ALIND MAHDI HADI

INTELLIGENT FUZZY SYSTEM FOR EARLY PROSTATE CANCER DIAGNOSIS

NEU 2020

INTELLIGENT FUZZY SYSTEM FOR EARLY PROSTATE CANCER DIAGNOSIS

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF APPLIED SCIENCES OF NEAR EAST UNIVERSITY

By
ALIND MAHDI HADI

In Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Engineering

NICOSIA 2020

INTELLIGENT FUZZY SYSTEM FOR EARLY PROSTATE CANCER DIAGNOSIS

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF APPLIED SCIENCES OF NEAR EAST UNIVERSITY

By
ALIND MAHDI HADI

In Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Engineering

NICOSIA 2020

Alind Mahdi HADI: INTELLIGENT FUZZY SYSTEM FOR EARLY PROSTATE CANCER DIAGNOSIS

Approval of Director of Graduate School of
Applied Sciences

Prof. Dr. Nadire ÇAVUŞ

We certify this thesis is satisfactory for the award of the degree of Masters of Science in Computer Engineering

Examining Committee in Charge:

Prof. Dr. Rahib H. Abiyev	Committee Chairman, Department of Computer Engineering, NEU
Assist. Prof. Dr. Elbrus Imanov	Supervisor, Department of Computer Engineering NEU
Assoc. Prof. Dr. Kamil Dimililer	Department of Automotive Engineering, NEU

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Alind Mahdi Hadi

Signature: 

Date: 28/12/2020

ACKNOWLEDGMENTS

My first acknowledgement goes to my mother and father and the entire family for always supporting me in my education and my life choices. Special thanks go to my wife and my children from always loving me unconditionally, my love for you will never die and I appreciate your effort. I will like to show my appreciation to my supervisor for his survivorship, guidance and care throughout my master program and through the thesis process. I will like to acknowledge the effort of the head of department for his works in the department. To the lecturers who taught me throughout the master program, thank you. Lastly, to my friends, thank you for your support.

To the victims of prostate cancer in Iraq and the world in general

ABSTRACT

One of the main reasons that kill men annual is prostate cancer. It is the second deadliest and common cancer in men after lung and skin cancer respectively. Prostate cancer in their benign and malignant state can cause damages to the glandular tissues in and around the prostate capsule. In worst cases, the prostate cancer aggressively grows and breaches the capsule and infiltrate nearby organs or separate and enter the blood stream where it is transported to other parts of the body to eventually grow and cause organ failure. Most of the time, prostate cancer grows without showing any major symptoms (early symptoms are similar to urinary tract problems) until it reaches advance stage. Most of the fatalities of prostate cancer can be associated to aggressive cancer and late diagnosis or treatment. Therefore, this study proposed an intelligent system diagnose prostate cancer.

In this study an intelligent fuzzy expert system was designed to for the purpose of diagnosis prostate early cancer among men in all races and age groups. The intelligent system is based on the Fuzzy Inference System (FIS) using MATLAB (Matrix laboratory) package and Graphics users Interface (GUI) to predict the presence or absence of prostate cancer. The systems are planned for detecting from early to malignant stage of prostate cancer using some important variables with a reference range. Important parameters such as PSA levels, age, biopsy, ethnicity, and so on were utilized as inputs. The output was designed in to three parts; positive, negative, and suspicious. The positive indicates diagnosis of prostate cancer considering all parameters while negative represent no cancer. Suspicious on the hand, indicates a problem that may be cause by another pathology instead prostate cancer, and may require further analysis. The study used the knowledge-based rules both in Mamdani and Sugeno type inference to process the inputted variables and provides a useful diagnosis as output.

Keywords: MATLAB (Matrix laboratory) software, Graphics users Interface (GUI); fuzzy inference system (FIS), and Prostate cancer.

OZET

Erkekleri her yıl öldüren ana nedenlerden biri prostat kanseridir. Erkeklerde sırasıyla akciğer ve cilt kanserinden sonra en ölümcül ve en yaygın ikinci kanserdir. İyi huylu ve kötü huylu prostat kanseri, prostat kapsülü içindeki ve etrafındaki glandüler dokulara zarar verebilir. En kötü durumlarda, prostat kanseri agresif bir şekilde büyür ve kapsülü ihlal eder ve yakındaki organlara sızar veya sonunda büyümesi ve organ yetmezliğine neden olmak için vücudun diğer bölgelerine taşındığı kan akışına girer ve girer. Çoğu zaman prostat kanseri, ileri aşamaya gelene kadar herhangi bir majör belirti göstermeden (erken belirtiler idrar yolu problemlerine benzer) büyür. Prostat kanserinin ölümlerinin çoğu, agresif kanser ve geç tanı veya tedaviyle ilişkilendirilebilir. Bu nedenle, bu çalışma prostat kanserini teşhis eden akıllı bir sistem önermiştir.

Bu çalışmada, tüm ırk ve yaş gruplarındaki erkeklerde prostat erken kanserinin teşhisi için akıllı bir bulanık uzman sistemi tasarlanmıştır. Akıllı sistem, prostat kanserinin varlığını veya yokluğunu tahmin etmek için MATLAB (Matrix laboratuvarı) paketi ve Grafik kullanıcı Arayüzü (GUI) kullanan Bulanık Çıkarım Sistemine (FIS) dayanmaktadır. Sistemler, bir referans aralığı ile bazı önemli değişkenler kullanılarak prostat kanserinin erken evresinden kötü huylu evresine kadar tespiti için planlanmıştır. PSA seviyeleri, yaş, biyopsi, etnik köken gibi önemli parametreler girdi olarak kullanıldı. Çıktı üç bölüm halinde tasarlandı; olumlu, olumsuz ve şüpheli. Pozitif, tüm parametreler dikkate alındığında prostat kanseri teşhisini gösterirken negatif, kanser olmadığını gösterir. Eldeki şüpheli, prostat kanseri yerine başka bir patolojinin neden olabileceği ve daha fazla analiz gerektirebilecek bir sorunu gösterir. Çalışma, girilen değişkenleri işlemek için hem Mamdani hem de Sugeno tipi çıkarımdaki bilgiye dayalı kuralları kullandı ve çıktı olarak yararlı bir tanı sağlıyor.

Anahtar Kelimeler: MATLAB (Matrix laboratuvarı) yazılımı, Grafik kullanıcı Arayüzü (GUI); bulanık çıkarım sistemi (FIS) ve Prostat kanseri.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF THE ABREVATIONS

AI:	Artificial Intelligence
AJCCpTNM:	American Joint Committee on Cancer
ANFIS:	Adaptive network-based fuzzy inference system
AUC:	Areas under the ROC curve
BPH:	Benign prostate hyperplasia
CHFS:	Cubic hesitant fuzzy set
CT:	Computed tomography
DRE:	Digital rectal exam
ED:	Extra-Prostatic Disease
FES:	Fuzzy expert system
FIS:	Fuzzy Inference System
FPSA:	Free PSA
GUI:	Graphics users Interface
OCD:	Organ-Confined Disease
PC:	Prostate cancer
PCC:	Prostate cancer cells
PCR:	Prostate cancer risk
PET:	Positron emission tomography
PGC:	Prostate gland cells
PSAP:	Prostatic specific acid phosphatase
PSO:	Particle swarm optimization
PSA:	Prostate specific antigen
PV:	Prostate volume
MATLAB:	Matrix laboratory
mL:	Milliliter
MRI:	Magnetic resonance imaging
ng:	Nanograms
ROC:	Curve
UCI:	University of California Irvine

US:	United states
TCGA:	The Cancer Genome Atlas

CHAPTER 1

INTRODUCTION

1.1. Background

Globally, cancer has been recorded as the second-high profile cause of deaths among men. Prostate cancer is also known to be the second and fourth most prevalent cancer cases in men and among all cancer types (regardless of gender) respectively (world cancer research fund (WCRF), 2018). According to world health organization (WHO), about 1.3 million of prostate cancer were recorded in the year 2018 (Bray et al., 2018). In fact, it occupies second place for high levels of annual mortality in men after heart diseases. Prostate cancer is more prevalent among men of older age, as report shows a mean of 66 years old (Rawla, 2019). The older the man grows, so do the prostate grow in size, however, this is benign cancer (nonmalignant) and is usually not harmful at this stage. However, prostate cancer becomes eminent when prostate of the cells begins to be uncontrollable replicated or observed uncontrollable apoptosis in the prostate gland, and breach the prostate capsule to nearby nodes, organs, and eventually the rest of the body.

There are several symptoms that are used as indicators for the diagnosis of prostate cancer. Such symptoms are important bioindicators that not only provide information for informed medical decisions about the presence and absence of prostate cancer, but whether something else is wrong with the prostate or urinary tract. Alongside, the symptoms are the risk factors which show the level of vulnerability of an individual to prostate cancer due to certain factors. Such factors include age, environment, ethnicity, region, diet, occupation, and gene heredity etc. Some vital early symptoms of prostate cancer are sometimes similar to symptoms associated to bowel or other urinary medical conditions. Therefore, symptoms should be used in combination with risk factors to better prediction of prostate cancer. Furthermore, various tests are carried out to further narrow down the possibility of both non-cancerous and cancerous tumor growth in the prostate gland. This can be achieved using DRE, PSA, and biopsy of sampled cells. DRE refers to digital rectal examination (DRE), while PSA stands for prostate specific antigen.

The mortality rate of localized prostate (diagnosed early) is almost 100%. This means on-time analysis of can be passionate about the prevention of its effects, its control, and its medication. An effective combination of symptoms, risk factors, and test result can provide a near accurate prediction of any cancer type. In DRE, the medical expat manually feels for any abnormal growth or lumps as a result of tumor either through the rectum by means of palpation. However, digital rectal exam (DRE) is mostly useful at middle and late stage of prostate cancer because they tumor has already gown and obstructing vital organ functions especially within the urinary system and rectum, it is therefore almost useless for on time detection of cancer cells for proper diagnosis of prostate cancer. If DRE seems to be unsuccessful, a better approach is the histopathology test that is useful in the measurement of the levels of PSA inside the blood stream of the patient.

Reports have shown that several men that do not have cancer cells in their prostate gland are mostly associated with PSA levels that are low or within the threshold of 4 ng/mL, and most patients with diagnosed with prostate cancer have higher levels of PSA inside the blood stream. However, levels of PSA levels in the blood are not always an effective method for identifying the presence or absence of prostate cancer, since other urinary pathologies or prostate problems can increase the levels of PSA in the blood. Therefore, biopsy is a better option where samples of the prostate cells are taken from multiple tissue locations of the prostate to the laboratory for analysis to check for presence of cancer cells. Among all the mentioned test, biopsy seems to be the most reliable, however, it has other challenges that my not detect the manifestation of very early prostate cancer growth. After prostate cancer has been detected, the analysis of Gleason score and prostate cancer staging is implemented in order to have knowledge of each stage the cancer has reached. This will indicate how the prostate cancer cells have grown aggressively to nearby organs/entire body or whether it has confined in the prostate capsule or has spread pass the prostate capsule.

All of these tests and processes for diagnosing prostate cancer have a high degree of uncertainty, thereby sometimes leading to illogical as well as uncertain evaluation of clinical outcomes to some certain degree. There is also high possibility that fuzzy and uncertain data in these existing tests for the evaluation of prostate cancer may be lost with

non-fuzzy information. Apart from this, there are also high possibilities of both human and errors arising from the instrumentation used for testing, thereby jeopardizing the final result. Although clinicians are doing their best, they are also humans and may sometimes affect the accuracy of results with errors. This jeopardizes accurate diagnosis of prostate cancer. Shukla, et al. (2009) believes that clinicians have high possibility of error sabotage in diagnosis processes.

To resolve or avoid this issue, intelligent expert systems can be used. These systems are based on artificial intelligence and help clinicians check accurate likelihood of prostate cancer from laboratory examinations data (Isa et al., 2010; Sarasvathi and Santhakumaran, 2011). According to Farokhzad, and Ebrahimi (2016), fuzzy expert system utilizes the concept of fuzzy logic, that provides benefits such as user friendly, high flexibility, tolerance of inexact data suitably, modelling of complicated non- linear functions, to act based on specialized experience, adjustment with common controlling procedures and its natural based language. On another description, intelligent expert systems can be categorized as one of the facets of artificial intelligence (AI) that actively encourages the utilization of techno-scientific human knowledge/skills for the purpose of providing solutions to partial or full day to day problems that occur in all the places that there is not a particular certainty of finding the algorithm.

According to Rajabi et al. (2019), any intelligent expert system that is developed utilizes intelligent processes (with high accuracy and precisions) that provides solutions to complex issues in order to get significant human information to clarify them. Although expert systems are intelligent and uses information and inferences levels for problem solving, it still relies on several factors in order to provide desired outcomes. The algorithms for developing intelligent expert systems should be carefully established otherwise it will ultimately not perform well. Another factor to consider is that, depending on the quality of the data or information provided to an expert system, that may significantly affect the outcome (Bridged, 2019; Hagendorff and Wezel, 2020). So, if the system is provided with quality data and instructions, then the likelihood of any error is significantly reduced and vice versa.

Furthermore, an intelligent expert system that utilizes the concept of fuzzy logic is referred to as fuzzy expert system or intelligent fuzzy expert system (FES). It can be a fixed set of information-based system that consist of the following characteristics; fuzzification, information database, inference rules, and defuzzification parts. Additionally, the fuzzy expert system does not use the Boolean logic (traditional logic of 0 and 1representing false and truth respectively), rather, it utilizes the fuzzy logic which uses partial truth that includes all the values ranging from 0 to 1 (Stanford Encyclopedia of Philosophy, 2017). That means instead of true or false, we may have completely true or completely false. As a result, the fuzzy logic concept takes in to consideration the data in the inference mechanism. This system is accepted to describe decision-making problems, everywhere there is no scientific algorithm occurs, while instead of, the solution to the problem could be possibly considered as heuristical, that emanates from a medical expert in the form of If-Then rules (Alaybeyoglu, and Mulayam, None). A fuzzy expert system is enough of a model for the provision of effective and efficient solution to real life problems that are usually associated with hesitation releasing from fuzziness, uncertainty or partiality

In modern times, the adoption and implementation of fuzzy expert systems (FES) have been conducted in several studies in academia across the globe (Rajabi et al., 2019). In this research, we will utilize MATLAB software due to its flexibility, plenty of available functions and its efficiency to simulate fuzzy logic. MATLAB software increases the precision of results and the contrast of the effectiveness of systems and delivers the best operation for system training in the shortest time probable.

Previous studies that are strongly related to my study include the study of Fu et al. (2018) and Mahanta and Panda (2018). The first study related to my work is the study of Mahanta and Panda (2018) who investigated risk factors associated to prostate cancer among 119 male patients using FES- the Mamdani (involving maximum-minimum values) inference method which is based on IF-THEN form. Among the 119 male patients participating in their study, 61 had positive biopsy results, while 58 patients had negative results. Their study took in to consideration patients' vital parameters such as age associated to the patient, the level of PSA in the patient, prostate volume, and % of free PSA (%FPSA) in the patient's blood stream. After successfully inputting the values for each parameter in to

the model, they obtained their output as prostate cancer risk. If the output for a given patient is ≥ 50%, there is a likelihood of abnormality with the prostate either benignant or malignant stage. Results shows that the fuzzy expert system model have 68.91% accurately (true) predicted the likelihood of prostate cancer among the patients. In comparison to the biopsy result, the model predicted 45 patients to have positive biopsy result and 37 negative results. Comparatively, this predictive result is better compared to the studies of result Saritas et. al (2013) with accurate (true) prediction at 64.71% for the same data set. The researchers also indicated the significance of the vital parameters in early detection of prostate cancer.

The second related study to my work is the study of Fu et al. (2018) who proposed an intelligent hybrid system for the prediction of prostate disorder. There model combined fuzzy logic cubic set and a hesitant fuzzy set. The hybrid model was designed to handle uncertain as well as hesitant fuzzy medical data of prostate risks grades. Furthermore, the study introduced an overall distance and similarity measure of cubic hesitant fuzzy set (CHFS). After that, the study the study also proposed an elaborate approach of evaluating risk by utilizing the cubic hesitant fuzzy set (CHFS) similarity measure. A total of 16 clinical prostate cancer patient cases were tested in the model for prediction. After successful analysis, the results show that this model is better in terms of evaluation performance compared to the evaluation method used by Ren et al. (2013).

1.2. Statement of problem

Prostate cancer kills a huge number of men annually, therefore making it an important threat to the health of men. It is important to continue to tackle the growing pandemic status of prostate cancer. Using artificial intelligence prostate cancer (a real-life problem) can be diagnosed effectively and efficiently, of course with the guidance of the medical experts. As it is with all real-life problems, ambiguity is usually associated with the results. Therefore, fuzzy expert systems can rise to the occasion. For this study, Fuzzy Inference System (FIS), MATLAB (Matrix laboratory) software function and Graphics users Interface (GUI) has helped doctors to embody these diseases. The researcher had a problem in using Laboratory-device, MRI, CT scan and PET scan reader results when

taking the result of test value after that need to record this value by manual in MATLAB (Matrix laboratory) software program until to get a type of diseases (Şahan, et al., 2007). Fuzzy expert systems have the capability of providing answers or a way out to real world problems in the medical field that may be ambiguous. The scope of the application of fuzzy expert systems have expanded to almost every field of discipline and researchers have presented scientific research works. One of such scientific research works is the study of Polat et al. (2006) that was focused on the analysis as well as forecast of a variety of medical conditions.

1.3. Research Aim and Objectives

This study focuses on the development of an intelligent, dynamic, as well as accurate fuzzy logic computational system that can be helpful in diagnosing prostate cells among men. The following objectives will be explored;

- To predict possible prostate cancer from the patient.
- To use MATLAB (Matrix laboratory) software to identify the stages of prostate cancer from patient laboratory data.
- To generate or establish rules from actual expert experience as well as medical records of prostate cancer patient cases that in order to test the performance of the system.
- To provide the percentage of correct prediction of the system compared to previous systems.

1.4. Significance of Study

The study is significant because it contributes to the growing literature on the application of fuzzy logic intelligent systems in the medical field. Specifically, it contributes to the mission of intelligently diagnosing diseases effectively and efficiently with a significant positive accuracy level as well as the level of precision.

CHAPTER 2

LITERATURE REVIEW

2.1. Introduction

Several medical and alternative studies have been conducted in various pathological disorders including breast, lung, and liver disorders using intelligent systems. Some very important intelligent systems include fuzzy logic, artificial neural networks, and genetic algorithms (Adeli and Neshat, 2010). However, no much researchers have explicitly explored the significance of intelligent systems in prostate disorders. Intelligent systems are designed to assist medical physicians with accurate prediction or likelihood to make clinical decision. Therefore, research on designing intelligent systems for improving the accuracy, effectiveness, and efficiency of predicting the likelihood of prostate cancer is very significant and relevant.

In the real world, nothing is certain or absolute. The various systems do not function by absolute values due to the influence of other systems or unwanted influence. In fact, no system function completely in an absolute manner, they usually function by some degree of vagueness (McCrady et al., 1998). It is up to the humans to create means to understand the natural values of functioning system by processing or filtering and other wise to get the true or absolute values. In order to diagnose any form of cancer, there are several factors that must be put in to consideration to not only provide an effective diagnosis, but also an efficient one. This statement is same when artificial intelligence approach is used for the diagnosis or prediction of cancer diseases. According to Tewari et al. (2001), in order to effectively predict any form of cancer, an intelligent computer system (computer intelligence) must be employed and that intelligent system must consider all vital variables or factors associated to that specific form of cancer.

It is widely known that adoption and implementation of computational intelligent, models have been utilized in the prediction of prostate cancer status among men, however, only a small percentage of such established models are capable of providing a reasonable outcome for the prediction of the prostate cancer pathological stages. In the area of

computational intelligence, models established for classification purposes are in fact used for prediction of various tasks and especially real-life problems such as prostate cancer. Han (2005) describes classification models in computer intelligence as an approach of analyzing data involving the extraction of classifier models that have descriptions of classes of data, which are used for the prediction of labels in categories (classes) or numeric values. In a computational indigent system where a classifier is utilized for the predication of numeric values, such operation is called a predictor. Both the operations of classification models and prediction of numbers are actively employed for the predicting. According to Cosma et al. (2016), a researcher should consider implementing classification models in the analysis of medical data and for the purpose of extracting a model for prediction in the medical scenario.

In the medical field, medical experts have to uncover the real values of our biological processes either in the form of numerical values, colors, smell etc. to make informed decision about patient's health status. However, the human biological processes can be understood as confined space consisting of different system after by bother external and internal influences. These systems interact with each other to affect individual absolute values. For instance, early diagnosis of a very specific neurological disease from only the known symptoms may be challenging for medical specialists, due to several factors including the fact that humans are same but different (symptoms can vary from patient to patient), presence of approximate and inaccurate data, environmental factors, genetic factors, life style of patients, and even chances (U.S. Department of Health and Human Services, 2019).

On the other hand, Mahanta and Panda (2018) explained that a specific symptom can be shared with similar diseases leading to disconnected diagnosis, while a specific disease can express diverse types of symptoms depending on the patient influenced by several factors.

Characteristics make these algorithms a suitable platform on which to base new strategies for diagnosing and staging prostate cancer. For example, not everyone diagnosed with prostate cancer will exhibit abnormal results in all tests, as a consequence of which, different test result combinations can lead to the same outcome. (Cosma et al. 2016).

Furthermore, medical doctors are humans too; this mean they sometimes get tired, confused, or simply make errors. This can significantly jeopardize the accuracy and precision of the result, thereby affecting effective and efficient diagnosis as well as the process of medical decision making (Mahanta and Panda, 2018 files). All of these arguments indicate the challenges with accurate and precise diagnosis of pathological problems in the medical field.

In diagnosis of prostate cancer, biomarkers specific to the prostate are usually the leading indication of the development of prostate cancer and the various stages. This biomarker is called the prostate-specific antigen (PSA). The value of PSA is obtained by running a PSA test. A patient with higher levels of PSA have higher indication of the presence of prostate cancer (Eggener et al, 2020). PSA marks any abnormal differentiation or apoptosis (tumor) and is very specific to the prostate gland (Fu et al., 2018). There are several factors to consider in combination with PSA levels; this means only PSA is not sufficient for the accurate and precise diagnosis of prostate cancer (Sternberg et al., 2014; Cella et al., 2015). Additionally, PSA level in person may contain uncertain values and a high degree of vagueness.

In analysis of prostate cancer, staging of the diseases is very crucial. One of the ways to achieve stages of prostate cancer is through the Gleason scoring system. This scoring system is used in histological grading of the condition of the prostate to determine both the stage and the prognosis. Furthermore, to identify the degree of prostate cancer spread within the prostate capsule, to neighboring organs, or to other parts of the body, the TNM staging are utilized (Fu et al., 2018). To expatiate on the TNM staging system, letter T represent the tumor size and the extent of development of progression from the primary tumor to the next stage. In the TNM, the T (T2) is crucially used to measure the degree of prostate cancer risk of a patient (Edge et al., 2010). On the other hand, letter M represent the summation of proximal lymph nodes that the primary cancer had spread (secondary tumor). Lastly, letter M stands for metastasization of the prostate cancer from the primary tumor to nearby or distant organs (National cancer institute, 2015).

In general, the diagnosis of prostate is performed by considering a combination of result value from PSA test, Biopsy, TNM staging, Gleason score value, and T2 staging score value (Partin et al., 1997; Fu et al., 2018). Some other risk factors considered include patient age, and the volume of the prostate (Saritas et al., 2003 files). Despite using a combination of one or more values obtained from these medical procedures, there is still some significant degree of uncertainty, vagueness, or simply fuzzy values. These could be due to errors in the process of testing, or from the testing instruments, error from the medical experts etc. whatever the error causing the uncertainty is, the accuracy of prostate cancer risk becomes significantly affected, leading to ineffective or mis diagnosis. One of such consequences is underestimating the risk (where there is cancer but the test does not indicate (false positive)) or over estimating (where there is no cancer but the test indicate there is cancer (false negative)), leading to lack of early treatment or administration of unnecessary treatments to the patient (Gospodarowicz et al., 2015). This has constituted to the reason for increased morbidity, disability, premature death and more costly health services (Abdulhaqq, 2019).

Studies have shown that the existing methods of evaluating the likelihood or predicting the risk of prostate cancer is inadequate (Edge et al., 2010; Ren et al., 2013; Fu et al., 2018). According to Fu et al. (2018), these challenges is blamed on the difficulty of depicting and evaluating the various factors and stages of prostate cancer using absolute and imprecise ranges. Some of the process of staging prostate cancer could possibly provide uncertainty in their resultant values due to the occurrence of some patients' risk factors or data belonging to diverse risk grades. Another issue is the prostate biopsy that increases the PSA level in the patient system. this can cause not only misdiagnosis, but also complications such as the spread of cancer from the primary tumor (Metlin et al., 1991). In order to handle the uncertainty in risk evaluation, staging, and general diagnosis of prostate cancer, using intelligent expert systems based on the working principles of fuzzy logic can help eliminate or significantly reduce these uncertainties (Adeli and Neshat, 2010; Mahanta and Panda, 2018).

2.2. Fuzzy Set Theory

Fuzzy set theories are mathematical tools works on the basis of fuzzy logic. The idea of fuzzy set theory was first conceived and made public in the year 1965 by Prof. L. A. Zadeh to handle the inefficiencies of traditional or binary logic (Aristotelian logic). After, its introduction, it gone through various development and applications in wide areas including medicine. The concept of fuzzy logic is based on the idea that the practical world systems are not absolute, therefore we must consider all of the values to properly represent them in studies. This consideration is referred to as multi-valued logic (Mahanta and Panda, 2018).

One of the very reasons why the concept of fuzzy logic has gain huge success and acceptance in several fields of discipline is not only due to its effectiveness to handle real and practical world systems, but it is also very user friendly without requiring any complex processes. In medicine and healthcare, fuzzy logic studies have applied the concepts in the processes of risk prediction, decision making, and general improvement in diagnosis of diseases such as breast, lung, and liver diseases. However, there are only a handful of researchers that have explored the application capabilities of fuzzy logic in improving prediction of risks in prostate cancer and other prostate related diseases (Seritas et al., 2013; Kar and Majumder, 2017).

According to Gorgulu and Akilli (2016 files), the concept of fuzzy logic is a facet of artificial intelligence that deals with linguistic variables represented using mathematical models of the real-world systems. Intelligent system has proven to perform almost similar to human capacity. However, there are several questions including ethical reasons that artificial intelligent systems cannot over take human intelligence. Intelligent systems are required to assist or complement areas that may be challenging to human intelligence. Samuel et al. (2013) indicated that one of the aims of fuzzy logic is to assist in solving problems similar to human capabilities. In the medical field, Ekong et al. (2012) and Vaghela et al. (2015) fuzzy logic address the problems of lack of precision and certainty by considering all values in medical data that Aristotelian logic cannot handle. The Aristotelian logic is only run by some degree of imprecision as well as uncertainty such as the pro vision of only two options for logical values either ranging from black to white, from true to false, and 1-0 (Gorgulu and Akilli, 2016).

Fuzzy logic considers all values in a data set and converts in to values between 0 and 1. For example, instead of handling a data set of just hot and cold, fuzzy logical can tell you the degrees of coldness (slightly cold, moderately cold, severely cold)or hotness (slightly hot, moderately hot, severely hot) of the data set. Thus, rendering it a very powerful tool for nearly accurate representation and prediction of the data set. Fuzzy set theories are usually designed as fuzzy logic and based decision support systems (Belard et al. 2016) to provide swift and cost-effective decision-making solution without compromising accuracy and precision (Gorgulu and Akilli, 2016).

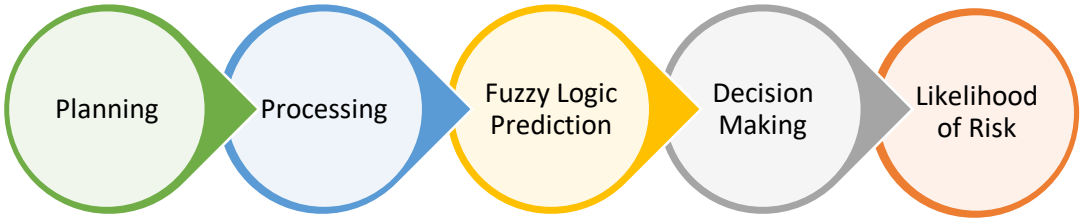


Figure 2.1: Basic intelligent fuzzy logic process for the prediction of likelihood of prostate cancer in sequential order (Gorgulu and Akilli, 2016)

The fuzzy logic process uses natural languages to tackle uncertain values (Chen and Bau 2013; Fraccaro 2015). For prostate cancer prediction, these uncertain values care obtained from PSA test, Biopsy, TNM staging, Gleason score value, and T2 staging score value to obtain the likelihood of prostate cancer risks. Different studies have developed fuzzy logic based computational intelligent and mathematical models. Of these fuzzy models include rule-based fuzzy expert system (FES) (Seritas et al., 2013), cubic hesitant fuzzy set (CHFS) (Fu et al., 2018), and adaptive network-based fuzzy inference system (ANFIS) (Rajabi et al., 2019) etc. This thesis uses ANFIS developed by Jang (1993) for the sole purpose of predicting the likelihood of cancer of the prostate and their staging. The ANFIS fuzzy model is suitable for this study because it has shown to be economical and provide accurate predictions due to it high sensitivity, flexibility, ability to endure error data properly, showing complex non-linear functions, considers real patients data and expert knowledge, as well as its ability to conform to routine controlling systems.

Furthermore, particle swarm optimization (PSO) is considered a very popular swarm intelligence technique and it has a crucial role in the influencing the progress of our analysis. According to Rajabi et al. (2019), PSO is responsible for tuning hyper-parameters of ANFIS like the number, kind of fuzzy membership functions and developing the fuzzy rules. ANFIS and PSO if run in MATLAB software increases the degree of precision from the outcome of the study. the MATLAB software is suitable because it is flexible, with diverse operational functions researchers can choose from, and its higher efficiency in the simulation of fuzzy logic. Additionally, the MATLAB software provides a platform for swift training of systems without compromising effectiveness. In general, the study used a combination of MATLAB software functions, Interference fussy system (IFS) and Graphics users Interface (GUI) (Farokhzad and Ebrahimi, 2016).

2.3. Previous work on Prostate Cancer Prediction

Most studies do not use a single concept to provide solutions to problems, rather, a combination of concepts provides a better and more accurate solution. This is called hybrid system approach. researchers have combined the concepts of fuzzy logic with neural networks to provide better results. Therefore, this section explores some of the previous studies that used hybrid fuzzy logic systems to predict the likelihood of prostate cancer using patients' vital medical data.

Gorgulu and Akilli (2016) uses their proposed fuzzy logic-based decision support systems for medical diagnosis applications. Their model was based on the If-Then rule base which considers the imputation of medical data, conversion of data in to fuzzy scale (fuzzification), inference, defuzzification of data, a meaningful output in form of assistive system decision. After successful application in both heart and prostate disorders, the mean success of the model was found to be 90%. Abiyev and Abizade (2016) predicted the likelihood of a different disorder called Parkinson's diseases using fuzzy neural system designed as an automated recognition system for the detection of the disorder in patients. Basically, there design was fabricated to differentiate between normal or healthy patients with patients with Parkinson's disease. This was achieved through the incorporation of fuzzy system and neural networks. Simulation was run on data collected from UCI

machine learning repository. Findings were obtained from their simulation using the proposed designed recognition system shows that the intelligent fuzy logic based neutral system designed provides a significant enhancement to previous recognition rates in designs established from previous studies. Other studies such as Abiyev et al. (2015) and Idoko et al. (2018) also found similar conclusion to Abiyev and Abizade (2016).

Furthermore, Cosma et al. (2016) proposed a neuro-fuzzy model which is a hybrid intelligent system the involves the incorporation of the concepts of neural system and fuzzy logic. The model was designed and proposed to make prognosis or prediction of prostate cancer according to the stages they exist. In order to achieve this, the study took in to consideration vital parameters such as Primary and secondary Gleason pattern, PSA levels, patients' age, as well as the clinical T stage. After inputting the vital parameters and training it, the intelligent model establishes the necessary fuzzy rules need for application on patients known medical data. A combination of the fuzzy rules and the test conducted to validate the data set is used as a prediction process of the various stages of prostate cancer. Furthermore, the study optimized the predictive performance of the intelligent hybrid system using adaptive neuro-fuzzy inference system. the result from this model when compared to previous models and the AJCCpTNM staging nomogram provided the highest area under the ROC curve (AUC) as well as the lowest incidence of false positives. The researcher therefore concluded that their model is an improved development of AJCCpTNM staging nomogram.

Ma'aitah, Abiyev, and Bus (2017D) suggested an intelligent fuzzy neural system to distinguish liver diseases, based on neural networks and fuzzy logic. After conducting 10 cross proof tests and utilizing a dataset extracted from the UCI repository, the show was executed. The researchers conducted two experiments with two precisions, 72%, and 97%, respectively. This revealed that their second experiment based on the suggested show gotten the optimal result. Using medical data collected from 299 patients, Tsao et al. (2014) proposed an artificial neural network model for the purpose of predicting the pathological stages of prostate disorder. All participants of the study had not undergone radical prostatectomy. After successful analyzing the data, the study found that their model was had excellent capabilities of predicting primary or localized tumor (cancer that

develops and spreads within an organ) prostate cancer compared to Logistic Regression model and Partin Tables.

In their study, Mahanta and Panda (2018) investigated the prostate cancer risk factors among 119 male patients using a fuzzy expert system- the Mamdani (involving maximum-minimum values) inference method which is based on the IF-THEN form. Among the 119 male patients participating in their study, 61 had positive biopsy results, while 58 patients had negative results. Their study took in to consideration patients' vital parameters such as age associated to the patient, the level of PSA in the patient, prostate volume, and % of free PSA (%FPSA) in the patient's blood stream. After successfully inputting the values for each parameter in to the model, they obtained their output as prostate cancer risk. If the output for a given patient is $\geq 50\%$, there is a likelihood of abnormality with the prostate either benignant or malignant stage. Results shows that the fuzzy expert system model have 68.91% accurately (true) predicted the likelihood of prostate cancer among the patients. In comparison to the biopsy result, the model predicted 45 patients to have positive biopsy result and 37 negative results. Comparatively, this predictive result is better compared to the studies of result Saritas et. al (2013) with accurate (true) prediction at 64.71% for the same data set. The researchers also indicated the significance of the vital parameters in early detection of prostate cancer.

Lastly in this literature review, Fu et al. (2018) proposed an intelligent hybrid system for the prediction of prostate disorder. There model combined fuzzy logic cubic set and a hesitant fuzzy set. The hybrid model was designed to handle uncertain as well as hesitant fuzzy medical data of prostate risks grades. Furthermore, the study introduced an overall distance and similarity measure of cubic hesitant fuzzy set (CHFS). After that, the study the study also proposed an elaborate approach of evaluating risk by utilizing the cubic hesitant fuzzy set (CHFS) similarity measure. A total of 16 clinical prostate cancer patient cases were tested in the model for prediction. After successful analysis, the results show that this model is better in terms of evaluation performance compared to the evaluation method used by Ren et al. (2013).

CHAPTER 3

THE PROSTATE AND PROSTATE CANCER

3.1. Anatomy of the Prostate Gland

The prostate gland has a very important role it plays in the male reproductive system. it is a biological factory for the production of a secretion that constitute the male discharged sperm. The prostate gland produces this secreted fluid that is part of the seminal fluids. Without the prostate gland secretion, then sperm cells may be immobile, underperform, and subsequently die. According to Brennhovd and Iversen (2016), the secretion from the prostate gland is a requirement needed for sperm cells to be able to not only performance, well, but to also survive, and move freely within the gland and the destination necessary. The prostate is built like a capsule as seen in Figure 3.1 in which prostate specific antigen (PSA) are produced.

PSA are secreted alkaline which are rich in enzymes and prostate glandons. Prostate gland lays below the bladder proximally, and can be seen in front of the rectum. This location is the reason why problems associated to prostate can be related to urinary tract infection and the ability of prostate tumors to be detected through the rectum respectively. The prostate gland is built like a capsule or a walnut shape. It encircles the upper part of the urethra. The prostate, bladder, urethra, seminal vessels and so any other organs work together under the urethmetic contraction during the processes of urination and ejaculation etc. by closing and opening one side to block the flow of a specific fluid (e.g. urine) in order to allow the flow and passage of the other (e.g. semen) and vice versa.

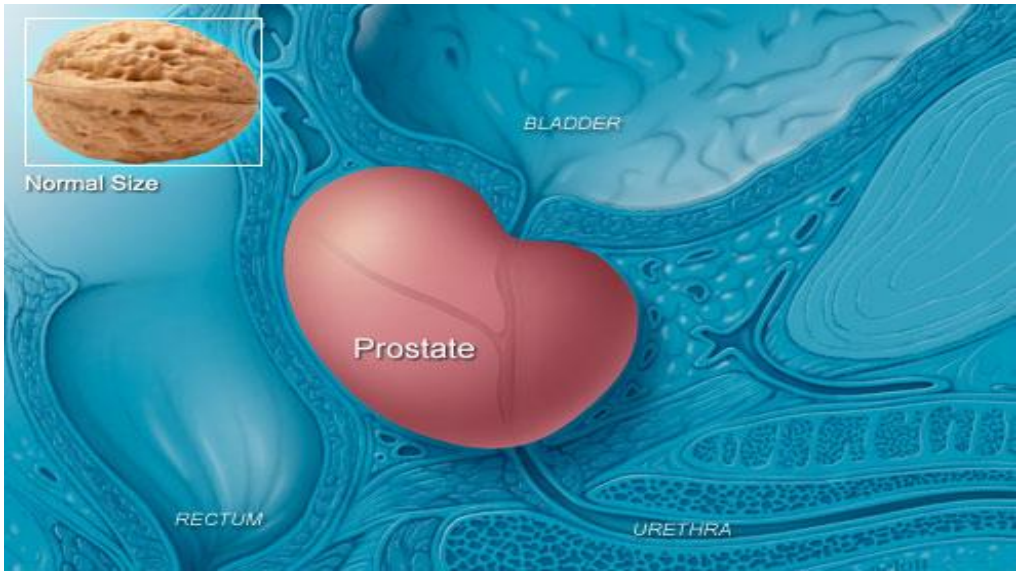


Figure 3.1: location of prostate gland (Brennhovd and Iversen, 2016)

As the man grows older, the prostate grows bigger in volume commonly known as benign prostate hyperplasia (BPH) (Chang, 2018). That means, new cells are growing faster than they are killed and begin to affect the normal functions of the urethra and the bladder. The worst case of prostate problem is the carcinoma, see Figure 3.2. Carcinoma involves uncontrollable division and multiplication of malignant cells where they spread to nearby organs such as the seminal vessels and other parts of the body (when they breach the capsule).

Prostate cancer in their benign and malignant state can cause damages to the glandular tissues in and around the prostate area which will result in leakage of PSA in the blood stream. It has always been a challenge in the early detection of cancer cells originating from the prostate gland (Brennhovd and Iversen, 2016), but currently, the presence of high value PSA in the blood is one of the most common indications of the possibility of the cancer. In most cases, the cancer develops for several years without any noticeable symptom, some develop to a volume that obstruct urination or they reach a volume that exposes them to palpation via the rectum. In the worst cases, the prostate cancer aggressively grows and breach the capsule and infiltrate nearby organs or the cancer cells will separate and enter the blood stream where it is transported to other parts of the body to eventually grow and cause organ failure in other parts of the body. Risk factors include smoking, age and family history. A diet high in red meat also plays a role, studies suggest

(Brennhovd, 2016). Various studies have shown that naïve black men (African Americans not blacks in Africa) are more likely to get prostate cancer than others men of other ethnic origin (Chang, 2018).

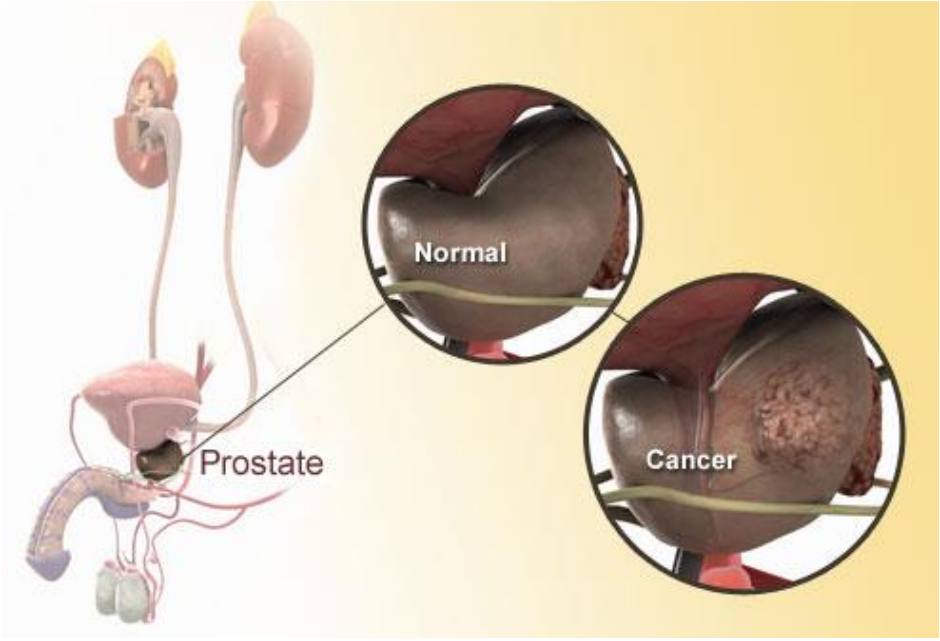


Figure 3.2: Visual difference between normal and cancer of the prostate (Brennhovd and Iversen, 2016)

3.2. Prostate cancer

Prostate cancer is the second and most common killer cancer in men after skin and lung cancer. A study in 2016 estimated that about 26,120 men with prostate can had lost their battle with the cancer (Nichols, 2019). Approximately, more than 80% of men above the age of 70 have high chance of getting the cancer. It is largely curable when diagnosed and treated early, however, patients with advanced prostate cancer have less favorable prognosis. Despite being one of the reasons for high mortality in among men annually, it is in fact, the second deadliest and common cancer in men, prostate cancer is still rocked with several challenges. Because prostate cancer in its early stage is asymptomatic, it always shows ambiguity in diagnosis. In some cases, prostate cancer in early stages may show basic symptoms that a very similar to urinary tract infections or other prostate related conditions that may be associated to cancer.

Prostate cancer early detection has always been a challenge and currently, the early symptoms of the cancer are quite difficult to detect (Fu et al., 2018). As mentioned earlier, early prostate cancer developments are usually asymptomatic which is a major reason why most men live with prostate cancer for several years without noticeable indications. Moreover, most of the symptoms associated with prostate cancer are usually common symptoms associated with urinary tract infections and other diseases in and around the bladder and urethra area (Chang, 2018). Part of this reason is that prostate cancer is slow growing and mostly symptom free.

The presence of high value PSA in the blood is one of the most common indications of the possibility of prostate cancer (Fu et al., 2018). PSA is a secreted substance that is manufactured within the prostate where in the event of abnormality with the prostate, there may be more than the normal level in the blood stream. However, the presence of high PSA level doesn't only or always indicate prostate cancer, it could also be an indication of infection or benign prostatic hyperplasia which is referred to as a noncancerous prostate expansion in size or other urinary or bladder diseases.

3.3. What is PSA (Prostate Specific Antigen)

PSA are biomarkers which exist as proteins or enzymes that are produced almost exclusively by the prostate gland by the activities of both health prostate cells or prostate cancer cells (PPC). This means that, PSA can be found or produced in other part of the body in small concentration, however, the prostate gland is the largest producer and is where majority of the concentration is naturally under normal conditions. As the name implies, prostate specific antigen only binds exclusively to specific antibodies in complementation (enzyme substrate antigen antibody) reaction.

One may ask, how is PSA helpful when prostate cancer cells needs to be detected? Well biomarker have to be linked directly or indirectly with a given diseases before they are considered biomarkers for that diseases. PSA is linked to prostate cancer in the sense that research have found that men with healthy prostate have low levels of PSA while there is reported high levels of PSA in the blood of men with prostate cancer or another prostate gland disorder (Fu et al., 2018). There is a threshold or baseline used to understand the

PSA level among men. The higher the PSA level is from the threshold the higher the risk of prostate cancer in the metastatic level and vice versa (Fu et al., 2018).

In biochemistry, PSA is referred to as kallikrein III, seminin, semenogelase, γ-seminoprotein and P-30 antigen. The antigen exists as a 34-kD glycoprotein and a serine protease (EC 3.4.21.77) enzyme, the gene of which is located on the 19th chromosome (19q13) in humans as shown in Figure 3.3. Histologically, prostate specific antigens are produced in the epithelial cells of the prostate gland which is one of the reasons why it is possible to collect sample cells/tissues from the prostate gland for biopsy and other immunochemistry analysis (Wong, 2016).

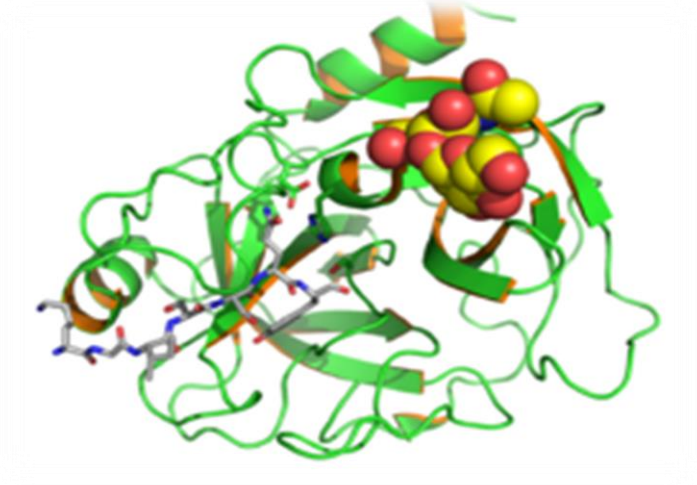


Figure 3.3: 34-kD glycoprotein (PSA) (Wong, 2016)

Consequently, when cells or tissue samples are collected, there is disruption of the epithelium tissue. In the absence of prostate cancer (maybe inflammation or benign prostatic hyperplasia), may allow escape and diffusion of some PSA into the tissue around the epithelium and subsequently entering in to the blood- increasing the level of PSA in the blood and making the condition more complex for effective diagnosis.

Prostate cancer cells (PCC) in the usual, are not capable of staining PSA effectively. This is as a result of the disturbance in the usual functions of the PSA. Compared to each healthy prostate gland cells (PGC), each PCC have lower production of PSA. However, due to the rapid, high and uncontrollable growth, production and multiplication of prostate cancer cells, their collective production of PSA is the reason for high PSA level in the

blood of men with prostate cancer. In most cases, PCC remain positive for the antigen, which is used to diagnosis prostate cancer in metastasis stage. Some high-grade, advanced or aggressive PCC may be entirely negative for PSA. For this reason, other histological is required for effective diagnosis of prostate cancer such the accurate and precise linking of one or more antibodies with the aim of improving effective detection such as linking of PSAP to CD57 (Chang, 2018).

3.4. Symptoms and Risk Factors of Prostate cancer

3.4.1. Earliest symptoms

These symptoms include; unusual weakness or interruption in the flow of urine, constant fullness of bladder and the urge to urinate (usually at night), hematuria, and contaminated seminal fluid with blood stains. The symptoms also include erectile dysfunction that occurs all of a sudden, experience of pain and sensation of burning or needle like feeling in the process of emptying the bladder from urine, pain or lack of comfort associate to enlargement of the prostate or bladder and other urinary areas when sitting. Furthermore, there these symptoms could also be linked to prostate disorders including inability to initiate urination or vice versa, partial or complete loss in bladder control, as well as increased viscosity of the urine which affect the flow of the urine (Live Science, 2010).

3.4.2. Risk factors

Risk factors associated to the development of prostate cancer are those conditions that directly or indirectly increases or influences the possibility of prostate cancer occurrence. There are several risk factors associated to prostate cancer such as age, ethnicity, lifestyle, and family history etc. an individual with two or more of these risk factors should be on regular surveillance (Live Science, 2010). However, there are several reported cases of individuals with several known risk factors who do not develop prostate cancer and other people who show no known risk factors that still develop prostate cancer. This is to show that the risk factors only influence the possibility of the cancer occurrence. Age and ethnicity are the strongest risk factors. A summary of risk factors of prostate cancer is provided in Table 3.1.

Table 3.1: Summary of risk factors

Risk Factors	
Critical risk factors	**General risk factors**
○ Ethnicity ▪ African American (70% VH) age 40 ▪ African (L) ▪ Asian ▪ Caucasian (M) age 50 ▪ Hispanic (L) ○ Age ▪ 10-20 (VL) ▪ 30-40 (L) ▪ 50-60 (M-H) ▪ 70-80 (VH) ▪ 90-100 (M-H) ○ Family History ▪ Familial prostate cancer (20%) ▪ Hereditary prostate cancer (5%) ✓ 3 or more1st^0 relatives ✓ 3 generations on the same side of the family ✓ 2 or more close relatives, before age 55 (2-3 x higher) ○ Region ▪ North America, northwestern Europe, Australia/New Zealand, and on Caribbean islands (VH) ▪ Asia, Africa, Central America, and South America (L)	○ Lifestyle ▪ Exercise ▪ Diet (H) ✓ saturated fat, red meat, high-fat dairy, vitamin D deficiency ▪ Smoking (H) ▪ Stress ✓ Height Obesity (M)

3.5. Stages of Prostate Cancer

Tagging of prostate is usually from a combination of various approaches to determine the pathological staging of the cancer. One of such approaches is the Gleason scoring system. the information from the Gleason scoring system can provide useful data regarding the rate at which prostate cancer have spread pass the capsule in stages (Fu et al., 2018). The Gleason scoring system can also provide useful information about the prognosis of the cancer over time. The system's score is evaluated using a grade range value from 2 to 10 which are the dominant histological grades used for the scoring (Zelefsky et al., 2011). The Gleason scores ranging from 2 grade to 10^{th} grade represent the extent of which the cancer cells have undergone differentiation. The lower, the grade score, the higher the extent of cancer cell differentiation and vice versa (Fu et al., 2018).

Furthermore, the information on whether the cancer cells have differentiated within or outside the prostate capsule is very significant in determining the progression of the disease and the necessary decision for therapy and prognosis. This is achieved using the T2 in the TNM scoring system. In general, the clinical staging of prostate cancer is dependent on useful information retrieved from various approaches such as the PSA levels in the patients, the Gleason score system, and the general rules of prostate cancer staging (Ren et al., 2013). These are approaches which are currently widely used for the diagnosis of prostate cancer among medical experts are utilized using exact values as shown in Table 3.2. However, they contain imprecise ranges and uncertainties that needs an effective approach to include all of the uncertainties to provide better evaluation (Fu et al., 2018).

Table 3.2: Evaluation table of prostate cancer risk grades (Ren et al., 2013)

Risk grade	**PSA (ng/ml)**	**Gleason score**	**T2**
Low-risk	<10	≤ 6	$\leq$T2a
Moderate-risk	10–20	7	T2b
High-risk	> 20	≥ 8	$\geq$T2c

It is important to note that;

T2a: The tumor is within the capsule or lobe and has spread 50% within it

T2b: The tumor is within the capsule or lobe and has spread more than 50% within one lobe

T2c: The tumor has spread within both lobes of the prostate

The following are summarized stages of prostate cancer in the TNT staging guide (James, 2018).

- Stage I: Early stage and slowly dividing and spreading cancer cells i.e. non aggressive
- Stage II: Early stage but cancer cells tend to divide and grow quickly within the prostate capsule
- Stage III: Cancer has spread outside the capsule
- Stage IV: Has spread to other parts of the body

3.6. Traditional PSA Test

After the checking the symptoms and various risk factors of a patient, medical physicians would collect blood samples of the patient for analysis using Immunoassay such as ELIZA (enzyme-linked immunosorbent assay). The analysis is to screen for prostate cancer possibilities by checking the concentration of PSA in the blood. As mentioned earlier, high levels of PSA may be an indication of prostate cancer and vice versa. Results from this test are expressed in nanograms of PSA per milliliter (ng/mL) in blood. Traditionally, the cut off or reference level PSA level is 4 ng/mL. A man with PSA levels higher than 4 ng/mL would require further analysis such as undergoing prostate biopsy and other histopathological analysis. It is important to note that traditional PSA test is usually observed alongside DRE (digital rectal exam) to feel for abnormality the size or volume of the prostate through the walls of the patient's rectum by inserting the finger (Kiefer, 2017).

3.7. Prostate Biopsy

Basically, multiple tissue samples of the prostate are collected from the prostate gland by putting needles that are hallow with ultrasound into the prostate gland to collect some cells or tissues. After that, the examiner withdraws the needles with a thin layer containing the prostate cells or tissues. When the pathway of the needle is through the rectum wall to reach the prostate gland, the process is referred to a transrectal biopsy as shown in Figure 3.4. It is the duty of the examiner who is a pathologist to examine the sampled prostate tissue or cells usually special microscope to check for abnormalities or cancer cells in the prostate gland. This is done after the PSA test in events of suspicious concentration of PSA level in the blood and the collective analysis of the patient's symptoms and risk factors.

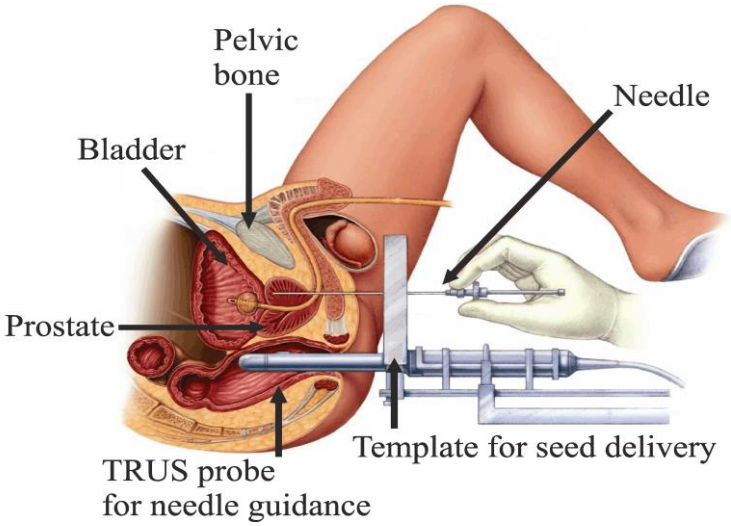


Figure 3.4: Transrectal biopsy (Swindle, 2020)

CHAPTER 4

Methodology

4.1. Introduction

The overall methodology used in this thesis took in to consideration the established aims and objectives presented in chapter one. A structural frame work for data analysis will be employed. This study utilized Fuzzy Inference System (FIS), MATLAB (Matrix laboratory) software function and Graphics users Interface (GUI) to make predictive diagnosis for prostate cancer patients. This will improve effective diagnosis of prostate cancer among men especially those with high degree of risks. In the health facility, the diagnosis of prostate cancer can be achieved first by undergoing some sort of screening either prostate-specific antigen (PSA) test or a digital rectal exam (DRE).

The PSA test basically test for the level of PSA of the patient whether it is within acceptable range while DRE is a physical examination to feel for lumps or abnormalities around the prostate. After that, a patient may provide urine sample for test, histopathology (biopsy), and imaging of the prostate (MRI, CT, PET, SPECT). The whole diagnostic processes of prostate cancer can be summarized in to four categories; screening (PSA and DRE), diagnosis (biopsy), the grade of cancer (Gleason score), and the stage of cancer (imaging and TNM staging). For each of the categories, there is a subtest that is performed independently taking into consideration some important parameters.

4.2. Diagnosis of Prostate Cancer

Generally, prostate cancer screening is first observed by considering all patients risk factors, after which DRE is performed. If the DRE examination does not provide useful information, the medical expert may perform the PSA test. PSA in the blood is measured in units called nanograms per milliliter (ng/mL). The chance of having prostate cancer goes up as the PSA level goes up, but there is no set cutoff point that can tell for sure if a man does or doesn’t have prostate cancer. If the results of a PSA blood test, DRE, or other tests suggest that you might have prostate cancer, your will most likely need a prostate biopsy. The medical exert takes few prostate tissue samples for histopathological analysis in the

laboratory. There are three possible results from the procedure of biopsy. If the result turns *positive*, then it means the analysis have detected cancer cells from the prostate tissue samples. However, if no cancer cells are detected, then it means the biopsy analysis is *negative*.

Moreover, if biopsy procedure neither indicated positive or negative, but there are abnormal activities in the prostate, then the result is considered *suspicious*. Patients with suspicious results are recommended to go through further with other tests that are not related to prostate cancer. Furthermore, the Gleason score indicates whether the cancer is aggressive or not and fast spreading or not. Lastly, imaging modalities can be taken and used alongside the TNM staging system to identified the stages of the cancer. After all these processes and clear and efficient medical information has been gathered, the medical expert may provide recommendation for treatment for the patient.

4.3. Structure of Fuzzy System used for Diagnosing Prostate Cancer

The structure of the fussy system is enumerated below;

1. The system uses a fuzzy sugeno model
2. It takes the inputs (such as age, biopsy, and PSA etc.) through a GUI interface
3. The GUI interface is designed using MATLAB module guide for GUI design
4. The user enters the data through the GUI
5. The result button handler (in Cancer Test.m)
6. This calls the processing function (in cancer.m)
7. This function runs the data through defuzzification process to get the result of the fuzzy model according to the rules defined in the model
8. The rules and the model were designed using fuzzy systems module in MATLAB

4.3.1 Diagnosing of Prostate Cancer by Screening, Biopsy, Gleason score, TNM staging (Input Variables)

In prostate cancer screening, DRE or PSA test is performed to check for abnormal appearance or levels of PSA in the patient. In the DRE, if the medical expert feels some types of inflammation, then the PSA test will be carried out. The PSA test have specific

range that shows lower levels of PSA (<), moderate levels of PSA (<&>) and higher levels of PSA (>). The result from PSA test are compared to the laboratory references range. This ranges are briefly sampled in fuzzy rule bases in the following subsections.

1. *Screening of Prostate cancer*

As mentioned earlier, the screening procedures usually involves a DRE or a PSA test. Results from this test may provide useful information that can show high or low likelihood of prostate cancer manifestation. However, these procedures do not necessary guarantee the presences or absence of prostate cancer especially when patients are asymptomatic. Therefore, PSA was made assigned as a membership function in the system with a trapmf type. Based on this PSA test the developed rule base is presented below;

IF PSA (VALUE) is lower than the range, patient has high possibility for negative prostate cancer
IF PSA (VALUE) is average, then result is equivocal for prostate cancer
IF PSA (VALUE) is higher than the range, then result is positive prostate cancer

Note: The input range of (PSA) is between 0 and 50. here we have very low, low, average, high, and very high are fuzzy linguistic values. Very low is in the interval [0, 4], low is [2, 8], the average is in the interval [4, 10], high is [8, 16], and very high is in the interval [12, 50] respectively. The PSA level should be as low as possible for negative diagnosis.

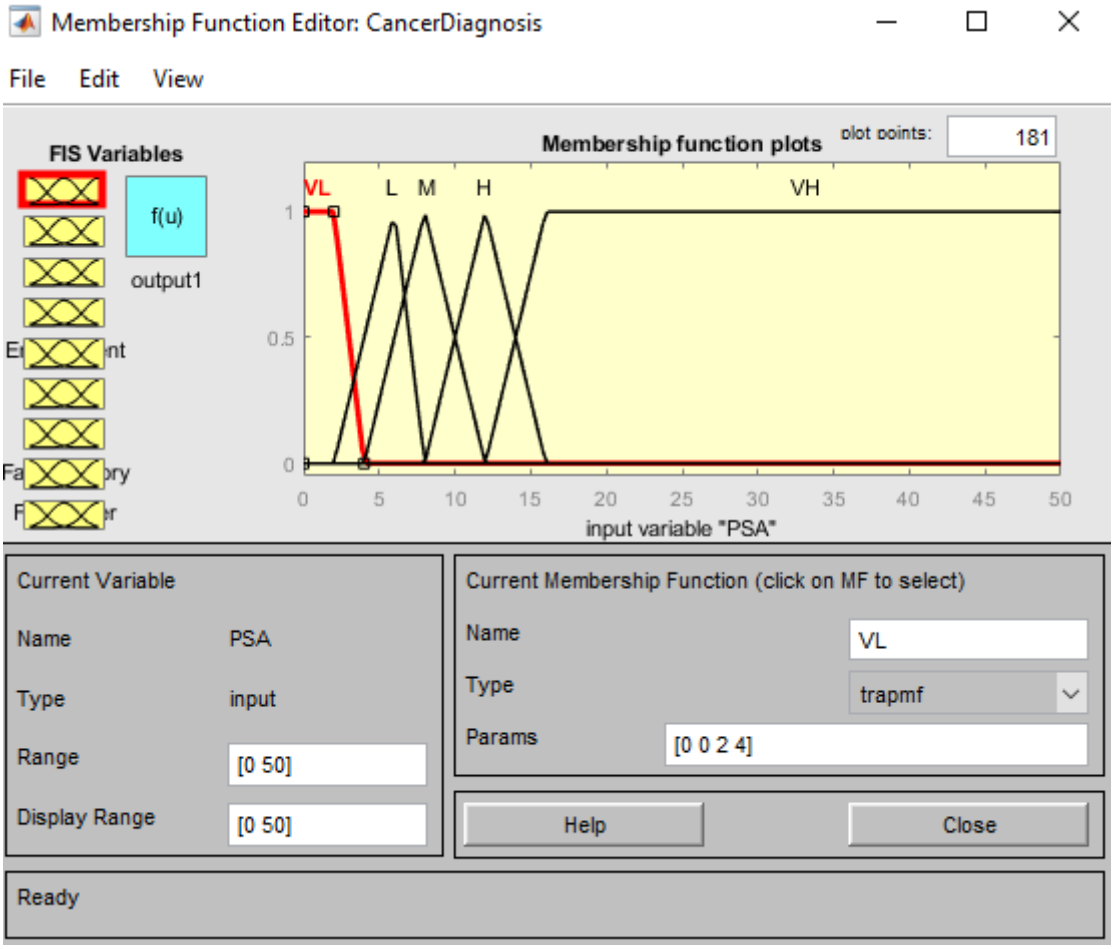


Figure 4.1: Membership plots for PSA test

2. *Diagnosing Prostate cancer using biopsy*

The medical expert fetch about 12 samples of tiny cylindrical cores of prostate tissues (containing prostate cells) from different location of the prostate gland. These samples are sent to the pathological laboratory for histopathological analysis. A microscope is used to investigate the presence or absence of prostate cancer cells. Based on this laboratory test the developed rule base is presented below;

IF biopsy result (VALUE) does not show cancer cells, patient is negative for prostate cancer

IF biopsy result (VALUE) shows no cancer but abnormality in the prostate, result is suspicious

IF biopsy result (VALUE) shows cancer cells, then patient is positive for prostate cancer

Note: The input range of (biopsy result) is between 0 and 1. Here no cancer cells (low), no cancer cells but abnormal prostate (average), and presence of cancer cells (high) are fuzzy linguistic values low that are in intervals [0,0.5], the average that is in intervals [0.2,0.8] and high that is in intervals [0.5,1] respectively.

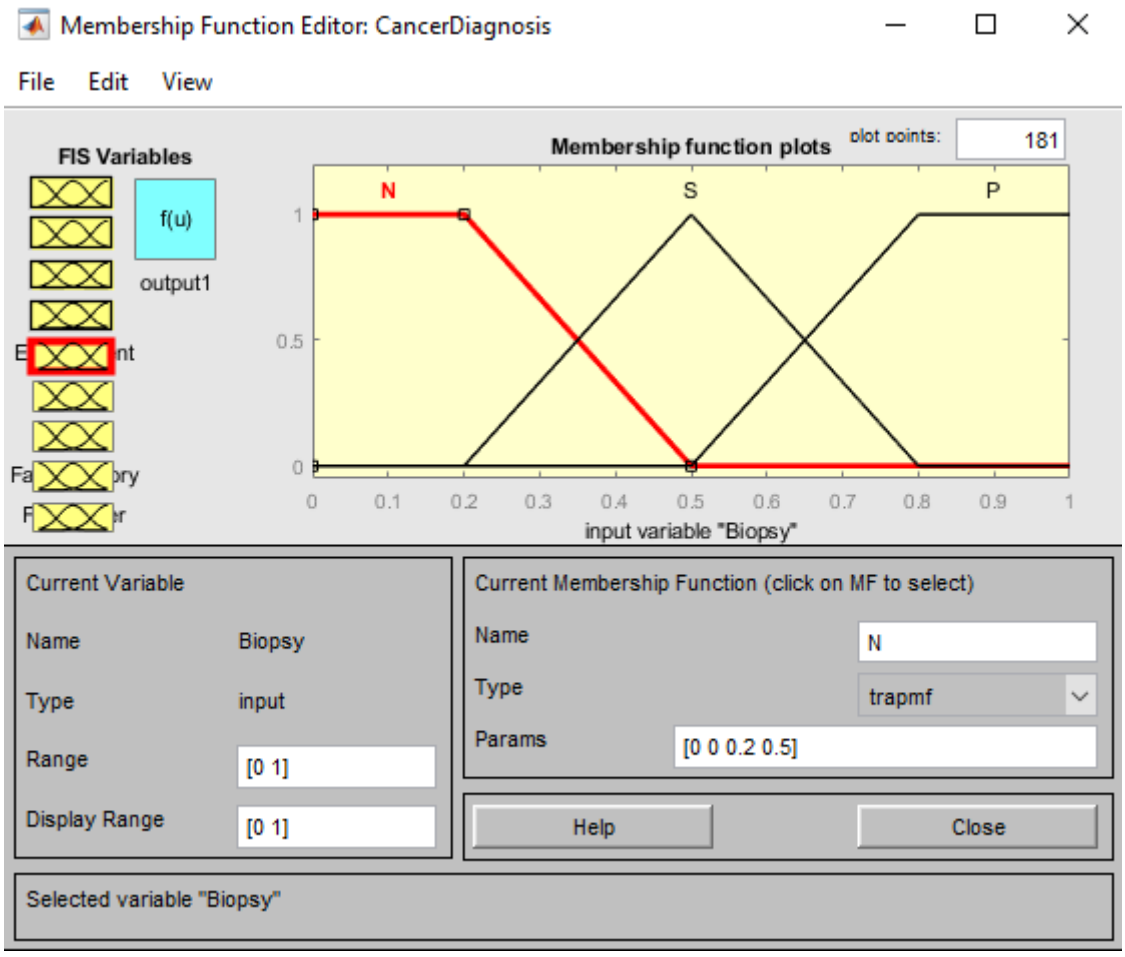


Figure 4.2: Membership plots for biopsy test

3. *Grading of prostate cancer by Gleason score*

If biopsy result from a patient shows presence of prostate cancer cells (positive), the cancer will be marked with a grade group using the Gleason score. Based on this Gleason score, the developed rule base is presented below;

IF prostate cancer cells (VALUE) resembles regular healthy cells, the Gleason score is 1

IF prostate cancer cells (VALUE) appears very irregular, then the Gleason score is 5
IF Gleason score (VALUE) range from 2 to 4, cancer have features between these extremes.

4. *Staging of prostate cancer by TNM staging*

the TNM scan is initiated to investigate the level of spread of prostate cancer within the prostate capsule, neighboring lymph nodes or tissues, or whether it has affected other organs and parts of the body. T stage indicates the spread of cancer only within the prostate capsule, N stage represent the extend of spread to lymph nodes, while M stage shows whether prostate cancer has metastasized to other parts of the patient's body. The T stage ranges from T1-T4. On the other hand, the N and M stage ranges from NX to N1 and MX to M1 respectively. Based on the result from prostate cancer staging, the developed rule base is presented in the following subsections for each stage and substage.

1. *T stage of prostate cancer*

Can be stages using the results from DRE, MRI or CT scan.

IF prostate cancer cells (VALUE) are within the prostate capsule, T stage

IF cancer (VALUE) can't be detected by DRE, only by biopsy, then T1 stage

IF cancer (VALUE) can be detected by DRE but within capsule, then stage is T2

IF cancer (VALUE) is detected by DRE but breaches capsule layer, then stage is T3

IF cancer (VALUE) has spread to nearby organs, stage is T4 (localized cancer).

2. *N stage of prostate cancer*

MRI and CT scans can be used to assign the N stage which is spread of cancer to lymph nodes.

IF cancer (VALUE) is not found in the lymph nodes, the stage is N0
IF cancer (VALUE) is not clear in the lymph nodes, the stage is NX
IF cancer (VALUE) is found in lymph nodes, the stage is N1 (advanced localized prostate cancer).

3. *M stage of prostate cancer*

Bone scans can be used to determine the M stage.

IF cancer (VALUE) is has not spread to other parts of the body, the stage is M0
IF cancer (VALUE) is not clear in other parts of the body, the stage is MX
IF cancer (VALUE) is found in other parts of the body, the stage is M1 (advanced prostate cancer).

4.3.2 Predicting the likelihood of prostate cancer using risk factors

There are several risk factors that increases the chance of getting prostate cancer. These risks include age, race or ethnicity, family history (genetic), lifestyle, and the environment etc. To predict the likelihood of prostate cancer, medical expert uses the risk factors in combination with the results from the tests discussed in previous sections. The word low refers to the (<) symbol and the word big refers to the (>) symbol and average refers to the (<&>). which are given and below, the fuzzy rule bases used for the diagnosis of prostate cancer are presented in subsections below.

1. *Age as a risk factor*

Age as a risk factor is a very important predictor for prostate cancer. Normally, the genetic information (family) history of the patient is one of the most important prediction of prostate cancer, however, study have shown that the chances of prostate increases after 50 years and this is for a patient without any family history. Most prostate cancer diagnosed were found to be individuals within the age of 65 and above. Therefore, age can not be under considered in the prediction of prostate cancer among men. Based on this Gleason score, the developed rule base is presented below;

IF age range (VALUE) is below 40 years, low risk prostate cancer
IF age range (VALUE) is between 50 years to 60 years, medium risk prostate cancer
IF age range (VALUE) is above 65 years and above, the high risk prostate cancer

Note: The input range of (age range) is between 0 and 100. here low, moderate, and high are fuzzy linguistic values. Very young is in the interval [0, 25], young is [15, 30], the medium age is in the interval [25, 40], old is [35, 50], and very old is in the interval [45, 100] respectively.

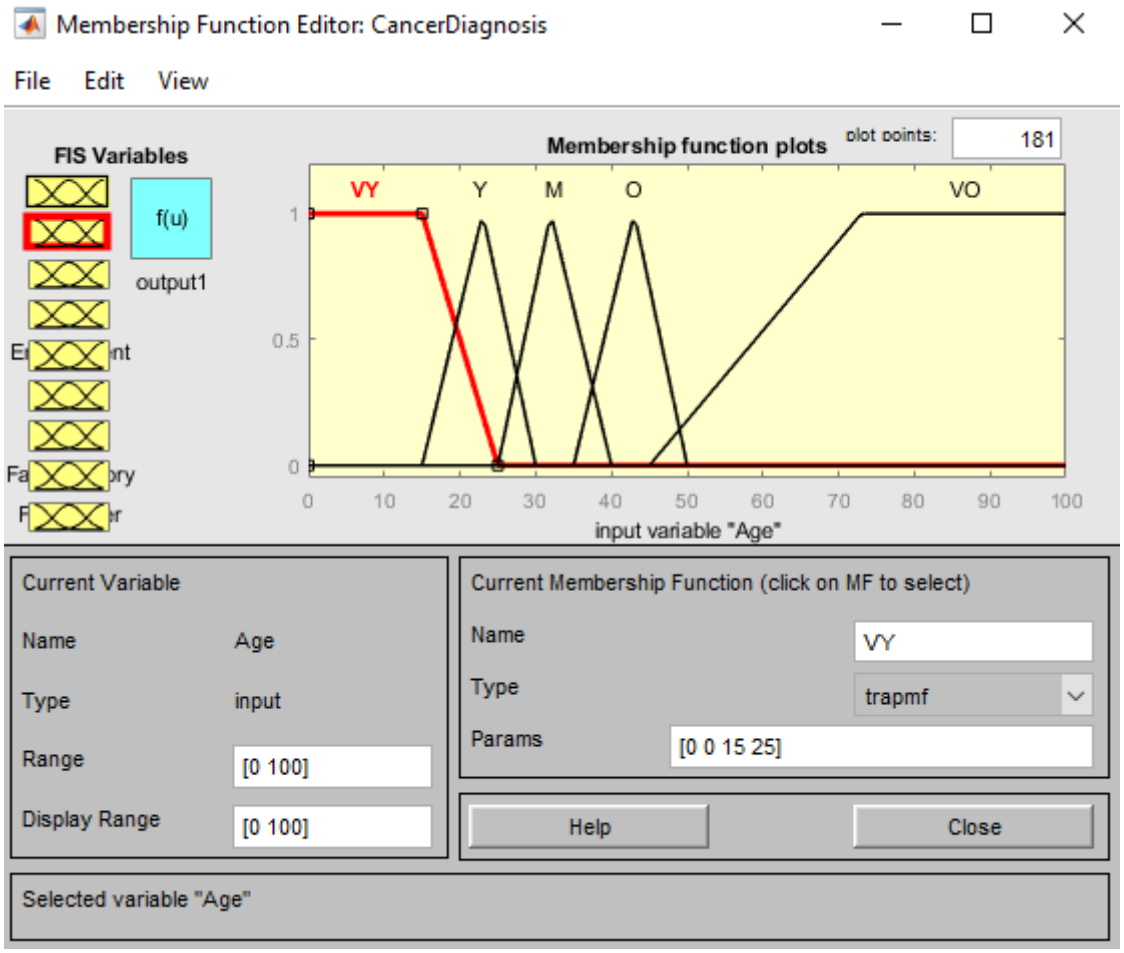


Figure 4.3: Membership plots for Age risk

2. *Ethnicity as a risk factor*

Prostate cancer develops more often in African-American men and in Caribbean men of African ancestry than in men of other races. And when it does develop in these men, they tend to be younger. Prostate cancer occurs less often in Asian-American and Hispanic/Latino men than in non-Hispanic whites. The reasons for these racial and ethnic differences are not clear. Based on this Gleason score, the developed rule base is presented below;

IF ethnicity (VALUE) is African American, very high (VH) risk prostate cancer
IF ethnicity (VALUE) is Caucasian, moderate (M) risk prostate cancer
IF ethnicity (VALUE) is African, low (L) risk prostate cancer
IF ethnicity (VALUE) is Hispanic, low (L) risk prostate cancer
IF ethnicity (VALUE) is Asian, very low (VL) risk prostate cancer.

Note: The input range of (ethnicity) is between 0 and 7. here from very low to very high are fuzzy linguistic values. Very low is in the interval [0, 2], low is in the interval [1, 4], the medium is in the interval [2, 5], high is [3, 6], and high is in the interval [4, 7] respectively.

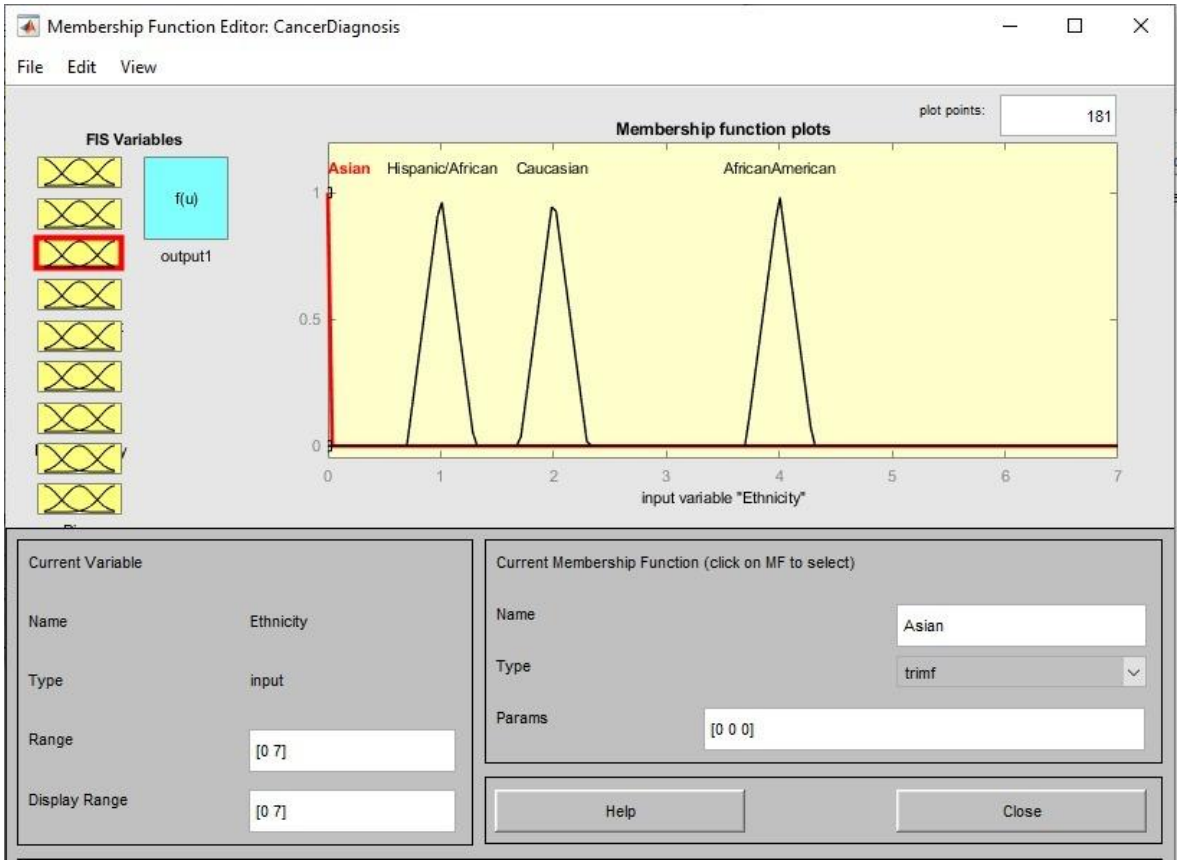


Figure 4.4: Membership plots for ethnicity risk

3. *Family History as a risk factor*

Prostate cancer seems to run in some families, which suggests that in some cases there may be an inherited or genetic factor. Still, most prostate cancers occur in men without a family history of it. Having a father or brother with prostate cancer more than doubles a man's

risk of developing this disease. (The risk is higher for men who have a brother with the disease than for those who have a father with it.) The risk is much higher for men with several affected relatives, particularly if their relatives were young when the cancer was found. Based on this Gleason score, the developed rule base is presented below;

IF patient (VALUE) has familial prostate cancer, risk is 20%
IF patient (VALUE) has hereditary prostate cancer, risk is 5%

Note: The input range of (family history) is between 0 and 7. here from very low to very high are fuzzy linguistic values. Familial is in the interval [1, 7] and hereditary is in the interval [0, 7] respectively.

4. *Environment (Region) as a risk factor*

Prostate cancer is most common in North America, northwestern Europe, Australia, and on Caribbean islands. It is less common in Asia, Africa, Central America, and South America. The reasons for this are not clear. More intensive screening for prostate cancer in some developed countries probably accounts for at least part of this difference, but other factors such as lifestyle differences (diet, etc.) are likely to be important as well. For example, Asian Americans have a lower risk of prostate cancer than white Americans, but their risk is higher than that of men of similar ethnic backgrounds living in Asia. Based on this Gleason score, the developed rule base is presented below;

IF patient (VALUE) is north America, northwest Europe, Australia/New Zealand, and Caribbean, risk is very high
IF patient (VALUE) is Asia, Africa, Central America, and South America risk is low

Note: The input range of (environment risk factor) is between 0 and 1. Here from low to high are fuzzy linguistic values. Low is in the interval [0,0] and high is in the interval [1, 1] respectively.

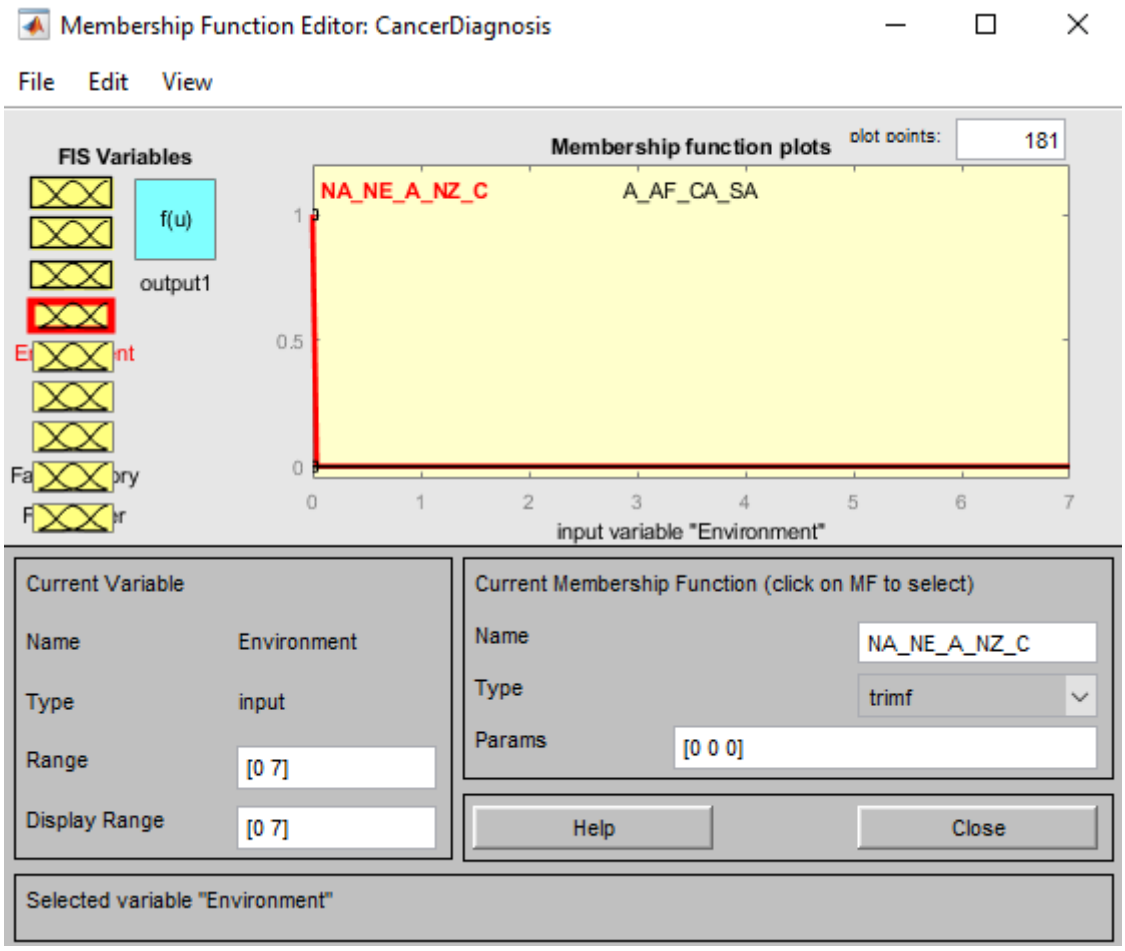


Figure 4.5: Membership plots for environment risk

5. ***Lifestyle as a risk factor***

Although lifestyle as a risk factor is does not contribute significantly to development of prostate cancer like age and gene, studies have shown some important direct and indirect connection with lifestyle choices and emergence of prostate cancer. Based on this Gleason score, the developed rule base is presented below;

IF patient (VALUE) is obese, prostate cancer risk is high

IF patient (VALUE) is a fire fighter, prostate cancer risk is high

IF patient (VALUE) eats high dairy products, prostate cancer risk is high

Note: The input range of (family history) is between 0 and 1 for obesity, firefighter, and high dairy products. here from low to very high are fuzzy linguistic values. Low is in the interval [0], and represent the absence of such variable. High is in the interval [1] respectively.

4.4. Diagnosing of Prostate (output Variables)

The output for this study should be three; diagnosis, Gleason score, and stage of the cancer. However, Gleason score, and stage depends on positive diagnosis, otherwise, there is no need for them. Note Diagnosis have positive (P), suspicious (S), and negative (N). What we want to do is to create rules that will combine these all the values for input variables associated for each patient to determine if a patient has prostate cancer or not or they are at risk of some prostate problems (suspicious). For example, a rule can be like this;

If (Age is 30 (young=L)) and (PSA is 3 (VL)) and (ethnicity is Asian (VL)) *and* (Environment is Asia (L)) then (Biopsy is suspicious (Average=M)) (1)

Then this patient may not have prostate cancer considering all of the inputs, however, biopsy result shows that there are no prostate cancer cells but result is suspicious, that means the patient needs to go for other tests such test to identify urinary tract infection. Therefore, this patient's diagnosis for prostate cancer is negative and there is no need for the Gleason score and staging. From this we can make several rules that will be useful for the FIS as shown in table 4.1.

Table 4.1: Rules of the FIS system for prostate cancer diagnosis

Inputs	**Output**				
Rules	Diagnosis			Gleason score	Stage
	P	S	N		
If (Age is L) and (PSA is VL) and (ethnicity is VL) and (Environment is L), and (Biopsy is A) (1)	x	✓	x	x	x
If (Age is L) and (PSA is VH) and (ethnicity is VL) and (Environment is L), and (Biopsy is A) (1)	x	✓	x	x	x
If (Age is A) and (PSA is L) and (ethnicity is VL) and (Environment is L) and (Biopsy is L) (1)	x	x	✓	x	x
If (Age is A) and (PSA is VH) and (ethnicity is VL) and (Environment is L) and (Biopsy is A) (1)	x	✓	x	x	x

If (Age is A) and (PSA is VH) and (ethnicity is VL) and (Environment is L) and (Biopsy is H) (1)	✓	x	x	✓	✓
If (Age is A) and (PSA is H) and (ethnicity is VH) and (Environment is VH) and (Biopsy is H) (1)	✓	x	x	✓	✓
If (Age is A) and (PSA is VH) and (ethnicity is L) and (Environment is L) and (Biopsy is A) (1)	x	✓	x	x	x
If (Age is A) and (PSA is VH) and (ethnicity is L) and (Environment is L) and (Biopsy is L) (1)	x	✓	x	x	x
If (Age is A) and (PSA is VH) and (ethnicity is L) and (Environment is L) and (Biopsy is H) (1)	✓	x	x	✓	✓
If (Age is A) and (PSA is VH) and (ethnicity is L) and (Environment is L), (obese=H), and (Biopsy is A) (1)	x	✓	x	x	x
If (Age is L) and (PSA is L) and (ethnicity is VH) and (Environment is VH), (obese=H) and (Biopsy is A) (1)	x	✓	x	x	x
If (Age is L) and (PSA is VH) and (ethnicity is VH and (Environment is VH), (familial=H), (obese=H) and (Biopsy is A) (1)	x	✓	x	x	x
If (Age is L) and (PSA is A) and (ethnicity is L) and (Environment is L) and (Biopsy is L) (1)	x	x	✓	x	x
If (Age is L) and (PSA is VH) and (ethnicity is L) and (Environment is L) and (Biopsy is A) (1	x	✓	x	x	x
If (Age is L) and (PSA is VL) and (ethnicity is L) and (Environment is L) and (Biopsy is H) (1	✓	x	x	✓	✓
If (Age is VL) and (PSA is L) and (ethnicity is L) and (Environment is L) and (Biopsy is A) (1)	x	✓	x	x	x
If (Age is VL) and (PSA is L) and (ethnicity is L) and (Environment is L) and (Biopsy is L) (1)	x	x	✓	x	x
If (Age is VL) and (PSA is L) and (ethnicity is L) and (Environment is L) and (Biopsy is H) (1)	✓	x	x	✓	✓
If (Age is A) and (PSA is A) and (ethnicity is VL) and (Environment is L), (L), and (Biopsy is L) (1)	x	x	✓	x	x

If (Age is A) and (PSA is VH) and (ethnicity is M) and (Environment is VH) and (Biopsy is A) (1)	✓	x	x	✓	✓
If (Age is A) and (PSA is L) and (ethnicity is VH) and (Environment is VH) and (Biopsy is A) (1)	x	✓	x	x	x
If (Age is A) and (PSA is A) and (ethnicity is M) and (Environment is VH) and (Biopsy is A) (1)	x	✓	x	x	x
If (Age is A) and (PSA is A) and (ethnicity is L) and (Environment is VH) and (Biopsy is A) (1)	x	✓	x	x	x
If (Age is A) and (PSA is A) and (ethnicity is VL) and (Environment is VH) and (Biopsy is H) (1)	✓	x	x	✓	✓
If (Age is A) and (PSA is A) and (ethnicity is L) and (Environment is VH) and (Biopsy is L) (1)	x	x	✓	x	x
If (Age is A) and (PSA is A) and (ethnicity is VL) and (Environment is VH), (obese=H), fire fighter=H), (high dairy products=H), and (Biopsy is L) (1)	x	✓	x	x	x
If (Age is A) and (PSA is A) and (ethnicity is L) and (Environment is VH), (obese=H), fire fighter=H), (high dairy products=H), and (Biopsy is H) (1)	✓	x	x	✓	✓
If (Age is L) and (PSA is A) and (ethnicity is M) and (Environment is VH) and (Biopsy is L) (1)	x	x	✓	x	x
If (Age is L) and (PSA is A) and (ethnicity is M) and (Environment is VH) and (Biopsy is A) (1)	x	✓	x	x	x
If (Age is L) and (PSA is A) and (ethnicity is M) and (Environment is VH), (obese=H), fire fighter=H), (high dairy products=H) and (Biopsy is A) (1)	✓	x	x	✓	✓
If (Age is A) and (PSA is L) and (ethnicity is M) and (Environment is VH), (obese=H), fire fighter=H), (high dairy products=H) and (Biopsy is H) (1)	✓	x	x	✓	✓
If (Age is A) and (PSA is VH) and (ethnicity is M) and (Environment is VH), (H), and (Biopsy is L) (1)	x	✓	x	x	x
If (Age is A) and (PSA is VH) and (ethnicity is M) and (Environment is VH), (obese=H), fire fighter=H), (high	✓	x	x	✓	✓

dairy products=H) and (Biopsy is A) (1)					
If (Age is A) and (PSA is L) and (ethnicity is M) and (Environment is VH), (obese=H), fire fighter=H), (high dairy products=H) and (Biopsy is A) (1)	x	✓	x	x	x
If (Age is A) and (PSA is L) and (ethnicity is M) and (Environment is VH), and (Biopsy is L) (1)	x	x	✓	x	x

Figure 4.6 shows a combination of all input variables fed in to the FIS system to produce an output. The input are crisp data that represent the test values and risk factors of the patients which are processed by the FIS system to produce a likely diagnosis for the patient.

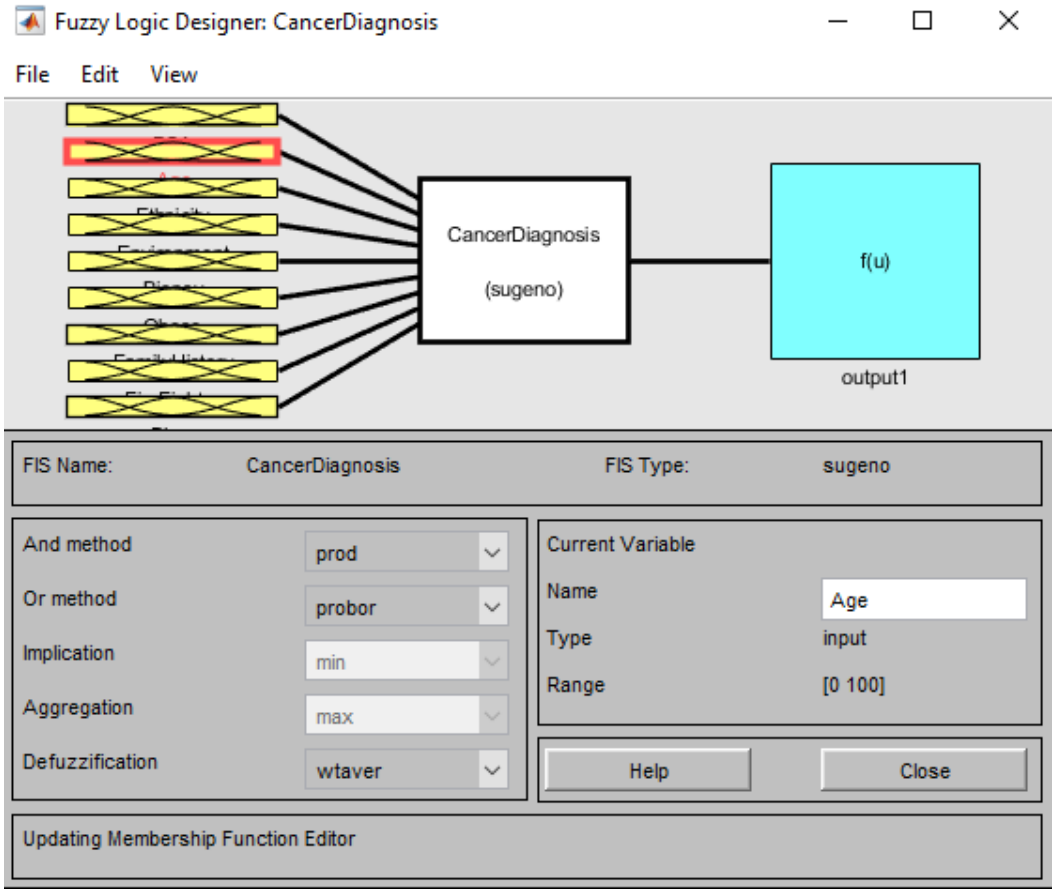


Figure 4.6: Input and output variables used in the FIS system

CHAPTER 5

RESULT AND DISCUSSION

5.1. Overview

This study utilized Fuzzy Inference System (FIS), MATLAB (Matrix laboratory) software function and Graphics users Interface (GUI) to make predictive diagnosis for prostate cancer patients for the purpose of improving effective diagnosis of prostate cancer among men especially those with high degree of risks. The predictive results from the system were based on the rules for the FIS system shown in table 4.1 in chapter four. The result (outputs) from the simulation of rules (inputs) which are a combination of all important variables important for the diagnosis of prostate cancer. These variables included PSA test, DRE, histopathology (biopsy), and the risk factors (age, ethnicity, environment, family history, and lifestyle). This is shown in the main page of the system as screen shot in figure 5.1. From the left side of the page, patients' data can be inputted accordingly to provide a meaningful information or diagnostic result (output) in the right-side dialog box showing result.

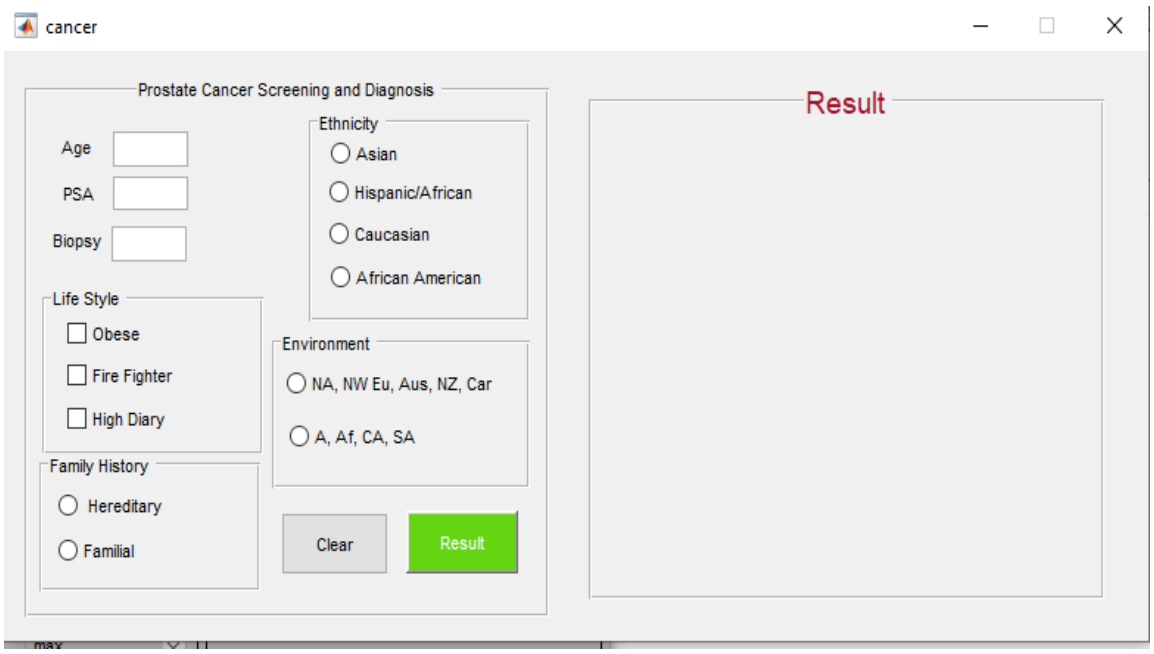


Figure 5.1: Main page of the FIS system designed for the purpose of diagnosing prostate cancer

The choice of using fuzzy Inference system (FIS) is for the purpose of determining the range as well as the parametric values of all variables used in the study. Additionally, the reason why graphic user interface (GUI) as adopted and implemented in this study is to provide an easier method of inputting prostate cancer patients data. How the whole system works is, patients' data are inputted in to the FIS model shown in figure 5.1 using the inputs options in the left side. After that, the user must click on result button in order to display the result as an output in the right dialog box. The FIS system operates a comparative analysis of the inputted prostate cancer patient's data in accordance to the established rules and range of each variable provided by medical experts as shown below;

If (Age is (-)) and (PSA (-)) and (ethnicity (-)) and (Environment (-)), and (Lifestyle (-)) then (Biopsy (-)) (1)
The FIS automatically compares the patient data to the rules and range of each variable.

5.2. Result from Prostate Cancer Diagnosis with FIS

The output for this study was categorized in to three parts; diagnosis of prostate cancer, Gleason score, and stage of the cancer. However, Gleason score, and stage depends on positive diagnosis, otherwise, there is no need for them. Note Diagnosis have positive (P), suspicious (S), and negative (N). If the output from the simulation of a patient's data comes out positive, it means the FIS system have detected the presence of cancer after comparing all the variables in the patient's data. Therefore, the FIS recommends the user to investigate the Gleason score and the stage of the prostate cancer using imaging technology such as MRI, CT, PET, and SPECT.

Moreover, if patient result comes out as suspicious, then it means the FIS system neither detected the presence or absence of prostate cancer, therefore, the result is ambiguous. Furthermore, a suspicious result mostly indicates that there is no prostate cancer, however, the patient may have other non-cancerous problems affecting his prostate such as urinary tract infection and kidney problems etc. Either way, a suspicious result after simulation, recommends the expert to perform to either repeat the various prostate cancer test or perform other tests such as the urinary test to find out the problem with the prostate. Lastly,

a negative result from a prostate cancer patient's simulation data (all variables) signify that the system didn't identify the presence of prostate cancer in the patient.

For this study, data for prostate cancer data was collected for 6 patients. The patients' data are provided in table 5.1. In order to follow the ethical rules, the identity of the patients was kept anonymous, therefore, the patients are indicated as patient 1 to patient 6.

Table 5.1: Medical data (in numerical values) from 6 prostate cancer patients used in the FIS simulation

	Age	**PSA ng/mL**	**Biopsy value**	**Family history**	**Ethnicity**	**Lifestyle**	**Environment**
Patient 1	34	23	1	None	Asian	none	Australia
Patient 2	27	44	0.5	None	Caucasian	O, F, H	N America
Patient 3	25	2	0.5	None	Asian	none	Africa
Patient 4	25	25	0.5	None	Asian	none	Asia
Patient 5	32	6	0	None	Asian	none	Asia
Patient 6	35	19	1	None	Asian	none	S America

The result from each simulation of patient's prostate cancer medical data are presented and discussed in the next sections.

5.3. FIS Simulation Result for Patient 1

Patient 1 is 34 years old which is within the age interval [25, 40] which is classified as medium age. His prostate specific antigen (PSA) level is 23 which is within the interval [12, 50] that is classified as very high level of PSA in ng/mL. Patient 1's biopsy is 1 (positive), which means prostate cancer cells were detected in him. record of his family shows that no one from his family have either hereditary or familial prostate cancer risk. Furthermore, Patient 1's ethnicity is Asian who are categorized under low risk of prostate cancer. The environment in which patient 1 resides is Australia, a high-risk environment for prostate cancer cases. Lastly, data of the patient's lifestyle shows no record of obesity, high diary intake, and his occupation was not firefighting. After all the data were inputted in to the FIS system, the result (output) shows that the prostate cancer is positive, and

requires further examination of Gleason score and the staging of the prostate cancer. The result is presented in figure 5.2.

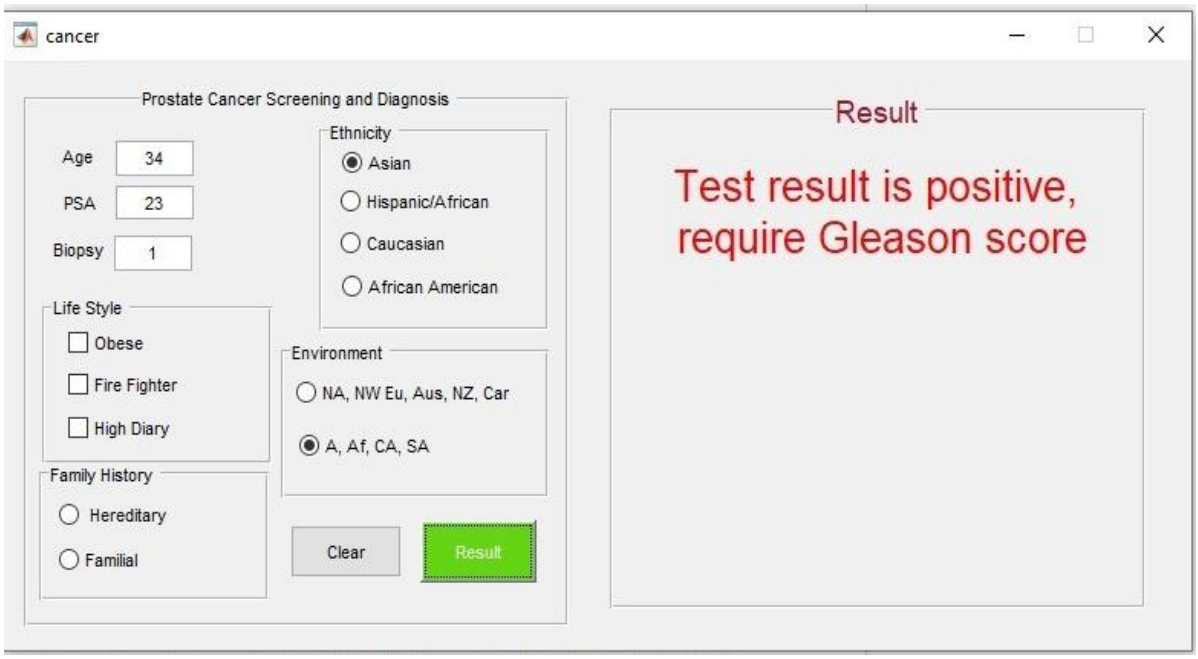


Figure 5.2: A Screen dialog showing Patient 1 simulation result (prostate cancer diagnosis)

5.4. FIS Simulation Result for Patient 2

Patient 2 is 27 years old which is within the age interval [15, 30] which is classified as young. His prostate specific antigen (PSA) level is 44 which is within the interval [12, 50] that is classified as very high level of PSA in ng/mL. Patient 2's biopsy is 0.5 (average), which means prostate cancer cells were not detected in him, therefore, the patient may need another biopsy test, or further analysis of urinary tract infection and kidney problems. Record of his family shows that no one from his family have either hereditary or familial prostate cancer risk.

Furthermore, Patient 2's ethnicity is Caucasian who are categorized under medium risk of prostate cancer. The environment in which patient 2 resides is North America, a high-risk environment for prostate cancer cases. Lastly, data of the patient's lifestyle shows record of obesity, high diary intake, and his occupation was firefighting. After all the data were inputted in to the FIS system, the result (output) shows that the prostate cancer is positive,

and requires further examination of Gleason score and the staging of the prostate cancer. The result is presented in figure 5.3.

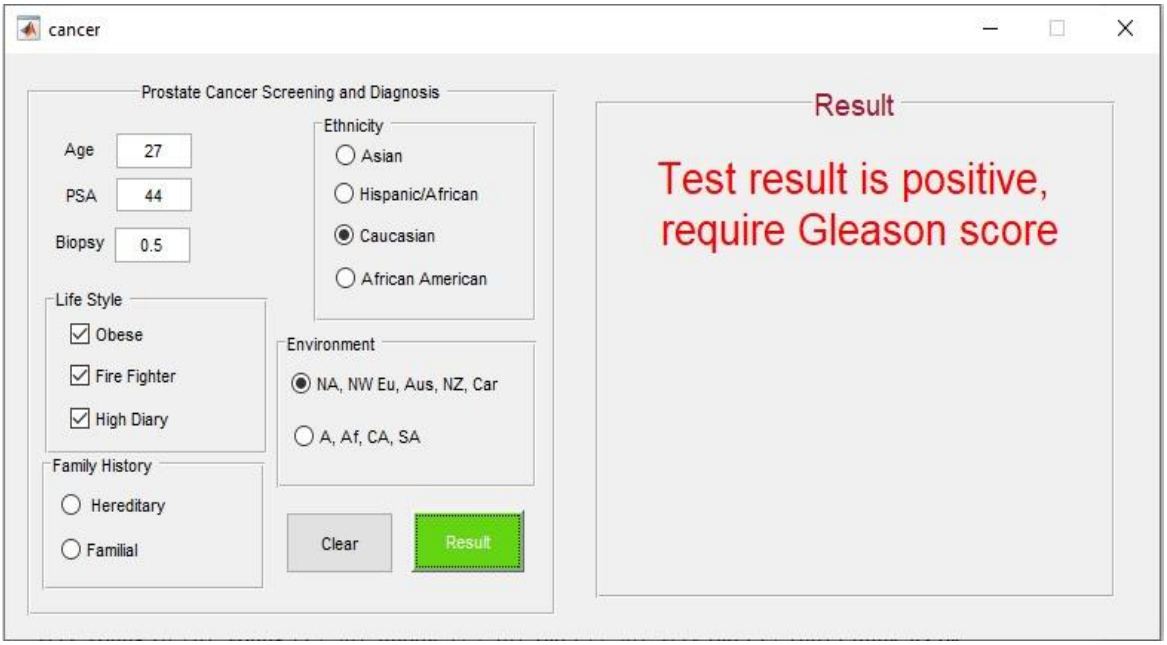


Figure 5.3: A Screen dialog showing Patient 2 simulation result (prostate cancer diagnosis)

5.5. FIS Simulation Result for Patient 3

Patient 3 is 25 years old which is within the age interval [15, 30] which is classified as young. His prostate specific antigen (PSA) level is 2 which is within the interval [2, 8] that is classified as low level of PSA in ng/mL. Patient 3's biopsy is 0.5 (positive), which means prostate cancer cells were not detected in him, therefore, the patient may need another biopsy test, or further analysis of urinary tract infection and kidney problems. Record of his family shows that no one from his family have either hereditary or familial prostate cancer risk. Furthermore, Patient 3's ethnicity is Asian who are categorized under low risk of prostate cancer. The environment in which patient 3 resides is Africa, a low risk environment for prostate cancer cases.

Lastly, data of the patient's lifestyle shows no record of obesity, high diary intake, and his occupation was not firefighting. After all the data were inputted in to the FIS system, the result (output) shows that the patient 3's result is suspicious. What this means is that, no prostate cancer risk was detected, however, there was problem detected in the prostate

gland that is not related with prostate cancer, therefore patient 3 is recommend to go for urinary tract infection and kidney pathological test to check the problem. Additional, since patient 3's result shows suspicious, then there is no need for further examination of Gleason score and the staging of the prostate cancer. The result is presented in figure 5.4.

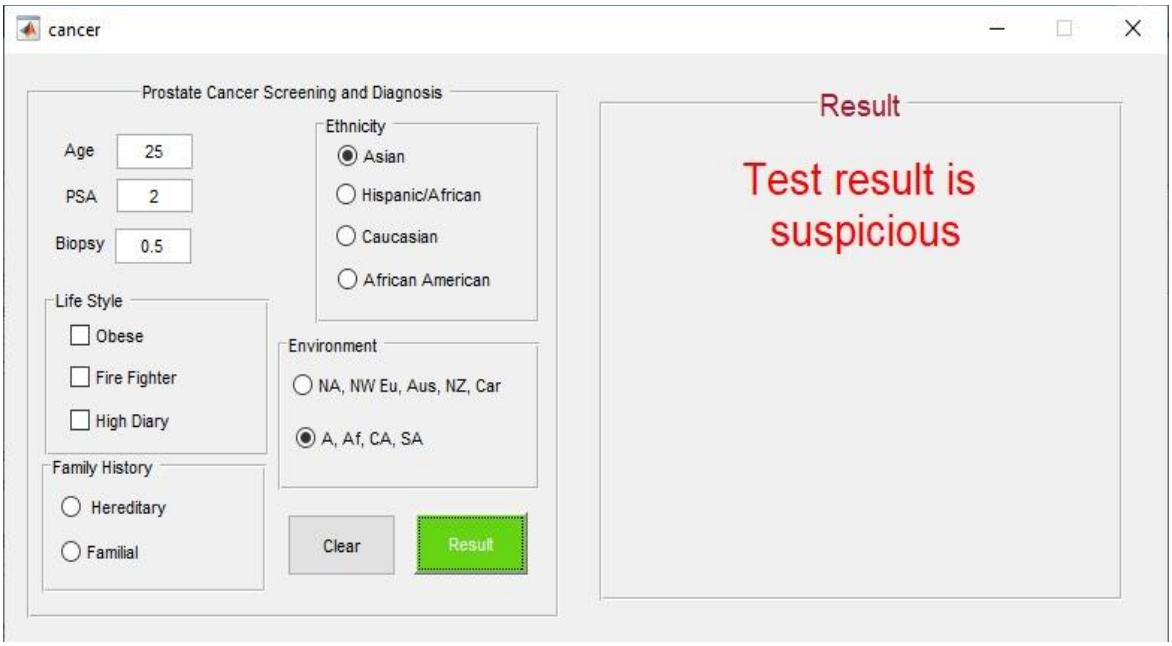


Figure 5.4: A Screen dialog showing Patient 3 simulation result (prostate cancer diagnosis)

5.6. FIS Simulation Result for Patient 4

Patient 4 is 25 years old which is within the age interval 15, 30] which is classified as young. His prostate specific antigen (PSA) level is 25 which is within the interval [12, 50] that is classified as very high level of PSA in ng/mL. Patient 4's biopsy is 0.5 (average), which means prostate cancer cells were not detected in him, therefore, the patient may need another biopsy test, or further analysis of urinary tract infection and kidney problems. Record of his family shows that no one from his family have either hereditary or familial prostate cancer risk. Furthermore, Patient 4's ethnicity is Asian who are categorized under low risk of prostate cancer.

The environment in which patient 4 resides is Asia, a low risk environment for prostate cancer cases. Lastly, data of the patient's lifestyle shows no record of obesity, high diary intake, and his occupation was not firefighting. After all the data were inputted in to the

FIS system, the result (output) shows that the patient 4's result is suspicious. What this means is that, no prostate cancer risk was detected, however, there was problem detected in the prostate gland that is not related with prostate cancer, therefore patient 4 is recommend to go for urinary tract infection and kidney pathological test to check the problem. Additional, since patient 4's result shows suspicious, then there is no need for further examination of Gleason score and the staging of the prostate cancer. The result is presented in figure 5.5.

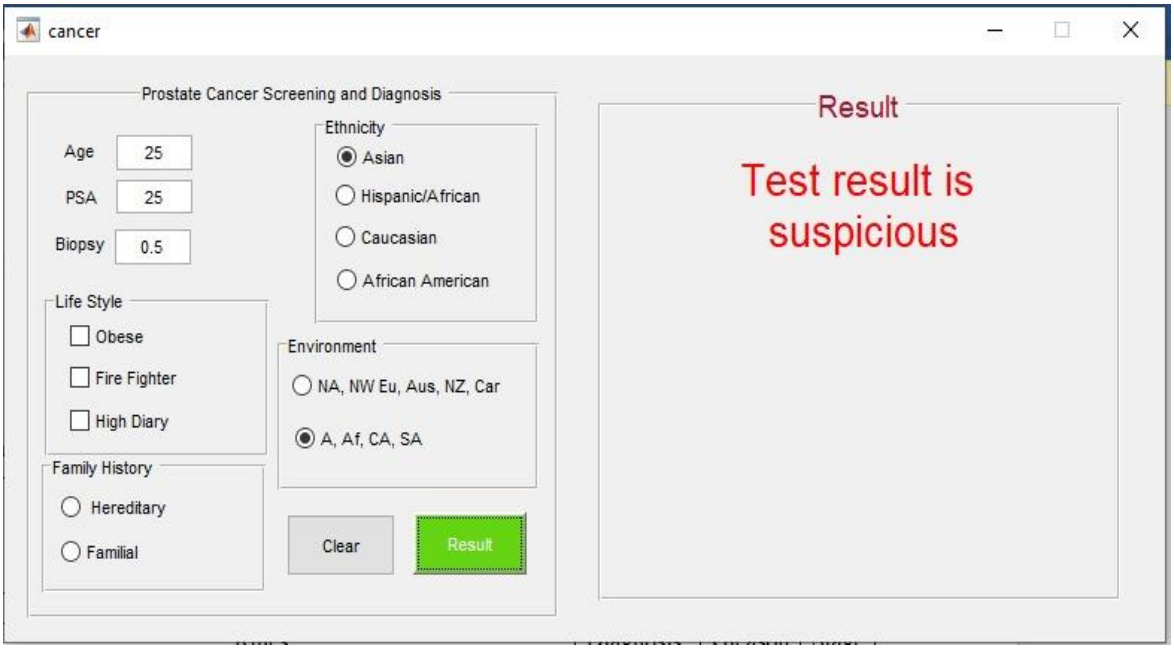


Figure 5.5: A Screen dialog showing Patient 4 simulation result (prostate cancer diagnosis)

5.7. FIS Simulation Result for Patient 5

Patient 5 is 32 years old which is within the age interval [25, 40] which is classified as medium age. His prostate specific antigen (PSA) level is 6 which is within the interval [2, 8] that is classified as low level of PSA in ng/mL. Patient 5's biopsy is 0 (negative), which means prostate cancer cells were not detected in him. Record of his family shows that no one from his family have either hereditary or familial prostate cancer risk. Furthermore, Patient 5's ethnicity is Asian who are categorized under low risk of prostate cancer. The environment in which patient 5 resides is Asia, a low risk environment for prostate cancer cases. Lastly, data of the patient's lifestyle shows no record of obesity, high diary intake, and his occupation was not firefighting. After all the data were inputted in to the FIS

system, the result (output) shows that the prostate cancer is negative, and does not require any further examination of Gleason score and the staging of the prostate cancer. The result is presented in figure 5.6.

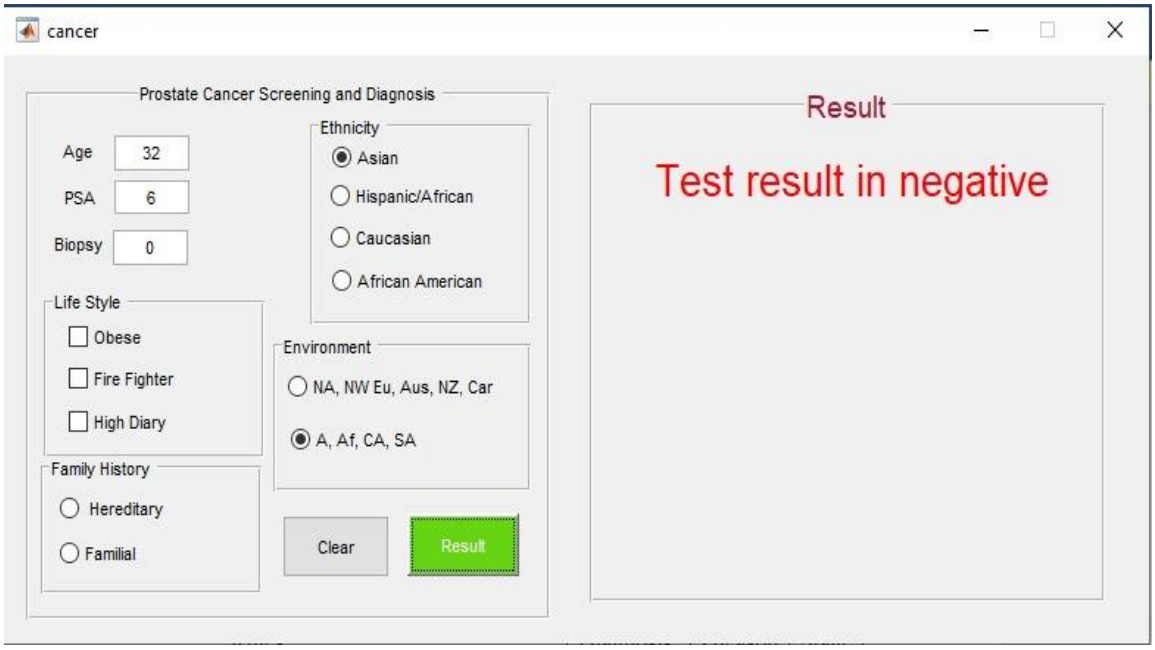


Figure 5.6: A Screen dialog showing Patient 5 simulation result (prostate cancer diagnosis)

5.8. FIS Simulation Result for Patient 6

Patient 6 is 35 years old which is within the age interval [25, 40] which is classified as medium age. His prostate specific antigen (PSA) level is 19 which is within the interval [12, 50] that is classified as very high level of PSA in ng/mL. Patient 6's biopsy is 1 (positive), which means prostate cancer cells were detected in him. Record of his family shows that no one from his family have either hereditary or familial prostate cancer risk. Furthermore, Patient 6's ethnicity is Asian who are categorized under low risk of prostate cancer.

The environment in which patient 6 resides is South America, a high-risk environment for prostate cancer cases. Lastly, data of the patient's lifestyle shows no record of obesity, high diary intake, and his occupation was not firefighting. After all the data were inputted in to the FIS system, the result (output) shows that the prostate cancer is positive, and requires further examination of Gleason score and the staging of the prostate cancer. The result is presented in figure 5.7. A summary of all patient results is tabulated in table 5.2

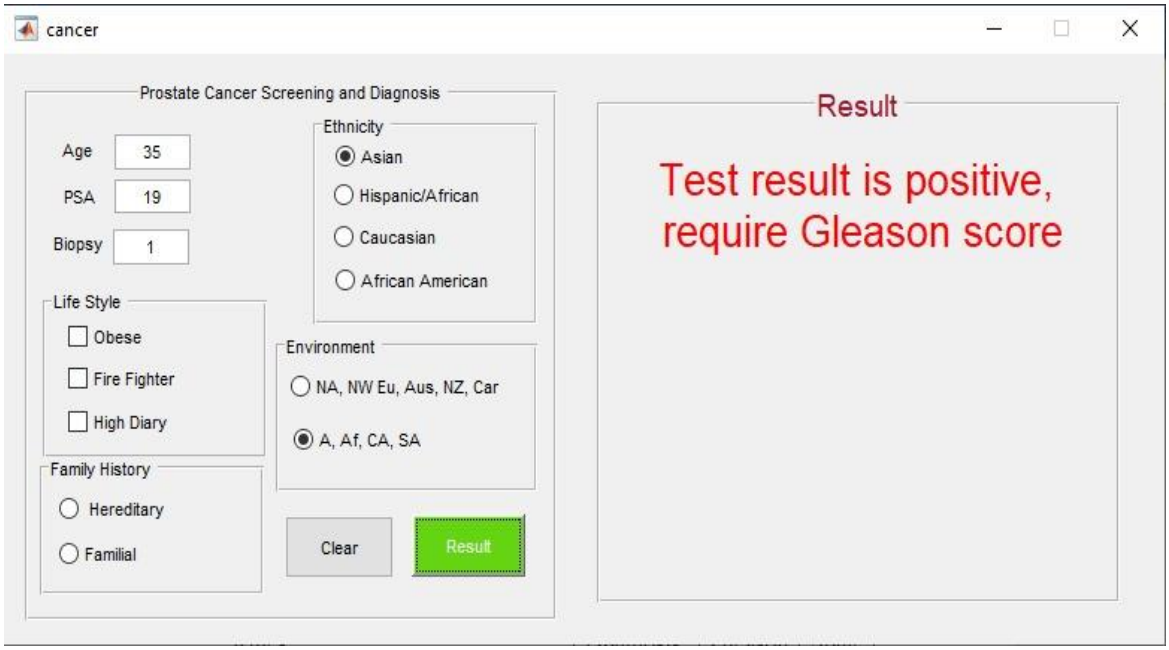


Figure 5.7: A Screen dialog showing Patient 6 simulation result (prostate cancer diagnosis)

Table 5.2: Summary of all patient simulation results (prostate cancer diagnosis)

Patients	**Diagnosis (output)**	**Recommendation**
Patient 1	Positive	Gleason score, staging
Patient 2	Positive	Gleason score, staging
Patient 3	Suspicious	Urinary tract, kidney test
Patient 4	Suspicious	Urinary tract, kidney test
Patient 5	Negative	None
Patient 6	Positive	Gleason score, staging

CHAPTER 6

CONCLUSION AND RECOMMENDATION

6.1. Conclusion

This study has successful provided effective diagnosis of prostate cancer patients using fuzzy inference system (FIS) simulation. The who process was coded and run using the phenomenal MATLAB (Matrix laboratory) software. Furthermore, Graphics users Interface (GUI) was utilized to make predictive diagnosis for six (6) prostate cancer patients for the purpose of improving effective diagnosis of prostate cancer among men especially those with high degree of risks. It is important to note that the predictive results from the system were based on the rules for the FIS system that was established after consulting with medical experts in the field of prostate cancer. The result (outputs) was achieved from the simulation of each patient's medical data (inputs) which are a combination of all-important variables important for the diagnosis of prostate cancer. These variables included PSA test, DRE, histopathology (biopsy), and the risk factors (age, ethnicity, environment, family history, and lifestyle).

The output from the FIS system after simulation was in three forms; diagnosis of prostate cancer, Gleason score, and stage of the cancer. The prostate cancer diagnosis has positive (P), suspicious (S), and negative (N). If the output from the simulation of a prostate cancer patient's medical data comes out positive, it means the FIS system have detected the presence of cancer after comparing or simulating all the variables in the patient's data. Moreover, if patient result comes out suspicious, it means the FIS system did not detected the presence of prostate cancer, therefore, the recommends further test. This is because a suspicious result indicates that there is no prostate cancer, however, the patient may have other non-cancerous problems affecting his prostate such as urinary tract infection and kidney problems etc.

A negative result from a prostate cancer patient's simulation of medical data (all variables) signify that the FIS system didn't detect the presence of prostate cancer in the patient. Lastly, Gleason score, and stage are dependent on positive diagnosis, otherwise, there is no

need for further test. Therefore, when the output of a patient's medical data come out as positive, the FIS recommends the medical expert to go ahead and examine the Gleason score and the stage of the prostate cancer using imaging technology such as MRI, CT, PET, and SPECT.

6.2. Recommendation

Although the study was successful, the following two recommendations can improve future studies;

- The simulation was done on medical data of only six (6) patients, future studies should expand the data set to include more medical data of patients. This will help improve the outcome that can be provided by the FIS system and it will improve the performance

- There are several factors and risk factors that may be directly or indirectly related to the diagnosis of prostate cancer, a full incorporation of all of them in to the system in order to provide not only and effective diagnosis, but also increase the efficiency of providing accurate prediction of the likelihood of prostate cancer.

REFERENCES

Abdulhaqq, N.I. (2019). A critical evaluation of cancer care in Nigeria, comparative analysis to other countries, and customization of standard cancer care system for Nigeria. *Master thesis, Institute of applied science, department of biomedical engineering, Near east university,* Nicosia, North Cyprus

Abiyev R.H., & Abizade, S. (2016). Diagnosing Parkinson's diseases using fuzzy neural system. *Journal of computational and mathematical methods in medicine*,3, 1-9.

Abiyev, R. H., Aliev, R., Kaynak, O., Turksen, I. B., & Bonfig, K. W. (2015). Fusion of Computational Intelligence Techniques and Their Practical Applications. *Computational. intelligence and neuroscience*, *201*, 463147.

Adeli A, & Neshat, M. (2010). A Fuzzy Expert System for Heart Disease Diagnosis. *In Proceedings of the International Multi Conference of Engineers and Computer Scientist* (pp. 17-19), Hong Kong, IMECS.

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. (2018). Global Cancer Statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin, in press as part of IARC's Global Cancer Observatory*, http://gco.iarc.fr/.

Brennhovd, B. (2016). Treatment of prostate cancer. In Oncology Encyclopedia. Oslo University Hospital Norway. Retrieved June 17, 2020 from http://health.costhelper.com/prostate-cancer-treatment.html

Oncology Encyclopedia. Cost helper. (2018). How Much Does Prostate Cancer Treatment Cost? Retrieved march 30, 2020 from http://health.costhelper.com/prostate-cancer-treatment.html

Brennhovd, B., & Iversen, J. (2016). Prostate cancer. In Oncology Encyclopedia. Norway: Oncology Encyclopedia. Retrieved June 17, 2020 from http://health.costhelper.com/prostate-cancer-treatment.html

Bridged, C. (2019). 10 Challenges In Building High-quality AI Training Data. Retrieved May 27th, 2020 from https://bridged.co/blog/ai-training-data-10-common-challenges/

Sternberg, C. N., DeBono,J. S., Chi, K. N., et al. (2014). Improved outcomes in elderly patients with metastatic castration-resistant prostate cancer treated with the

androgen receptor inhibitor enzalutamide: results from the phase III AFFIRM trial, *Annals of Oncology*, 25 (2), 429–434.

Chang, L. (2018). Treatment Options for Prostate Cancer. Retrieved May 20, 2020 from https://www.webmd.com/prostate-cancer/prostate-cancer-treatment-options#2

Chen R.C., & Bau, C.T. (2013). An ontological approach for guideline-based decision support system. *International Journal of Computer, Consumer and Control (IJ3C)*, 2(3), 27-34.

Cosma, G., Acampora, G., & Brown, D. et al. (2016). Prediction of Pathological Stage in Patients with Prostate Cancer: A Neuro-Fuzzy Model. *PLOSONE*, DOI: 10.1371/journal.pone.0155856

Cella, D., Ivanescu, C., Holmstrom, S. et al. (2015). Impact of enzalutamide on quality of life in men with metastatic castration-resistant prostate cancer after chemotherapy: additional analyses from the AFFIRM randomized clinical trial, *Annals of Oncology*, 26(1), 179–185.

Ekong, V.E., Inyang, U.G., & Onibere, E.A. (2012). Intelligent decision support system for depression diagnosis based on neuro-fuzzy-CBR hybrid. *Mod Appl Sci*, 6, 79-88.

Farokhzad, M. R., & Ebrahimi, I. (2016). A novel adaptive neuro fuzzy inference system for the diagnosis of liver disease. *International journal of academic research in computer engineering*, 1(1), 61-66.

Fraccaro, P., O'Sullivan, D., Plastiras, P., O'Sullivan, H., Dentone, C., Di Biago, A., Weller, P. (2015). Behind the screens: Clinical decision support methodologies – A review. *Health Policy and Technology*, 4(1), 29–38.

Fua, J., Yeb, J., & Cuib, W. (2018). An evaluation method of risk grades for prostate cancer using similarity measure of cubic hesitant fuzzy sets. *Journal of Biomedical Informatics*, 87, 131–137.

Gorgulu, O., and Akilli, A. (2016). Use of Fuzzy Logic Based Decision Support Systems in Medicine. *Journal of Ethno-Medicine*, 10(4), 393- 403

Gospodarowicz, M., Trypuc, J., & Cruz, A.D., et al. (2015). Cancer Services and the Comprehensive Cancer Center. In: Gelband H, Jha P, Sankaranarayanan R, et al., (Eds.). *Cancer: Disease Control Priorities*, Third Edition (V 3), Chapter 11 (pp. 124-172). Washington (DC)

Guess, B. (2006). Chemotherapy for Prostate Cancer: Why Bother?. Retrieved April 26, 2020 from https://pcri.org/chemotherapy-for-prostate-cancer-why-bother/

Hagendorff, T., Wezel, K. (2020). 15 challenges for AI: or what AI (currently) can't do. *AI & Soc*, 35, 355–365.

Han J. (2005). Data Mining: concepts and techniques. San Francisco, CA, USA: *Morgan Kaufmann Publishers Inc*.

Idoko, J.B., Arslan, M., & Aliyev, R. et al. (2018). Fuzzy neural system application to differential diagnosis of erythemato-squamous diseases. *Cyprus journal of medical sciences*, 3(2), 90-97.

Jang, J.R. (1993). ANFIS: adaptive-network-based fuzzy inference system, in IEEE Transactions on Systems, *Man, and Cybernetics*, 23(3), 665-685.

James, M., & Slater, M.D. P. (2018). Prostate Cancer Stages & Survival Rates. Retrieved March 25, 2020 from https://protons.com/proton-treatments/prostate-cancer/diagnosis-prostate-cancer

Kar, S., & Majumder, D.D. (2017). An Investigative Study on Early Diagnosis of Prostate Cancer Using Neuro-Fuzzy Classification System for Pattern Recognition. *Int. J. Fuzzy Syst.* 19, 423–439.

Kiefer, D. (2017). Digital Rectal Exam. Retrieved April 1, 2020 from https://www.healthline.com/health/digital-rectal-exam

Zadeh, L. (1965). Information and Control 8, 338, ISSN 0019-9958. Retrieved June 14, 2020 from http://www.sciencedirect.com/science/article/pii/S001999586590241X

Live Science. (2010). Prostate Cancer Symptoms and Treatment. Retrieved April 1, 2020 from https://www.livescience.com/9934-prostate-cancer-symptoms-treatment.html

Zelefsky, M.J., Eastham, J.A., Sartor, A.O. (2011). Cancer of the prostate, in: V.T. DeVita, Jr., T.S. Lawrence, S.A. Rosenberg (Eds.), *Cancer: Principles and Practice of Oncology. 9th ed*, (pp. 1220–1271). Philadelphia, Pa: Lippincott Williams & Wilkins.

Ma'aitah, M.K.S., Aliyev, R., & Bus, I.J. (2017). Intelligent classification of liver disorder using fuzzy neural system. *International journal of advanced computer science and applications*, 8(12), 25-31.

Mahant, J., & Panda, S. (2018). Fuzzy expert system for prediction of prostate cancer. *Journal of New Mathematics and Natural Computation*, 16(1), 163-176

McCrady, B.S., Epstein, E.E., & Kahler, C.W. (1998). Families of Alcoholics. *Journal of comprehensive clinical psychology*, 9, 199-218

MD Anderson Cancer Center. (2018). What is Proton Therapy? Retrieved April 10th, 2020 from https://www.mdanderson.org/patients-family/diagnosis-treatment/care-centers-clinics/proton-therapy-center/what-is-proton-therapy.html

Metlin C., Lee F., Drago J., et al. (1991). The American Cancer Society National Prostate Cancer Detection, Project: Findings on the detection of Early Prostate Cancer in 2425 Men, *Cancer*, 67, 2949-2958.

National cancer institute. (2015). Cancer Staging. Retrieved April 2, 2020 from https://www.cancer.gov/about-cancer/diagnosis-staging/staging

Nichols, H. (2019). What are the leading causes of death in the US? Retrieved March 13, 2020 from https://www.medicalnewstoday.com/articles/282929

Prostate Cancer foundation. (2018). What Are the Side Effects of Prostate Cancer Treatment? Retrieved April 20, 2020 from https://www.pcf.org/c/side-effects/

Prostate cancer free. (2018). Prostate Cancer Treatment Side Effects Treatment vs Side Effects. Retrieved May 4th, 2020 from https://prostatecancerfree.org/prostate-cancer-treatment-side-effects/

Prostate Cancer UK. (2016). chemotherapy. Retrieved April 26, 2020 from https://prostatecanceruk.org/prostate-information/treatments/chemotherapy

Rajabi, M., Sadeghizadeh, H., Mola-Amini, Z., & Ahmadyrad, N. (2019). Hybrid adaptive neuro-fuzzy inference system for diagnosing the liver disorders. *Arxiv preprint arxiv:*1910.12952.

Rawla P. (2019). Epidemiology of Prostate Cancer. *World J Oncol*. 10(2), 63-89.

Ren, S., F. Wang, F., & Shen, J. et al. (2013). Long non-coding RNA metastasis associated in lung adenocarcinoma transcript1 derived mini RNA as a novel plasma-based biomarker for diagnosing prostate cancer, *Eur J. Cancer*, S0959–8049(13), 360–362.

Edge, S.B., Byrd, D.R., Compton, C.C. (2010). *AJCC Cancer Staging Manual*, 7th ed., New York, NY: Springer.

Samuel, O.W., Omisore, M.O., & Ojokoh, B.A. (2013). A web-based decision support system driven by fuzzy logic for the diagnosis of typhoid fever. *Expert Syst Appl*, 40, 4164-4171.

Saritas, I., Allahverdi, N., & Sert, I.U. (2003). A Fuzzy Expert System Design for Diagnosis of Prostate Cancer. *In International Conference on Computer Systems and Technologies – Computer System Technology, Proceedings of the 4th international conference on Computer systems and technologies: e-Learning*, 345–351

Saritas, I.N., Allahverdi, N. & Sert, U. (2013). A fuzzy approach for determination of prostate cancer, *International Journal of Intelligent Systems and Applications in Engineering*, 1(1), 01-07.

Scott, E., Eggener, R., Bryan, R., Andrew, J., Armstrong, A.J., Todd, M. et al. (2020). *Journal of Clinical Oncology*, 38(13), 1474-1494.

Stanford Encyclopedia of Philosophy. (2017). Fuzzy Logic. Retrieved August 23rd, 2020 from https://plato.stanford.edu/entries/logic-fuzzy/

Swindle, P. (2020). Trans perineal ultrasound guided prostatic biopsy (TPUS). Retrieved from April 25th, 2020 https://qprostate.com.au/prostate-cancer/diagnosis/transperineal-ultrasound-guided-prostatic-biopsy-tpus/

Tewari, A., Porter, C., Peabody, J., Crawford, E.D., Demers, R., Johnson, C.C. et al. (2001). Predictive modeling techniques in prostate cancer. *Molecular Urology*, 5(4), 147–152.

Tsao, C.W., Liu, C.Y., & Cha, T.L., et al. (2014). Artificial neural network for predicting pathological stage of clinically localized prostate cancer in a Taiwanese population. *Journal of the Chinese Medical Association*, 77(10), 513–518.

U.S. Department of Health and Human Services. (2019). Diagnostic Errors. Retrieved June 21, 2020 from https://psnet.ahrq.gov/primer/diagnostic-errors

UCLA Prostate Cancer Program. (2018). Radiation therapy prostate cancer - Prostate Cancer Program, UCLA Urology - Los Angeles, CA. Retrieved March 3, 2020 from http://urology.ucla.edu/prostate-cancer/radiation-therapy-prostate-cancer

Vaghela, C., Bhatt, N., Mistry, D. (2015). A survey on various classification techniques for clinical decision support system. *International Journal of Computer Applications*, 116(23), 14-17.

Wong, M.K.S. (2016). Kallikrein. *Handbook of Hormones*; Comparative Endocrinology for Basic and Clinical Research. Edited by: Yoshio Takei, Hironori Ando and

Kazuyoshi Tsutsui. Retrieved June 27th, 2020 from https://www.sciencedirect.com/topics/medicine-and-dentistry/plasma-kallikrein

World cancer research fund. (2018). Prostate cancer statistics. Retrieved March 30th, 2020 from https://www.wcrf.org/dietandcancer/cancer-trends/prostate-cancer-statistics

APPENDICES

APPENDIX 1: Ethical Approval letter

ETHICAL APPROVAL DOCUMENT

Date:01/12/2020

To the Graduate School of Applied Sciences

For the thesis project entitled as “Intelligent Fuzzy System for Early Prostate Cancer Diagnosis”, the researchers declare that they did not collect any data from human/animal or any other subjects. Therefore, this project does not need to go through the ethics committee evaluation.

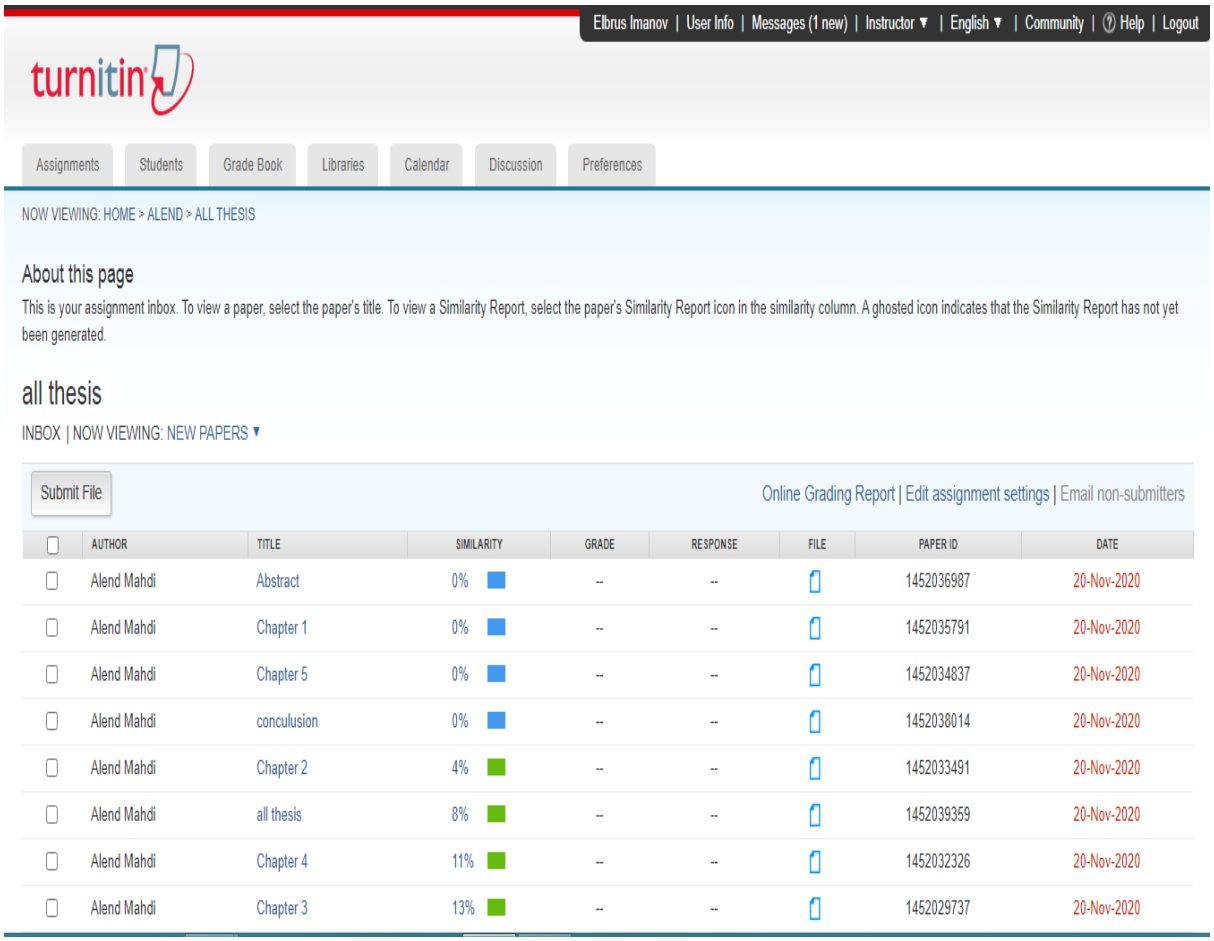
Title: Assist. Prof. Dr

Name Surname: Elbrus Imanov

Signature:

Role in the Research Project: Supervisor

APPENDIX 2: Similarity Report



Elbrus Imanov | User Info | Messages (1 new) | Instructor ▼ | English ▼ | Community | Help | Logout

turnitin

Assignments | Students | Grade Book | Libraries | Calendar | Discussion | Preferences

NOW VIEWING: HOME > ALEND > ALL THESIS

About this page

This is your assignment inbox. To view a paper, select the paper's title. To view a Similarity Report, select the paper's Similarity Report icon in the similarity column. A ghosted icon indicates that the Similarity Report has not yet been generated.

all thesis

INBOX | NOW VIEWING: NEW PAPERS ▼

Submit File

Online Grading Report | Edit assignment settings | Email non-submitters

AUTHOR	TITLE	SIMILARITY	GRADE	RESPONSE	FILE	PAPER ID	DATE
Alend Mahdi	Abstract	0%	--	--		1452036987	20-Nov-2020
Alend Mahdi	Chapter 1	0%	--	--		1452035791	20-Nov-2020
Alend Mahdi	Chapter 5	0%	--	--		1452034837	20-Nov-2020
Alend Mahdi	conculusion	0%	--	--		1452038014	20-Nov-2020
Alend Mahdi	Chapter 2	4%	--	--		1452033491	20-Nov-2020
Alend Mahdi	all thesis	8%	--	--		1452039359	20-Nov-2020
Alend Mahdi	Chapter 4	11%	--	--		1452032326	20-Nov-2020
Alend Mahdi	Chapter 3	13%	--	--		1452029737	20-Nov-2020

Title of Thesis: **Intelligent Fuzzy System for Early Prostate Cancer Diagnosis**

Supervisor, Assist.Prof.Dr. Elbrus Imanov