**T.C.**
**REPUBLIC OF TURKEY**
**HACETTEPE UNIVERSITY**
**GRADUATE SCHOOL HEALTH SCIENCES**

# SPECIFYING THE BOUNDARIES OF GRAY ZONE IN DIAGNOSTIC TESTS WITH INFORMATION CRITERIA

**Ebru ÖZTÜRK**

**Program of Biostatistics**

**DOCTOR OF PHILOSOPHY THESIS**

**ANKARA**

**2023**

**T.C.**
**REPUBLIC OF TURKEY**
**HACETTEPE UNIVERSITY**
**GRADUATE SCHOOL HEALTH SCIENCES**

# SPECIFYING THE BOUNDARIES OF GRAY ZONE IN DIAGNOSTIC TESTS WITH INFORMATION CRITERIA

**Ebru ÖZTÜRK**

**Program of Biostatistics**

**DOCTOR OF PHILOSOPHY THESIS**

**ADVISOR OF THE THESIS**

**Assoc. Prof. Dr. Jale KARAKAYA KARABULUT**

**ANKARA**

**2023**

**SPECIFYING THE BOUNDARIES OF GRAY ZONE IN DIAGNOSTIC TESTS WITH INFORMATION CRITERIA**

**Ebru ÖZTÜRK**

**Supervisor: Assoc. Prof. Dr. Jale KARAKAYA KARABULUT**

This thesis study has been approved and accepted as a PhD dissertation in "Biostatistics Program" by the assesment committee, whose members are listed below, on July 26, 2023.

| | |
|---|---|
| **Chairman of the Committee :** | *Prof. Dr. Mehtap AKÇİL OK*<br>*Başkent University* |
| **Member :** | *Assoc. Prof. Dr. Osman DAĞ*<br>*Hacettepe University* |
| **Member :** | *Assist. Prof. Dr. Sevilay KARAHAN*<br>*Hacettepe University* |
| **Member :** | *Assist. Prof. Dr. Dinçer GÖKSÜLÜK*<br>*Erciyes University* |
| **Member :** | *Assist. Prof. Dr. Ayhan PARMAKSIZ*<br>*Istanbul Health & Technology University* |

This dissertation has been approved by the above committee in conformity to the related issues of Hacettepe University Graduate Education and Examination Regulation.

*Prof. Müge YEMİŞCİ ÖZKAN, MD, PhD*

**Director**

## YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan ***"Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge"*** kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü / Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir. [(1)]
- Enstitü / Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ... ay ertelenmiştir. [(2)]
- Tezimle ilgili gizlilik kararı verilmiştir. [(3)]

08 /08/2023

Ebru ÖZTÜRK

*i*

---

[i] *"Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge"*

*(1) Madde 6. 1. Lisansüstü tezle ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez **danışmanının** önerisi ve **enstitü anabilim dalının** uygun görüşü üzerine **enstitü** veya **fakülte yönetim kurulu** iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.*

*(2) Madde 6. 2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç imkanı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez **danışmanının** önerisi ve **enstitü anabilim dalının** uygun görüşü üzerine **enstitü** veya **fakülte yönetim kurulunun** gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.*

*(3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, **tezin yapıldığı kurum** tarafından verilir *. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, **ilgili kurum ve kuruluşun önerisi** ile **enstitü** veya **fakültenin** uygun görüşü üzerine **üniversite yönetim kurulu** tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.*

*Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir*

****** Tez **danışmanının** önerisi ve **enstitü anabilim dalının** uygun görüşü üzerine **enstitü** veya **fakülte yönetim kurulu tarafından karar verilir.***

**ETHICAL DECLARATION**

In this thesis study, I declare that all the information and documents have been obtained in the base of the academic rules and all audio-visual and written information and results have been presented according to the rules of scientific ethics. I did not do any distortion in data set. In case of using other works, related studies have been fully cited in accordance with the scientific standards. I also declare that my thesis study is original except cited references. It was produced by myself in consultation with supervisor (Title, Name SURNAME) and written according to the rules of thesis writing of Hacettepe University Institute of Health Sciences.

*Ebru ÖZTÜRK*

**ACKNOWLEDGMENTS**

First, I would like to express my sincere gratitude to my advisor Assoc. Prof. Dr. Jale Karakaya Karabulut for her endless support, guidance, motivation, and patience during the thesis study. Without her guidance and feedback, this study would not have been completed.

I also would like to thank my former advisor Prof. Dr. Ergun Karaağaoğlu and the thesis committee members Assoc. Prof. Dr. Osman Dağ and Assist. Prof. Dr. Dinçer Göksülük for their contribution to this study.

I would like to present my grateful thanks to my examining committee members: Prof. Dr. Mehtap Akçil Ok, Assist. Prof. Dr. Sevilay Karahan and Assist. Prof. Dr. Ayhan Parmaksız for their valuable time to review my study and suggestions.

I would also like to thank my dear friends and colleagues Merve Kaşıkcı and Hanife Avcı for their nice friendship and kind support. I'm also thankful to Assist. Prof. Dr. Duygu Aydın Haklı and Dr. Merve Başol for their kind friendship, unconditional support and wise advice. Moreover, I would thank all members of the Hacettepe University Department of Biostatistics.

My special thanks go to my dear friends Büşra Karakaya, Merve Acıelma, and Şeyma Gürkan for their warm friendship, support, and inspiration.

I would like to represent my deepest thanks to my dear family, Hacer Öztürk, İsmail Öztürk, Burcu Öztürk Şahin, Hatice Öztürk Ferah, Arif Ferah, and Çağlar Şahin for their support and patience. I'm grateful to my little nieces Defne Ferah and Pelin Ferah for their inspirational souls and for making my life more enjoyable and colorful.

## ABSTRACT

**Öztürk, E. Specifying the Boundaries of Gray Zone in Diagnostic Tests with Information Criteria, Hacettepe University Graduate School Health Scıences, Department of Biostatistics Doctor of Philosophy Thesis, Ankara, 2023.** The decision-making process in medicine is a crucial subject due to the classification of subjects as healthy or diseased. Mostly, it concludes with binary outcomes, such as whether the person has a condition or not. The various information from subjects is taken, such as the complaints, family history, symptoms, or laboratory tests (known as diagnostic tests) to rule-in or rule-out the disease. Due to their advantages such as being cost-effective or rapid, even if some diagnostic tests cannot perfectly discriminate subjects they are commonly used in clinics. One method for assessing quantitative diagnostic tests to diagnose subjects is to specify an optimal cut-off point. Yet, this may cause issues on quantitative diagnostic tests with a single cut-off value as the distributions of diseased and healthy subjects overlap. Forcing the subjects in this overlapped area one of the classes causes the false negative or false positive rates. To deal with this issue, there are some approaches called a gray zone or middle inconclusive area in which subjects are classified diseased, non-diseased, and neither diseased nor non-diseased. In this thesis, we aim to propose a new solution to find the boundaries of the gray zone based on the information theory approach. We intend to compare and evaluate the performance of this proposed solution against existing methods (“grey zone” and “uncertain interval” approaches). The proposed algorithm was based on joint entropy. In the simulation scenarios, we considered effect size, sample size, the homogeneity of variance and prevalence of the disease. To compare the results of the proposed methods with existing algorithms, the length of the gray zone was examined under the condition of fixed area under the receiver operating curve in out of the gray zone. In simulations, the suggested approach mostly produced the lowest gray zone length with equal variances. In some simulation scenarios, it outperformed for unequal variances. However, it has the benefit that the suggested algorithm has no previous knowledge.

**Keywords:** Diagnostic tests, medical decision, gray zone, information criteria

## ÖZET

**Öztürk, E. Tanı Testlerinde Gri Alan Sınırlarının Bilgi İçeriği Yaklaşımı İle Belirlenmesi, Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Programı Doktora Tezi, Ankara, 2023.** Tıpta karar verme, bireylerin sağlıklı veya hastalıklı olarak sınıflandırılması nedeniyle çok önemli bir konudur. Çoğunlukla, bireyin bir rahatsızlığı olup olmadığı gibi ikili sonuçlarla sonuçlanır. Hastalık tanısı koymak ya da dışlamak için bireylerden şikayetleri, aile öyküsü, semptomları veya laboratuvar testleri (tanı testleri) gibi çeşitli bilgiler alınır. Maliyet-etkinliği ya da hızlı olması gibi avantajları nedeniyle bazı tanı testleri bireyleri mükemmel olarak sınıflamasa da kliniklerde yaygın olarak kullanılmaktadır. Sıralı ve sürekli sayısal yanı testlerini değerlendirmenin bir yöntemi, optimal bir kesim noktası belirlemektir. Ancak, hasta ve sağlıklı bireylerin dağılımları örtüşmesi durumunda tek kesim noktasını ele alarak ikili sınıflandırmak bazı sorunlara neden olabilir. Bireyleri bu örtüşen alanda sınıflardan birine zorlamak yanlış negatif veya yanlış pozitife neden olur. Bu durumu ele almak için literatürde, bireylerin hastalıklı, sağlıklı ve hasta ya da sağlıklı değil olarak sınıflandırıldığı gri bölge veya orta sonuçsuz alan adı verilen bazı yaklaşımlar vardır. Bu tezde, bilgi teorisi yaklaşımına dayalı olarak gri bölgenin sınırlarını bulmak için yeni bir algoritma önermeyi amaçlıyoruz. Ayrıca önerilen bu algoritmayı literatürde mevcut yaklaşımlarla (“gri bölge” ve “belirsiz aralık” yaklaşımları) karşılaştırmayı ve değerlendirmeyi hedefliyoruz. Önerilen algoritma birleşik entropiye dayanıyordu. Simülasyon senaryolarında, etki büyüklüğü, örneklem büyüklüğü, varyansın homojenliği ve hastalık prevalansı faktörlerini dikkate aldık. Önerilen yöntemlerin sonuçlarını mevcut algoritmalarla karşılaştırmak için gri bölgenin uzunluğu, gri bölge dışında sabit bir alıcının işlem karakteristiği eğrisi altında incelendi. Simülasyonlarda, önerilen yaklaşım eşit varyans için çoğunlukla en düşük gri bölge uzunluğunu üretti. Bazı simülasyon senaryolarında, eşit olmayan varyans için daha iyi sonuçlar elde etti. Bununla birlikte, önerilen algoritmanın bir ön bilgiye ihtiyacı olmaması avantaj sağlamaktadır.

**Anahtar Kelimeler:** Tanı testleri, tıpta karar verme, gri bölge, bilgi kriteri

## TABLE OF CONTENT

## LIST OF ABBREVIATIONS

| | |
|---|---|
| **AUROC** | Area Under Receiver Operating Characteristics Curve |
| **FN** | False Negative |
| **FP** | False Positive |
| **GZ** | Grey Zone |
| **JE-UB** | Joint Entropy based Algorithm for Uncertain Boundaries |
| **JE-UB-ED** | Joint Entropy based Algorithm for Uncertain Boundaries with Euclidean Distance |
| **JE-UB-KB** | Joint Entropy based Algorithm for Uncertain Boundaries with Kernel Based Youden's J Index |
| **LL** | Lower Limit |
| **LR(-)** | Negative Likelihood Ratio |
| **LR(+)** | Positive Likelihood Ratio |
| **NPV** | Negative Predictive Value |
| **OVL** | Overlap Coefficient |
| **P(D)** | Prevalence of Disease |
| **PPV** | Positive Predictive Value |
| **ROC** | Receiver Operating Characteristics |
| **SLSQP** | Sequential Least Squares Programming Optimizer |
| **TG-ROC** | Two-Graph Receiver Operating Characteristics |
| **TN** | True Negative |
| **TP** | True Positive |
| **UI** | Uncertain Interval |
| **UL** | Upper Limit |

# LIST OF FIGURES

## LIST OF TABLES

# 1. INTRODUCTION

The majority of decision-making process in medicine concludes with a binary status, such as whether the individual should receive treatment or not, or whether they have a condition. Deciding on this binary status is even more difficult and complex with various information such as patients' complaints, family history, and the occurrence of different symptoms. Based on this information, some diseases should have been ruled out or ruled in. However, the information that is based on the knowledge of the patient doesn't significantly reduce uncertainty about the decision. To deal with this issue, laboratory tests, also known as diagnostic tests, are developed. Therefore, the uncertainty for decision-making in the clinic will decrease by combining the results of diagnostic tests with the information obtained from the patient. Physician will ask for the right diagnostic test based on the patient's symptoms and complaints in order to diagnose the disease.

Some of the diagnostic tests are referred to as the reference standard (or gold standard) tests. Optimally, it is expected that the reference standard test gives 100% correctness for the subjects whether they have the disease or not. Even though, some of the diagnostic tests might not discriminate subjects perfectly, they are widely used in clinics as they might be cost-effective, easy to apply, and rapid for diagnosis (1).

The types of variables in diagnostic tests might be nominal, ordinal, or quantitative. The distributions of the two groups which consisted of diseased and non-diseased subjects for ordinal and quantitative diagnostic tests overlapped since they are not gold standard tests. Thus, one of the approaches for statistical disease detection for ordinal and quantitative (continuous) diagnostic tests for a binary response is specifying an optimal cut-off point through several methods such as Youden's J statistic or cost-benefit method. By using this cut-off point, one can discriminate subjects whether they are diseased or not. However, this optimal value might not separate subjects such as classifying some diseased subjects as non-diseased or vice versa. Therefore, the decision about the status of the disease is uncertain on the overlapped range of distributions of diseased and non-diseased

groups. The issue is represented in Figure 1.1. Suppose the blue and red shaded areas represent the non-diseased and diseased subjects, respectively. From Figure 1.1., it might be seen the overlapped area of two distributions. The optimal cut-off point was pointed in Figure 1.1. as “C”. The decision about this overlapped area, even if the optimal cut-off point “C” is determined, will be uncertain. This overlapped area is not sufficient to discriminate whether the disease is absent or present.

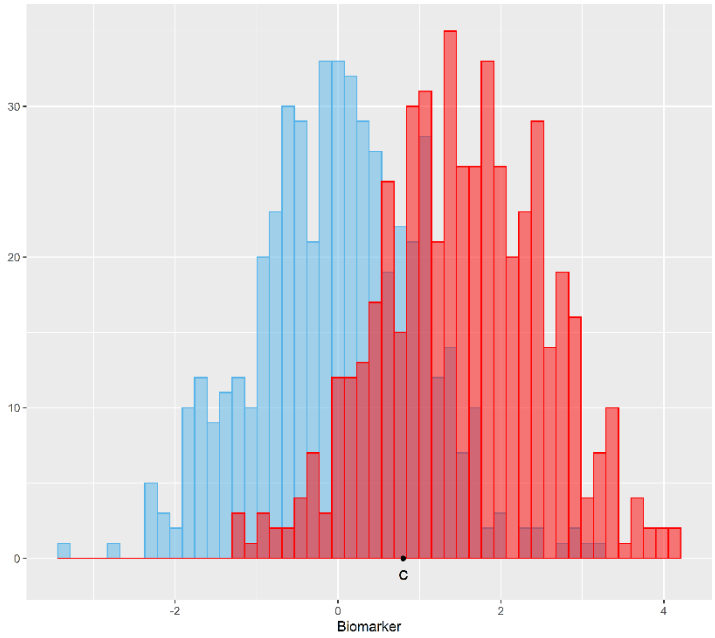



**Figure 1.1.** The Illustration of the Problem.

In order to deal with this uncertainty in this overlapped area, there exist different approaches in the literature. One of the approaches is based on specifying the boundaries of the middle inconclusive area, also known as the gray zone, is discussed. Moreover, there exist different methods to determine the boundaries of the gray zone such as by using likelihood ratios (2) or balancing true positive to false negative and true negatives to false positives within uncertain area (3).

This research mainly is concerned with constructing the boundaries of the gray zone for quantitative diagnostic tests. The main objectives of this thesis are:

- To propose a method for constructing the boundaries of the gray zone for quantitative diagnostic tests based on information theory

- To compare the performance of existing methods with the proposed ones.

The organization of this thesis is as follows:

- Chapter 1: The introduction and aims of this thesis were explained.
- Chapter 2: The approaches about the uncertain area were briefly reviewed.
- Chapter 3: The material and method were explained in detail.
- Chapter 4: The results were presented.
- Chapter 5: The discussion of the results was given.
- Chapter 6: The final remarks and future studies were given.

## 2. LITERATURE REVIEW

The binary status decision (diseased or non-diseased) in medicine might not be sufficient. Especially, in the overlapped area of the distributions of diseased and non-diseased subjects, forcing the subject as diseased or non-diseased might cause a false decision. To overcome this issue, there are different approaches in the literature.

One of the approach to deal with the binary status decision is related to fuzzy function and Receiver Operating Characteristics (ROC). The main aims of ROC analysis are finding the cut-off value to discriminate between healthy and diseased subjects and showing the test performances with respect to optimal cut-off point. The fuzzy functions concern the degree of membership. For instance, the subject may involve in either a diseased or non-diseased class. Therefore, this subject has partial membership and is between two classes. Campbell et al. (4) extend the ROC analysis to multi-labeled membership using a fuzzy algorithm called fuzzy ROC. Evangelista et al. (5) developed the fuzzy ROC for unsupervised learning-based nonparametric ensemble techniques. ROC analysis with the binary response is commonly used for evaluating diagnostic tests. The classic ROC methodology and the fuzzy sets theory are combined with a new approach known as a fuzzy-rule-based system. They implemented this algorithm to predict the pathological level of prostate cancer (6). The fuzzy ROC algorithm for visualizing the bounds of the fuzzy ROC and for examining the performance of the gray area are identified (7).

Another solution for the binary status decision is constructing a gray zone which indicates classifying subjects as diseased, non-diseased, and neither diseased nor non-diseased. The background and the approaches of the middle inconclusive area are presented in subsection 2.1 in details.

### 2.1. The Background of Gray Zone

There are several methods for handling the gray zone. Feinstein (8) has desired to highlight how only two zones specified by binary models—the person is

either classed as having a disease or not— are inadequate for making medical decisions. As a result, it was determined that creating three zones including an inconclusive zone was not unusual for clinical and statistical techniques. Moreover, the author has figured out the inconclusive area for diagnosing myocardial infarction with creatinine kinase by using the likelihood ratios and Bayesian approach (8). A new solution to get sensitivity and specificity that is written as chance-corrected sensitivity and specificity was proposed for the matrix with dimension 3x2 to take the middle zone into account (9). Simel et al. (10) present that likelihood ratios give a general methodology for binary, ordinal, and continuous diagnostic tests. Moreover, they highlight that for likelihood ratios, additional information about a patient might be used via logistic regression.

One of the earliest approaches to observe the intermediate region, the "two-graph receiver operating characteristic" (TG-ROC), was defined (11). It is a graphic that shows sensitivity and specificity plotted against the thresholds. This method reveals the two thresholds with pre-selected sensitivity and specificity (95% or 90%). Therefore, it guarantees that sensitivity and specificity are at least 95% or 90% (the other values of sensitivity and specificity may be observed graphically) outside of the intermediate region.

By using likelihood ratios, Coste and Pouchot (2) outline the boundaries of the "grey zone"[1] (GZ) in their investigation study. They also showed that using the pre-test probabilities to obtain post-test probabilities and likelihood ratios. Thus, they applied the tuberculin skin test to decide the starting antituberculous therapy in HIV-seropositive patients and the reticulocyte hemoglobin content test to diagnose iron deficiency in children. Coste et al. (12) showed the application of the gray zone to diagnose heart failure in acute dyspneic patients by using brain natriuretic peptide. The details of this approach are presented in Material and Methods.

[1] The term "grey zone" was adopted for Coste and Pouchot's technique (2) to avoid misunderstanding between it and "gray zone". As a result, the phrase "gray zone" which refers to the middle inconclusive area was used as general term.

An alternative approach in order to find out the inconclusive area is proposed by Landsheer (3) and is known as the "uncertain interval" (UI) approach. The two decision thresholds are used by the author to investigate this uncertain area and to present an alternative trichotomization strategy. Basically, this methodology is based on finding the boundaries of the gray zone while balancing true positive to false negative and true negative to false positive at the same time in the overlapping area of the distributions of diseased and non-diseased subjects. In this study, the predicted probabilities of risk of prostate cancer in the data (13) were obtained and the uncertain interval was specified based on these probabilities. The details of this approach are also presented in Material and Methods.

Another clinical application of gray zone was presented (14) to find out the uncertainty for the no-reflow phenomenon by using systemic inflammation index in patients with ST-elevation myocardial infarction admitted for primary percutaneous coronary intervention (PCI).

### 2.2. The Background of Information Criteria

The information theory is presented by Shannon (15) for communication purposes, such as data transmission rate and optimal data compression. Information theory is also one of the approaches for evaluating the performance of diagnostic tests. However, unlike other methods, the information provided by the diagnostic test is obtained from entropy (1). Surprisal, entropy, relative entropy, known as Kullback–Leibler divergence, and mutual information are common terms that are placed in the information theory. Benish (16) used relative entropy, which is the measure of the distance between two distributions, as a measure of diagnostic accuracy. Channel capacity in the diagnostic tests is a statement about the sensitivity and specificity of the diagnostic tests (17). It represents the maximum value of the mutual information and is used to evaluate the performance of the diagnostic tests. For the evaluation and comparison of diagnostic tests, information theory approach is beneficial for the researchers (18).

## 3. MATERIAL AND METHODS

The details of the "grey zone" approach (2), the "uncertain interval" approach (3), and the proposed algorithm by using information statistics were presented in this section. The simulation scenarios, including the data generation process and the properties of the real data set, were given in this section.

### 3.1. The Evaluation of Diagnostic Tests

In this section, some of the features and performance measures used in the evaluation of diagnostic tests are explained.

For the binary outcome of diagnostic tests, the results of diagnostic tests are presented in Table 3.1. Suppose D and Y represent the true disease status and the diagnostic test results, respectively.

$$D = \begin{cases} D=1, \text{ for the diseased subjects} \\ D=0, \text{ for the non-diseased subjects} \end{cases}$$

and

$$Y = \begin{cases} Y=1, \text{ for the positive test result} \\ Y=0, \text{ for the negative test result} \end{cases}$$

**Table 3.1.** The Cross Table of the disease status and diagnostic test.

| | D=0 | D=1 |
|---|---|---|
| Y=0 | True negative | False negative |
| Y=1 | False positive | True positive |

In order to evaluate the diagnostic tests, four decisions are appeared. True negative (TN) represents the negative test result among non-diseased subjects, while true positive (TP) represents the positive test result among diseased ones. The

negative test result among the diseased subjects and the positive test result among the non-diseased subjects are false negative (FN) and false positive (FP), respectively.

The common performance measures for the diagnostic tests are accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio, LR(+), and negative likelihood ratio, LR(-) (19).

Accuracy is the correct classification of subjects among all subjects in the data set.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (3.1.)$$

Sensitivity is the proportion of positive test results among the diseased subjects.

$$Sensitivity = P(Y = 1 \mid D = 1) = \frac{TP}{TP + FN} \qquad (3.2.)$$

Specificity is the proportion of the negative test results among the non-diseased subjects.

$$Specificity = P(Y = 0 \mid D = 0) = \frac{TN}{TN + FP} \qquad (3.3.)$$

Positive predictive value and negative predictive value are predictive values related to how accurately the test result reflects the true disease status. Positive predictive value is the proportion of diseased subjects among the positive test results.

$$PPV = P(D = 1 \mid Y = 1) = \frac{TP}{TP + FP} \qquad (3.4.)$$

The negative predictive value is the proportion of non-diseased subjects among the negative test results.

$$NPV = P(D=0 \mid Y=0) = \frac{TN}{TN+FN} \qquad (3.5.)$$

Another approach to describe the performance of the diagnostic test is to use likelihood ratios, which are the ratios of the likelihood of the observed test result in populations with and without disease.

$$LR(+) = \frac{P(Y=1 \mid D=1)}{P(Y=1 \mid D=0)} = \frac{Sensitivity}{1-Specificity} \qquad (3.6.)$$

$$LR(-) = \frac{P(Y=0 \mid D=1)}{P(Y=0 \mid D=0)} = \frac{1-Sensitivity}{Specificity} \qquad (3.7.)$$

The likelihood ratios can also be used for calculating post-test odds by using pre-test odds.

$$post\text{-}test\ odds\ for\ positive\ test\ result\ (Y=1) = pre\text{-}test\ odds \times LR(+) \qquad (3.8.)$$

$$post\text{-}test\ odds\ for\ negative\ test\ result\ (Y=0) = pre\text{-}test\ odds \times LR(-) \qquad (3.9.)$$

While accuracy, sensitivity, specificity, PPV, and NPV are probabilities, LR(+) and LR(-) are likelihood ratios whose scale is between 0 to ∞. LR(+) >1 suggests that a positive test is more likely to occur in a subject who is diseased than in a subject who is not diseased, while; LR(-)≤1 suggests that a negative test is more likely to occur in a non-diseased subject than in a diseased subject.

In addition to above-mentioned performance measures, information theory is used to evaluate diagnostic test performance. Some basic definitions of the terms belonging to information theory are explained in this subsection. The surprisal function, u, measures how unlikely an event is to occur (18). Assume that $d_1$, $d_2$ ,…,$d_k$ and $p_i$ are the mutually exclusive true diagnoses of the diseases and the probability of the true diagnosis of the $d_i$.

$$u_i = -\log_a p_i; \qquad (3.10.)$$

where i is the interested disease. In Equation 3.10, the base of the logarithm (a) can be chosen arbitrarily since it can be changed by multiplying by a constant. However, most of the time, it is specified as two, which provides measurements in units of bits (binary digits). It can be written as in Equation 3.11.

$$u_i = -\log_2 p_i; \qquad (3.11.)$$

where i is the event. The entropy which is represented as H(D) is the expected value of the surprisal. H gives a measurement of a diagnostic' degree of uncertainty (16).

$$H(D) = -\sum_{i=1}^{k} p_i \times \log_2 p_i \qquad (3.12.)$$

The joint entropy is the entropy of the joint probability distribution (20). Let $d_1, d_2, \ldots, d_k$ and $y_1, y_2, \ldots, y_m$ are the mutually exclusive true diagnosis of the diseases and the diagnosis of disease based on test results. Moreover, $p(d_i, y_j)$ shows the joint probability of the d and y.

$$H(D,Y) = -\sum_{j=1}^{m}\sum_{i=1}^{k} p(d_i, y_j) \times \log_2 p(d_i, y_j) \qquad (3.13.)$$

### 3.2. The Gray Zone Approaches

#### 3.2.1. The Grey Zone Approach[1]

The boundaries of the middle inconclusive area are defined by using positive LR(+) and LR(-) - in grey zone approach (2). The algorithm of this approach is finding the boundaries of the gray zone based on pre-selected LR(+) and LR(-) values. Based on the post-test probabilities or sensitivity and specificity of the outside of the gray

[1] The name "grey zone" was used to Coste and Pouchot's technique to avoid misunderstanding between it and "gray zone" (2). As a result, the phrase "gray zone"—which refers to the uncertain area—was established.

zone may be determined and LR(+) and LR(-) are calculated using Equations 3.6., 3.7., 3.8., and 3.9. The algorithm for the grey zone approach is based on pre-selected sensitivity and specificity of the outside of the gray zone:

**Step 1:** Specify the sensitivity and specificity of the diagnostic test out of the gray zone.

**Step 2:** Calculate the LR(+) and LR(-) with respect to Equations 3.6 and 3.7.

**Step 3:** Find the LR(+) and LR(-) for cut-off values which is the sorted diagnostic test values from lowest to highest.

**Step 4:** Find the closest values in Step 2.

The algorithm for the grey zone approach based on the post-test odds for the outside of the gray zone:

**Step 1:** Give the pre-test probability to calculate the pre-test odds. If there is no information for pre-test probabilities, the prevalence of the diseases may be considered as pre-test probability.

**Step 2:** Provide the post-test probabilities to calculate post-test odds.

**Step 3:** Calculate the LR(+) and LR(-) with respect to Equations 3.8. and 3.9.

**Step 4:** Find the LR(+) and LR(-) for cut-off values which is the sorted diagnostic test values from lowest to highest.

**Step 5:** Find the closest values in Step 3.

### 3.2.2. Uncertain Interval Approach

The inconclusive area has been identified by using a different method (3). By utilizing the two decision thresholds based on previously chosen levels of sensitivity and specificity in this middle inconclusive area, this strategy relies on a different trichotomization mechanism. The algorithm of this approach is basically based on the

balancing true negative to false positive rates and true positive to false negative rates in the gray zone which is illustrated in Figure 3.1.

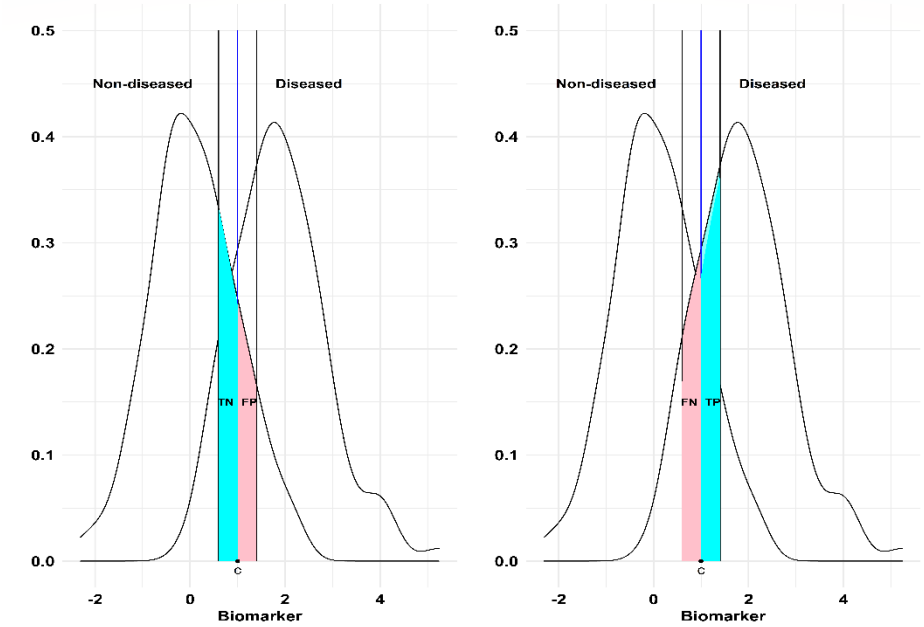



**Figure 3.1.** The Visualization of Uncertain Interval Approach.

The steps of the uncertain interval approach as follows:

**Step 1:** Find the intersection of two distributions by Youden's J statistic or kernel density estimation (In this study, for the uncertain interval algorithm, the intersection was specified with kernel density estimation since it was stated that kernel estimation is slightly more useful than Youden's J index (3).

**Step 2:** Based on the pre-selected values of sensitivity and specificity in the gray zone, specify the ratio of true negative to false positive and true positive to false negative.

**Step 3:** By using sequential least squares programming optimizer (SLSQP) which is nonlinearly constrained gradient-based optimization find the areas simultaneously around the intersection with the true negative balanced by false positive and the true positive balanced by false negative.

### 3.2.3. Proposed Algorithm of Uncertain Boundaries with Joint Entropy

The proposed algorithm is based on joint entropy. The proposed algorithm was named a "**J**oint **E**ntropy based algorithm for **U**ncertain **B**oundaries" and shortened to "**JE-UB**". The algorithm was constructed as follows:

**Step 1:** Find the initial values for the lower and upper boundaires of the gray zone that is based on the overlapped area of diseased and non-diseased distributions:

- $c_{01}$ : the minimum value of the overlapped area
- $c_{02}$ : the maximum value of the overlapped area

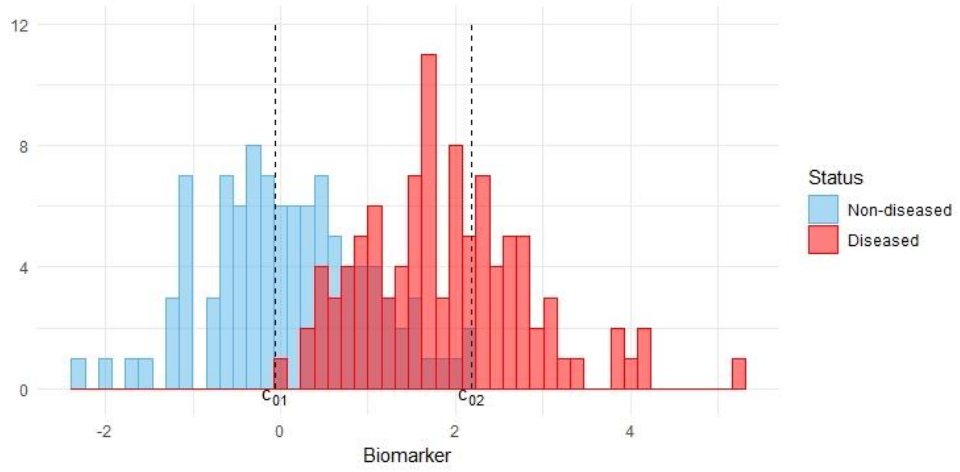



**Figure 3.2.** The First Step of the JE-UB Algorithm.

**Step 2:** Find the intersection point of two distributions whether maximizing the Youden's J statistic based on kernel smoothing or the Euclidean distance of observations[1] in $c_{01}$ and $c_{02}$ : C

The Youden's J statistic based on kernel smoothed densities was chosen for determining the intersection point since, in different scenarios under different

[1] In order to differentiate which method was used for finding the cut-off point in Step 2, using **K**ernel **S**moothing for Youden's J statistic and the **E**uclidean **D**istance of observations are shortened and added to "**JE-UB-KB**" and "**JE-UB-ED**", respectively.

distributions, the Youden's J statistic based on kernel smoothed densities showed good properties (21).

The Euclidean distances from one observation to another were calculated. The summation of these distances was taken, and the overall distance measure in each observation was constructed. The point where this sum was at its minimum was accepted as the intersection. Suppose p and q are the two points in the Euclidian n space; the formulation for this distance (d) is (22):

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \qquad (3.14.)$$

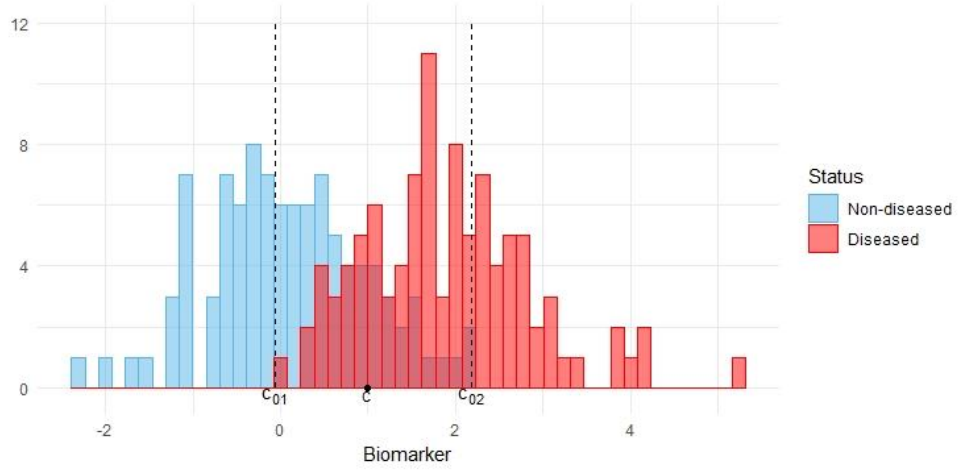



**Figure 3.3.** The Second Step of the JE-UB Algorithm.

**Step 3:** Find the point between $c_{01}$ and C which gives the maximum value of joint entropy: $c_L$

**Step 4:** Find the point between $c_{02}$ and C which gives the maximum value of joint entropy: $c_U$

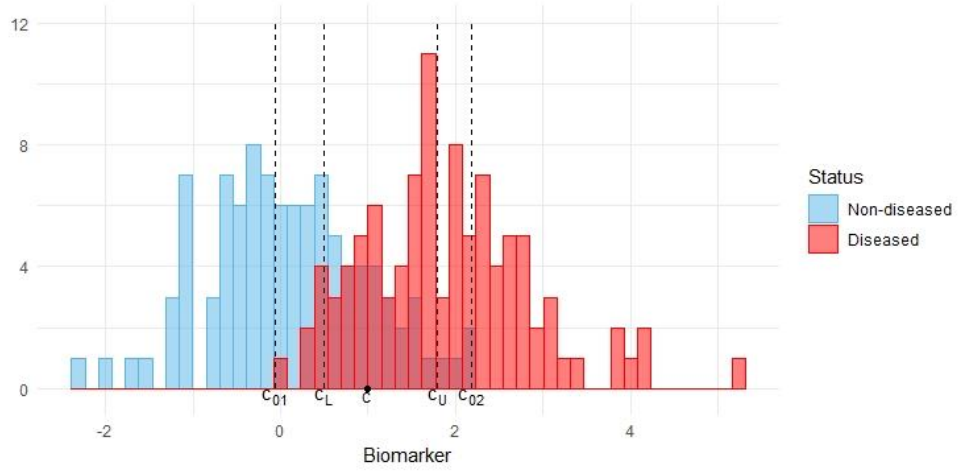



**Figure 3.4.** The General Framework of JE-UB Algorithm.

### 3.3. The Simulation Studies

In order to compare the JE-UB algorithm with the existing algorithms, a comprehensive simulation study was conducted by using the R language environment. Suppose that $Y_C$ and $Y_D$ are the distributions of diagnostic tests of the non-diseased group and diseased group, respectively. The factors which are considered for the simulation scenarios are sample size (n which is the total sample sizes of $Y_C$ and $Y_D$), the prevalence of the disease, P(D), and the ratio of variances (the ratio of variances of a diseased group to non-diseased group). Thus, in order to reveal how much the distributions are close to each other or how much the distributions are overlapped, the effect size (d) of $Y_C$ and $Y_D$ was used. The results of simulation scenarios, which consist of all possible combinations, are evaluated:

1. Sample size (n): 50, 100, and 200
2. Prevalence of the disease (P(D)): 0.2, 0.5, and 0.8
3. The ratio of the variances: 1 (equal variances) and 3 (unequal variances) (23)
4. Effect size (d): 0.5 (medium), 0.8 (large), and 1.2 (large) (24)

For all the combinations of total of 54 scenarios, the distribution of $Y_C$ was distributed as a normal distribution with 0 mean and 1 standard deviation. In addition to those scenarios, for the prevalence of the disease is 0.8 and the ratio of variances

3, in order to prevent the underrepresentation of the distribution of $Y_C$, the prevalence of the disease is 0.8 and the ratio of variances 1/3 is also considered. Therefore, total of 63 scenarios were evaluated. Each simulation scenario was repeated 1000 times.

One of the difficult issues for the comparisons of the gray zone is if the length of the gray zone (the difference between upper and lower boundaries of the gray zone) increases, then the out-of-the-performance measures (such as accuracy, sensitivity, and specificity) will increase or vice versa. Therefore, the question has arisen, whether the narrower length of the gray zone is important or the out of the gray zone is important. For this issue, in simulation studies, the area under receiver operating characteristics curve (AUROC) out of the gray zone was fixed with a 1% interval, and the gray zone's narrowest length was chosen as the best algorithm.

For each effect size considered in the simulation study, the AUROC values were chosen differently. For effect sizes equal to 0.5, 0.8, and 1.2, the AUROC was specified as 0.69, 0.77, and 0.85 with 1% intervals based on the relationship between effect size and AUROC (25). The AUROC values are approximately, 0.64, 0.72, and 0.8. Therefore, the AUROC was fixed higher than 5% of those values with a 1% interval. For the performance measures, the boundaries and the length of the gray zone, the percentage of subjects that are in gray zone, the AUROC, accuracy, sensitivity, and specificity of the outside the gray zone were evaluated which is illustrated in Table 3.2.

**Table 3.2.** The Values for AUROC with Effect Sizes.

| Effect Size | Interval of AUROC |
|---|---|
| d=0.5 | 68 % to 70 % |
| d=0.8 | 76 % to 78 % |
| d=1.2 | 84 % to 86 % |

For the uncertain interval algorithm, the optimum boundaries were searched between 0.51 to 0.65 of sensitivity and specificity in the gray zone. For the multiple boundaries in the interval of AUROC, the boundaries of the gray zone were chosen based on the minimum length of the gray zone. For the grey zone algorithm, the optimum boundaries were started to search for the sensitivity and specificity optimal cut-off point based on Youden's J index to the value of sensitivity and specificity of the out of the gray zone 1. Then the likelihood ratios were calculated based on Equations 3.6 and 3.7. For multiple boundaries, the same structure of uncertain intervals is followed. For the JE-UB algorithm, for the multiple boundaries, the boundaries were selected with respect to the maximum joint entropy.

All analysis was conducted on R environment (26). The distributions of $Y_C$ and $Y_D$ were generated with "stats" package (26) in R. Thus, the "UncertainInterval" package (27) was used for the grey zone and uncertain interval approaches. The performance measures were obtained by using the "caret" package (28).

### 3.4. The Real Data Set

To reveal the application of the real data set, the one of the versions of the Pima Indian Data Set was used. The National Institute of Diabetes and Digestive and Kidney Diseases is the original source of this data set (29). The main objective of this data set is to predict whether a patient has diabetes or not. All patients at this facility are Pima Indian women who are at least 21 years old. The data set in this thesis is obtained from "MASS" package (30) in R which is the small part of this larger data set. The variables of the data set are: number of pregnancy, age, diastolic blood pressure, triceps skin fold thickness, body mass index, diabetes pedigree function, glucose and class (a patient has the diabetes or not). For this study, we consider glucose, and class variables to find the boundaries of the gray zone.

# 4. RESULTS

The results of both simulation studies and real data sets are presented in this section.

## 4.1. The Simulation Results

The mean values of the boundaries and length of the gray zone, the percentage of the subject who is classified as gray zone and the percentage of diseased and non-diseased groups, the AUROC, accuracy, sensitivity, and specificity of the out of the gray zone of the 1000 replications were given in Tables 4.1., 4.2., 4.3., and 4.4. For Figures 4.1., 4.2., and 4.3.[1], the y-axis shows the length of the gray zone while in the x-axis, the methods were placed. Thus, the sample sizes are represented with the lines while the prevalence of the disease were split in grids. For Figures 4.4.,4.5., and 4.6., the same structure is followed except for the prevalence of the disease was fixed as 0.8 so there is no grid. All Figures were generated by using the “ggplot2” package (31) in R.

[1] For Tables 4.1., 4.2., 4.3., and 4.4., and Figures 4.1., 4.2., 4.3., 4.4., 4.5., and 4.6., the uncertain interval, grey zone, joint entropy based algorithm for uncertain boundaries with kernel smoothed densities for Youden’s j index, and joint entropy based algorithm for uncertain boundaries with the Euclidean distance of observations were represented as "UI", "GZ", "JE-UB-KB" and "JE-UB-ED". For Tables 4.1., 4.2., 4.3., and 4.4., LL and UL shows the lower and upper limit of the gray zone, respectively.

**Table 4.1.** The Simulation Results for d=0.5.

| P(D) | The Ratio of Variances | n | Methods | Boundaries | | Proportion of Gray Zone | | | Out of in the Gray Zone | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | LL-UL | Length | Total % | Diseased % | Control % | AUROC | Accuracy | Sensitivity | Specificity |
| P(D)=0.2 | Equal Variances | 50 | UI | -0.451 - 0.821 | 1.272 | 41.972 | 19.913 | 80.087 | 0.681 | 0.637 | 0.757 | 0.606 |
| | | | GZ | -0.264 - 1.036 | 1.300 | 41.772 | 21.196 | 78.804 | 0.687 | 0.711 | 0.666 | 0.707 |
| | | | JE-UB-KB | -0.173 - 0.864 | 1.038 | 37.376 | 20.853 | 79.147 | 0.689 | 0.700 | 0.673 | 0.705 |
| | | | JE-UB-ED | -0.222 - 0.604 | 0.827 | 31.366 | 20.831 | 79.169 | 0.687 | 0.652 | 0.748 | 0.627 |
| | | 100 | UI | -0.500 - 0.984 | 1.484 | 47.389 | 20.279 | 79.721 | 0.680 | 0.657 | 0.721 | 0.639 |
| | | | GZ | -0.439 - 1.102 | 1.540 | 49.064 | 20.392 | 79.608 | 0.685 | 0.701 | 0.670 | 0.700 |
| | | | JE-UB-KB | -0.265 - 0.833 | 1.098 | 38.665 | 20.618 | 79.382 | 0.686 | 0.678 | 0.699 | 0.672 |
| | | | JE-UB-ED | -0.305 - 0.653 | 0.957 | 34.672 | 20.726 | 79.274 | 0.684 | 0.643 | 0.751 | 0.617 |
| | | 200 | UI | -0.655 - 1.146 | 1.801 | 56.222 | 20.065 | 79.935 | 0.678 | 0.665 | 0.704 | 0.652 |
| | | | GZ | -0.641 - 1.295 | 1.936 | 60.153 | 20.321 | 79.679 | 0.685 | 0.708 | 0.662 | 0.709 |
| | | | JE-UB-KB | -0.406 - 0.897 | 1.304 | 44.485 | 20.327 | 79.673 | 0.683 | 0.667 | 0.710 | 0.656 |
| | | | JE-UB-ED | -0.435 - 0.819 | 1.254 | 43.111 | 20.421 | 79.579 | 0.682 | 0.649 | 0.737 | 0.627 |
| | Unequal Variances | 50 | UI | 0.430 - 1.600 | 1.170 | 26.926 | 21.726 | 78.274 | 0.680 | 0.813 | 0.465 | 0.895 |
| | | | GZ | 0.107 - 1.510 | 1.403 | 35.610 | 19.770 | 80.230 | 0.686 | 0.788 | 0.544 | 0.829 |
| | | | JE-UB-KB | -0.231 - 1.074 | 1.304 | 40.780 | 15.562 | 84.438 | 0.690 | 0.716 | 0.647 | 0.732 |
| | | | JE-UB-ED | -0.175 - 0.740 | 0.915 | 31.226 | 16.204 | 83.796 | 0.687 | 0.678 | 0.700 | 0.674 |
| | | 100 | UI | 0.591 - 1.650 | 1.059 | 22.539 | 20.573 | 79.427 | 0.675 | 0.828 | 0.423 | 0.927 |
| | | | GZ | 0.146 - 1.692 | 1.546 | 36.512 | 19.402 | 80.598 | 0.683 | 0.812 | 0.497 | 0.870 |
| | | | JE-UB-KB | -0.306 - 1.172 | 1.478 | 43.177 | 14.802 | 85.198 | 0.688 | 0.720 | 0.627 | 0.748 |
| | | | JE-UB-ED | -0.185 - 0.890 | 1.075 | 33.798 | 15.386 | 84.614 | 0.684 | 0.693 | 0.657 | 0.711 |
| | | 200 | UI | 0.580 - 1.723 | 1.143 | 23.723 | 20.328 | 79.672 | 0.670 | 0.832 | 0.401 | 0.939 |
| | | | GZ | 0.149 - 1.705 | 1.556 | 36.112 | 19.470 | 80.530 | 0.681 | 0.815 | 0.485 | 0.877 |
| | | | JE-UB-KB | -0.250 - 1.234 | 1.484 | 43.128 | 14.913 | 85.087 | 0.685 | 0.731 | 0.592 | 0.779 |
| | | | JE-UB-ED | -0.181 - 1.065 | 1.246 | 37.514 | 15.396 | 84.604 | 0.682 | 0.715 | 0.610 | 0.753 |

**Table 4.1. (Continued).** The Simulation Results for d=0.5.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| P(D)=0.5 | Equal Variances | 50 | UI | -0.468 - 0.904 | 1.373 | 44.498 | 49.930 | 50.070 | 0.678 | 0.680 | 0.702 | 0.654 |
| | | | GZ | -0.439 - 0.922 | 1.361 | 45.642 | 49.634 | 50.366 | 0.686 | 0.693 | 0.693 | 0.679 |
| | | | JE-UB-KB | -0.252 - 0.777 | 1.029 | 38.220 | 49.492 | 50.508 | 0.688 | 0.686 | 0.683 | 0.692 |
| | | | JE-UB-ED | -0.201 - 0.737 | 0.937 | 35.582 | 49.424 | 50.576 | 0.686 | 0.684 | 0.680 | 0.692 |
| | | 100 | UI | -0.575 - 1.069 | 1.645 | 52.351 | 49.997 | 50.003 | 0.677 | 0.680 | 0.682 | 0.672 |
| | | | GZ | -0.550 - 1.103 | 1.653 | 53.985 | 50.075 | 49.925 | 0.685 | 0.692 | 0.672 | 0.698 |
| | | | JE-UB-KB | -0.323 - 0.848 | 1.172 | 41.338 | 50.075 | 49.925 | 0.684 | 0.683 | 0.681 | 0.688 |
| | | | JE-UB-ED | -0.293 - 0.829 | 1.122 | 39.952 | 50.058 | 49.942 | 0.684 | 0.682 | 0.677 | 0.690 |
| | | 200 | UI | -0.698 - 1.205 | 1.902 | 59.641 | 49.971 | 50.029 | 0.678 | 0.683 | 0.677 | 0.680 |
| | | | GZ | -0.748 - 1.251 | 1.999 | 62.799 | 49.914 | 50.086 | 0.684 | 0.695 | 0.684 | 0.684 |
| | | | JE-UB-KB | -0.458 - 0.956 | 1.414 | 48.174 | 49.894 | 50.106 | 0.682 | 0.682 | 0.683 | 0.682 |
| | | | JE-UB-ED | -0.452 - 0.952 | 1.404 | 47.891 | 49.903 | 50.097 | 0.682 | 0.682 | 0.682 | 0.682 |
| | Unequal Variances | 50 | UI | 0.508 - 1.731 | 1.223 | 27.182 | 49.570 | 50.430 | 0.670 | 0.669 | 0.396 | 0.943 |
| | | | GZ | 0.066 - 1.496 | 1.429 | 35.656 | 47.280 | 52.720 | 0.684 | 0.686 | 0.508 | 0.861 |
| | | | JE-UB-KB | -0.319 - 1.145 | 1.464 | 40.718 | 42.340 | 57.660 | 0.687 | 0.683 | 0.614 | 0.761 |
| | | | JE-UB-ED | -0.173 - 1.050 | 1.223 | 35.346 | 42.030 | 57.970 | 0.686 | 0.678 | 0.600 | 0.772 |
| | | 100 | UI | 0.543 - 1.728 | 1.185 | 25.686 | 49.350 | 50.650 | 0.671 | 0.670 | 0.389 | 0.952 |
| | | | GZ | 0.205 - 1.612 | 1.407 | 32.944 | 48.540 | 51.460 | 0.682 | 0.684 | 0.466 | 0.898 |
| | | | JE-UB-KB | -0.225 - 1.205 | 1.430 | 38.104 | 41.423 | 58.577 | 0.685 | 0.674 | 0.580 | 0.790 |
| | | | JE-UB-ED | -0.081 - 1.104 | 1.185 | 32.803 | 42.088 | 57.912 | 0.683 | 0.670 | 0.569 | 0.796 |
| | | 200 | UI | 0.561 - 1.715 | 1.154 | 24.795 | 49.752 | 50.248 | 0.671 | 0.670 | 0.392 | 0.949 |
| | | | GZ | 0.242 - 1.586 | 1.343 | 31.512 | 48.409 | 51.591 | 0.681 | 0.683 | 0.466 | 0.896 |
| | | | JE-UB-KB | -0.110 - 1.211 | 1.322 | 34.998 | 41.429 | 58.571 | 0.683 | 0.670 | 0.560 | 0.806 |
| | | | JE-UB-ED | -0.056 - 1.132 | 1.188 | 32.163 | 41.817 | 58.183 | 0.681 | 0.668 | 0.562 | 0.800 |

**Table 4.1. (Continued).** The Simulation Results for d=0.5.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| P(D)=0.8 | Equal Variances | 50 | UI | -0.372 - 0.872 | 1.244 | 40.414 | 79.532 | 20.468 | 0.679 | 0.661 | 0.648 | 0.710 |
| | | | GZ | -0.441 - 0.675 | 1.116 | 37.870 | 78.579 | 21.421 | 0.687 | 0.713 | 0.721 | 0.653 |
| | | | JE-UB-KB | -0.346 - 0.719 | 1.065 | 38.648 | 78.985 | 21.015 | 0.689 | 0.699 | 0.703 | 0.676 |
| | | | JE-UB-ED | -0.062 - 0.765 | 0.827 | 32.002 | 78.664 | 21.336 | 0.687 | 0.646 | 0.619 | 0.755 |
| | | 100 | UI | -0.528 - 0.994 | 1.522 | 48.209 | 79.776 | 20.224 | 0.676 | 0.667 | 0.658 | 0.695 |
| | | | GZ | -0.589 - 0.943 | 1.531 | 49.382 | 79.685 | 20.315 | 0.686 | 0.702 | 0.704 | 0.669 |
| | | | JE-UB-KB | -0.323 - 0.822 | 1.145 | 40.291 | 79.494 | 20.506 | 0.686 | 0.673 | 0.664 | 0.709 |
| | | | JE-UB-ED | -0.160 - 0.864 | 1.024 | 36.769 | 79.243 | 20.757 | 0.684 | 0.639 | 0.610 | 0.758 |
| | | 200 | UI | -0.528 - 0.994 | 1.522 | 24.105 | 79.776 | 20.224 | 0.676 | 0.667 | 0.658 | 0.695 |
| | | | GZ | -0.589 - 0.943 | 1.531 | 24.691 | 79.685 | 20.315 | 0.686 | 0.702 | 0.704 | 0.669 |
| | | | JE-UB-KB | -0.323 - 0.822 | 1.145 | 20.146 | 79.494 | 20.506 | 0.686 | 0.673 | 0.664 | 0.709 |
| | | | JE-UB-ED | -0.160 - 0.864 | 1.024 | 18.385 | 79.243 | 20.757 | 0.684 | 0.639 | 0.610 | 0.758 |
| | Unequal Variances | 50 | UI | 0.361 - 1.745 | 1.384 | 31.224 | 80.195 | 19.805 | 0.676 | 0.518 | 0.412 | 0.941 |
| | | | GZ | -0.166 - 1.124 | 1.290 | 30.396 | 76.747 | 23.253 | 0.685 | 0.635 | 0.597 | 0.774 |
| | | | JE-UB-KB | -0.366 - 0.959 | 1.325 | 32.756 | 74.954 | 25.046 | 0.688 | 0.667 | 0.652 | 0.724 |
| | | | JE-UB-ED | -0.031 - 1.040 | 1.071 | 27.430 | 74.896 | 25.104 | 0.687 | 0.623 | 0.584 | 0.789 |
| | | 100 | UI | 0.506 - 1.737 | 1.231 | 26.847 | 80.020 | 19.980 | 0.674 | 0.503 | 0.388 | 0.960 |
| | | | GZ | 0.114 - 1.391 | 1.277 | 29.190 | 78.434 | 21.566 | 0.684 | 0.577 | 0.504 | 0.864 |
| | | | JE-UB-KB | -0.198 - 1.100 | 1.298 | 31.120 | 73.429 | 26.571 | 0.686 | 0.625 | 0.589 | 0.782 |
| | | | JE-UB-ED | 0.077 - 1.100 | 1.024 | 25.346 | 73.538 | 26.462 | 0.684 | 0.592 | 0.542 | 0.826 |
| | | 200 | UI | 0.542 - 1.727 | 1.185 | 25.724 | 79.774 | 20.226 | 0.674 | 0.500 | 0.385 | 0.962 |
| | | | GZ | 0.263 - 1.514 | 1.251 | 28.180 | 79.354 | 20.646 | 0.682 | 0.551 | 0.463 | 0.902 |
| | | | JE-UB-KB | -0.061 - 1.166 | 1.227 | 28.941 | 73.497 | 26.503 | 0.683 | 0.601 | 0.555 | 0.812 |
| | | | JE-UB-ED | 0.135 - 1.137 | 1.002 | 24.322 | 73.686 | 26.314 | 0.682 | 0.581 | 0.526 | 0.838 |

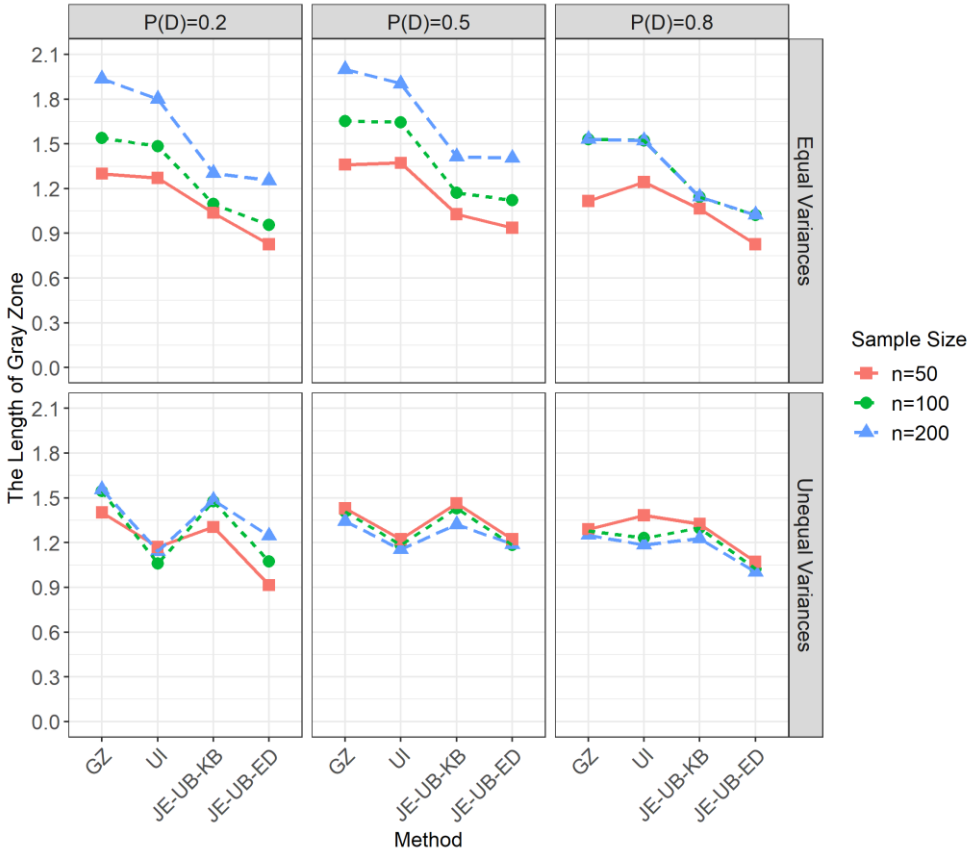



**Figure 4.1.** The Simulation Results for d=0.5.

For the effect size 0.5, the results of simulation scenarios were presented in Table 4.1. and Figure 4.1. With respect to similar performances out of the gray zone of the four methods,

- For the equal variances, the proposed algorithm which is **J**oint **E**ntropy based algorithm for **U**ncertain **B**oundaries with both **K**ernel **S**moothed densities for Youden's J statistic (JE-UB-KB) and **E**uclidian **D**istance (JE-UB-ED) gives the smallest length of the gray zone for sample sizes 50, 100, and 200 and prevalence of the disease 0.2, 0.5 and 0.8.

- For the unequal variances, for the prevalence of the disease 0.2 for the sample size 50, the JE-UB-ED algorithm gives the minimum length of the gray zone while for the sample sizes 100 and 200, the uncertain interval algorithm gives the minimum length of the gray zone.

- For the unequal variances, for the prevalence of the disease 0.5 for the sample sizes 50 and 100, the JE-UB-ED algorithm and uncertain interval gives the minimum length of the gray zone while for the sample size 200, the uncertain interval algorithm gives the minimum length of the gray zone.

- For the unequal variances, for the prevalence of the disease 0.8 for all sample sizes the JE-UB-ED algorithm and uncertain interval gives the minimum length of the gray zone.

**Table 4.2.** The Simulation Results for d=0.8.

| P(D) | The Ratio of Variances | n | Methods | Boundaries | | Proportion of Gray Zone | | | Out of in the Gray Zone | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | LL-UL | Length | Total % | Diseased % | Control % | AUROC | Accuracy | Sensitivity | Specificity |
| P(D)=0.2 | Equal Variances | 50 | UI | -0.123 - 1.005 | 1.127 | 37.064 | 19.998 | 80.002 | 0.762 | 0.752 | 0.780 | 0.744 |
| | | | GZ | -0.091 - 1.140 | 1.231 | 39.282 | 20.467 | 79.533 | 0.766 | 0.784 | 0.747 | 0.785 |
| | | | JE-UB-KB | -0.054 - 1.007 | 1.061 | 36.966 | 20.035 | 79.965 | 0.768 | 0.772 | 0.761 | 0.775 |
| | | | JE-UB-ED | -0.119 - 0.802 | 0.922 | 33.688 | 19.473 | 80.527 | 0.765 | 0.736 | 0.816 | 0.715 |
| | | 100 | UI | -0.292 - 1.130 | 1.423 | 45.915 | 20.216 | 79.784 | 0.761 | 0.750 | 0.780 | 0.742 |
| | | | GZ | -0.323 - 1.225 | 1.548 | 49.166 | 20.256 | 79.744 | 0.767 | 0.772 | 0.765 | 0.768 |
| | | | JE-UB-KB | -0.200 - 1.000 | 1.200 | 41.319 | 20.141 | 79.859 | 0.765 | 0.749 | 0.792 | 0.739 |
| | | | JE-UB-ED | -0.251 - 0.893 | 1.143 | 39.863 | 20.237 | 79.763 | 0.764 | 0.726 | 0.825 | 0.702 |
| | | 200 | UI | -0.450 - 1.264 | 1.713 | 54.306 | 20.125 | 79.875 | 0.762 | 0.755 | 0.775 | 0.749 |
| | | | GZ | -0.485 - 1.355 | 1.839 | 57.257 | 20.275 | 79.725 | 0.766 | 0.774 | 0.762 | 0.770 |
| | | | JE-UB-KB | -0.333 - 1.051 | 1.383 | 46.770 | 20.096 | 79.904 | 0.763 | 0.738 | 0.804 | 0.722 |
| | | | JE-UB-ED | -0.352 - 1.014 | 1.367 | 46.381 | 20.170 | 79.830 | 0.762 | 0.729 | 0.817 | 0.707 |
| | Unequal Variances | 50 | UI | 0.522 - 1.585 | 1.063 | 24.984 | 20.157 | 79.843 | 0.747 | 0.850 | 0.577 | 0.918 |
| | | | GZ | 0.033 - 1.560 | 1.527 | 39.522 | 17.929 | 82.071 | 0.766 | 0.831 | 0.669 | 0.863 |
| | | | JE-UB-KB | -0.183 - 1.199 | 1.382 | 42.094 | 15.508 | 84.492 | 0.768 | 0.781 | 0.745 | 0.791 |
| | | | JE-UB-ED | -0.155 - 1.025 | 1.180 | 37.976 | 15.915 | 84.085 | 0.766 | 0.760 | 0.772 | 0.761 |
| | | 100 | UI | 0.557 - 1.657 | 1.100 | 24.427 | 20.404 | 79.596 | 0.744 | 0.859 | 0.553 | 0.935 |
| | | | GZ | -0.005 - 1.671 | 1.676 | 41.411 | 17.877 | 82.123 | 0.764 | 0.840 | 0.654 | 0.875 |
| | | | JE-UB-KB | -0.227 - 1.227 | 1.454 | 42.772 | 14.872 | 85.128 | 0.766 | 0.779 | 0.738 | 0.793 |
| | | | JE-UB-ED | -0.205 - 1.119 | 1.324 | 39.995 | 15.429 | 84.571 | 0.764 | 0.765 | 0.754 | 0.773 |
| | | 200 | UI | 0.512 - 1.748 | 1.237 | 26.696 | 20.153 | 79.847 | 0.737 | 0.861 | 0.529 | 0.944 |
| | | | GZ | -0.170 - 1.673 | 1.843 | 45.969 | 17.110 | 82.890 | 0.761 | 0.832 | 0.668 | 0.855 |
| | | | JE-UB-KB | -0.252 - 1.336 | 1.588 | 45.321 | 15.069 | 84.931 | 0.763 | 0.787 | 0.710 | 0.816 |
| | | | JE-UB-ED | -0.246 - 1.301 | 1.547 | 44.445 | 15.251 | 84.749 | 0.762 | 0.782 | 0.716 | 0.808 |

**Table 4.2. (Continued).** The Simulation Results for d=0.8.

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P(D)=0.5 | Equal Variances | 50 | UI | -0.291 - 1.052 | 1.343 | 44.356 | 50.023 | 49.977 | 0.757 | 0.757 | 0.767 | 0.746 |
| | | | GZ | -0.305 - 1.139 | 1.444 | 47.946 | 50.065 | 49.935 | 0.767 | 0.771 | 0.761 | 0.772 |
| | | | JE-UB-KB | -0.168 - 1.012 | 1.180 | 42.050 | 49.389 | 50.611 | 0.767 | 0.765 | 0.756 | 0.777 |
| | | | JE-UB-ED | -0.145 - 0.992 | 1.137 | 40.982 | 49.470 | 50.530 | 0.766 | 0.765 | 0.756 | 0.777 |
| | | 100 | UI | -0.384 - 1.190 | 1.574 | 50.618 | 49.949 | 50.051 | 0.760 | 0.761 | 0.761 | 0.760 |
| | | | GZ | -0.432 - 1.248 | 1.681 | 53.960 | 49.855 | 50.145 | 0.766 | 0.771 | 0.763 | 0.769 |
| | | | JE-UB-KB | -0.243 - 1.066 | 1.309 | 44.559 | 49.815 | 50.185 | 0.764 | 0.764 | 0.760 | 0.769 |
| | | | JE-UB-ED | -0.233 - 1.062 | 1.295 | 44.189 | 49.847 | 50.153 | 0.764 | 0.764 | 0.758 | 0.770 |
| | | 200 | UI | -0.505 - 1.308 | 1.813 | 57.469 | 49.946 | 50.054 | 0.762 | 0.765 | 0.763 | 0.762 |
| | | | GZ | -0.591 - 1.379 | 1.970 | 61.257 | 49.772 | 50.228 | 0.766 | 0.775 | 0.769 | 0.763 |
| | | | JE-UB-KB | -0.323 - 1.137 | 1.461 | 49.084 | 49.997 | 50.003 | 0.762 | 0.763 | 0.759 | 0.766 |
| | | | JE-UB-ED | -0.322 - 1.137 | 1.460 | 49.063 | 49.991 | 50.009 | 0.762 | 0.763 | 0.758 | 0.766 |
| | Unequal Variances | 50 | UI | 0.455 - 1.694 | 1.239 | 29.036 | 49.814 | 50.186 | 0.742 | 0.742 | 0.539 | 0.945 |
| | | | GZ | -0.075 - 1.544 | 1.619 | 40.412 | 45.615 | 54.385 | 0.764 | 0.763 | 0.657 | 0.870 |
| | | | JE-UB-KB | -0.206 - 1.267 | 1.473 | 40.306 | 42.420 | 57.580 | 0.766 | 0.761 | 0.713 | 0.819 |
| | | | JE-UB-ED | -0.125 - 1.249 | 1.373 | 38.284 | 42.566 | 57.434 | 0.765 | 0.759 | 0.701 | 0.829 |
| | | 100 | UI | 0.481 - 1.740 | 1.259 | 28.165 | 49.988 | 50.012 | 0.744 | 0.744 | 0.531 | 0.957 |
| | | | GZ | 0.002 - 1.614 | 1.612 | 38.589 | 45.536 | 54.464 | 0.762 | 0.762 | 0.634 | 0.891 |
| | | | JE-UB-KB | -0.126 - 1.284 | 1.409 | 37.450 | 41.100 | 58.900 | 0.763 | 0.754 | 0.686 | 0.840 |
| | | | JE-UB-ED | -0.074 - 1.277 | 1.351 | 36.225 | 41.300 | 58.700 | 0.763 | 0.752 | 0.679 | 0.847 |
| | | 200 | UI | 0.452 - 1.793 | 1.342 | 29.532 | 50.108 | 49.892 | 0.740 | 0.740 | 0.522 | 0.958 |
| | | | GZ | -0.067 - 1.624 | 1.690 | 40.480 | 44.995 | 55.005 | 0.760 | 0.760 | 0.638 | 0.883 |
| | | | JE-UB-KB | -0.115 - 1.371 | 1.485 | 38.539 | 41.607 | 58.393 | 0.761 | 0.749 | 0.667 | 0.856 |
| | | | JE-UB-ED | -0.111 - 1.365 | 1.476 | 38.363 | 41.644 | 58.356 | 0.761 | 0.749 | 0.667 | 0.855 |

**Table 4.2. (Continued).** The Simulation Results for d=0.8.

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P(D)=0.8 | Equal Variances | 50 | UI | -0.217 - 0.892 | 1.109 | 36.122 | 79.841 | 20.159 | 0.761 | 0.760 | 0.759 | 0.762 |
| | | | GZ | -0.278 - 0.853 | 1.131 | 37.118 | 78.878 | 21.122 | 0.767 | 0.781 | 0.786 | 0.747 |
| | | | JE-UB-KB | -0.190 - 0.879 | 1.070 | 37.696 | 79.398 | 20.602 | 0.768 | 0.769 | 0.769 | 0.767 |
| | | | JE-UB-ED | 0.021 - 0.939 | 0.918 | 34.428 | 80.028 | 19.972 | 0.766 | 0.734 | 0.712 | 0.819 |
| | | 100 | UI | -0.334 - 1.083 | 1.417 | 45.801 | 79.721 | 20.279 | 0.760 | 0.752 | 0.747 | 0.774 |
| | | | GZ | -0.390 - 1.126 | 1.516 | 48.723 | 79.765 | 20.235 | 0.766 | 0.767 | 0.763 | 0.770 |
| | | | JE-UB-KB | -0.186 - 1.023 | 1.208 | 41.775 | 79.636 | 20.364 | 0.766 | 0.746 | 0.733 | 0.800 |
| | | | JE-UB-ED | -0.074 - 1.073 | 1.146 | 40.318 | 79.684 | 20.316 | 0.764 | 0.722 | 0.693 | 0.834 |
| | | 200 | UI | -0.461 - 1.246 | 1.708 | 54.139 | 79.935 | 20.065 | 0.762 | 0.758 | 0.753 | 0.770 |
| | | | GZ | -0.511 - 1.287 | 1.799 | 56.805 | 79.851 | 20.149 | 0.766 | 0.767 | 0.760 | 0.772 |
| | | | JE-UB-KB | -0.242 - 1.145 | 1.387 | 46.595 | 79.917 | 20.083 | 0.763 | 0.737 | 0.719 | 0.807 |
| | | | JE-UB-ED | -0.208 - 1.169 | 1.376 | 46.417 | 79.929 | 20.071 | 0.762 | 0.728 | 0.705 | 0.819 |
| | Unequal Variances | 50 | UI | 0.299 - 1.695 | 1.395 | 32.384 | 81.052 | 18.948 | 0.759 | 0.651 | 0.576 | 0.941 |
| | | | GZ | -0.099 - 1.126 | 1.224 | 28.776 | 75.904 | 24.096 | 0.765 | 0.738 | 0.718 | 0.811 |
| | | | JE-UB-KB | -0.233 - 1.100 | 1.333 | 32.008 | 75.781 | 24.219 | 0.768 | 0.754 | 0.745 | 0.790 |
| | | | JE-UB-ED | 0.048 - 1.222 | 1.174 | 29.430 | 77.757 | 22.243 | 0.765 | 0.721 | 0.692 | 0.837 |
| | | 100 | UI | 0.392 - 1.731 | 1.338 | 30.258 | 79.847 | 20.153 | 0.750 | 0.621 | 0.536 | 0.964 |
| | | | GZ | -0.019 - 1.430 | 1.449 | 33.005 | 76.737 | 23.263 | 0.765 | 0.694 | 0.649 | 0.880 |
| | | | JE-UB-KB | -0.118 - 1.261 | 1.379 | 32.473 | 74.240 | 25.760 | 0.766 | 0.712 | 0.682 | 0.849 |
| | | | JE-UB-ED | 0.106 - 1.357 | 1.251 | 29.933 | 74.663 | 25.337 | 0.763 | 0.681 | 0.634 | 0.893 |
| | | 200 | UI | 0.431 - 1.768 | 1.337 | 29.685 | 79.752 | 20.248 | 0.744 | 0.609 | 0.520 | 0.967 |
| | | | GZ | -0.030 - 1.569 | 1.599 | 35.412 | 76.906 | 23.094 | 0.761 | 0.681 | 0.629 | 0.893 |
| | | | JE-UB-KB | -0.006 - 1.393 | 1.399 | 32.363 | 74.289 | 25.711 | 0.763 | 0.682 | 0.640 | 0.885 |
| | | | JE-UB-ED | 0.092 - 1.440 | 1.348 | 31.309 | 74.475 | 25.525 | 0.762 | 0.668 | 0.618 | 0.905 |

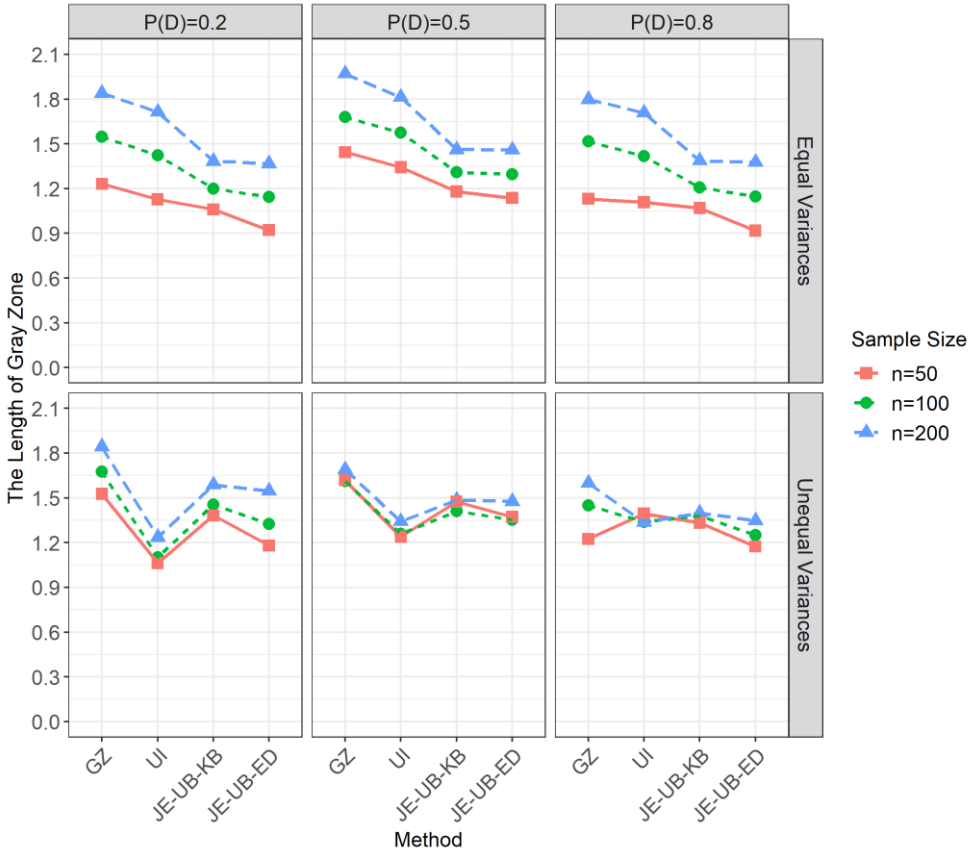



**Figure 4.2.** The Simulation Results for d=0.8.

The results of the simulation for an effect size equal to 0.8 were given in Table 4.2. and Figure 4.2. According to these results,

- The same results with effect size 0.5 and equal variances were obtained. The the JE-UB algorithms give the minimum length of the gray zone.

- For unequal variances with all sample sizes of the prevalence of disease 0.2 and 0.5, the uncertain interval algorithm has the minimum length of the gray zone. It was followed with the the JE-UB-ED algorithm.

- For the prevalence of the disease 0.8 with unequal variances, for sample sizes 50 and 100, the JE-UB-ED algorithm shows the minimum length of the gray zone, on the other hand; for sample size 200, the uncertain interval algorithm reveals the minimum length of gray zone.

**Table 4.3.** The Simulation Results for d=1.2.

| P(D) | The Ratio of Variances | n | Methods | Boundaries | | Proportion of Gray Zone | | | Out of in the Gray Zone | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | LL-UL | Length | Total % | Diseased % | Control % | AUROC | Accuracy | Sensitivity | Specificity |
| P(D)=0.2 | Equal Variances | 50 | UI | 0.113 - 1.252 | 1.139 | 35.180 | 18.897 | 81.103 | 0.842 | 0.850 | 0.829 | 0.855 |
| | | | GZ | 0.066 - 1.398 | 1.333 | 38.922 | 19.804 | 80.196 | 0.845 | 0.870 | 0.814 | 0.875 |
| | | | JE-UB-KB | 0.051 - 1.285 | 1.234 | 39.120 | 18.062 | 81.938 | 0.847 | 0.859 | 0.827 | 0.866 |
| | | | JE-UB-ED | 0.006 - 1.213 | 1.206 | 39.248 | 17.040 | 82.960 | 0.846 | 0.851 | 0.839 | 0.853 |
| | | 100 | UI | -0.041 - 1.287 | 1.328 | 41.146 | 20.143 | 79.857 | 0.840 | 0.835 | 0.850 | 0.831 |
| | | | GZ | -0.101 - 1.394 | 1.495 | 45.242 | 20.490 | 79.510 | 0.845 | 0.851 | 0.840 | 0.851 |
| | | | JE-UB-KB | -0.079 - 1.220 | 1.299 | 41.980 | 19.414 | 80.586 | 0.845 | 0.831 | 0.869 | 0.821 |
| | | | JE-UB-ED | -0.109 - 1.152 | 1.262 | 41.352 | 19.208 | 80.792 | 0.843 | 0.819 | 0.885 | 0.801 |
| | | 200 | UI | -0.153 - 1.369 | 1.522 | 46.445 | 20.256 | 79.744 | 0.840 | 0.836 | 0.847 | 0.833 |
| | | | GZ | -0.273 - 1.477 | 1.750 | 52.401 | 20.136 | 79.864 | 0.846 | 0.846 | 0.852 | 0.840 |
| | | | JE-UB-KB | -0.167 - 1.218 | 1.385 | 44.419 | 19.407 | 80.593 | 0.842 | 0.817 | 0.886 | 0.798 |
| | | | JE-UB-ED | -0.179 - 1.193 | 1.372 | 44.199 | 19.373 | 80.627 | 0.842 | 0.812 | 0.894 | 0.790 |
| | Unequal Variances | 50 | UI | 0.629 - 1.666 | 1.037 | 22.958 | 20.507 | 79.493 | 0.826 | 0.897 | 0.708 | 0.945 |
| | | | GZ | 0.169 - 1.700 | 1.531 | 37.670 | 17.329 | 82.671 | 0.848 | 0.889 | 0.784 | 0.912 |
| | | | JE-UB-KB | -0.036 - 1.394 | 1.430 | 40.770 | 14.422 | 85.578 | 0.847 | 0.857 | 0.830 | 0.865 |
| | | | JE-UB-ED | -0.042 - 1.289 | 1.331 | 39.496 | 14.270 | 85.730 | 0.846 | 0.847 | 0.843 | 0.849 |
| | | 100 | UI | 0.610 - 1.741 | 1.131 | 23.817 | 20.842 | 79.158 | 0.819 | 0.898 | 0.689 | 0.949 |
| | | | GZ | 0.004 - 1.723 | 1.719 | 42.146 | 16.312 | 83.688 | 0.845 | 0.882 | 0.789 | 0.901 |
| | | | JE-UB-KB | -0.146 - 1.381 | 1.527 | 43.219 | 14.114 | 85.886 | 0.845 | 0.844 | 0.843 | 0.846 |
| | | | JE-UB-ED | -0.151 - 1.329 | 1.480 | 42.429 | 14.266 | 85.734 | 0.844 | 0.837 | 0.853 | 0.834 |
| | | 200 | UI | 0.552 - 1.836 | 1.284 | 26.341 | 20.937 | 79.063 | 0.816 | 0.901 | 0.675 | 0.956 |
| | | | GZ | -0.141 - 1.736 | 1.877 | 46.908 | 15.561 | 84.439 | 0.843 | 0.876 | 0.797 | 0.890 |
| | | | JE-UB-KB | -0.171 - 1.444 | 1.615 | 44.707 | 14.258 | 85.742 | 0.842 | 0.847 | 0.830 | 0.854 |
| | | | JE-UB-ED | -0.177 - 1.431 | 1.608 | 44.632 | 14.308 | 85.692 | 0.842 | 0.844 | 0.833 | 0.851 |

**Table 4.3. (Continued).** The Simulation Results for d=1.2.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| P(D)=0.5 | Equal Variances | 50 | UI | -0.049 - 1.235 | 1.284 | 40.832 | 49.868 | 50.132 | 0.839 | 0.838 | 0.841 | 0.836 |
| | | | GZ | -0.095 - 1.316 | 1.410 | 44.192 | 49.679 | 50.321 | 0.846 | 0.848 | 0.840 | 0.852 |
| | | | JE-UB-KB | -0.039 - 1.242 | 1.281 | 42.074 | 48.847 | 51.153 | 0.847 | 0.847 | 0.839 | 0.855 |
| | | | JE-UB-ED | -0.024 - 1.230 | 1.254 | 41.508 | 48.767 | 51.233 | 0.847 | 0.846 | 0.839 | 0.855 |
| | | 100 | UI | -0.117 - 1.314 | 1.431 | 44.343 | 49.888 | 50.112 | 0.840 | 0.841 | 0.841 | 0.839 |
| | | | GZ | -0.192 - 1.416 | 1.609 | 49.003 | 50.062 | 49.938 | 0.846 | 0.850 | 0.842 | 0.851 |
| | | | JE-UB-KB | -0.047 - 1.283 | 1.330 | 42.527 | 49.754 | 50.246 | 0.844 | 0.844 | 0.835 | 0.853 |
| | | | JE-UB-ED | -0.042 - 1.280 | 1.321 | 42.366 | 49.804 | 50.196 | 0.844 | 0.843 | 0.835 | 0.853 |
| | | 200 | UI | -0.191 - 1.403 | 1.593 | 48.787 | 50.154 | 49.846 | 0.840 | 0.841 | 0.839 | 0.841 |
| | | | GZ | -0.296 - 1.531 | 1.827 | 54.453 | 50.265 | 49.735 | 0.846 | 0.852 | 0.840 | 0.851 |
| | | | JE-UB-KB | -0.107 - 1.321 | 1.428 | 45.109 | 49.965 | 50.035 | 0.842 | 0.842 | 0.840 | 0.844 |
| | | | JE-UB-ED | -0.107 - 1.322 | 1.429 | 45.121 | 49.962 | 50.038 | 0.842 | 0.842 | 0.840 | 0.844 |
| | Unequal Variances | 50 | UI | 0.536 - 1.752 | 1.216 | 27.424 | 51.072 | 48.928 | 0.825 | 0.826 | 0.690 | 0.959 |
| | | | GZ | 0.004 - 1.615 | 1.611 | 38.254 | 43.378 | 56.622 | 0.846 | 0.845 | 0.791 | 0.900 |
| | | | JE-UB-KB | -0.072 - 1.423 | 1.495 | 38.212 | 40.725 | 59.275 | 0.847 | 0.844 | 0.816 | 0.877 |
| | | | JE-UB-ED | -0.016 - 1.442 | 1.458 | 37.300 | 41.029 | 58.971 | 0.846 | 0.842 | 0.806 | 0.886 |
| | | 100 | UI | 0.536 - 1.813 | 1.277 | 27.390 | 51.008 | 48.992 | 0.822 | 0.823 | 0.679 | 0.966 |
| | | | GZ | 0.001 - 1.707 | 1.706 | 39.157 | 43.303 | 56.697 | 0.845 | 0.843 | 0.775 | 0.914 |
| | | | JE-UB-KB | -0.039 - 1.480 | 1.519 | 37.513 | 39.866 | 60.134 | 0.843 | 0.836 | 0.793 | 0.893 |
| | | | JE-UB-ED | -0.014 - 1.489 | 1.503 | 37.100 | 40.035 | 59.965 | 0.843 | 0.835 | 0.789 | 0.897 |
| | | 200 | UI | 0.485 - 1.903 | 1.418 | 29.663 | 51.023 | 48.977 | 0.818 | 0.819 | 0.666 | 0.969 |
| | | | GZ | -0.126 - 1.751 | 1.877 | 42.995 | 42.464 | 57.536 | 0.844 | 0.842 | 0.780 | 0.907 |
| | | | JE-UB-KB | -0.041 - 1.559 | 1.601 | 38.496 | 40.694 | 59.306 | 0.841 | 0.833 | 0.781 | 0.901 |
| | | | JE-UB-ED | -0.039 - 1.561 | 1.599 | 38.462 | 40.704 | 59.296 | 0.841 | 0.833 | 0.780 | 0.902 |

**Table 4.3. (Continued).** The Simulation Results for d=1.2.

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P(D)=0.8 | Equal Variances | 50 | UI | -0.054 - 1.090 | 1.144 | 35.052 | 80.480 | 19.520 | 0.841 | 0.852 | 0.859 | 0.824 |
| | | | GZ | -0.102 - 1.081 | 1.183 | 36.148 | 79.717 | 20.283 | 0.846 | 0.862 | 0.869 | 0.823 |
| | | | JE-UB-KB | -0.057 - 1.117 | 1.174 | 37.434 | 80.526 | 19.474 | 0.847 | 0.859 | 0.865 | 0.829 |
| | | | JE-UB-ED | 0.004 - 1.194 | 1.190 | 39.018 | 82.265 | 17.735 | 0.845 | 0.851 | 0.853 | 0.838 |
| | | 100 | UI | -0.096 - 1.241 | 1.337 | 41.017 | 79.704 | 20.296 | 0.841 | 0.836 | 0.834 | 0.848 |
| | | | GZ | -0.146 - 1.305 | 1.451 | 44.308 | 79.645 | 20.355 | 0.846 | 0.842 | 0.837 | 0.856 |
| | | | JE-UB-KB | -0.019 - 1.275 | 1.294 | 41.617 | 80.160 | 19.840 | 0.846 | 0.830 | 0.818 | 0.874 |
| | | | JE-UB-ED | 0.049 - 1.317 | 1.268 | 41.529 | 80.529 | 19.471 | 0.844 | 0.817 | 0.798 | 0.890 |
| | | 200 | UI | -0.163 - 1.359 | 1.522 | 46.584 | 79.840 | 20.160 | 0.840 | 0.836 | 0.832 | 0.847 |
| | | | GZ | -0.249 - 1.494 | 1.743 | 52.374 | 80.054 | 19.946 | 0.846 | 0.841 | 0.831 | 0.861 |
| | | | JE-UB-KB | -0.018 - 1.392 | 1.411 | 45.200 | 80.633 | 19.367 | 0.843 | 0.816 | 0.796 | 0.889 |
| | | | JE-UB-ED | 0.007 - 1.409 | 1.401 | 45.115 | 80.747 | 19.253 | 0.842 | 0.811 | 0.788 | 0.896 |
| | Unequal Variances | 50 | UI | 0.306 - 1.819 | 1.513 | 33.302 | 81.923 | 18.077 | 0.839 | 0.778 | 0.734 | 0.944 |
| | | | GZ | -0.105 - 1.335 | 1.440 | 30.206 | 76.190 | 23.810 | 0.846 | 0.847 | 0.844 | 0.848 |
| | | | JE-UB-KB | -0.196 - 1.359 | 1.555 | 32.926 | 77.173 | 22.827 | 0.846 | 0.858 | 0.863 | 0.830 |
| | | | JE-UB-ED | -0.190 - 1.612 | 1.802 | 38.154 | 80.783 | 19.217 | 0.845 | 0.855 | 0.858 | 0.833 |
| | | 100 | UI | 0.438 - 1.841 | 1.403 | 30.370 | 80.405 | 19.595 | 0.829 | 0.743 | 0.685 | 0.974 |
| | | | GZ | -0.013 - 1.519 | 1.531 | 32.366 | 74.350 | 25.650 | 0.844 | 0.808 | 0.786 | 0.903 |
| | | | JE-UB-KB | -0.043 - 1.476 | 1.518 | 32.688 | 74.183 | 25.817 | 0.845 | 0.813 | 0.795 | 0.895 |
| | | | JE-UB-ED | 0.021 - 1.593 | 1.572 | 33.891 | 75.601 | 24.399 | 0.842 | 0.803 | 0.779 | 0.906 |
| | | 200 | UI | 0.471 - 1.868 | 1.398 | 29.779 | 80.160 | 19.840 | 0.823 | 0.731 | 0.669 | 0.977 |
| | | | GZ | -0.034 - 1.673 | 1.708 | 35.544 | 74.665 | 25.335 | 0.843 | 0.797 | 0.770 | 0.917 |
| | | | JE-UB-KB | 0.085 - 1.658 | 1.573 | 33.574 | 74.771 | 25.229 | 0.843 | 0.783 | 0.751 | 0.935 |
| | | | JE-UB-ED | 0.148 - 1.703 | 1.555 | 33.419 | 75.307 | 24.693 | 0.841 | 0.775 | 0.738 | 0.945 |

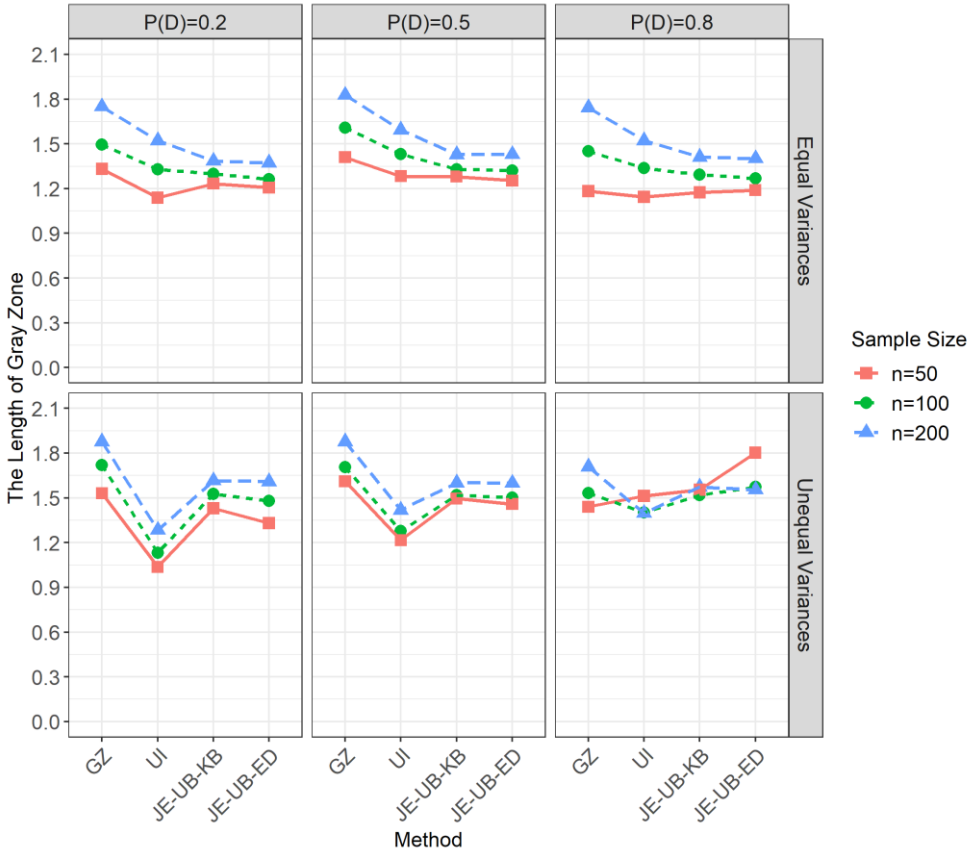



**Figure 4.3.** The Simulation Results for d=1.2.

For effect size 1.2, the results of the simulation study were represented in Table 4.3. and Figure 4.3.

- For equal variances, the minimum length of the gray zone is obtained from the the JE-UB-ED algorithm except for the prevalence of disease 0.2 and 0.8 with sample size 50. In this scenario, the minimum length is recorded for an uncertain interval algorithm.

- For unequal variances, for all the cases, the minimum length of the gray zone is found for the uncertain interval algorithm.

In all 54 scenarios, a similar pattern is found for unequal variances for the prevalence 0.2 and 0.5. However, for the prevalence of 0.8, this pattern was not observed. One of the reasons is when the prevalence is 0.8 and the variance of the diseased group is 3 times higher than the non-diseased group, there exists an underrepresentation of the non-diseased group. To research this issue, the additional 9 scenarios including the prevalence is 0.8, and the variance of the non-diseased group is 3 times higher than the diseased group with all sample sizes and effect sizes. The results are shown in Tables 4.4. and Figures 4.4., 4.5., and 4.6.

**Table 4.4.** The Simulation Results for P(D) = 0.8 and Var(YC)=3 x Var(YD).

| d | n | Methods | Boundaries | | Proportion of Gray Zone | | | Out of in the Gray Zone | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | LL-UL | Length | Total % | Diseased % | Control % | AUROC | Accuracy | Sensitivity | Specificity |
| 0.5 | 50 | UI | -0.848 - 0.144 | 0.992 | 22.406 | 78.675 | 21.325 | 0.680 | 0.820 | 0.910 | 0.449 |
| | | GZ | -0.925 - 0.368 | 1.293 | 31.048 | 77.892 | 22.108 | 0.686 | 0.812 | 0.883 | 0.489 |
| | | JE-UB-KB | -0.352 - 0.981 | 1.333 | 41.590 | 84.501 | 15.499 | 0.690 | 0.715 | 0.731 | 0.649 |
| | | JE-UB-ED | -0.011 - 0.901 | 0.912 | 31.418 | 84.143 | 15.857 | 0.687 | 0.674 | 0.670 | 0.703 |
| | 100 | UI | -0.950 - 0.101 | 1.051 | 22.129 | 79.574 | 20.426 | 0.672 | 0.829 | 0.932 | 0.412 |
| | | GZ | -0.945 - 0.419 | 1.363 | 32.688 | 79.720 | 20.280 | 0.681 | 0.812 | 0.891 | 0.471 |
| | | JE-UB-KB | -0.472 - 1.049 | 1.521 | 44.816 | 85.050 | 14.950 | 0.688 | 0.720 | 0.752 | 0.625 |
| | | JE-UB-ED | -0.232 - 0.930 | 1.161 | 36.105 | 84.717 | 15.283 | 0.684 | 0.697 | 0.719 | 0.649 |
| | 200 | UI | -0.997 - 0.118 | 1.115 | 23.252 | 79.681 | 20.319 | 0.668 | 0.832 | 0.940 | 0.396 |
| | | GZ | -0.895 - 0.413 | 1.308 | 32.055 | 80.326 | 19.674 | 0.679 | 0.810 | 0.891 | 0.466 |
| | | JE-UB-KB | -0.519 - 0.966 | 1.486 | 43.088 | 84.830 | 15.170 | 0.686 | 0.732 | 0.783 | 0.589 |
| | | JE-UB-ED | -0.376 - 0.924 | 1.300 | 39.043 | 84.599 | 15.401 | 0.682 | 0.716 | 0.757 | 0.607 |
| 0.8 | 50 | UI | -0.446 - 0.561 | 1.007 | 23.048 | 79.365 | 20.635 | 0.748 | 0.853 | 0.923 | 0.574 |
| | | GZ | -0.459 - 1.003 | 1.461 | 37.194 | 80.631 | 19.369 | 0.767 | 0.835 | 0.877 | 0.656 |
| | | JE-UB-KB | -0.051 - 1.308 | 1.359 | 41.328 | 84.262 | 15.738 | 0.768 | 0.779 | 0.787 | 0.750 |
| | | JE-UB-ED | 0.168 - 1.296 | 1.128 | 36.610 | 83.977 | 16.023 | 0.766 | 0.751 | 0.744 | 0.788 |
| | 100 | UI | -0.558 - 0.598 | 1.155 | 25.106 | 79.766 | 20.234 | 0.742 | 0.860 | 0.939 | 0.545 |
| | | GZ | -0.475 - 1.105 | 1.579 | 40.413 | 82.206 | 17.794 | 0.762 | 0.832 | 0.876 | 0.648 |
| | | JE-UB-KB | -0.123 - 1.401 | 1.524 | 44.616 | 84.678 | 15.322 | 0.766 | 0.779 | 0.794 | 0.738 |
| | | JE-UB-ED | -0.028 - 1.381 | 1.409 | 42.248 | 84.423 | 15.577 | 0.764 | 0.767 | 0.777 | 0.752 |
| | 200 | UI | -0.614 - 0.625 | 1.239 | 26.641 | 79.642 | 20.358 | 0.737 | 0.862 | 0.944 | 0.531 |
| | | GZ | -0.463 - 1.185 | 1.648 | 42.915 | 82.794 | 17.206 | 0.759 | 0.826 | 0.865 | 0.654 |
| | | JE-UB-KB | -0.189 - 1.385 | 1.574 | 45.293 | 84.845 | 15.155 | 0.763 | 0.786 | 0.814 | 0.713 |
| | | JE-UB-ED | -0.158 - 1.385 | 1.542 | 44.715 | 84.746 | 15.254 | 0.762 | 0.781 | 0.807 | 0.718 |

**Table 4.4. (Continued).** The Simulation Results for P(D) = 0.8 and Var(YC)=3 x Var(YD).

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1.2 | 50 | UI | 0.022 - 1.076 | 1.053 | 23.144 | 79.036 | 20.964 | 0.823 | 0.897 | 0.946 | 0.701 |
| | | GZ | 0.072 - 1.586 | 1.514 | 38.168 | 82.525 | 17.475 | 0.846 | 0.878 | 0.899 | 0.793 |
| | | JE-UB-KB | 0.323 - 1.734 | 1.411 | 40.596 | 84.442 | 15.558 | 0.848 | 0.855 | 0.862 | 0.834 |
| | | JE-UB-ED | 0.410 - 1.757 | 1.346 | 39.956 | 84.998 | 15.002 | 0.846 | 0.847 | 0.848 | 0.844 |
| | 100 | UI | -0.052 - 1.100 | 1.152 | 23.815 | 79.022 | 20.978 | 0.818 | 0.898 | 0.950 | 0.685 |
| | | GZ | 0.015 - 1.721 | 1.706 | 42.806 | 83.713 | 16.287 | 0.843 | 0.875 | 0.894 | 0.793 |
| | | JE-UB-KB | 0.314 - 1.870 | 1.555 | 43.991 | 85.392 | 14.608 | 0.845 | 0.843 | 0.843 | 0.848 |
| | | JE-UB-ED | 0.364 - 1.874 | 1.510 | 43.225 | 85.314 | 14.686 | 0.844 | 0.837 | 0.833 | 0.855 |
| | 200 | UI | -0.142 - 1.160 | 1.302 | 26.583 | 79.071 | 20.929 | 0.816 | 0.902 | 0.957 | 0.674 |
| | | GZ | 0.004 - 1.843 | 1.839 | 46.853 | 84.427 | 15.573 | 0.843 | 0.872 | 0.886 | 0.800 |
| | | JE-UB-KB | 0.250 - 1.874 | 1.624 | 44.784 | 85.609 | 14.391 | 0.843 | 0.848 | 0.856 | 0.829 |
| | | JE-UB-ED | 0.259 - 1.878 | 1.619 | 44.744 | 85.595 | 14.405 | 0.842 | 0.846 | 0.854 | 0.830 |

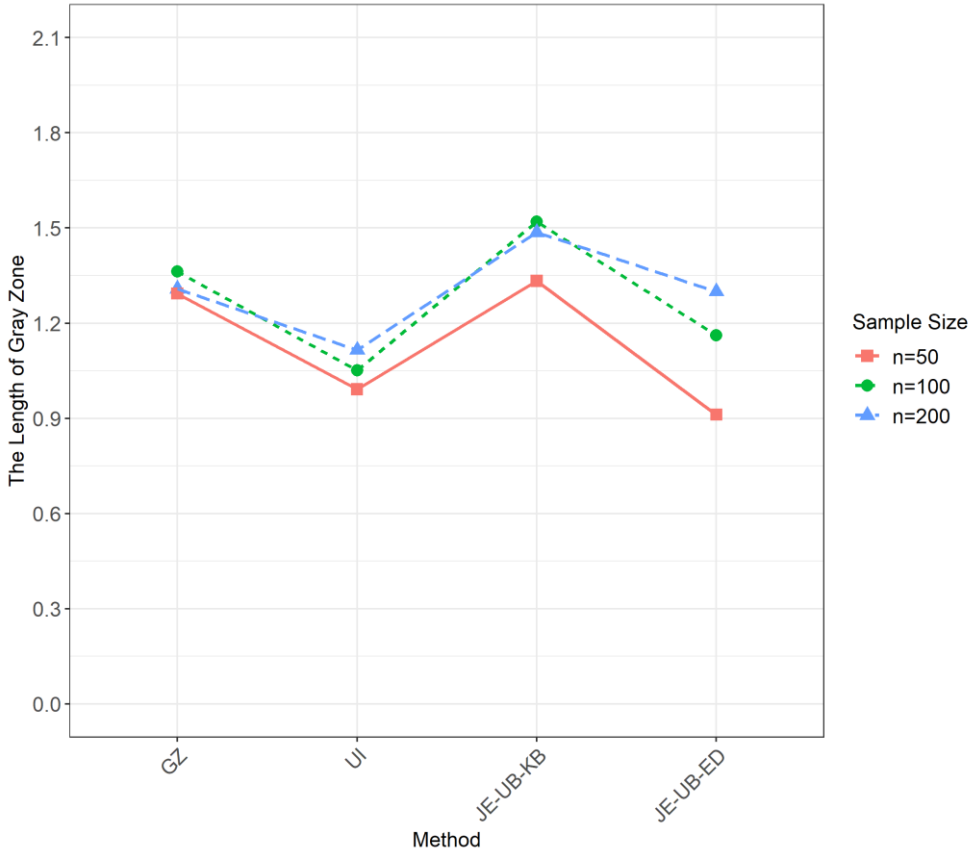



**Figure 4.4.** The Simulation Results for P(D) = 0.8, Var(YC)=3 x Var(YD) and d=0.5.

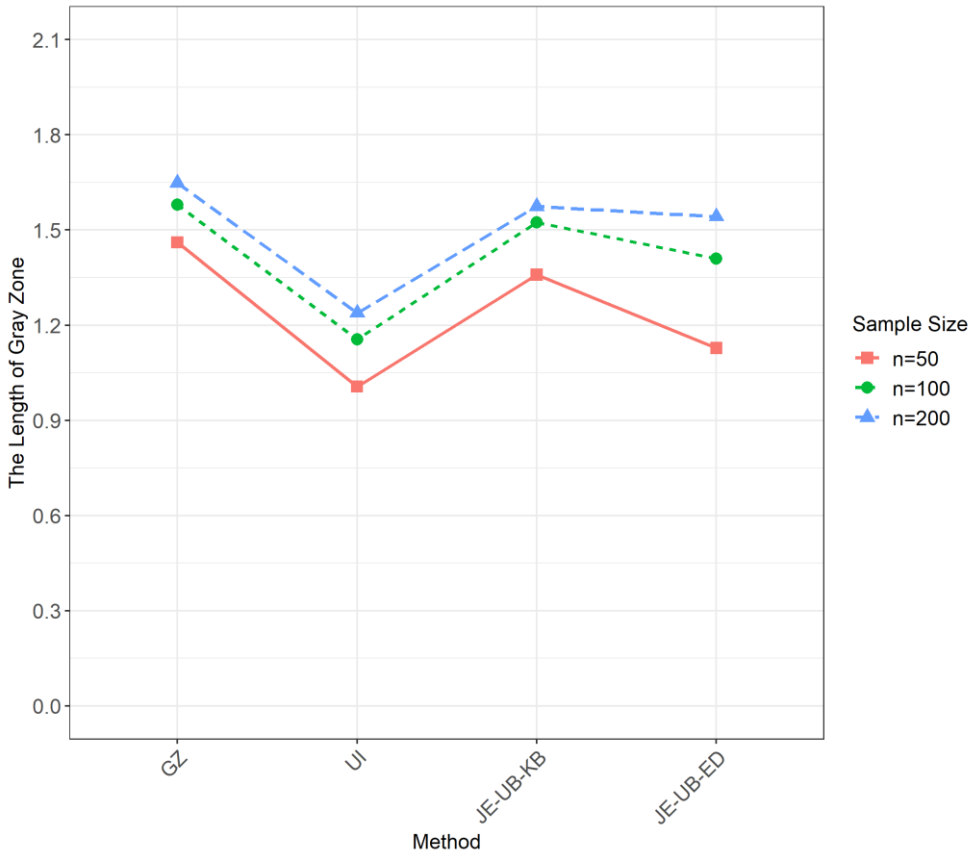



**Figure 4.5.** The Simulation Results for P(D) = 0.8, Var(YC)=3 x Var(YD) and d=0.8.

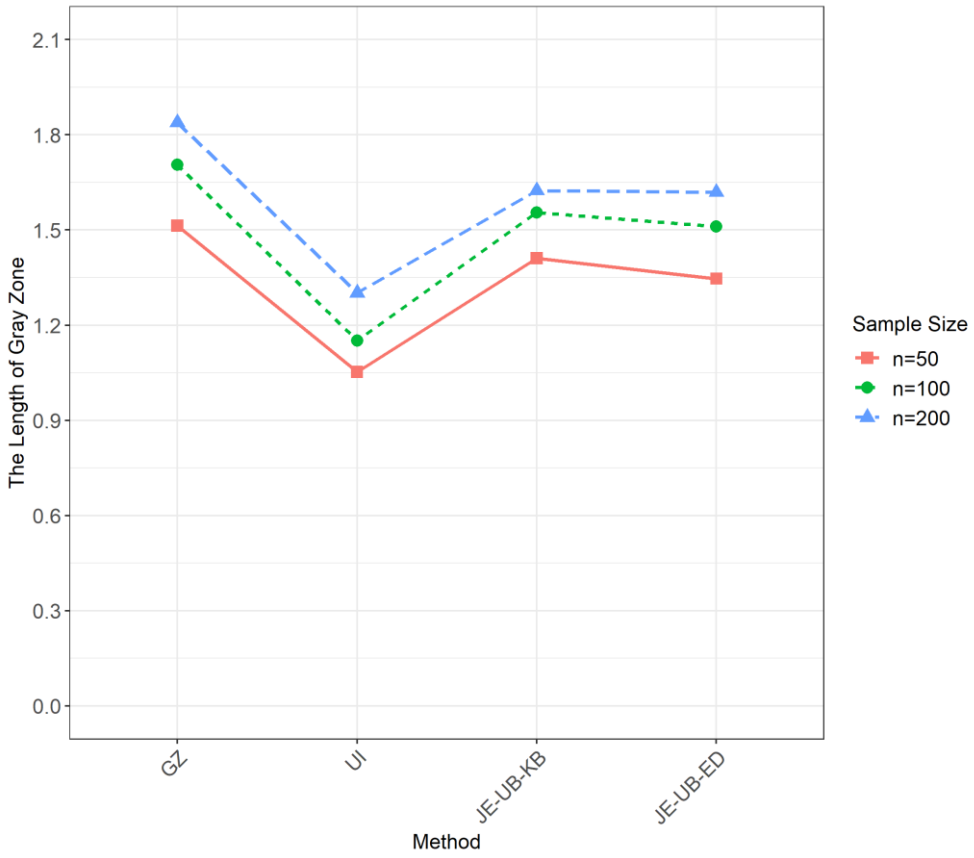



**Figure 4.6.** The Simulation Results for P(D) = 0.8, Var(YC)=3 x Var(YD) and d=1.2**.**

According to the results in Table 4.4. and Figures 4.4., 4.5., and 4.6.:

- Similar patterns of unequal variances for the prevalence of 0.2 and 0.8 are obtained. In order to show this pattern more clearly, in the Appendices (A. Figure 1.1., A. Figure 1.2., A. Figure 1.3.), these results are plotted with effect sizes 0.2 and 0.5.

- In most of the results in Table 4.4., Figures 4.4., 4.5., and 4.6., the uncertain interval has the minimum length of the gray zone. For the smallest sample size in scenarios with effect size 0.5, the JE-UB-ED algorithm gives the minimum length of the gray zone.

### 4.2. The Results of the Real Data Set

A total of 532 complete data were used. The glucose variable was used to construct the boundaries of the gray zone. A total of 33.721% of subjects are in the diabetic group. The mean and standard deviation of each group with effect size is given in Table 4.5. Thus, the histograms of the two groups are plotted in Figure 4.8. The ROC curve was plotted in Figure 4.9.

**Table 4.5.** The Summary Statistics of the Glucose.

| | Diabetic Group ($n_1$=177) | Non-diabetic Group ($n_2$=355) | Effect Size |
|---|---|---|---|
| Glucose | 143.199±31.265 | 110.017±24.287 | 1.183 |

The effect size is large in Table 4.5. The AUROC is 0.794 in Figure 4.9.

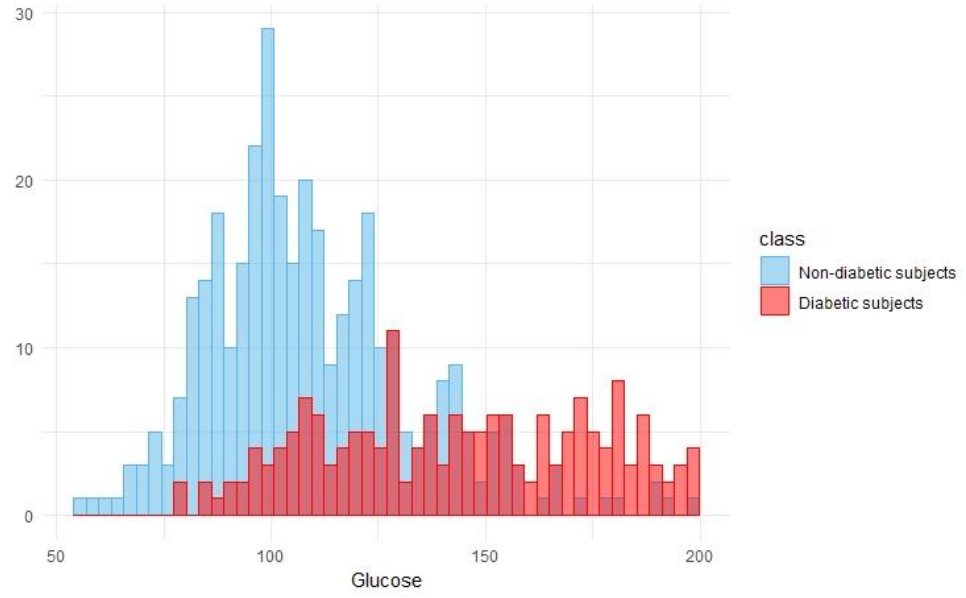



**Figure 4.8.** The Histograms of Diabetic and Non-Diabetic Groups.

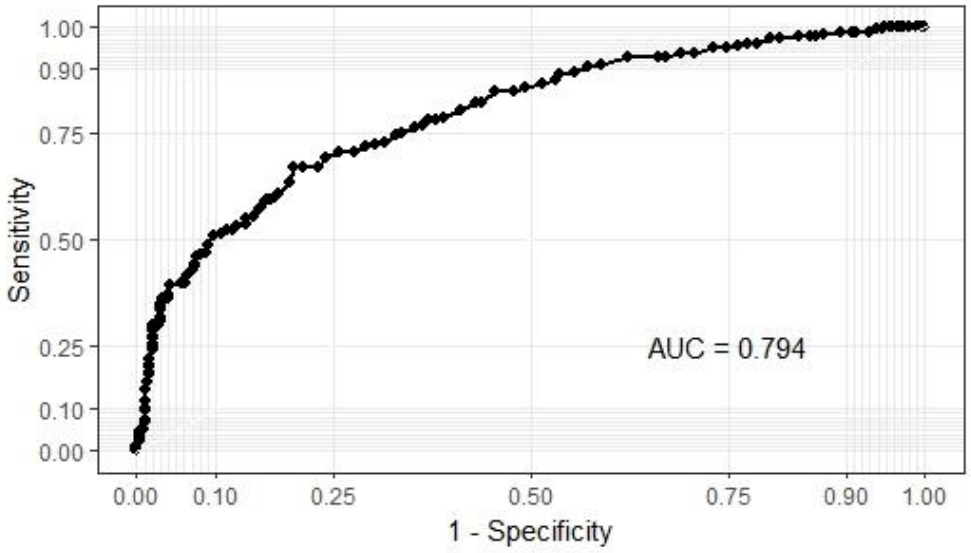



**Figure 4.9.** The ROC Curve of Glucose Variable.

**Table 4.6.** The Results of Real Data Set.

| Methods | Boundaries | | Proportion of Gray Zone | | | Out of in the Gray Zone | | |
|---|---|---|---|---|---|---|---|---|
| | LL-UL | Length | Total % | Diseased % | Control % | Accuracy | Sensitivity | Specificity |
| UI | 109 - 148 | 39 | 38.910 | 34.783 | 65.217 | 0.834 | 0.743 | 0.882 |
| GZ | 101 - 144 | 43 | 47.368 | 29.762 | 70.238 | 0.829 | 0.843 | 0.829 |
| JE-UB-KB | 113 - 129 | 16 | 19.925 | 30.189 | 69.811 | 0.761 | 0.738 | 0.772 |
| JE-UB-ED | 113 - 122 | 9 | 10.150 | 25.926 | 74.074 | 0.716 | 0.767 | 0.689 |

The pre-selected values of LR (+) and LR (-) are 5 and 0.22 are calculated based on PPV=0.7 and NPV=0.9 for the grey zone approach. Thus for the pre-selected values of sensitivity and specificity in the gray zone 0.55 (default value). The main reason that we fixed the performance measure out of the gray zone is illustrated in Table 4.6. In the real data set, since the pre-selected values are used, for different pre-selected values different boundaries will appear.

The effect size is approximately 1.2 and the ratio of variance is approximately 1.657. The real data set results were renewed based on the fixed value of AUROC in the out-of-gray zone of 0.85. The results were presented in Table 4.7. and Figure 4.10.

**Table 4.7.** The Results of Data Set Based on Simulation Scenario.

| Methods | Boundaries | | Proportion of Gray Zone | | | Out of in the Gray Zone | | | |
|---|---|---|---|---|---|---|---|---|---|
| | LL-UL | Length | Total % | Diseased % | Control % | AUROC | Accuracy | Sensitivity | Specificity |
| UI | 104 - 166 | 62 | 54.135 | 36.458 | 63.542 | 0.841 | 0.889 | 0.722 | 0.959 |
| GZ | 96 - 150 | 54 | 59.399 | 28.481 | 71.519 | 0.840 | 0.833 | 0.874 | 0.806 |
| JE-UB-KB | 100 - 146 | 46 | 51.504 | 28.472 | 71.528 | 0.843 | 0.837 | 0.863 | 0.822 |
| JE-UB-ED | 100 - 146 | 46 | 51.504 | 28.472 | 71.528 | 0.843 | 0.837 | 0.863 | 0.822 |

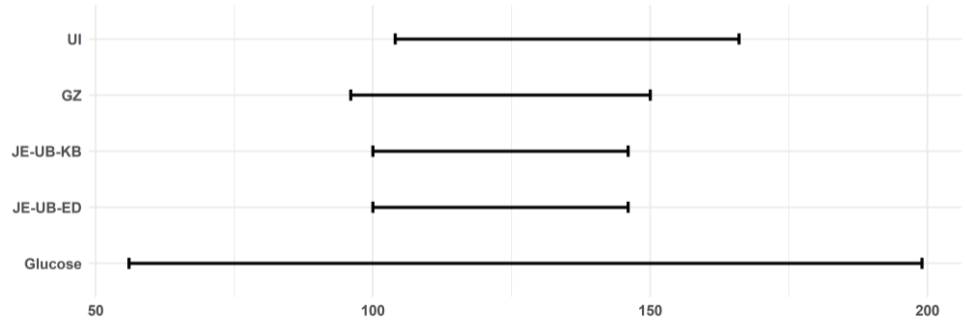



**Figure 4.10.** The Boundaries and Length of Gray Zone Approaches.

The minimum length of the gray zone is obtained for the JE-UB. Although the uncertain interval approach had the largest length of the gray zone, the total percentage of subjects who are in the gray zone is less than in the grey zone algorithm. The reason is that the number of subjects in 96 to 150 is more than 104 to 166. Thus, the specificity out of the gray zone is high in uncertain intervals but the sensitivity is low. For the JE-UB algorithms and the grey zone approach, the performance measures out of the gray zone are more balanced.

## 5. DISCUSSION

In statistical decision-making, diagnostic tests have an important role. There are a variety of diagnostic procedures that can offer information for medical decision-making (32). To rule out or rule in certain diseases by using diagnostic tests, which produce ordinal or quantitative results is specifying optimal cut-off points to discriminate subjects. To reach this aim, there exist several strategies based on some performance measures such as choosing optimal cut-off point based on maximizing sensitivity, specificity, the test statistics such that maximally selected chi-square statistic or minimum p-value and parametric approaches or parametric modeling by means of polynomials of covariates gives approximately solution (33). With respect to this cut-off value, no perfect discrimination can be made unless the test is the gold standard, thus yielding false positives and false negatives. Therefore, the middle inconclusive area (gray zone) in the overlapped area of diseased and non-diseased distributions is offered. The benefits of the gray zone approach are that false decisions decrease for patients in the gray zone, and true decisions increase outside of the gray zone (34).

There are different approaches to finding the boundaries of the gray zone. In this study, we take the grey zone (2) and uncertain interval (3) approaches into account. The grey zone method is based on pre-test probabilities while the uncertain interval approach is based on balancing TP to FN and TN to FP in the gray zone by using the SLSQP optimization. The simulation results for TG-ROC and uncertain interval approaches were given (3). In this study, we excluded TG-ROC since it is offered for the pre-selected values of sensitivity and specificity of 0.95 or 0.90. As a result, it ensures that sensitivity and specificity are a minimum of 95% or 90% outside of the intermediate region (other sensitivity and specificity values may be shown graphically). This algorithm has strict restrictions. Moreover, Landsheer (3) reported that the gray zone in TG-ROC does not clearly explain and outside of the gray zone does not meet expectations. Also, The TG-ROC simulation findings reveal that

improving decision-making may not always come from choosing just the most discriminative scores and excluding intermediate levels.

When evaluating and comparing diagnostic tests, information statistics are useful (18). Our main aim is to propose an algorithm to find the borders of the gray zone by using the information theory. We also desire to highlight and show the application of the information theory to the decision-making process. The second aim of this study is to compare the proposed algorithm with the grey zone and uncertain interval algorithms with a simulation study. In addition to the simulation study, we also showed the application of these approaches on the real data set which is taken from the Pima Indian Data Set (29). We explained the results of the real data set in detail.

Both methods for the grey zone and uncertain interval approaches were used in the studies (14, 34). However, these studies did not compare the performance comparisons of the algorithms with a simulation study. The TG-ROC and uncertain interval comparisons were given by simulation study (3). The simulation scenarios are based on Somaza's (35) study which is the separation coefficient (the mean differences of two distributions divided by the standard deviation of the distribution of non-diseased groups) and asymmetry coefficient (the ratio of the standard deviations of two distributions). On the other hand, in this thesis, we used the effect size to standardize how much the distributions are overlapped. The medium and large effect sizes (0.5, 0.8, and 1.2) were considered, yielding the AUROC of the diagnostic test 0.64, 0.72, and 0.80. Therefore, we considered the discriminative ability of the tests as poor to fair (36). Yet, we also deal with the equality of the variances. Landsheer (3) regarded sample size and prevalence as a total of 1000 and 0.5, 0.2, and 0.1, respectively. However, we take in our study different sample sizes of 50,100, and 200, with the prevalence of the disease as 0.2, 0.5, and 0.8. We also desire to research the disease with high prevalence.

The major difficulty in the comparisons of the three approach is when the length of the gray zone increases, the performance of out of the gray zone increases.

For comparing these three methods, the AUROC out of the gray zone was fixed differently under different scenarios. The minimum length of the gray zone was accepted as the best algorithm.

In the light of the simulation results, we have found that:

- For equal variances, the proposed algorithms (JE-UB-KB and JE-UB-ED) gives the minimum length of the gray zone in most of the scenarios.

- For unequal variances and effect size 0.5, the length of the gray zone for the uncertain interval and the JE-UB-ED algorithms is narrower compared to the grey zone and the JE-UB-KB algorithms.

- For unequal variances, and the effect sizes 0.8 and 1.2, in some cases the uncertain interval gives the best results and for some cases, the JE-UB-ED algorithm gives the best results.

- In most of the scenarios, when the sample size increases, the length of the grey zone also increases. The sample size is a significant part of this study, as Landsheer (3) mentioned that how a larger sample size is needed for an uncertain interval approach is not clear. We accept total sample sizes of 50, 100, and 200 as small, medium, and large, respectively.

- For equal variances, similar patterns are observed among algorithms with respect to different effect sizes, sample sizes, and prevalence.

- For unequal variances, while similar patterns are shown for the algorithms for the prevalence of 0.2 and 0.5, similar patterns are not observed for the prevalence of 0.8. The reason is that when the prevalence is 0.8 and the variance of $Y_D$ is three times that of $Y_C$, the representation of the non-diseased group is low. To show this idea, the prevalence of 0.8 and the variance ratio of $Y_D$ to $Y_C$ of 1/3 are also investigated. Then a similar pattern with a prevalence of 0.2 and 0.8 is found.

For the final remark, in this study, we propose an algorithm to find the boundaries of the gray zone based on joint entropy without prior knowledge. The performance of the proposed algorithm was as good as or better compared to the other algorithms for the poor to fair discriminative ability of the diagnostic tests.

## 6. CONCLUSION

The main aim of this study is to propose a new algorithm to find the boundaries of the gray zone based on the information statistics for quantitative diagnostic tests. Moreover, we also compared the proposed algorithm (JE-UB) with existing algorithms through a compressive simulation study.

In general, for simulation scenarios with equal variances, the JE-UB algorithm gives the minimum length of the gray zone while for unequal variances, in some cases the JE-UB algorithm gives better results for other cases the uncertain interval algorithm gives better results. As a result, the JE-UB algorithm is superior in some scenarios. For the rest of the scenarios, it is in the second place. We can conclude that the JE-UB algorithm is as good as or better compared to existing algorithms. Besides, the main advantage of the algorithm we propose is that there is no prior information to run the algorithm, unlike others.

As a final remark, the main contribution of constructing a gray zone can be summarized such that discriminates subjects into the diseased, non-diseased, and gray zone (neither diseased nor non-diseased) groups. This might provide clinicians to lead other biomarkers, tests, or medical imaging or to take precautions for the subjects in the gray zone.

For future studies:

- For all the simulation scenarios, the distributions were generated from normal distributions. The skewed distributions should be also examined in future studies. Yet, the cases in which skewed distributions are transformed into normal distributions might also be observed.

- For determining how distributions are close to each other, the effect size was used in this study. For this aim, the overlap coefficient (OVL) that gives the similarity of two distributions in the overlapped area (37) may also be used in future studies.

- In this study, a diagnostic test and class variables are only used. It is also significant to specify the boundaries of a gray zone for subgroups such as gender or age categories. Thus, creating a gray zone could provide clinicians with the tools to construct a decision-making algorithm. To rule in or rule out the disease for the patients in the gray zone, for instance, they may use more biomarkers or more advanced tests.

- Additional to those issues, the R package and RShiny web application will be developed and they will be publicly available to researchers.

## 7. REFERENCES


**1.** Karaagaoglu E, Karakaya J, Kılıçkap M. Tani Testlerinin Değerlendirilmesinde İstatistiksel Yöntemler. 1st ed. Ankara: Detay Yayincilik; 2016.

**2.** Coste J, Pouchot J. A grey zone for quantitative diagnostic and screening tests. International Journal of Epidemiology. 2003; 32(2):304–13.

**3.** Landsheer JA. Interval of Uncertainty: An Alternative Approach for the Determination of Decision Thresholds, with an Illustrative Application for the Prediction of Prostate Cancer. Hurst R, editor. PLOS ONE. 2013; 11(11):e0166007.

**4.** Campbell G, Levy D, Bailey JJ. Bootstrap comparison of fuzzy ROC curves for ECG-LVH algorithms using data from the Framingham heart study. Journal of Electrocardiology. 1990;23:132–7.

**5.** Evangelista PE, Embrechts MJ, Bonissone P, Szymanski BK. Fuzzy ROC curves for unsupervised nonparametric ensemble techniques. Proceeding of 2005 IEEE International Joint Conference on Neural Networks; 2005 Jul 31 – August 4, Montreal, QC, Canada. p. 3040-3045).

**6.** Castanho MJ, Barros LC, Yamakami A, Vendite LL. Fuzzy receiver operating characteristic curve: an option to evaluate diagnostic tests. IEEE Transactions on Information Technology in Biomedicine. 2007; 11(3):244-50.

**7.** Parmigiani G. Receiver operating characteristic curves with an indeterminacy zone. Pattern recognition letters. 2020; 136:94-100.

**8.** Feinstein AR. The inadequacy of binary models for the clinical reality of three-zone diagnostic decisions. Journal of Clinical Epidemiology. 1990; 43(1):109–13.

**9.** Jamart J. Chance-corrected sensitivity and specificity for three-zone diagnostic tests. Journal of Clinical Epidemiology. 1992; 45(9):1035–8.

**10.** Simel DL, Samsa GP, Matchar DB. Likelihood ratios for continuous test results—making the clinicians' job easier or harder? Journal of Clinical Epidemiology. 1993; 46(1):85–93.

**11.** Greiner M, Sohr D, Göbel P. A modified ROC analysis for the selection of cut-off values and the definition of intermediate results of serodiagnostic tests. Journal of immunological methods. 1995; 185(1):123-32.

**12.** Coste J, Jourdain P, Pouchot J. A Gray Zone Assigned to Inconclusive Results of Quantitative Diagnostic Tests: Application to the Use of Brain Natriuretic Peptide for Diagnosis of Heart Failure in Acute Dyspneic Patients. Clinical Chemistry. 2006; 52(12):2229–35.

**13.** Hosmer DW, Lemeshow S. Applied logistic regression. 1st ed. New York Toronto: Wiley; 2000.

**14.** Öztürk E, Esenboğa K, Kurtul A, Kılıçkap M, Karaagaoglu E, Karakaya J. Measurement of Uncertainty in Prediction of No-Reflow Phenomenon after

Primary Percutaneous Coronary Intervention Using Systemic Immune Inflammation Index: The Gray Zone Approach. Diagnostics. 2023; 13(4):709–720.

**15.** Shannon CE, Weaver W. The Mathematical Theory of Communication. The Bell System Technical Journal. 1948; 27(3), 379-423.

**16.** Benish WA. Relative Entropy as a Measure of Diagnostic Information. Medical Decision Making. 1999; 19(2):202–6.

**17.** Benish WA. The channel capacity of a diagnostic test as a function of test sensitivity and test specificity. Statistical Methods in Medical Research. 2012; 24(6):1044–52.

**18.** Benish WA. A Review of the Application of Information Theory to Clinical Diagnostic Testing. Entropy. 2020; 22(1):97.

**19.** Pepe MS. The statistical evaluation of medical tests for classification and prediction. Oxford University Press; 2003.

**20.** Learned-Miller, E. G. Entropy and mutual information [Internet]. 2003 [cited: 03.04.2023]. Available from: https://people.cs.umass.edu/~elm/Teaching/Docs/mutInf.pdf

**21.** Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. Biometrical Journal. 2005; 47(4):458–72.

**22.** Adams RA, Essex, C. Calculus: A Complete Course.7th ed. Pearson Canada; 2010.

**23.** Alpar R. Uygulamalı istatistik ve geçerlik-güvenirlik: spor, sağlık ve eğitim bilimlerinden örneklerle. 7th ed. Detay Yayıncılık; 2010.

**24.** Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Routledge New York; 1988.

**25.** Salgado JF. Transforming the Area under the Normal Curve (AUC) into Cohen's d, Pearson's r pb, Odds-Ratio, and Natural Log Odds-Ratio: Two Conversion Tables. The European Journal of Psychology Applied to Legal Context. 2018; 10(1):35–47.

**26.** R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

**27.** Landsheer H. UncertainInterval: Uncertain Interval Methods for Three-Way Cut-Point Determination in Test Results; 2021, R package version 0.7.0.

**28.** Kuhn M. caret: Classification and Regression Training; 2022, R package version 6.0-91.

**29.** Smith JW, Everhart JE, Dickson WC, Knowler WC, Johannes RS. Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. Greenes RA, editor, Proceedings of the Symposium on Computer Applications in Medical

Care; 1988 Nov. 6-9; Washington, USA. Los Alamitos, CA: IEEE Computer Society Press; 1988. p. 261-265

**30.** Venables WN, Ripley BD. Modern applied statistics with S-Plus. Springer New York; 2002.

**31.** Wickham H, Sievert C, Springer. ggplot2: Elegant Graphics for Data Analysis. Springer; 2016.

**32.** Kattan MW. Encyclopedia of Medical Decision Making. Thousand Oaks Sage Publications; 2009.

**33.** Mi G, Li W, Nguyen TS. Characterize and Dichotomize a Continuous Biomarker. Fang L, Su C, editors Statistical Methods in Biomarker and Early Clinical Development. , South San Francisco, CA, USA, Springer; 2019.

**34.** Landsheer JA. The Clinical Relevance of Methods for Handling Inconclusive Medical Test Results: Quantification of Uncertainty in Medical Decision-Making and Screening. Diagnostics. 2018; 8(2).

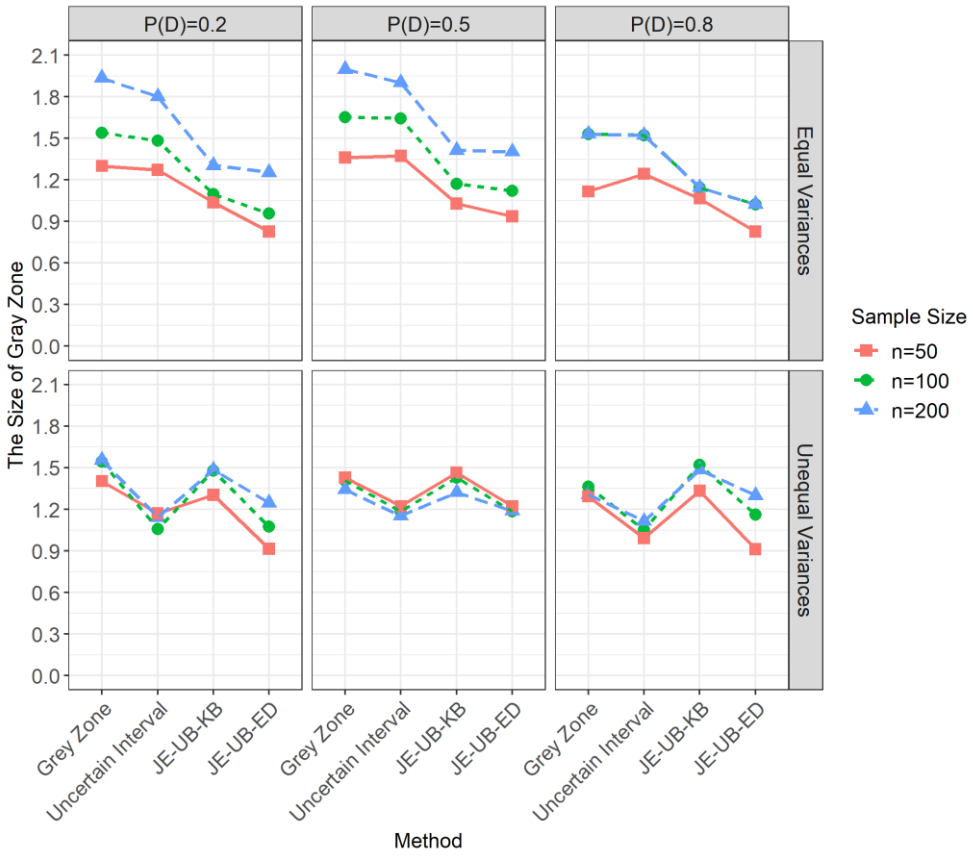
**35.** Somaza, E. Classifying Binormal Diagnostic Tests Using Separation – Asymmetry Diagrams with Constant-performance Curves. Medical Decision Making. 1994; 14: 157-168.

**36.** Li F, He H. Assessing the Accuracy of Diagnostic Tests. Shanghai Arch Psychiatry. 2018 ;30(3):207-212.

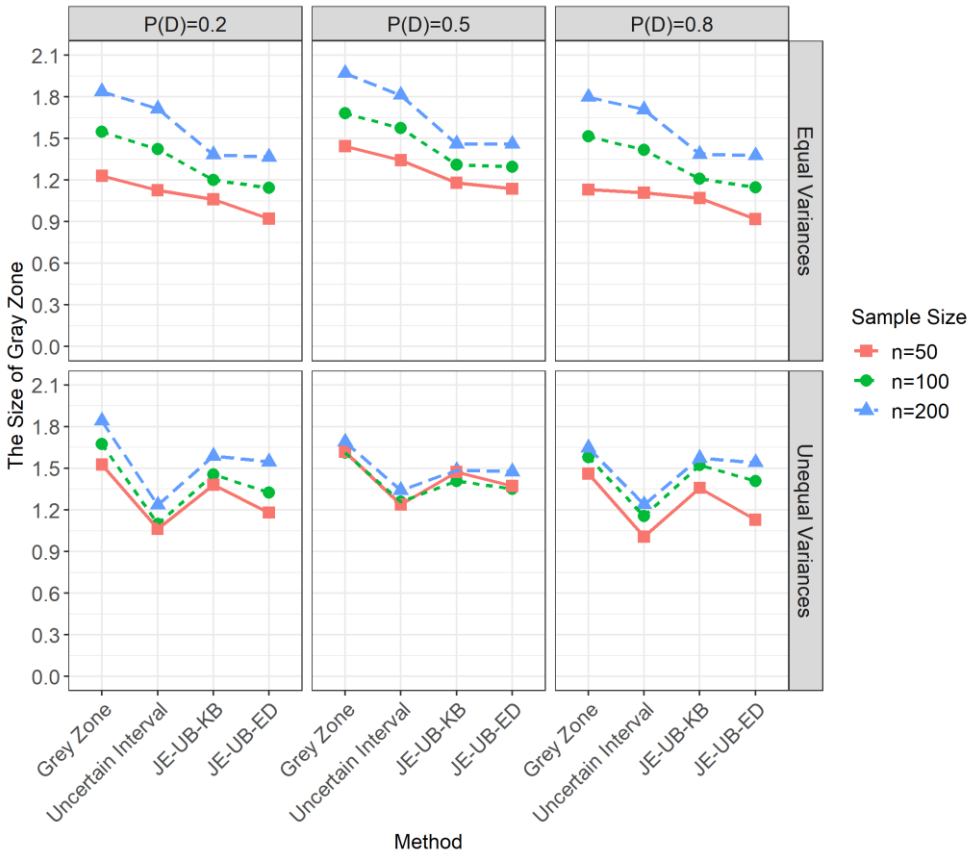
**37.** Franco-Pereira AM, Nakas CT, Reiser B, Carmen Pardo M. Inference on the overlap coefficient: The binormal approach and alternatives. Statistical methods in medical research. 2021; 30(12):2672-84.

# 8. APPENDICES

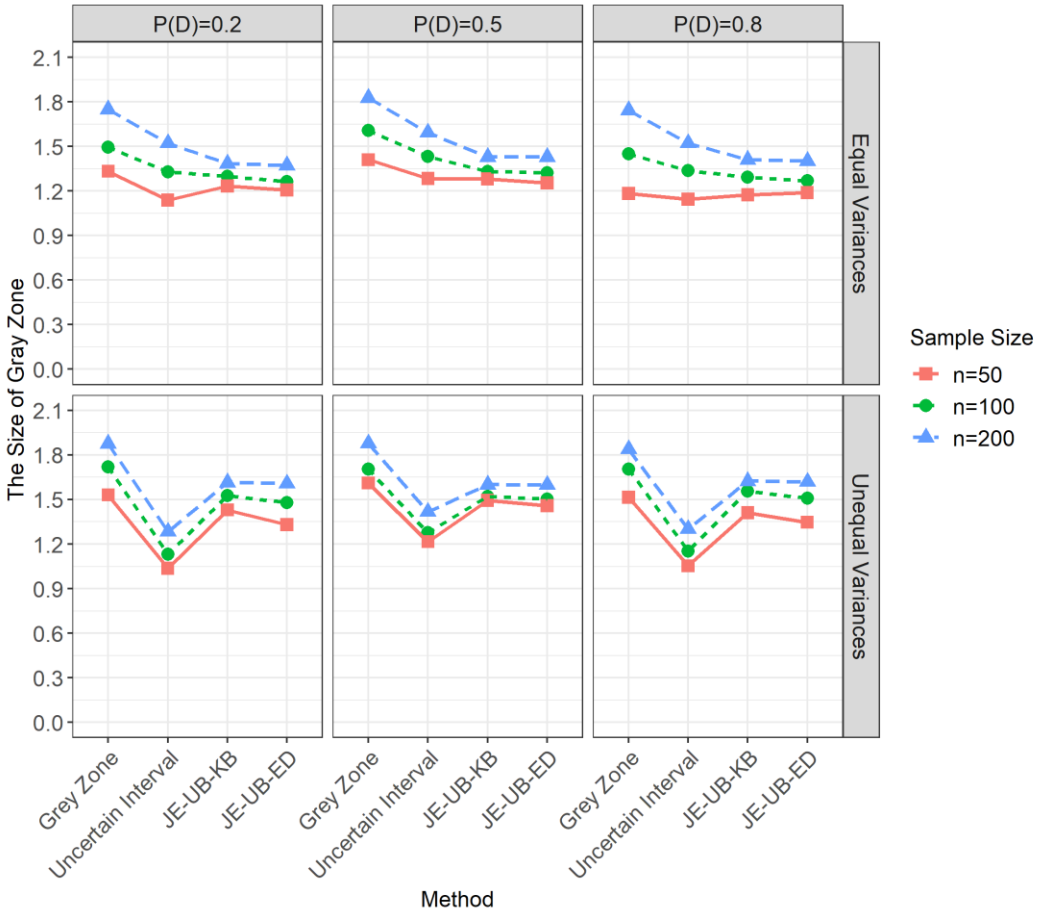
## Appendix 1: The Simulation Results for P(D) = 0.8 and Var(YC)=3 x Var(YD) with Other Scenarios




**A. Figure 1.1.** The Simulation Results for d=0.5, for P(D) =0.2 and 0.5 $Var(Y_D)=3 \times Var(Y_C)$, for P(D) = 0.8, $Var(Y_C)=3 \times Var(Y_D)$.




**A. Figure 1.2.** The Simulation Results for d=0.8, for P(D) =0.2 and 0.5 $Var(Y_D)=3 \times Var(Y_C)$, for P(D) = 0.8, $Var(Y_C)=3 \times Var(Y_D)$.




**A. Figure 1.3.** The Simulation Results for d=1.2, for P(D) =0.2 and 0.5 $Var(Y_D)=3$ x $Var(Y_C)$, for P(D) = 0.8, $Var(Y_C)=3$ x $Var(Y_D)$.

**Appendix 2: The Originality Report of Thesis Study**

turnitin

## Dijital Makbuz

Bu makbuz ödevinizin Turnitin'e ulaştığını bildirmektedir. Gönderiminize dair bilgiler şöyledir:

Gönderinizin ilk sayfası aşağıda gönderilmektedir.

| | |
|---|---|
| Gönderen: | Ebru Öztürk |
| Ödev başlığı: | SPECIFYING THE BOUNDARIES OF GRAY ZONE IN DIAGNOSTI... |
| Gönderi Başlığı: | SPECIFYING THE BOUNDARIES OF GRAY ZONE IN DIAGNOSTI... |
| Dosya adı: | ebru_ozturk_phd_thesis_d_zetlmeler_turnitin.pdf |
| Dosya boyutu: | 2.5M |
| Sayfa sayısı: | 54 |
| Kelime sayısı: | 13,695 |
| Karakter sayısı: | 56,834 |
| Gönderim Tarihi: | 28-Tem-2023 05:22ÖS (UTC+0300) |
| Gönderim Numarası: | 2138059234 |


SPECIFYING THE BOUNDARIES OF GRAY ZONE IN DIAGNOSTIC TESTS WITH INFORMATION CRITERIA

Thesis Title: Specifying the Boundaries of Gray Zone in Diagnostic Tests with Information Criteria

Student Name and Surname: Ebru Öztürk

Total Page Number: 54

SPECIFYING THE BOUNDARIES OF GRAY ZONE IN DIAGNOSTIC TESTS WITH INFORMATION CRITERIA

ORİJİNALLİK RAPORU

| %8 | %4 | %6 | %1 |
|---|---|---|---|
| BENZERLİK ENDEKSİ | İNTERNET KAYNAKLARI | YAYINLAR | ÖĞRENCİ ÖDEVLERİ |

BİRİNCİL KAYNAKLAR

| | | |
|---|---|---|
| 1 | Ebru Ozturk, Kerim Esenboga, Alparslan Kurtul, Mustafa Kilickap, Ergun Karaagaoglu, Jale Karakaya. "Measurement of Uncertainty in Prediction of No-Reflow Phenomenon after Primary Percutaneous Coronary Intervention Using Systemic Immune Inflammation Index: The Gray Zone Approach", Diagnostics, 2023<br>Yayın | %3 |
| 2 | "Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy", Wiley, 2023<br>Yayın | <%1 |
| 3 | researchspace.ukzn.ac.za<br>İnternet Kaynağı | <%1 |
| 4 | Joël Coste, Patrick Jourdain, Jacques Pouchot. "A Gray Zone Assigned to Inconclusive Results of Quantitative Diagnostic Tests: Application to the Use of Brain Natriuretic Peptide for Diagnosis of Heart Failure in Acute Dyspneic Patients", Clinical Chemistry, 2006<br>Yayın | <%1 |

# 9. CURRICULUM VITAE

**Ebru Öztürk**