a1.vec*b2.vec)/(2*((a1.vec*a2.vec)-(c1.vec*c2.vec))))
ramda<-((sum(f2.vector)*sum(d1.vector)-
sum(f1.vector)*sum(d2.vector)+sum(d2.vector)-
sum(d1.vector))/(sum(e1.vector)*sum(d2.vector)-sum(e2.vector)*sum(d1.vector)))
mu<-((sum(f2.vector)*sum(e1.vector)-sum(f1.vector)*sum(e2.vector)+sum(e2.vector)-
sum(e1.vector))/(sum(d1.vector)*sum(e2.vector)-sum(d2.vector)*sum(e1.vector)))
x1.vec<-ramda+2*b1.vec
x2.vec<-mu+2*b2.vec
weight.fun<-function(k){
  wmat <- matrix(NA,nrow=k,ncol=2)
  for (j in 1: k)
  {
A<-matrix(c(2*a1.vec[j],-2*c2.vec[j],-2*c1.vec[j],2*a2.vec[j]),nrow=2)
x<-c(x1.vec[j],x2.vec[j])
  wmat[j,]<-rbind(solve(A,x))
}
wmat<-matrix(rbind(wmat[1:k,]),k,2)
w1.vec<-wmat[,1]
w2.vec<-wmat[,2]
list(w1.vec=w1.vec,w2.vec=w2.vec)
}
wmat <- matrix(NA,nrow=n,ncol=2)
temp<-weight.fun(n)
opw1.vec<-temp[[1]]

```

```

opw1.vec[which(opw1.vec[1:n]>=1)]<-NA
opw2.vec<-temp[[2]]
opw2.vec[which(opw2.vec[1:n]>=1)]<-NA
sum(w1.vec)
sum(w2.vec)
opsigmahat<-sum(a1.vec*opw1.vec^2-
2*b1.vec*opw1.vec)+sum(a2.vec*opw2.vec^2-2*b2.vec*opw2.vec)-
sum(2*c1.vec*opw1.vec*opw2.vec)+var(v1 vec1[v1deltavec==1])/sum(v1deltavec)+var(
v2vec1[v2deltavec==1])/sum(v2deltavec)-2*rhov1v2hat/(D1*D2)
w1.vec<- m1.vec/sum(m1.vec)
w2.vec<- m2.vec/sum(m2.vec)
simp1sigmahat<-sum(a1.vec*w1.vec^2-2*b1.vec*w1.vec)+sum(a2.vec*w2.vec^2-
2*b2.vec*w2.vec)-
sum(2*c1.vec*w1.vec*w2.vec)+var(v1 vec1[v1deltavec==1])/sum(v1deltavec)+var(v2ve
c1[v2deltavec==1])/sum(v2deltavec)-2*rhov1v2hat/(D1*D2)
w1.vec<- rep(1/length(m1.vec), length(m1.vec))
w2.vec<- rep(1/length(m2.vec), length(m2.vec))
simp2sigmahat<-sum(a1.vec*w1.vec^2-2*b1.vec*w1.vec)+sum(a2.vec*w2.vec^2-
2*b2.vec*w2.vec)-
sum(2*c1.vec*w1.vec*w2.vec)+var(v1 vec1[v1deltavec==1])/sum(v1deltavec)+var(v2ve
c1[v2deltavec==1])/sum(v2deltavec)-2*rhov1v2hat/(D1*D2)
list ( w1.vec=opw1.vec,w2.vec=opw2.vec,
sigmau1sqhat=ru1u1hat,sigmau2sqhat=ru2u2hat,
sigmav1sqhat=rv1hat,sigmav2sqhat=rv2hat,
rhou1u1hat=rhou1u1hat,rhou2u2hat=rhou2u2hat,
rhou1v1hat=rhou1v1hat,rhou2v2hat=rhou2v2hat, rhou1u2hat=rhou1u2hat,
rhov1v2hat=rhov1v2hat,rhou1v2hat=rhou1v2hat,rhou2v1hat=rhou2v1hat,
a1.vec=a1.vec,a2.vec=a2.vec,b1.vec=b1.vec,b2.vec=b2.vec,c1.vec=c1.vec,c2.vec=c2.vec
, opsigmahat= opsigmahat, smp1sigmahat=simp1sigmahat,
simp2sigmahat=simp2sigmahat )

```

```

transform.fun<-function(xmat, yvec, xdeltamat, ydeltavec)
{
  mvec <- apply(xdeltamat, 1, sum)
  n <- nrow(xmat)
  r <- ncol(xmat)
  D <- sum(ydeltavec)
  yvec1 <- yvec[ydeltavec == 1]
  xlong <- rep(xmat, D)
  ylong <- rep(yvec1, rep(n * r, D))
  xyarray <- array(ifelse(xlong - ylong <= 0, 1, 0), dim = c(n, r, D))
  xdeltalong <- rep(xdeltamat, D)
  ydeltalong <- rep(rep(1, D), rep(n * r, D))
  xydeltaarray <- array(xdeltalong * ydeltalong, dim = c(n, r, D))
  mvec <- apply(xdeltamat, 1, sum)
  wvec <- mvec/sum(mvec)

```

```

wmmarray <- array(rep(rep(wvec/mvec, r), D), dim = c(n, r, D))
xyarray1 <- array(ifelse(xlong - ylong > 0, 1, 0), dim = c(n, r, D))

umat <- apply(xyarray1, c(1, 2), sum)/sum(ydeltavec)
vvec <- apply(xyarray * wmmarray * xydeltaarray, c(3), sum)
vvec1 <- replace(ydeltavec, ydeltavec > 0, vvec)

list(umat=umat, vvec=vvec1, udeltamat=xdeltamat, vdeltavec=ydeltavec)
}
#compute wu's weights
weight.fun1 <- function(xmat, yvec, xdeltamat, ydeltavec)
{
  temp <- transform.fun(xmat, yvec, xdeltamat, ydeltavec)
  umat <- temp[[1]]
  vvec <- temp[[2]]
  udeltamat <- temp[[3]]
  vdeltavec <- temp[[4]]
  p <- ncol(umat)
  ruuhat <- matrix(rep(0, p^2), ncol=p)
  rhouhat <- matrix(rep(0, p^2), ncol=p)
  for(i in 1:p){
    for(j in 1:p){
      temp <- matrix(c(umat[udeltamat[,i]*udeltamat[,j]==1,i],
        umat[udeltamat[,i]*udeltamat[,j]==1,j]), ncol=2)
      ruuhat[i,j] <- var(temp)[1,2]
      rhouhat[i,j] <- var(temp)[1,2]/sqrt(var(temp)[1,1]*var(temp)[2,2])
    }
  }
  ruvhat <- c(rep(0, p-1), 0)
  rhouvhat <- c(rep(0, p-1), 0)
  for(i in 1:(p-1)){
    temp <-
      matrix(c(umat[udeltamat[,i]*vdeltavec==1,i], vvec[udeltamat[,i]*vdeltavec==1]), ncol=2)
    ruvhat[i] <- var(temp)[1,2]
    rhouvhat[i] <- var(temp)[1,2]/sqrt(var(temp)[1,1]*var(temp)[2,2])
  }
  mvec <- apply(udeltamat, 1, sum)
  avec <- rep(0, length(mvec))
  bvec <- rep(0, length(mvec))
  for(i in 1:length(mvec)){
    avec[i] <- sum(ruuhat[1:mvec[i], 1:mvec[i]])
    avec[i] <- avec[i]/mvec[i]^2

    bvec[i] <- vdeltavec[i]*sum(ruvhat[1:mvec[i]])/(sum(vdeltavec)*mvec[i])
  }
  temp <- bvec/avec

```



```

c <- (1 - sum(temp))/sum(avec^(-1))
opwvec <- c * avec^(-1) + temp
opsigmahat<-sum(avec*opwvec^2-
2*bvec*opwvec)+var(vvec[vdeltavec==1])/sum(vdeltavec)

simp1wvec <- mvec/sum(mvec)
simp1sigmahat<-sum(avec*simp1wvec^2-
2*bvec*simp1wvec)+var(vvec[vdeltavec==1])/sum(vdeltavec)
simp2wvec <- rep(1/length(mvec), length(mvec))
simp2sigmahat<-sum(avec*simp2wvec^2-
2*bvec*simp2wvec)+var(vvec[vdeltavec==1])/sum(vdeltavec)
list(opwvec=opwvec,simp1sigmahat=simp1sigmahat,opsigmahat=opsigmahat,simp2sigmahat=simp2sigmahat,ruuhat=ruuhat,ruvhat=ruvhat,rhouhat=rhouhat,rhouvhat=rhouvhat)
)
### compute theta(AUC)#####

auc.fun<-function(x1mat,x2mat,y1vec,y2vec,
x1deltamat,x2deltamat,y1deltavec,y2deltavec)
{
n <- nrow(x1mat)
r1 <- ncol(x1mat)
r2 <- ncol(x2mat)
D1 <- sum(y1deltavec)
D2 <- sum(y2deltavec)
m1.vec <- apply(x1deltamat, 1, sum)
m2.vec <- apply(x2deltamat, 1, sum)
y1vec1 <- y1vec[y1deltavec == 1]
y2vec1 <- y2vec[y2deltavec == 1]
x1long <- rep(x1mat, D1)
x2long <- rep(x2mat, D2)
y1long <- rep(y1vec1, rep(n * r1, D1))
y2long <- rep(y2vec1, rep(n * r2, D2))
xyarray_1 <- array(ifelse(x1long - y1long<= 0,1, 0), dim = c(n, r1, D1))
xyarray_2 <- array(ifelse(x2long - y2long<= 0,1, 0), dim = c(n, r2, D2))
x1deltalong <- rep(x1deltamat, D1)
x2deltalong <- rep(x2deltamat, D2)
y1deltalong <- rep(rep(1, D1), rep(n * r1, D1))
y2deltalong <- rep(rep(1, D2), rep(n * r2, D2))
xydeltaarray_1 <- array(x1deltalong * y1deltalong, dim = c(n, r1, D1))
xydeltaarray_2<- array(x2deltalong * y2deltalong, dim = c(n, r2, D2))
#weight scheme 1:the number of the visits
w1.vec <- m1.vec/sum(m1.vec)
w2.vec <- m2.vec/sum(m2.vec)
wmarray_1 <- array(rep(rep(w1.vec/m1.vec, r1), D1), dim = c(n, r1, D1))
wmarray_2 <- array(rep(rep(w2.vec/m2.vec, r2), D2), dim = c(n, r2, D2))
v1vec <- apply(xyarray_1 * wmarray_1 * xydeltaarray_1, c(3), sum)

```

```

v2vec <- apply(xyarray_2 * wmarray_2 * xydeltaarray_2, c(3), sum)
auc1_1 <- sum(v1vec)/D1
auc1_2 <- sum(v2vec)/D2
#weight scheme 2: the number of the subjects
w1.vec <- rep(1/n, n)
w2.vec <- rep(1/n, n)
wmarray_1 <- array(rep(rep(w1.vec/m1.vec, r1), D1), dim = c(n, r1, D1))
wmarray_2 <- array(rep(rep(w2.vec/m2.vec, r2), D2), dim = c(n, r2, D2))
v1vec <- apply(xyarray_1 * wmarray_1 * xydeltaarray_1, c(3), sum)
v2vec <- apply(xyarray_2 * wmarray_2 * xydeltaarray_2, c(3), sum)
auc2_1 <- sum(v1vec)/D1
auc2_2 <- sum(v2vec)/D2

temp1 <- weight.fun1(x1mat,y1vec, x1deltamat,y1deltavec)
temp2 <- weight.fun1(x2mat,y2vec, x2deltamat,y2deltavec)
w1.vec<-temp1[[1]]
w2.vec<-temp2[[1]]
wmarray_1 <- array(rep(rep(temp1[[1]]/m1.vec, r1), D1), dim = c(n, r1, D1))
wmarray_2 <- array(rep(rep(temp2[[1]]/m2.vec, r2), D2), dim = c(n, r2, D2))
v1vec <- apply(xyarray_1 * wmarray_1 * xydeltaarray_1, c(3), sum)
v2vec <- apply(xyarray_2 * wmarray_2 * xydeltaarray_2, c(3), sum)
auc3_1 <- sum(v1vec)/D1
auc3_2 <- sum(v2vec)/D2
temp <- weight.fun(x1mat,x2mat,y1vec,y2vec, x1deltamat,
x2deltamat,y1deltavec,y2deltavec)
w1.vec<-temp[[1]]
w2.vec<-temp[[2]]
wmarray_1 <- array(rep(rep(temp[[1]]/m1.vec, r1), D1), dim = c(n, r1, D1))
wmarray_2 <- array(rep(rep(temp[[2]]/m2.vec, r2), D2), dim = c(n, r2, D2))
v1vec <- apply(xyarray_1 * wmarray_1 * xydeltaarray_1, c(3), sum)
v2vec <- apply(xyarray_2 * wmarray_2 * xydeltaarray_2, c(3), sum)
auc4_1 <- sum(v1vec)/D1
auc4_2 <- sum(v2vec)/D2
aucdiff1<-auc1_1-auc1_2
aucdiff2<-auc2_1-auc2_2
aucdiff3<-auc3_1-auc3_2
aucdiff4<-auc4_1-auc4_2
aucvec<-c(auc1_1,auc1_2,auc2_1,auc2_2,auc3_1,auc3_2,auc4_1,auc4_2)
aucdiff<-c(aucdiff1, aucdiff2, aucdiff3,aucdiff4)
list(aucdiff=aucdiff)
}

var.fun<-function(x1mat,x2mat,y1vec,y2vec,
x1deltamat,x2deltamat,y1deltavec,y2deltavec)
{

```

```

temp <- weight.fun(x1mat,x2mat,y1vec,y2vec, x1deltamat,
x2deltamat,y1deltavec,y2deltavec)
sigma1sqhat <- temp[[3]]
sigma2sqhat <- temp[[4]]
sigma1sqhat <- temp[[5]]
sigma2sqhat <- temp[[6]]
rho1u1hat <- temp[[7]]
rho2u2hat <- temp[[8]]
rho1v1hat <- temp[[9]]
rho2v2hat <- temp[[10]]
rho1u2hat <- temp[[11]]
rho1v2hat <- temp[[12]]
rho1v2hat <- temp[[13]]
rho2v1hat <- temp[[14]]
a1.vec <- temp[[15]]
a2.vec <- temp[[16]]
b1.vec <- temp[[17]]
b2.vec <- temp[[18]]
c1.vec <- temp[[19]]
c2.vec <- temp[[20]]
opw1.vec<-temp[[1]]
opw2.vec<-temp[[2]]
opsigmahat<-sum(a1.vec*opw1.vec^2-2*b1.vec*opw1.vec)+sum(a2.vec*opw2.vec^2-
2*b2.vec*opw2.vec)-
sum(2*c1.vec*opw1.vec*opw2.vec)+var(v1vec1[v1deltavec==1])/sum(v1deltavec)+var(
v2vec1[v2deltavec==1])/sum(v2deltavec)-2*rhov1v2hat/(D1*D2)
w1.vec<- m1.vec/sum(m1.vec)
w2.vec<- m2.vec/sum(m2.vec)
simp1sigmahat<-sum(a1.vec*w1.vec^2-2*b1.vec*w1.vec)+sum(a2.vec*w2.vec^2-
2*b2.vec*w2.vec)-
sum(2*c1.vec*w1.vec*w2.vec)+var(v1vec1[v1deltavec==1])/sum(v1deltavec)+var(v2ve
c1[v2deltavec==1])/sum(v2deltavec)-2*rhov1v2hat/(D1*D2)
w1.vec<- rep(1/length(m1.vec), length(m1.vec))
w2.vec<- rep(1/length(m2.vec), length(m2.vec))
simp2sigmahat<-sum(a1.vec*w1.vec^2-2*b1.vec*w1.vec)+sum(a2.vec*w2.vec^2-
2*b2.vec*w2.vec)-
sum(2*c1.vec*w1.vec*w2.vec)+var(v1vec1[v1deltavec==1])/sum(v1deltavec)+var(v2ve
c1[v2deltavec==1])/sum(v2deltavec)-2*rhov1v2hat/(D1*D2)
temp <- weight.fun1(x1mat,y1vec,x1deltamat, y1deltavec)
w1.vec<-temp[[1]]
temp <- weight.fun1(x2mat,y2vec,x2deltamat, y2deltavec)
w2.vec<-temp[[1]]
simp3sigmahat<-sum(a1.vec*w1.vec^2-2*b1.vec*w1.vec)+sum(a2.vec*w2.vec^2-
2*b2.vec*w2.vec)-
sum(2*c1.vec*w1.vec*w2.vec)+var(v1vec1[v1deltavec==1])/sum(v1deltavec)+var(v2ve
c1[v2deltavec==1])/sum(v2deltavec)-2*rhov1v2hat/(D1*D2)

```

```

var<-c( simp1sigmahat, simp2sigmahat, simp3sigmahat,opsigmahat)
list( var=var)
}
##data simulation
library(MASS)
sigma.fun<-function(sigmaxsq,rho,lambda)
{
  rxx<- matrix( c(
    sigmaxsq,rho,  rho,  rho,  rho,  rho,  lambda, lambda, lambda, lambda,
lambda, lambda,
    rho,  sigmaxsq,rho,  rho,  rho,  rho,  lambda, lambda, lambda, lambda,
lambda, lambda,
    rho,  rho,  sigmaxsq,rho,  rho,  rho,  lambda, lambda, lambda, lambda,
lambda, lambda,
    rho,  rho,  rho,  sigmaxsq,rho,  rho,  lambda, lambda, lambda, lambda,
lambda, lambda,
    rho,  rho,  rho,  rho, sigmaxsq, rho,  lambda, lambda, lambda, lambda,
lambda, lambda,
    rho,  rho,  rho,  rho,  rho,  sigmaxsq, lambda, lambda, lambda, lambda,
lambda, lambda,
    lambda, lambda, lambda, lambda, lambda, lambda, sigmaxsq, rho,  rho,  rho,
rho,  rho,
    lambda, lambda, lambda, lambda, lambda, lambda,  rho,  sigmaxsq, rho,  rho,
rho,  rho,
    lambda, lambda, lambda, lambda, lambda, lambda,  rho,  rho,  sigmaxsq, rho,
rho,  rho,
    lambda, lambda, lambda, lambda, lambda, lambda,  rho,  rho,  rho,
sigmaxsq,rho,  rho,
    lambda, lambda, lambda, lambda, lambda, lambda,  rho,  rho,  rho,  rho,
sigmaxsq,rho,
    lambda, lambda, lambda, lambda, lambda, lambda,  rho,  rho,  rho,  rho,
rho,  sigmaxsq
  ),
  byrow=T,ncol=12)
}

data.fun<-function(n,rate1,rate2,sigmaxsq,rho,lambda)
{
  sigma<-sigma.fun(sigmaxsq,rho, lambda)
  mu<-c(0,0,0,0,0,0,0,0,0,0,0,0)
  lis <- lapply(1:1, mvrnorm,n=n,mu=mu,Sigma=sigma)
  temp<-lis[[1]]
  x1mat<-temp[,1:6]
  x2mat<-temp[,7:12]
  temp1.dat<-rexp(n,rate=rate1)
  y1deltavec<-ifelse(temp1.dat<5,1,0)
}

```

```

psi1<-sum(y1deltavec)/length(y1deltavec)
x1deltamat<-matrix(rep(1,6*n),ncol=6)
x1deltamat[,1]<-rep(1,n)
x1deltamat[,2]<-ifelse(temp1.dat<=1,0,1)
x1deltamat[,3]<-ifelse(temp1.dat<=2,0,1)
x1deltamat[,4]<-ifelse(temp1.dat<=3,0,1)
x1deltamat[,5]<-ifelse(temp1.dat<=4,0,1)
x1deltamat[,6]<-ifelse(temp1.dat<=5,0,1)
x1deltamat<-matrix(c(x1deltamat[,1:6]),ncol=6)
y1vec<-rep(0,n)
for ( i in 1:n) {
if (temp1.dat[i]<=1){y1vec[i]=x1mat[i,2]+1 } else
if (temp1.dat[i]<=2){y1vec[i]=x1mat[i,3]+1 } else
if (temp1.dat[i]<=3){y1vec[i]=x1mat[i,4]+1 } else
if (temp1.dat[i]<=4){y1vec[i]=x1mat[i,5]+1 } else
if (temp1.dat[i]<=5){y1vec[i]=x1mat[i,6]+1 } else
if (temp1.dat[i]>5) {y1vec[i]=x1mat[i,6]+1 }
}
temp2.dat<-rexp(n,rate=rate2)
y2deltavec<-ifelse(temp2.dat<5,1,0)
psi2<-sum(y2deltavec)/length(y2deltavec)
x2deltamat<-matrix(rep(1,6*n),ncol=6)
x2deltamat[,1]<-rep(1,n)
x2deltamat[,2]<-ifelse(temp2.dat<=1,0,1)
x2deltamat[,3]<-ifelse(temp2.dat<=2,0,1)
x2deltamat[,4]<-ifelse(temp2.dat<=3,0,1)
x2deltamat[,5]<-ifelse(temp2.dat<=4,0,1)
x2deltamat[,6]<-ifelse(temp2.dat<=5,0,1)
x2deltamat<-matrix(c(x2deltamat[,1:6]),ncol=6)
y2vec<-rep(0,n)
for ( i in 1:n) {
if (temp2.dat[i]<=1){y2vec[i]=x2mat[i,2]+1 } else
if (temp2.dat[i]<=2){y2vec[i]=x2mat[i,3]+1 } else
if (temp2.dat[i]<=3){y2vec[i]=x2mat[i,4]+1 } else
if (temp2.dat[i]<=4){y2vec[i]=x2mat[i,5]+1 } else
if (temp2.dat[i]<=5){y2vec[i]=x2mat[i,6]+1 } else
##if (temp2.dat[i]>5) {y2vec[i]=0}##
if (temp2.dat[i]>5) {y2vec[i]=x2mat[i,6]+1 }
}
list(x1mat=x1mat,x2mat=x2mat,y1vec=y1vec,y2vec=y2vec,
x1deltamat=x1deltamat,x2deltamat=x2deltamat,y1deltavec=y1deltavec,
y2deltavec=y2deltavec,psi1=psi1,psi2=psi2)
}
sim.fun<-function(repl,n,rate1,rate2,sigmaxsq,rho,lambda)
{
auc<-matrix(rep(0,repl*4),ncol=4)

```

```

var<-matrix(rep(0,repl*4),ncol=4)
out<-rep(0,8)
i=0
for(i in 1:repl)
{
  sim.dat<-data.fun(n,rate1,rate2,sigmaxsq,rho,lambda)
  x1mat<-sim.dat[[1]]
  x2mat<-sim.dat[[2]]
  y1vec<-sim.dat[[3]]
  y2vec<-sim.dat[[4]]
  x1deltamat<- sim.dat[[5]]
  x2deltamat<- sim.dat[[6]]
  y1deltavec<-sim.dat[[7]]
  y2deltavec<-sim.dat[[8]]
  temp<-auc.fun(x1mat,x2mat,y1vec,y2vec, x1deltamat,
x2deltamat,y1deltavec,y2deltavec)
  auc[i,]<-temp[[1]]
}
mean<-apply(auc,2,mean, na.rm=T)[1:4]
var<-apply(auc,2,var, na.rm=T)[1:4]
truediff<-c(0,0,0,0)
bias<-mean-truediff
mse<-var+bias^2
out<-c(bias,mse)
print(out)
}
sim50<-sim.fun(2000,50,.05,0.10,1,0.3,0.1)
sim50<-sim.fun(2000,50,.05,0.10,1,0.5,0.1)
sim50<-sim.fun(2000,50,.05,0.10,1,0.7,0.1)
sim50<-sim.fun(2000,50,.05,0.10,1,0.9,0.1)
sim50<-sim.fun(2000,50,.05,0.10,1,0.3,0.3)
sim50<-sim.fun(2000,50,.05,0.10,1,0.5,0.3)
sim50<-sim.fun(2000,50,.05,0.10,1,0.7,0.3)
sim50<-sim.fun(2000,50,.05,0.10,1,0.9,0.3)
sim50<-sim.fun(2000,50,.05,0.10,1,0.3,0.5)
sim50<-sim.fun(2000,50,.05,0.10,1,0.5,0.5)
sim50<-sim.fun(2000,50,.05,0.10,1,0.7,0.5)
sim50<-sim.fun(2000,50,.05,0.10,1,0.9,0.5)

sim50<-sim.fun(2000,50,.05,0.10,1,0.3,0.7)
sim50<-sim.fun(2000,50,.05,0.10,1,0.5,0.7)
sim50<-sim.fun(2000,50,.05,0.10,1,0.7,0.7)
sim50<-sim.fun(2000,50,.05,0.10,1,0.9,0.7)
sim50<-sim.fun(2000,50,.10,0.10,1,0.3,0.1)
sim50<-sim.fun(2000,50,.10,0.10,1,0.5,0.1)

```

```

sim50<-sim.fun(2000,50,.10,0.10,1,0.7,0.1)
sim50<-sim.fun(2000,50,.10,0.10,1,0.9,0.1)
sim50<-sim.fun(2000,50,.10,0.10,1,0.3,0.3)
sim50<-sim.fun(2000,50,.10,0.10,1,0.5,0.3)
sim50<-sim.fun(2000,50,.10,0.10,1,0.7,0.3)
sim50<-sim.fun(2000,50,.10,0.10,1,0.9,0.3)
sim50<-sim.fun(2000,50,.10,0.10,1,0.3,0.5)
sim50<-sim.fun(2000,50,.10,0.10,1,0.5,0.5)
sim50<-sim.fun(2000,50,.10,0.10,1,0.7,0.5)
sim50<-sim.fun(2000,50,.10,0.10,1,0.9,0.5)
sim50<-sim.fun(2000,50,.10,0.10,1,0.3,0.7)
sim50<-sim.fun(2000,50,.10,0.10,1,0.5,0.7)
sim50<-sim.fun(2000,50,.10,0.10,1,0.7,0.7)
sim50<-sim.fun(2000,50,.10,0.10,1,0.9,0.7)
sim50<-sim.fun(2000,50,.10,0.15,1,0.3,0.1)
sim50<-sim.fun(2000,50,.10,0.15,1,0.5,0.1)
sim50<-sim.fun(2000,50,.10,0.15,1,0.7,0.1)
sim50<-sim.fun(2000,50,.10,0.15,1,0.9,0.1)
sim50<-sim.fun(2000,50,.10,0.15,1,0.3,0.3)
sim50<-sim.fun(2000,50,.10,0.15,1,0.5,0.3)
sim50<-sim.fun(2000,50,.10,0.15,1,0.7,0.3)
sim50<-sim.fun(2000,50,.10,0.15,1,0.9,0.3)
sim50<-sim.fun(2000,50,.10,0.15,1,0.3,0.5)
sim50<-sim.fun(2000,50,.10,0.15,1,0.5,0.5)
sim50<-sim.fun(2000,50,.10,0.15,1,0.7,0.5)
sim50<-sim.fun(2000,50,.10,0.15,1,0.9,0.5)
sim50<-sim.fun(2000,50,.10,0.15,1,0.3,0.7)
sim50<-sim.fun(2000,50,.10,0.15,1,0.5,0.7)
sim50<-sim.fun(2000,50,.10,0.15,1,0.7,0.7)
sim50<-sim.fun(2000,50,.10,0.15,1,0.9,0.7)

```

```

####read the thyroid data into R ####
library(foreign)
case<- read.xport("M:/PHD/dissertation/manuscript 7--compare correlated
group/case_tpo.xport")
control<- read.xport("M:/PHD/dissertation/manuscript 7--compare correlated
group/control_tpo.xport")

detach("package:foreign")
names(case) #see the variable names
names(case)<-tolower(names(case))
names(control) #see the variable names
names(control)<-tolower(names(control))
x1mat<-as.matrix(case[,2:4])
x1deltamat<- as.matrix(case[,5:7])

```

```

y1deltavec<-as.vector(case[,8])
y1vec<-as.vector(case[,9])
x2mat<-as.matrix(control[,2:4])
x2deltamat<- as.matrix(control[,5:7])
y2deltavec<-as.vector(control[,8])
y2vec<-as.vector(control[,9])
Auc<-auc.fun(x1mat,x2mat,y1vec,y2vec, x1deltamat, x2deltamat,y1deltavec,y2deltavec)
var<-var.fun(x1mat,x2mat,y1vec,y2vec, x1deltamat, x2deltamat,y1deltavec,y2deltavec)
weight.fun(x1mat,x2mat,y1vec,y2vec, x1deltamat, x2deltamat,y1deltavec,y2deltavec)

### 95% confidence interval###
theta.ci<-function(x1mat,x2mat,y1vec,y2vec, x1deltamat,
x2deltamat,y1deltavec,y2deltavec)
{
  n<-nrow(x1mat)
  auc<-auc.fun(x1mat,x2mat,y1vec,y2vec, x1deltamat, x2deltamat,y1deltavec,y2deltavec)
  var<-var.fun(x1mat,x2mat,y1vec,y2vec, x1deltamat, x2deltamat,y1deltavec,y2deltavec)
  cileftvec<-auc[[1]]-1.96*sqrt(var[[1]]/n)
  cirightvc<-auc[[1]]+1.96*sqrt(var[[1]]/n)
  ci<-c(cileftvec,cirightvc)
  print(ci)
}
ci<-theta.ci(x1mat,x2mat,y1vec,y2vec, x1deltamat, x2deltamat,y1deltavec,y2deltavec)

#####optimum cutoff,sensitivity, specificity, weight: 1/n#####
sen.spe.fun<-function(xmat, xdeltamat, yvec, ydeltavec)
{
  n <- nrow(xmat)
  r <- ncol(xmat)
  D <- sum(ydeltavec)
  yvec1 <- yvec[ydeltavec == 1]
  x<-as.vector(xmat*xdeltamat)[xmat*xdeltamat>0]
  y<-as.vector(yvec*ydeltavec)[yvec*ydeltavec>0]
  xy<-c(x,y)
  c<-unique(rev(sort(xy)))
  temp1<-t(matrix(rep(yvec1,length(c)), nrow=D))
  temp2<-matrix(rep(c,D), ncol=D)
  temparray<- array(ifelse(temp1- temp2 <= 0,1,0), dim = c(length(c),D))
  #use above for FPIR and ratio;
  TPF<-apply(temparray,1, sum)/D
  temp3<-t(matrix(rep(x,length(c)), nrow=length(x)))
  temp4<-matrix(rep(c,length(x)), ncol=length(x))
  temparray2<- array(ifelse(temp3- temp4 <= 0,1,0), dim = c(length(c),length(x)))
  #use above for FPIR and ratio;
  FPF <-apply(temparray2, 1, sum)/length(x)
  sen<-TPF

```



```

spe<-1-FPF
sum<-sen+spe
data<-matrix(c(c,sen,spe,sum),ncol=4)
cutoff<-data[sum==max(sum)]
print(cutoff)
library(ggplot2)
plot(FPF,TPF,type="l", col="blue", xlab="1-Specificity", ylab="Sensitivity")
plot.data<-matrix(c(TPF, FPF),ncol=2)
list(data.plot=plot.data)
}
#####plot roc curves#####
install.packages("ggplot2")
sen.spe_case<-sen.spe.fun(x1mat, x1deltamat, y1vec, y1deltavec)
case.plot<-sen.spe_case$data.plot
sen.spe_control<-sen.spe.fun(x2mat, x2deltamat, y2vec, y2deltavec)
control.plot<-sen.spe_control$data.plot
library(ggplot2)
h_range <- range(0, case.plot[,1], control.plot[,1])
v_range <- range(0, case.plot[,2], control.plot[,2])
plot(case.plot[,2],case.plot[,1],type="l", col="blue", ylim=h_range,xlim=v_range,
xlab="1-Specificity", ylab="Sensitivity")
title(main="ROC Curves For TSH (Cases VS Controls)", font.main=4)
lines(control.plot[,2],control.plot[,1],lty=3,col="red" )
legend(locator(1), c("Breast Cancer Group","Control Group"), cex=0.8,
col=c("blue","red"), lty=c(1,3))

```

## Appendix B: SAS codes for comparing the correlated AUCs.

Following are the SAS codes that were needed for comparing the prognostic accuracy of different biomarkers for the manuscript in Section Three.

```
%include "roc.sas";
%include "rocplot.sas";
%include " cutoff.sas";

libname ROC 'C:\Documents and Settings\Xup\Desktop\PHD\disseration\manuscript 1-
baseline ROC comparison';
data roc_low(keep=maskid ins_trt risktype age bmi gender age_at_draw baseica baseiaa
baseica512_res basegad_res basemiaa_res numabs basefastins basefastgluc basehomar
basefpir baseratio base_ogtt_fasting base_ogtt_twohour base_peak_c_pep
base_auc_c_pep basehb iddm race randage iddmtime);
    set roc.low_ivgtt;
base_auc_c_pep=10*(base_ogtt_pep_0+(4*base_ogtt_pep_30)+(2*base_ogtt_pep_60)+(
4*base_ogtt_pep_90)+(base_ogtt_pep_120));
    age=age_at_draw/365.25;
    run;
proc sort data=roc_low;
    by maskid age_at_draw;
    run;
data roc;
    set roc_low;
    by maskid;
    if first.maskid;
    run;
proc means data=roc n mean std;
    class iddm;
    var age basefastgluc basefastins basefpir basehomar baseratio base_ogtt_fasting
base_ogtt_twohour base_peak_c_pep base_auc_c_pep basehb;
    run;
proc ttest data=roc;
    class iddm;
    var age basefastgluc basefastins basefpir basehomar baseratio base_ogtt_fasting
base_ogtt_twohour base_peak_c_pep base_auc_c_pep basehb;
    run;
proc logistic data=roc;
    model iddm(event='1') = baseica / outroc=or roceps=0;
    output out=icapred p=pica;
    ods output association=assoc;
    run;
data _null_;
    set assoc;
    if label2='c' then call symput("area",cvalue2);
    run;
```

```

title2 "Approximate area under curve = &area";
%rocplot(out=icapred, outroc=or, p=pica, id=ica)
title2;

data joint;
  set _rocplot;
  length index $ 20;
  Index='ICA titer';
  run;
proc logistic data=roc;
  model iddm(event='1') = baseiaa / outroc=or roceps=0;
  output out=iaapred p=piaa;
  ods output association=assoc;
  run;
data _null_;
  set assoc;
  if label2='c' then call symput("area",cvalue2);
  run;

title2 "Approximate area under curve = &area";
%rocplot(out=iaapred, outroc=or, p=piaa, id=iaa)
title2;
data iaa;
  set _rocplot;
  length index $ 20;
  Index='IAA titer';
  run;
data joint;
  set joint iaa;
  run;

proc logistic data=roc;
  model iddm(event='1') = baseica512_res / outroc=or roceps=0;
  output out=ica512pred p=pica512;
  ods output association=assoc;
  run;
data _null_;
  set assoc;
  if label2='c' then call symput("area",cvalue2);
  run;
title2 "Approximate area under curve = &area";
%rocplot(out=ica512pred, outroc=or, p=pica512, id=ica512)
title2;
data ica512;
  set _rocplot;

```

```

length index $ 20;
Index='ICA512 titer';
Run;
data joint;
set joint ica512;
run;

proc logistic data=roc;
model iddm(event='1') = basegad_res / outroc=or roceps=0;
output out=gadpred p=pgad;
ods output association=assoc;
run;
data _null_;
set assoc;
if label2='c' then call symput("area",cvalue2);
run;

title2 "Approximate area under curve = &area";
%rocplot(out=gadpred, outroc=or, p=pgad, id=gad)
title2;
data gad;
set _rocplot;
length index $ 20;
Index='GAD65 titer';
run;
data joint;
set joint gad;
run;

proc logistic data=roc;
model iddm(event='1') = basemiaa_res / outroc=or roceps=0;
output out=miaapred p=pmiaa;
ods output association=assoc;
run;
data _null_;
set assoc;
if label2='c' then call symput("area",cvalue2);
run;
title2 "Approximate area under curve = &area";
%rocplot(out=miaapred, outroc=or, p=pmiaa, id=miaa)
title2;
data miaa;
set _rocplot;
length index $ 20;
Index='MIAA titer';
run;

```

```

data joint;
  set joint miaa;
  run;

proc logistic data=roc;
  model iddm(event='1') = numabs / outroc=or roceps=0;
  output out=abspred p=pabs;
  ods output association=assoc;
  run;
data _null_;
  set assoc;
  if label2='c' then call symput("area",cvalue2);
  run;
  title2 "Approximate area under curve = &area";
  %rocplot(out=abspred, outroc=or, p=pabs, id=abs)
  title2;

data abs;
  set _rocplot;
  length index $ 20;
  Index='Number of Abs';
  run;
data joint;
  set joint abs;
  run;
%roc(data= icapred iaapred ica512pred gadpred miaapred abspred,
  var= pica paaa pica512 pgad pmiaa pabs,
  response=iddm)

/*fast insulin*/
  title "ROC plot for Fasting Insulin";
proc logistic data=roc;
  model iddm(event='1') = basefastins / outroc=or roceps=0;
  output out=inspred p=pins;
  ods output association=assoc;
  run;
data _null_;
  set assoc;
  if label2='c' then call symput("area",cvalue2);
  run;
  title2 "Approximate area under curve = &area";
  %rocplot(out=inspred, outroc=or, p=pins, id=ins)
  title2;
data joint;
  set _rocplot;
  length index $ 20;

```

```

Index='Fasting Insulin';
run;

/*fast glucose*/
title "ROC plot for Fasting Glucose";
proc logistic data=roc;
  model iddm(event='1') =basefastgluc / outroc=or roceps=0;
  output out=glucpred p=pgluc;
  ods output association=assoc;
run;
data _null_;
  set assoc;
  if label2='c' then call symput("area",cvalue2);
run;
title2 "Approximate area under curve = &area";
%rocplot(out=glucpred, outroc=or, p=pgluc, id=gluc)
title2;
data glucose;
  set _rocplot;
  length index $ 20;
  Index='Fasting Glucose';
run;
data joint;
  set joint glucose;
run;

/* FPIR*/
title "ROC plot for FPIR";
proc logistic data=roc;
  model iddm(event='1') = basefpir / outroc=or roceps=0;
  output out=fpirpred p=pfpir;
  ods output association=assoc;
run;
data _null_;
  set assoc;
  if label2='c' then call symput("area",cvalue2);
run;

title2 "Approximate area under curve = &area";
%rocplot(out=fpirpred, outroc=or, p=pfpir, id=fpir)
title2;
data fpir;
  set _rocplot;
  length index $ 20;
  Index='FPIR';
run;

```

```

data joint;
  set joint fpir;
  run;

/* HOMAR*/
  title "ROC plot for HOMAR";
proc logistic data=roc;
  model iddm(event='1') = basehomar/ outroc=or roceps=0;
  output out=homarpred p=phomar;
  ods output association=assoc;
  run;
data _null_;
  set assoc;
  if label2='c' then call symput("area",cvalue2);
  run;

  title2 "Approximate area under curve = &area";
  %rocplot(out=homarpred, outroc=or, p=phomar, id=homar)
  title2;
data homar;
  set _rocplot;
  length index $ 20;
  Index='HOMA-IR';
  run;
data joint;
  set joint homar;
  run;

/* FPIR/HOMAR*/
  title "ROC plot for FPIR/HOMAR";
proc logistic data=roc;
  model iddm(event='1') = baseratio / outroc=or roceps=0;
  output out=ratiopred p=pratio;
  ods output association=assoc;
  run;
data _null_;
  set assoc;
  if label2='c' then call symput("area",cvalue2);
  run;

  title2 "Approximate area under curve = &area";
  %rocplot(out=ratiopred, outroc=or, p=pratio, id=ratio)
  title2;
data ratio;
  set _rocplot;
  length index $ 20;

```

```

Index='FPIR/HOMA-IR';
run;
data joint;
set joint ratio;
run;
/*fasting glucose*/
title "ROC plot for Fasting Glucose--OGTT";
proc logistic data=roc;
model iddm(event='1') = base_ogtt_fasting / outroc=or roceps=0;
output out=ogglucpred p=poggluc;
ods output association=assoc;
run;
data _null_;
set assoc;
if label2='c' then call symput("area",cvalue2);
run;

title2 "Approximate area under curve = &area";
%rocplot(out=ogglucpred, outroc=or, p=poggluc, id=oggluc)
title2;
data oggluc;
set _rocplot;
length index $ 20;
Index='Fasting Glucose--OGTT';
run;
/*2 hour glucose*/
title "ROC plot for two-hour Glucose--OGTT";
proc logistic data=roc;
model iddm(event='1') = base_ogtt_twohour / outroc=or roceps=0;
output out=oggluc2hpred p=poggluc2h;
ods output association=assoc;
run;
data _null_;
set assoc;
if label2='c' then call symput("area",cvalue2);
run;

title2 "Approximate area under curve = &area";
%rocplot(out=oggluc2hpred, outroc=or, p=poggluc2h, id=oggluc2h)
title2;
data oggluc2h;
set _rocplot;
length index $ 20;
Index='2-hour Glucose--OGTT';
run;
data joint1;

```



```

set oggluc oggluc2h;
run;

/* peak c peptide*/
title "ROC plot for Peak C Peptide";
proc logistic data=roc;
  model iddm(event='1') = base_peak_c_pep / outroc=or roceps=0;
  output out=peakcpred p=ppeakc;
  ods output association=assoc;
run;
data _null_;
  set assoc;
  if label2='c' then call symput("area",cvalue2);
run;

title2 "Approximate area under curve = &area";
%rocplot(out=peakcpred, outroc=or, p=ppeakc, id=peakc)
title2;

data peakcpep;
  set _rocplot;
  length index $ 20;
  Index='Peak C Peptide';
run;
data joint1;
  set joint1 peakcpep;
run;
/* peak c peptide*/
title "ROC plot for AUC C Peptide";
proc logistic data=roc;
  model iddm(event='1') = base_auc_c_pep / outroc=or roceps=0;
  output out=auccpred p=paucc;
  ods output association=assoc;
run;
data _null_;
  set assoc;
  if label2='c' then call symput("area",cvalue2);
run;

title2 "Approximate area under curve = &area";
%rocplot(out=auccpred, outroc=or, p=paucc, id=aucc)
title2;
data auccpep;
  set _rocplot;
  length index $ 20;
  Index='AUC C Peptide';

```

```

run;
data joint1;
  set joint1 auccpep;
run;

/* hb1ac*/
title "ROC plot for HBA1C";
proc logistic data=roc;
  model iddm(event='1') = basehb / outroc=or roceps=0;
  output out=hbpred p=phb;
  ods output association=assoc;
run;
data _null_;
  set assoc;
  if label2='c' then call symput("area",cvalue2);
run;

title2 "Approximate area under curve = &area";
%rocplot(out=hbpred, outroc=or, p=phb, id=hb)
title2;
data hb1ac;
  set _rocplot;
  length index $ 20;
  Index='HBA1c';
run;
data joint2;
  set joint joint1 hb1ac;
run;
data jointnew;
  set oggluc2h ratio;
run;

/* Compare areas under the ROC curves of the indices using the method presented by
DeLong, et. al. (1988).
*/
%roc(data= oggluc2hpred fpirpred homarpred ratiopred,
  var= poggluc2h pfpir phomar pratio,
  response=iddm)

%roc(data= glucpred inspred fpirpred homarpred ratiopred ogglucpred oggluc2hpred
peakcpred auccpred hbpred,
  var= pgluc pins pfpir phomar pratio poggluc poggluc2h ppeakc paucc phb ,
  response=iddm)

ods html;
ods graphics on;

```

```

*title1 h = 3 "Figure: ROC Curves for Various Metabolic Indexes";
axis1 c = black value = (h = 2.2) minor = none
order = (0 to 1 by 0.1) width = 2
label=(angle=90 h = 2.8 'Sensitivity');
axis2 c = black value = (h = 2.2)
order = (0 to 1 by 0.1) width = 2
minor = none label= ( h = 2.8 "1-Specificity");
legend value = (height = 1.8) cborder = black
label = none
position = (TOP inside RIGHT)
mode = share;
goptions reset = symbol rotate = landscape
device = cgmmwwc gsfname = event
ctext = black cback = white
gsfmode = replace ftext=swissb
gsflen=80 display autofeed
gunit = pct;
  symbol1 i=join v=dot c=black line=1 h=2.0 ;
  symbol2 i=join v=triangle c=black line=1 h=2.5;
  symbol3 i=join v=circle c=black line=1 h=2.0 ;
  symbol4 i=join v=star c=black line=1 h=3.5;
proc gplot data=jointnew;
  * title "ROC Curves for Various Metabolic Indexes";
  label index="Index";
  plot _sensit_ * _1mspec_ = Index /
    vaxis=axis1 haxis=axis2;
  run;
  quit;
  title;

  title2;
  * footnote;

ods html close;
ods graphics off;

symbol1 i=join v=circle c=black line=1;
symbol2 i=join v=dot c=black line=1;
symbol3 i=join v=triangle c=black line=1;
  symbol4 i=join v=square c=black line=1;
symbol5 i=join v=star c=black line=1;

proc gplot data=joint;
  title "Metabolic Indexed Derived from IVGTT";
  label index="Index";

```

```

plot _sensit_ * _1mspec_ = Index /
  vaxis=0 to 1 by .1 haxis=0 to 1 by .1;
run;
quit;
title;
title2;
footnote;

symbol1 i=join v=circle c=black line=2;
symbol2 i=join v=star c=black line=2;
symbol3 i=join v=triangle c=black line=2;
symbol4 i=join v=dot c=black line=2;
proc gplot data=joint1;
  title "Metabolic Indexes Derived from OGTT";
  label index="Index";
  plot _sensit_ * _1mspec_ = Index /
    vaxis=0 to 1 by .1 haxis=0 to 1 by .1;
  run;
  quit;
  title;
  title2;
  footnote;
/* find the optimal cutoff */;
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=baseica);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=baseiaa);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=baseica512_res);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=basegad_res);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=basemiaa_res);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=numabs);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=basefastgluc);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=basefastins);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=basefpir);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=basehomar);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=baseratio);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=base_ogtt_fasting);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=base_ogtt_twohour);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=base_peak_c_pep);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=base_auc_c_pep);
%findcut(ds=roc, time=iddmtime, stat=iddm, cutvar=basehb);

%macro ppv(var, cutoff);
  data temp;
  set roc;
  if &var = . then status=.;
  else if &var ge &cutoff then status=1;
  else status=0;

```

```

run;

proc freq data=temp;
table iddm*status;
run;
%mend ppv;
%ppv(basefastgluc,5.33);
%ppv(basefastins,10.23);
%ppv(basefpir,156);
%ppv(basehomar,2.64);
%ppv(baseratio,49.22);
%ppv(base_ogtt_fasting,88);
%ppv(base_ogtt_twohour,114);
%ppv(base_peak_c_pep,5.3);
%ppv(base_auc_c_pep,595);
%ppv(basehb,5.1);

data newroc;
set roc;
if baseratio ge 49.22 then group="FPIR/HOMA-IR>=49.22";
else if baseratio lt 49.22 then group="FPIR/HOMA-IR<49.22" ;
if base_ogtt_twohour ge 114 then status=1 ;
else if base_ogtt_twohour lt 114 then status=0;
if baseratio ge 49.22 then status1=0;
else if baseratio lt 49.22 then status1=1 ;
time=idmtime/365.25;
run;

***KM curves;
ods listing close;
ods output ProductLimitEstimates = s_est
CensoredSummary = cen_sum
HomTests = homtest
Quartiles = quart;
proc lifetest data = newroc;
time time*iddm(0);
strata group;
id maskid;
run;
ods output close;
ods listing;
data s_est;
set s_est;
retain pr_s;
if iddm = 1 then pr_s = survival;
surviva = .;

```

```

if iddm = 0 then do;
surviva = pr_s;
survival = .;
end;
run;
proc sort data = cen_sum;
by trt;
where control_var is missing;
run;
data _null_;
set cen_sum end = last;
pctevt = 100 - pctcens;
call symput('groupf'||compress(_n_),
compress(put(trt, trt.)));
call symput('ng'||compress(_n_),
compress(put(total, 8.)));
call symput('ngevt'||compress(_n_),
compress(put(failed, 8.)));
call symput('pgevt'||compress(_n_),
compress(put(pctevt, 8.2)));
call symput('ngcen'||compress(_n_),
compress(put(censored, 8.)));
call symput('pgcen'||compress(_n_),
compress(put(pctcens, 8.2)));
run;
proc sort data = homtest;
where test = 'Log-Rank';
by test;
run;
data _null_;
set homtest;
call symput('logrp', compress(put(probchisq,
pvalue6.4)));
run;
proc sort data = quart;
by group;
where percent = 50;
run;
data _null_;
set quart;
call symput('gmed'||compress(_n_),
compress(put(estimate, 8.2)));
call symput('lmit'||compress(_n_),
compress(put(lowerlimit, 8.2)));
call symput('ulmit'||compress(_n_),
compress(put(upperlimit, 8.2)));

```

```

run;
ods html;

title1 h = 3 "Figure 3.2: Kaplan-meier curves by the level of IVGTT FPIR/HOMA-IR";
filename event "C:\Documents and settings\xup\Desktop\PHD\disseration\manuscript 1-
baseline ROC comparison\lifetest_event.cgm";
filename censor "C:\Documents and Settings\xup\Desktop\PHD\disseration\manuscript
1- baseline ROC comparison\lifetest_censo.cgm";
axis1 c = black value = (h = 2.2) minor = none
order = (0 to 1 by 0.1) width = 2
label=(angle=90 h = 2.4 'Kaplan-Meier Estimate');
axis2 c = black value = (h = 2.2)
order = (0 to 7 by 1) width = 2
minor = none label= ( h = 2.4 "Time from Randomization to Disease Onset or Last
Follow-up (Years)");
legend value = (height = 1.5) cborder = black
label = none
position = (TOP inside RIGHT)
mode = share;
goptions reset = symbol rotate = landscape
device = cgmmwwc gsfname = event
cstext = black cback = white
gsfmode = replace ftext=swissb
gsflen=80 display autofeed
gunit = pct;
proc gplot data = s_est gout = work.mygraf;
note height = 1.3 j = left font = swissbu
move = (75, 78) "Log Rank P-Value : &logrp";
symbol1 V = none I = steplj L = 1 c = black
H = 1.5 W = 3.5;
symbol2 V = none I = steplj L = 2 c = black
H = 1.5 W = 3.5;
plot survival*time = group
/ name = 'g1'
haxis=axis2
vaxis=axis1
legend = legend ;
format survival 8.1 time 8.;
run;
ods html close;

* *****%findcut macro *;
options mprint symbolgen macrogen mlogic;
%macro findcut(ds=, time=, stat=, cutvar=);

```

```

PROC FORMAT;
VALUE pf -99='>0.30';
proc sort data=&ds; by &cutvar;

data difage; set &ds; by &cutvar;
title3 "step1: check no. of disticnt &cutvar";
if first.&cutvar;
proc sort; by &cutvar;
data ttt; set difage; by &cutvar;
cut=_N_;
keep &cutvar cut;
proc sort; by descending &cutvar;
data cut; set ttt; by descending &cutvar;
if _N_=1 then do;
nocut=cut; retain nocut;
end;
somecut=&cutvar;
drop &cutvar;
output;
proc sort; by somecut;
data dim; set cut; by somecut;
if last.somecut; dummy=1;
keep nocut dummy;
proc sort; by dummy;
proc sort data=&ds; by descending &time &stat;
data step1; set &ds; by descending &time &stat;
title3 "check step1";
if _N_=1 then do;
norisk=0; retain norisk;
end;
if first.&time then do;
nodeath=0; retain nodeath;
end;
norisk=norisk+1;
if &stat=1 then nodeath=nodeath+1;
if last.&time and &stat=1 then output;
data step1; set step1;
keep &time norisk nodeath;
proc sort; by &time;
data dummy; set &ds;
dummy=1;
proc sort; by dummy;
data double; merge dummy dim; by dummy;
do cut=1 to nocut;
output;
end;

```



```

proc sort; by cut;
proc sort data=cut; by cut;
data comb; merge double cut; by cut;
  keep obsno &time &stat &cutvar somecut cut;
proc sort; by cut descending &time &stat;
proc sort data=comb; by cut descending &time &stat;
data step2; set comb; by cut descending &time &stat;
  title3 "check step2";
  if first.cut then do;
    noriskc=0; retain noriskc;
  end;
  if first.&time then do;
    nodeathc=0; retain nodeathc;
  end;
  if &cutvar>=somecut then noriskc=noriskc+1;
  if &stat=1 and &cutvar>=somecut then nodeathc=nodeathc+1;
  if last.cut or (last.&time and &stat=1) then output;
data step2; set step2;
  keep &time cut somecut noriskc nodeathc;
  proc sort; by &time;
data step3; merge step2 step1; by &time;
  title3 "step3";
  r1 = noriskc; r2 = norisk - r1; r = r1 + r2; d = nodeathc;
  v = r1*r2*d*(r - d)/(r**2*(r-1));
  sik=nodeathc-nodeath*noriskc/norisk;
proc sort; by cut somecut;
proc univariate noprint; var sik v; by cut;
  output out=step4 sum=sk sumv;
data step4; set step4;
  title3 "step4";
  abs_sk=abs(sk);
  z = abs_sk/sqrt(sumv);
  nominal_p = 2*(1 - probnorm(z));
  if nominal_p > .30 then nominal_p = -99;
  dummy=1;
  format nominal_p pf.;
proc sort; by dummy;
proc univariate noprint; var abs_sk;
  output out=step5 max=maxsk;
data diftm; set &ds;
  title3 "step5";
  if &stat=1;
proc sort; by &time;
data ttt; set diftm; by &time;
  if first.&time;
proc univariate noprint; var &time;

```

```

output out=deathtim N=nodeath;
data deathtim; set deathtim;
dummy=1;
keep nodeath dummy;
data square; set deathtim;
do i=1 to nodeath;
do j=1 to i;
frac=1/(nodeath-j+1);
keep i j frac;
output;
end;
end;
proc sort; by i j;
data sums2; set square; by i j;
if first.i then do;
sumi=0; retain sumi;
end;
sumi=sumi+frac;
if last.i then do;
sumi=(1-sumi)*(1-sumi);
output;
end;
proc univariate noprint; var sumi;
output out=d2 sum=s2 n=n2;
data step5; set d2;
ssquare=(1.0/(n2-1))*s2;
dummy=1;
proc sort; by dummy;
data step5; merge step4 step5; by dummy;
title3 "step 5";
q=abs_sk/(sqrt(ssquare)*sqrt(n2-1));
if q>1 then p=2*exp(-2*q*q);
if q<=1 then p=-99;
format p pf.;
proc sort; by cut;
data cut; set cut;
&cutvar=somecut;
keep cut &cutvar;
proc sort; by cut;
data temp; merge step5 cut; by cut;
if _N_=1 then do;
maxsk=0; maxcut=0; retain maxsk maxcut;
end;
if abs_sk>maxsk then do;
maxsk=abs_sk;
maxcut=cut;

```

```

end;
dummy=1;
output;
proc sort; by dummy;
data temp; set temp; by dummy;
  cut=maxcut;
  if last.dummy then output;
  keep cut;
proc sort; by cut;
data final; merge step5 cut temp(in=in1); by cut;
  TITLE3 "Final Result";
  pick="  ";
  if in1 then Pick="<====";
  label cut="Cutpoint"
        sk="O-E"
        abs_sk="ABS(sk)"
        Q="Q statistics"
  p="Adjusted P-value"
        nominal_p = "Nominal P-value"
        sumv = Variance;
proc print label;
  var cut &cutvar sk sumv z nominal_p p pick;
run;
%mend findcut;
%findcut;

/* %ROC macro*/

%macro roc(version, data=, var=, response=, contrast=, details=no,
  alpha=.05);
%let opts = %sysfunc(getoption(notes))
  _last_=%sysfunc(getoption(_last_));
options nonotes;
if %sysevalf(&sysver >= 8.2) %then %do;
  filename _ver url 'http://ftp.sas.com/techsup/download/stat/versions.dat' termstr=crlf;
  data _null_;
  infile _ver;
  input name:$15. ver;
  if upcase(name)="&sysmacroname" then call symput("_newver",ver);
  run;
%if &syserr ne 0 %then
  %put ROC: Unable to check for newer version;
%else %if %sysevalf(&_newver > &_version) %then %do;
  %put ROC: A newer version of the ROC macro is available.;
  %put %str(      ) You can get the newer version at this location.;
  %put %str(      ) http://support.sas.com/ctx/samples/index.jsp;

```

```

%end;
%end;
title "The ROC Macro";
title2 " ";
%let error=0;
%if &data= %then %do;
    %put ERROR: Specify DATA= containing the OUT= data sets of models to be
    compared;
    %goto exit;
%end;
%if &var= %then %do;
    %put ERROR: Specify predictor or XBETA variables in the VAR= argument;
    %goto exit;
%end;
%if &response= %then %do;
    %put ERROR: Specify response variable in the RESPONSE= argument;
    %goto exit;
%end;
%let i=1;
%do %while (%scan(&data,&i) ne %str() );
    %let data&i=%scan(&data,&i);
    %let i=%eval(&i+1);
%end;
%let ndata=%eval(&i-1);
data _comp(keep = &var &response);
%if &data=%str() or &ndata=1 %then set;
%else merge;
    &data;
    if &response not in (0,1) then call symput('error',1);
run;
%if &error=1 %then %do;
    %put ERROR: Response must have values 0 or 1 only.;
    %goto exit;
%end;
proc iml;
    start mwcomp(psi,z);
        rz = ranktie( z[,1] );
        nx = sum( z[,2] );
        ny = nrow(z)-nx;
        loc = loc( z[,2]=1 );
        psi = j(nrow(z),1,0);
        psi[loc] = (rz[loc] - ranktie(z[loc,1]))/ny;
        loc = loc( z[,2]=0 );
        psi[loc] = ( nx+ranktie(z[loc,1])-rz[loc])/nx;
        free rz loc nx ny;
    finish;

```

```

start mwvar(t,v,nx,ny,z);
k = ncol(z)-1;
ind = z[k+1];
v = j(k,k,0); t=v; nx=v; ny=v;
do i=1 to k;
  do j=1 to i;
    who = loc( (z[,i]^=.)#(z[,j]^=.) );
    run mwcomp(psi,(z[,i]||ind)[who,]);
    run mwcomp(psj,(z[,j]||ind)[who,]);
    inow = ind[who,];
    m = inow[+];
    n = nrow(psi)-m;
    nx[i,j] = m; ny[i,j] = n;
    mi = (psi#inow)[+] / m;
    mj = (psj#inow)[+] / m;
    t[i,j] = mi; t[j,i] = mj;
    psii = psi-mi; psij = psj-mj;
    v[i,j] = (psi#psij#inow)[+] / (m#(m-1))
      + (psi#psij#(1-inow))[+] / (n#(n-1));
    v[j,i] = v[i,j];
  end;
end;
free psii psij inow ind who;
finish;
use _comp var {&var &response};
read all into data [colname=names];
run mwvar(t,v,nx,ny,data);
vname = names[1:(ncol(names)-1)];
manwhit = vecdiag(t);
%if &contrast= %then %do;
%put ROC: No contrast specified. Pairwise contrasts of all;
%put %str( ) curves will be generated.;
call symput('col',char(ncol(data)-1));
%if &col=1 %then %str(l=1); %else %do;
l=(j(&col-1,1)||-i(&col-1))
%do i=&col-2 %to 1 %by -1;
  //(j(&i,&col-&i-1,0)||j(&i,1)||-i(&i))
%end;
%end;
call symput('maxrow',char(comb(max(nrow(l),2),2)));
%end;
%else %do;
l = { &contrast };
call symput('maxrow',char(nrow(l)));
%end;

```

```

lt=l*manwhit;
lv=l*v*1';
c = ginv(lv);
chisq = lt`c*lt;
df = trace(c*lv);
p = 1 - probchi( chisq, df );
stderr=sqrt(vecdiag(v));
arealcl=manwhit-probit(1-&alpha/2)*stderr;
areaucl=manwhit+probit(1-&alpha/2)*stderr;
areastab=putn(manwhit||stderr||arealcl||areaucl,'7.4');
sediff=sqrt(vecdiag(lv));
diffcl=lt-probit(1-&alpha/2)*sediff;
diffucl=lt+probit(1-&alpha/2)*sediff;
diffchi=(lt##2)/vecdiag(lv);
diffp=1-probchi(diffchi,1);
%if %upcase(%substr(&details,1,1)) ne N %then %do;
print t [label='Pairwise Deletion Mann-Whitney Statistics' colname=vname
rowname=vname];
%end;
print areastab [label=
"ROC Curve Areas and %sysevalf(100*(1-&alpha))% Confidence Intervals"
rowname=vname colname={'ROC Area' 'Std Error' 'Confidence' 'Limits'}];
rname='Row1':"Row&maxrow";
%if %upcase(%substr(&details,1,1)) ne N %then %do;
print v [label='Estimated Variance Matrix' colname=vname rowname=vname];
print nx [label='X populations sample sizes' colname=vname rowname=vname];
print ny [label='Y populations sample sizes' colname=vname rowname=vname];
print lv [label='Variance Estimates of Contrast' rowname=rname
colname=rname];
%end;
print l [label='Contrast Coefficients' rowname=rname colname=vname];
fdiffchi=putn(diffchi,'9.4');
fdiffp=putn(diffp,'pvalue. ');
diffs=putn(lt||sediff||diffcl||diffucl,'7.4');
diffstab=diffs||fdiffchi||fdiffp;
print diffstab [label=
"Tests and %sysevalf(100*(1-&alpha))% Confidence Intervals for Contrast Rows"
rowname=rname colname={'Estimate' 'Std Error' 'Confidence' 'Limits'
'Chi-square' 'Pr > ChiSq'}];
c2=putn(chisq,'9.4');
df2=putn(df,'3. ');
p2=putn(p,'pvalue. ');
ctest=c2||df2||p2;
print ctest [label='Contrast Test Results'
colname={'Chi-Square' ' DF' 'Pr > ChiSq'}];
%global pvalue;

```

```

    call symput('pvalue',p2);
quit;
%exit;
options &opts;
title; title2;
%mend;

/* %ROC PLOT macro*/

%macro rocplot ( version, outroc=, out=, p=, id=, plottype=high, font=swissb,
                size=3, color=black, plotchar=dot,
                roffset=4, grid=n, mindist=0.02, round=1e-8 );
%let _version=1.1;
%if &version ne %then %put ROC PLOT macro Version &_version;
%let opts = %sysfunc(getoption(notes));
%if &version ne debug %then %str(options nonotes.);
%if %sysevalf(&sysver >= 8.2) %then %do;
    %let _notfound=0;
    filename _ver url 'http://ftp.sas.com/techsup/download/stat/versions.dat' termstr=crlf;
    data _null_;
        infile _ver end=_eof;
        input name:$15. ver;
        if upcase(name)="&sysmacroname" then do;
            call symput("_newver",ver); stop;
        end;
        if _eof then call symput("_notfound",1);
    run;
%if &syserr ne 0 or &_notfound=1 %then
    %put &sysmacroname: Unable to check for newer version;
%else %if %sysevalf(&_newver > &_version) %then %do;
    %put &sysmacroname: A newer version of the &sysmacroname macro is available.;
    %put %str(      ) You can get the newer version at this location.;
    %put %str(      ) http://support.sas.com/ctx/samples/index.jsp;
%end;
%end;
%let nomatch=0;
%if %quote(&id)= %then %do;
    %put ERROR: At least one ID= variable is required.;
    %goto exit;
%end;
%if %quote(&p)= %then %do;
    %put ERROR: The P= option is required.;
    %goto exit;
%end;
%if %quote(&outroc) ne %then %do;
    %if %sysfunc(exist(&outroc)) ne 1 %then %do;

```

```

    %put ERROR: OUTROC= data set not found.;
    %goto exit;
%end;
%end;
%else %do;
    %put ERROR: The OUTROC= option is required.;
    %goto exit;
%end;
%if %quote(&out) ne %then %do;
    %if %sysfunc(exist(&out)) ne 1 %then %do;
        %put ERROR: OUT= data set not found.;
        %goto exit;
    %end;
%end;
%else %do;
    %put ERROR: The OUT= option is required.;
    %goto exit;
%end;
data _outroc;
    set &outroc;
    _prob_=round(_prob_,&round);
run;
data _out;
    set &out;
    _obs_=_n_;
    if &p = . then delete;
    _prob_=round(&p , &round);
run;
proc sort data=_out nodupkey;
    by _prob_;
run;
proc sort data=_outroc nodupkey;
    by _prob_;
run;
data _rocplot;
    _inout=0; _inroc=0;
    merge _outroc(in=_inroc) _out(in=_inout);
    by _prob_;
    if not(_inout and _inroc) then do;
        call symput('nomatch',1);
        delete;
    end;
    length _id $ 200;
    _sens_=put(_sensit_,5.3);
    _spec_=put(-_1mspec_+1,5.3);
    _id=trim(left(%scan(&id,1)))

```



```

    %let i=2;
    %do %while (%scan(&id,&i) ne %str() );
||'|trim(left(%scan(&id,&i)))
    %let i=%eval(&i+1);
    %end; ;
run;
%if &nomatch=1 %then %do;
%put ROCPLOT: Some predicted values in OUT= did not match predicted values;
%put %str(    in OUTROC=. Verify that you used the ROCEPS=0 option in);
%put %str(    PROC LOGISTIC.);
%end;
data _add00;
    _sensit_=0; _1mspec_=0;
run;
%if %upcase(%substr(&plotype,1,1))=L %then %do;
data _rocplot;
    set _rocplot _add00;
run;
proc plot data=_rocplot;
    plot _sensit_*_1mspec_ /
        haxis=0 to 1 by .1 vaxis=0 to 1 by .1;
run; quit;
footnote "Point labels are values of &i";
proc plot data=_rocplot;
    plot _sensit_*_1mspec_ $ _id /
        haxis=0 to 1 by .1 vaxis=0 to 1 by .1;
run; quit;
%end;
%if %upcase(%substr(&plotype,1,1))=H %then %do;
data _anno;
    length function style color $ 8 position $ 1 text $ 200;
    retain function 'label' xsys ysys '2' hsys '3'
        size &size position '6' style "&font"
        color "&color" xprev xprevtop 1.4;
    set _rocplot(keep=_sensit_ _1mspec_ _id) end=eof;
    by _1mspec_ notsorted;
    x=_1mspec_; y=_sensit_+.025;
    if last._1mspec_ then do;
    if xprev-x>=&mindist then do;
        text=" ||trim(left(_id));
        xprev=x;
    end;
    else text="";
end;
else if first._1mspec_ then do;
    if xprevtop-x>=&mindist then do;

```

```

        text=trim(left(_id));
        xprevtop=x;
    end;
    else text="";
    position=4;
end;
else text="";
angle=350-x*80; output;
position='6'
if eof then do;
    x=0; y=0; function='move'; output;
    x=1; y=1; function='draw'; line=1; hsys='1'; size=0.25; output;
end;
run;
data _rocplot;
    set _rocplot _add00;
run;
symbol1 i=join v=&plotchar c=darkblue l=1;
footnote "Point labels are values of &id";
axis1 offset=(&roffset,&roffset)pct order=(0 to 1 by .1);
axis2 offset=(&roffset,&roffset)pct order=(0 to 1 by .1) label=(angle=90);
proc gplot data=_rocplot;
    plot _sensit_*_1mspec_=1 / vaxis=axis2
        haxis=axis1 annotate=_anno
        %if %upcase(%substr(&grid,1,1))=Y %then %do;
            vref=0 to 1 by .1 href=0 to 1 by .1
            cvref=cxd4d4d4 chref=cxd4d4d4
        %end;
run;
quit;
%end;
footnote;
%exit;
options &opts;
%mend;

```

## Appendix C: Permission for Use of Published Article.



Christine N. Taylor  
Associate  
Rights and Permissions  
Publications  
703-549-1500 ext. 1635 - Tel  
703-683-2890 - Fax  
permissions@diabetes.org

June 21, 2012

Ping Xu  
Pediatrics Epidemiology Center  
University of South Florida  
3650 Spectrum Blvd., Suite 100  
Tampa, FL 33612

Permission Request Number: *CT062112-USF*

Dear Ping Xu:

We are pleased to grant permission to you to reproduce the following article:

✓ Prognostic Performance of Metabolic Indexes in Predicting Onset of Type 1 Diabetes

From: *Diabetes Care*®, Vol. 33, 2010; 2508-2513  
For use in: **Dissertation**

The permission granted is for a one-time use only for English language as described in this agreement, and is subject to the following conditions:

1. Permission is granted for use in the following formats – **Print & Electronic**.
2. Each copy containing our material that you reproduce or distribute must bear the following copyright notice:  
"Copyright 2010 American Diabetes Association  
From *Diabetes Care*®, Vol. 33, 2010; 2508-2513  
Reprinted by permission of the *American Diabetes Association*."
3. The article will be reprinted exactly as it was originally published.
4. Permission is **non-exclusive** and the American Diabetes Association reserves the right to grant the same permission to others.
5. This permission agreement is valid for a period of twelve (12) months beginning from the date of this letter.

Sincerely,

*Christine Taylor*

Christine N. Taylor

AGREED: \_\_\_\_\_

DATE: \_\_\_\_\_

*06-28-2012*  
I elect not to use this material.

Unless specifically noted here, the permission granted does not include the use of the American Diabetes Association logo or the cover logo of the journal(s). Permission to reproduce material does not permit the above-named entity or the company sponsoring the resulting product to act as an agent of the American Diabetes Association or of the journal(s).

National Office  
1701 North Beauregard Street  
Alexandria, VA 22311  
Tel: 703-549-1500

Diabetes Information  
call 1-800-DIABETES (1-800-342-2383)  
cnifm www.diabetes.org  
The Association gratefully accepts gifts through your will.

The Mission of the American  
Diabetes Association is to prevent and  
cure diabetes and to improve the lives  
of all people affected by diabetes.

## **About the Author**

Ping Xu was born and raised in Southern China, where she earned her Bachelor's Degree in English Education from Jiangxi Normal University, Nanchang, China. She moved to the states in 2001 and received a M.P.H. degree in Biostatistics from University of South Florida in 2005. She entered the Ph.D. program in the Department of Epidemiology and Biostatistics at the University of South Florida in 2006. In her career, she has been working at the Pediatric Epidemiology Center within the College of Medicine, at the University of South Florida since 2005. She has served as a biostatistician in a variety of clinical studies, providing statistical support. She has approximately twenty peer-reviewed publications in multiple areas of research including type 1 diabetes investigations and cancer control research.