ASSESSMENT OF THE ERRORS INVOLVED IN PCR TESTS

by

Ayşe Nur Göz

Submitted to Graduate School of Natural and Applied Sciences
in Partial Fulfillment of the Requirements
for the Degree of Master of Science in
Biomedical Engineering

Yeditepe University

2022

ASSESSMENT OF THE ERRORS INVOLVED IN PCR TESTS

APPROVED BY:

Prof. Dr. Mustafa Özilgen ……
(Thesis Supervisor)
(Yeditepe University)

Assoc. Prof. Dr. Ayşegül Doğan ……
(Yeditepe University)

Assist. Prof. Dr. Zeynep Büşra Bolat ……
(University of Health Science)

DATE OF APPROVAL: ..../..../20..

I hereby declare that this thesis is my own work and that all information in this thesis has been obtained and presented in accordance with academic rules and ethical conduct. I have fully cited and referenced all material and results as required by these rules and conduct, and this thesis study does not contain any plagiarism. If any material used in the thesis requires copyright, the necessary permissions have been obtained. No material from this thesis has been used for the award of another degree.

I accept all kinds of legal liability that may arise in case contrary to these situations.

Name, Last name Ayşe Nur GÖZ

Signature …..…….……..

ACKNOWLEDGEMENTS

First of all, I would like to thank my thesis advisor Prof. Dr. Mustafa Özilgen, for being with me with all his support when I thought everything was over. Being with his students at all times, his great knowledge, and his kind and helpful personality were valuable to me in my thesis process and personal life.

I also would like to express my endless thanks and gratitude to my dear family and husband who have always given me their love and support throughout my life, and who have always encouraged and guided me to pursue my dreams.

Finally, I would like to thank my close friends who never stopped supporting me during this process.

ABSTRACT

ASSESSMENT OF THE ERRORS INVOLVED IN PCR TESTS

PCR tests are the diagnostic method that is accepted as the gold standard in the diagnosis of COVID-19 disease. There are many different kits with differences such as targeted gene regions and kit content used in performing these tests around the world. There are many factors that affect the accuracy of PCR test results. In general, some of these may be caused by the suspected patient, the laboratory environment, the specialist, or the diagnostic kit used. These factors cause false positive and false negative results to occur. In this study, data from thirty-one articles related to the polymerase chain reaction (PCR) method used in the diagnosis of SARS-COV-2 were used. In total, approximately 188,201 disease test data were obtained. Of these tests, 1472 were identified as discordant test results (false positive/false negative). Of these results, 1025 were accepted as false negative and 447 as false positive. The approximate false positive rate (FPR) calculated according to these data is 0.56%, and the false negative rate (FNR) is 0.95%. This study was carried out by making calculations and researches on approximate data in order to obtain a general result. As a result of the research, it is thought that false positive results are caused by pipetting errors, contamination and false negative results are caused by improper sampling and sampling in the early stages of the disease.

ÖZET

PCR TESTLERİNDEKİ HATALARIN DEĞERLENDİRİLMESİ

PCR testleri, COVID-19 hastalığının tanı ve teşhisinde altın standart olarak kabul edilen tanı yöntemidir. Dünya genelinde bu testlerin yapılmasında kullanılan hedeflenen gen bölgeleri ve kit içeriği gibi farklılıklar gösteren birçok farklı kit bulunmaktadır PCR test sonuçlarının doğruluğunu etkileyen birçok faktör vardır. Genelde bunların bir kısmı şüpheli hastadan, laboratuvar ortamından, uzmandan veya kullanılan tanı kitinden kaynaklanabilir. Bu faktörler yanlış pozitif ve yanlış negatif sonuçların ortaya çıkmasına neden olur. Bu çalışmada SARS-COV-2 tanısında kullanılan polimeraz zincir reaksiyonu (PCR) yöntemi ile ilgili otuz bir makaleden elde edilen veriler kullanılmıştır. Toplamda yaklaşık 188,201 hastalık testi verisi elde edildi. Bu testlerden 1472'si uyumsuz test sonuçları olarak tanımlandı (yanlış pozitif/yanlış negatif). Bu sonuçlardan 1025'i yanlış negatif, 4472'si yanlış pozitif olarak kabul edildi. Bu verilere göre hesaplanan yaklaşık yalancı pozitiflik oranı (FPR) %0.56, yalancı negatif oranı (FNR) ise %0.95'tir. Bu çalışma genel bir sonuç elde etmek için yaklaşık veriler üzerinde hesaplamalar ve araştırmalar yapılarak gerçekleştirilmiştir. Araştırma sonucunda yanlış pozitif sonuçların pipetleme hatalarından, kontaminasyondan ve yanlış negatif sonuçların hastalığın erken evrelerinde uygunsuz örnekleme ve örneklemeden kaynaklandığı düşünülmektedir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

mRNA	messenger RNA
ICTV	International Committee on Taxonomy of Viruses
WHO	World Health Organization
SARS-COV	Severe Acute Respiratory Syndrome Coronavirus
MERS-COV	Middle East Respiratory Sydrome Coronavirus
CoVS	Coronaviruses
ACE2	Angiotensin-converting enzyme 2
ORFs	Open Reading Frames
S	Spike
E	Envelope
N	Nucleocapsid
M	Membrane
aa	Amino acids
POCT	Point-of-care Testing
ELISA	Enzyme-linked immunosorbent assay
PCR	Polymerase Chain Reaction
Q-RT PCR	Quantitative Real Time PCR
RT-PCR	Reverse Transcriptase PCR
CT	Computed Tomography
BAL	Bronchoalveolar lavage
VTM	Viral transport medium
UV	Ultraviolet
cDNA	Complemantary DNA
Ct	Cycle Threshold
α	Significance, Type I error
β	Power of Tests, Type II error
δ	Effect size
H0	Null hypothesis
H1	Alternative Hypothesis
PC	Positive Control
NC	Negative Control

TPR	True Positive Rate
TP	True Positive
P	Positives
FN	False Negative
TNR	True Negative Rate
TN	True Negative
N	Negative
FP	False Positive
FPR	False Positive Rate
FNR	False Negative Rate

1. INTRODUCTION

1.1. GENERAL INFORMATION

Viruses and bacteria are known to cause serious and dangerous diseases in humans. Because these microorganisms have a difficulty in diagnosis, higher transmission rate, and lack of advanced healthcare systems, they can cause serious impacts and pandemics [1]. Viruses are defined as infectious microbes that have genetic material (DNA or RNA) wrapped in a protein coat. They cannot reproduce alone, they can only reproduce within living organisms. They also don't have their own methabolism[2]. It finds a host cell it can infect and multiplies through it. In this way, it harms the host organism. They use the metabolism, organelles, and molecules of the host cell and use the host cell to replicate themselves [3]. It has the capacity to affect all living forms. Viruses are studied by virology, a subspecialty of microbiology. There are millions of different virus types, less than 7000 of which have been described in detail in the literature[4]. They are structurally composed of genetic material (Long DNA or RNA molecules), a capsid layer is known as a protein coat, and an outside envelope of lipids. The capsid layer is formed by the proteins encoded by the viral genome. The shape of the capsid layer is very important for morphological differentiation. [5,6]. Most viruses have virion structures that are too small to be seen with a microscope. Virions, defined as complete infectious virus particles, are responsible for delivering the DNA or RNA genome to the host cell for the genome to be recognized by the host cell. It basically consists of nucleic acid and a protein coat. Their structural shapes are simple helical structures, icosahedral structures, prolate, enveloped, and complex. Some biologists consider viruses living forms because they carry genetic material, evolve through natural selection and reproduce,. However, they lack basic features, such as cell structure, that are often important criteria for defining life[7,8]. The life cycle of viruses usually takes place in six basic steps. These are respectively attachment, penetration, uncoating, replication, assembly, and release.Viruses spread in many different ways. They spread from disease-carrying organisms by air, fecal-oral route, hand-to-mouth contact, food or water, sexual contact, and contact with infected blood. Viruses are divided into DNA and RNA viruses according to the genetic material they contain DNA and RNA. Among them, the most common are RNA viruses. Besides containing only DNA and only RNA, they can contain both at different

steps of their life cycle. Shapely, they have three group as linear, circular and segmented. There are 3 groups of helical structures. These are double-stranded, double-stranded with regions of single-strandedness and single-stranded. Single-stranded genomes occur in two separated and unpaired nucleic acids, while double-stranded genomes consist of two paired and complementary nucleic acids. According to their genome orientation, they are divided into three groups. These are called positive sense (+), negative sense (-), and ambisense (+/-)[9]. Because positive sense viral RNA is in the identical orientation as messenger RNA (mRNA), it can be easily read by the host cell. In the negative sense, viral RNA acts as a complement to messenger RNA (mRNA) and therefore they need RNA polymerase enzyme. In general, DNA viruses have been found to have a larger genome than RNA viruses. The main purpose of classification is to name viruses by using their diversity and similarities and to provide identification by grouping them accordingly. While classifying viruses, features such as host organisms, replication mode, nucleic acid type, morphology, and the type of disease they cause are taken into consideration[10]. Two systems are used to classify according to these features. These are the Baltimore classification and the ICTV (International Committee on Taxonomy of Viruses) classification, respectively[11]. ICTV is a system that makes the species naming of viruses and aims to place the species into genera, subfamilies, families, and orders. The ordering and taxonomic appendices of the classification are as follows: Realm (-viria), Subrealm (-vira), Kingdom (-virae), Subkingdom (-virites) ,Phylum (-viricota) ,Subphylum (-viricotina) ,Class (-viricetes), Subclass (-viricetidae), Order (-virales), Suborder (-virineae), Family (-viridae), Subfamily (-virinae), Genus (-virus) ,Subgenus (-virus) and Species [12]. ICTV has identified 6 realms, 10 kingdoms, 17 phyla, 2 subphyla, 39 classes, 65 orders, 8 suborders, 233 families, 168 subfamilies, 2,606 genera, 84 subgenera, 10,434 species of viruses as of 2021. In the Baltimore classification system developed by David Baltimore, viruses are classified according to nucleic acid types (DNA or RNA), strandedness (number of strands, single-stranded or double-stranded), replication and sense strategies. In this direction, viruses are divided into seven groups[13]. These are, in order: I: dsDNA viruses: Adenoviruses, Herpesviruses, Poxviruses, II: ssDNA viruses (+ strand or "sense") DNA: Parvoviruses, III: dsRNA viruses: Reoviruses, IV:(+)ssRNA viruses (+ strand or sense) RNA: Coronaviruses, Picornaviruses, Togaviruse,V: (−)ssRNA viruses (− strand or antisense) RNA: Orthomyxoviruses, Rhabdoviruses,VI: ssRNA-RT viruses (+ strand or sense) RNA with DNA intermediate in life-cycle: Retroviruses, and VII: dsDNA-RT viruses DNA with RNA

intermediate in life-cycle: Hepadnaviruses. DNA viruses are viruses that contain DNA as genetic material and need DNA polymerase enzymes to ensure their replication. Although they are predominantly double-stranded, some may be single-stranded. These virus strains are included in Group I and Group II in the Baltimore classification. While viruses belonging to Group I have double-stranded DNA structures, viruses belonging to Group II have single-stranded DNA structures [14]. The virus families belonging to these groups are as follows: Adenoviridae, Papovaviridae, Parvoviridae, Herpesviridae, Poxviridae, Anelloviridae, and Pleolipoviridae. Viruses that contain RNA as their genetic material. Although they are predominantly single-stranded, some are double-stranded [15]. They are found in Group III, Group IV, and Group V according to the Baltimore classification. Group III has a double-stranded RNA genome, Group IV has a positive-sense single-stranded RNA genome, and Group V has a negative-sense single-stranded RNA genome. Group IV includes common and known viruses such as Hepatitis A, enterovirus, rhinovirus, and SARS virus. The SARS-COV-2 virus known as Covid-19, which causes a serious pandemic, is in Group IV, the Coronaviridae family. Coronaviruses are defined as zoonotic viruses that can be transmitted from animals to humans or from humans to animals. For example, SARS-COV(Severe Acute Respiratory Syndrome Coronavirus) is transmitted from civet cats to humans. Likewise, MERS-COV(Middle East Respiratory Syndrome Coronavirus) is transmitted from dromedary camels to humans. According to research and relationships between other coronavirus types, SARS-COV-2 is thought to have been transmitted to humans from bats. In recent years, Middle East respiratory syndrome coronavirus (MERS-CoV), severe acute respiratory syndrome coronavirus (SARS-CoV), and SARS-CoV-2 with highly pathogenic findings have been identified. These diseases emerged and started to spread in 2002/2003, 2012, and 2019, respectively. These three diseases emerged as deadly epidemics. However, among these, SARS-COV-2 has overtaken SARS and MERS in terms of the number of people it infects and causes death and the areas it affects. Coronavirus was first discovered in 2019 and it was not reported in humans. The emergence of the first officially known case was recorded on 8 December 2019 [16]. Then it was renamed as β- coronavirus on 12th January 2020, 2019-novel Coronavirus by World Health Organization. Lastly, on 11 February 2020, by the Coronavirus Study Group of the International Committee on Taxonomy, as it is used today, it was named SARS-COV-2 [17].The new coronavirus (SARS-CoV-2) which causes COVID- 19 is a virus that was first detected in the city of Wuhan, China, and causes severe respiratory tract infection. According to research, it is a

virus that was identified on January 13, 2020, and entered the literature. It was declared a global epidemic by the World Health Organization (WHO) and Public Health Emergency of International Concern on January 30, 2020, respectively [18]. As of October 18, 2022, there are 622.389.418 confirmed cases of COVID-19, including 6.548.492 deaths, reported by WHO[19]. On January 10, 2020, the genome sequence of the new coronavirus disease was published on the Virological website. Genome sequences made in different centers were collected and defined in the Global Initiative for Sharing All Influenza Data database on 12 January 2020. In the following days, on March 11, 2020, the World Health Organization officially declared the current COVID-19 outbreak as a pandemic[20]. In addition to showing similar symptoms to SARS and MERS, it has specific symptoms such a, cough, dyspnea, myalgia, fever, fatigue, sputum, sneezing, pneumonia, and changeable leukocyte counts are seen in people infected with COVID-19.

According to the phylogenetic studies, it has been observed that the lineage of SARS-COV-2 is the same as the coronaviruses identified in bats and pangolins. Despite being phylogenetically significantly related, SARS-COV-2 differs from all other described coronaviruses, including pangolins and bats. All around the world, there are seven coronaviruses (CovS) that infect humans [21]. These are grouped into four genera as alpha (α), beta (β), gamma (γ), delta (δ) CoV. Alpha- coronaviruses are includes HCoV-229E and NL63, and beta- coronaviruses are includes MERS-CoV, SARS-CoV, HCoV-OC43, and HCoV-HKU1. Among these species, alpha and beta coronaviruses tend to directly infect humans, while delta coronaviruses and gamma coronaviruses tend to infect bird species in general. SARS-COV-2 belongs to beta coronavirus genera and the Orthocoronavirinae subfamily of the Coronaviridae family.

1.2. STRUCTURE OF SARS-COV-2

There are definite percentages of similarities between SARS-COV-2, MERS, and SARS-COV as genome sequences. These percentages are 79% and 50%, respectively. The length of the proteins encoded by SARS-COV-2 is similar to that of SARS-COV. There is more than 90% amino acid similarity between SARS-CoV and SARS-CoV-2 except for the S gene which encodes a surface protein for entry into host cells [22].

SARS-COV-2 viral RNA genome is surrounded by a nucleocapsid with helical symmetry. In general, CovS structurally have a positive-sense, single-stranded, nucleocapsid, enveloped, and helical appearance. CovS are generally in pleomorphic form. The viral envelope that SARS-CoV-2 has consists of a lipid bilayer. Each coronavirus has an average diameter of 60-140 nm. The single-stranded RNA genome of SARS-COV-2 contains 9860 amino acids and 29891 nucleotides for coding. The process of entering the coronavirus into the cell begins with the preparation of the S protein. For this, proteins containing extensions called spicules first bind to cellular receptors [23]. SARS-COV-2 uses the ACE2 (angiotensin-converting enzyme 2) receptor and the TMPRSS2 serine which is an enzyme to initiate Protein S for this entry process. ACE2 receptor controls the transition between species and humans. The entry of the virus into the cell is shown in Figure 1.1 in detail [24].

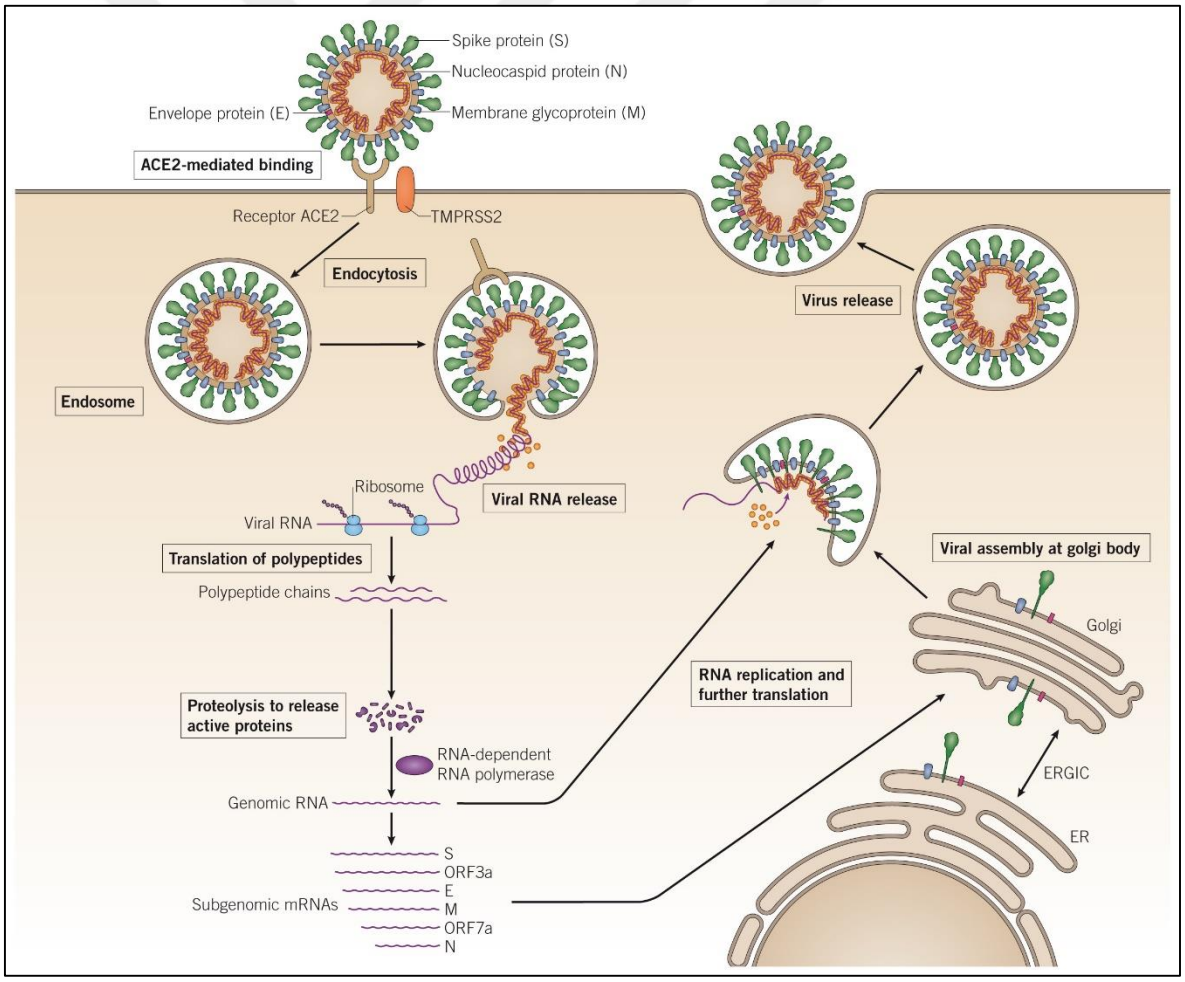

Figure 1.1. The cellular entry of SARS-COV-2

The S protein of the SARS-CoV genome contains 1,255 full-size amino acids, while the S-protein specific to SARS-COV-2 has 1,273 full-size amino acids. SARS-CoV-2 and SARS-

CoV have 73% amino acid similarity in their receptor binding sites. One of the definite genomic features of SARS -COV-2 is the addition of an extra four amino acid residues (PRRA) at the intersection of the S1 and S2 subunits of the S protein. There are 6 open reading frames (ORFs) in the SARS-COV-2 genome at least. Generally, SARS-COV-2 and CovS genomes are similar as both contain almost 10 open reading frames (ORFs). The first ORF(s) (ORF1a, ORF1b) which are almost two-thirds of the viral RNA are converted into two large polyproteins. These two ORFs guide the encoding of 16 non-structural proteins. The other ORFs (one-third of viral RNA) encode the main structural proteins spike (S), membrane (M), nucleocapsid (N), envelope (E) and proteins, and several accessory proteins(ORF8). Respectively the main structural proteins, the envelope (E) takes charge viral release and assembly process, the S protein takes part in the entry of the virus into the host cells and determines host tropism and transmission, N protein is responsible for viral replication. ORF 8, which belongs to SARS-COV-2 and is in the accessory protein category, is a protein with 121 amino acids (aa) from the N-terminal signal sequence in general terms. After the successful entry of the virus into the cell, the viral RNA genome migrates into the cytoplasm and is converted into two polyproteins and a structural protein in the cytoplasm. According to studies, the deletion of the ORF8 protein from the genomic sequence may be an indicator of the transition between species and the adaptation that took place in humans [25]. Considering the similarity rates between coronavirus species, it was seen that Alphacoronaviruses and Betacoronaviruses, which are defined as bat viruses, have more than 90% sequence similarity with SARS-COV-2 in all ORFs throughout the entire genome [26]. Considering all these similarities, the possibility that bat coronavirus may be the source of SARS-COV-2 becomes stronger. Besides bats, pangolins are other hosts that can be a source of SARS-COV-2. Among the coronaviruses related to pangolins, the similarity between Guangdong pangolin coronaviruses and SARS-CoV-2 is quite high. On the contrary, there is less than 85.5% genome sequence similarity between Guangxi pangolin coronaviruses and SARS-CoV-2. The National Center for Bioinformation of China aligned 77,801 genome sequences to study the genetic variations that occur with different SARS-COV-2 strains. Based on these studies, 15,018 different mutations have been identified so far [27].

1.3. DIAGNOSIS OF SARS-COV-2

Since the rise of SARS-COV-2, also known as COVID-19, many different methods have been used to diagnosis of the disease. These diagnostic methods vary according to the rate of spread of the disease, the rate of infection, and its effectiveness. Mainly detection of nucleic acid, CT chest scan as radiological imaging, serological testing for antibodies, enzyme-linked immunosorbent assay (ELISA)) blood culture, immune identification technology (Point-of-care Testing (POCT) of IgM/IgG, etc. are used in the diagnosis of the disease [28]. These various diagnostic methods support each other in making the correct diagnosis of the disease. In the diagnostic phase, polymerase chain reaction (PCR) as a molecular diagnostic test, antigen-antibody measurement as a serological test, and chest computed tomography (CT) as supportive radiological imaging are applied. Although each diagnostic method is of great importance in the detection of the disease, the diagnosis of SARS-COV-2 by PCR is the gold standard in this regard. Kary Mullis developed PCR in the 1980s [29]. In general terms, PCR is a molecular photocopying method that detects the genetic material of specific organisms such as viruses and bacteria by copying/amplifying. Due to the copying/amplifying feature specified in the definition of PCR, millions of copies of the targeted genome are obtained when the PCR process is completed. The sensitivity and accuracy of this method are quite high. It is also the most reliable way to diagnose COVID-19. There are many different PCR techniques recorded in the literature. The main ones can be listed as Fast-cycling PCR, Quantitative real-time PCR (Q-RT PCR), Multiplex PCR, Real-time PCR, Reverse Transcriptase PCR (RT-PCR), Nested PCR, Touch down PCR, , Long-range PCR, etc[30]. Among these, Reverse Transcriptase PCR (RT-PCR) is the most used type of PCR in the diagnosis of COVID-19. The basis of PCR is based on the synthesis of a new DNA strand that will complete a defined target template strand. DNA polymerases and primers take part in this process. DNA template, DNA polymerase, Primers, Nucleotides, and RT-PCR components are involved in the PCR process. The PCR method is used in areas such as diagnosing genetic disorders, detecting bacteria or viruse and DNA fingerprinting. Large amounts of DNA samples are required for analyzes performed in these areas. Many different types of samples can be collected to perform COVID-19 PCR testing. These sample types can be saliva, anal swabs, urine and stool, tears, conjunctival secretions, respiratory secretions, nasopharyngeal exudate or bronchial aspirate, oropharyngeal specimen, sputum, endotracheal aspirate, bronchoalveolar lavage (BAL) fluid. However,

among these, the most suitable sample type for PCR is the samples collected from the upper respiratory tract (nose and throat) via swab [31].

1.4. SAMPLE COLLECTION AND PCR PREPARATION PROCESS

The first step of the PCR test is the collection of samples from suspicious individuals. These samples are collected from the nasopharyngeal system, usually from the upper respiratory tract. A long, thin rod with a flexible, synthetic fiber tip and plastic shaft, called a swab, is used for sample collection. There are two types of swabs used in sample collection, nasal swabs, and nasopharyngeal swabs. The nasal/nose swab collects the sample just from the tip of the nostrils due to its structure and function. Nasopharyngeal swabs, on the other hand, advance deep into the nasal cavity and provide sample collection [32]. After collecting the appropriate samples, the swabs are placed in sterile tubes with viral transport medium (VTM) and stored at 4ºC. These samples, which are kept in the refrigerator, must be processed for up to 72 hours. To prevent degradation that may occur as a result of long-term storage, it is appropriate to keep the samples at -70ºC and below during this period [33]. Otherwise, viral load loss and degradation may occur in the stored samples. The occurrence of these negative situations may cause inaccurate results. After the tubes of the collected samples are transported from the sample collection areas to the laboratory with appropriate bags and safety equipment, preparations for the PCR process begin. First of all, the samples are arranged in an appropriate order in line with the systems of the hospitals and laboratories in a way that does not confuse them, and they are made suitable for working in sequence. In order not to cause contamination of the samples taken into the study collectively, firstly, the samples are isolated one by one. Each tube is first vortexed with the laboratory equipment vortex. This vortexing process is aimed at taking a homogeneous section from the VTM/swab combination in the tube. Homogeneous liquids taken from each sample are transferred to appropriate small falcon tubes or well plates. This process takes place in specially ventilated biosafety cabinets. According to the recommendations of the Centers for Disease Control and Prevention and the World Health Organization, biosafety cabinets to be used in testing for SARS-COV-2 should have three levels of protection. In this way, expert personnel protection, protection of existing samples and protection of the laboratory environment against the risk of contamination will be ensured [34]. These three protected cabins are referred to as Class II cabins in the literature [35]. These cabinets should be used

in any process that produces aerosols or droplets, such as vortexing, pipetting, and centrifugation, which is also applied in PCR tests. The COVID-19 diagnostic kit, which is used simultaneously, is prepared in biosafety cabinets according to the procedure determined by the manufacturers. Then, the prepared diagnostic kit and isolated patient samples are combined in biosafety cabinets with laboratory pipettes at determined rates. These combined samples are in the well plates used in the PCR device. Each well represents one patient. Usually, 96 well-plates are used. After the combining process is completed, the top of the well-plate is covered with a special adhesive transparent large tape called a sealer and placed in the PCR device. The process is started by selecting the appropriate working protocol of the diagnostic kit from the software program of the device. The PCR process takes an average of 40-60 minutes. The PCR process basically consists of three simple steps. These are denaturation, annealing and extension, respectively [36]. Reverse transcriptase PCR (RT-PCR) method will be used in the diagnosis of SARS-COV-2. In this method, unlike others, fluorescent dyes are used to determine the amount of genetic material. The first molecular process is reverse transcription. This process is to obtain double-stranded DNA (complementary DNA(cDNA)) strands from single-stranded RNA. The double-stranded structure is then separated from each other. This process is called denaturation. Then the primers come into play. The primers are attached to the ends of the separated DNA strands. Primers are designed specifically for viral DNA to bind to the genetic sequence of viral DNA. The primers then bind to the respective regions and new complementary DNA (cDNA) chains are extended along the template strand. During this process, fluorescent dyes adhere to the DNA at the point where there are suitable matches, informing that successful replication has taken place. This phenomenon is called hybridization. At the end of the transaction, while a copy of the viral DNA is formed thanks to the polymerase enzyme, radiation occurs by separating the primer-probe on the one hand. This process is called polymerization. The last two stages, hybridization and polymerization cycles, are repeated 20-30 times. Thus, millions of DNA copies with SARS-COV-2 viral RNA are obtained.

1.5. FACTORS AFFECTING THE EVALUATION OF PCR TEST RESULTS

The PCR method, which has effective results in the diagnosis and detection of SARS-COV-2, is very sensitive. The accuracy, precision, and reliability of the tests performed with this method are directly related to its sensitivity. Many parameters affect the process and results,

from the first stage of PCR tests to the last stage of evaluation and conclusion. The first of these parameters is taking samples suitable for the PCR process. The main element in PCR tests are samples taken from infected individuals. Failure to take appropriate samples from suspicious individuals (inability to enter the oral and nasal cavities properly, epithelial tissue not being able to be obtained) affects the results of PCR tests. Since the appropriate sample cannot be obtained, a sufficient amount of DNA cannot be provided for the process to take place. Another parameter that affects the results after proper sample collection is the sample collection period. Sampling from an infected person before the incubation period of the disease or when the severity and efficacy of the disease decreases affect the results. Results from people who are newly infected and have not completed the incubation period may be negative. Thermal cycler PCR devices are used for SARS-COV-2 PCR tests. These devices have the principle of working by heating and cooling. Technically, the device has lids that allow heating and cooling to be performed. In these devices, technical errors such as lid heating errors and connection errors are usually encountered. Lid errors caused by the device directly affect the results as they prevent proper heating. There are many PCR kits used in the diagnosis of SARS-COV-2 worldwide. The working principle, accuracy, and sensitivity of each kit are different. During the factory production of these kits, problems may occur because they are not stored under appropriate conditions. In addition, the fact that the kits used are not prepared in the environment and conditions determined by the manufacturer for the PCR process is among the parameters that affect the results. Another parameter affecting the results is the problems arising from the expert personnel working in the laboratory. Problems caused by people can be situations such as putting the samples in the wrong well, causing contamination during the procedures, mistakes made in the software parts connected to the device, incorrect preparation of the kit, etc. There may be problems arising from sample collection as well as problems arising from the processes after the collection of samples. The first of these is to keep the collected samples in suitable or unsuitable conditions for a long time (usually 24 hours or more for the PCR process). Waiting for samples for a long time may cause viral load loss and deterioration of sample tubes. These problems directly affect the results. Many laboratory materials are used in the performance of COVID-19 PCR tests. These are clear/white well plates, adjustable single and multiple pipettes, pipette tips compatible with pipettes, and Eppendorf. The compatibility of these materials may differ between devices and kits. For example, the signals generated as a result of the radiation provided by the kits during the operation of the device may differ depending

on the quality of the well plates and whether they are transparent or white. During the isolation of the samples, the preparation of the kits, and the combination of the kits and patients, samples and kits are used in the volumes determined by the companies. Pipettes and pipette tips are used to measure appropriate volumes. Other problems that directly affect the results are that the pipettes are not properly calibrated and that the pipettes do not work in harmony with the pipette tips (incomplete or excessive drafting). All laboratory processes take place in biosafety cabinets. The cleaning of these cabinets, equipment, and laboratory is provided by ultraviolet (UV) lamp. Problems in the airflows, filter, and UV lamp life of the cabinets can cause contamination by directly affecting the preparation of kits, sample isolation, and sample-kit combination processes in the cabinet. Likewise, the lack of cleanliness in the laboratory can cause serious contamination. Contaminations in the laboratory and biosafety cabinets directly affect the results.

1.6. EVALUATION OF PCR TEST RESULT

PCR analyzes are divided into two qualitative and quantitative. The purpose of quantitative analysis is to determine the amount of targeted DNA. The aim of the qualitative analysis is whether the target DNA is present in the collected sample. PCR analysis used in the diagnosis of COVID-19 is quantitative [37]. At the end of the whole PCR process, the interpretation of the results begins as a result of the signals obtained by copying and amplifying the viral genome millions of times. The obtained signals are reflected on the software screen in the form of sigmoidal graphs. The graphics are read and interpreted as specified by the manufacturer of the kit. After the PCR process is completed, the tests are analyzed by the specialist physician or authorized laboratory personnel. There are three definitive results in COVID-19 tests. These are positive, negative, and inappropriate samples. There are 2 analysis channels in the kits. These are called FAM(PCR dye marker) and HEX(PCR dye marker) channels. The FAM channel informs us about the amount of targeted viral genome. The HEX channel is called the internal channel. This channel gives information about whether the samples are taken properly and whether there is enough RNA/DNA in the collected sample to enable the study to take place. The names of these channels may change depending on the design of the kits and the marker dyes used. The most commonly used channels (dyes) are FAM, HEX, ROX, CY5, and CY5-5.

The main elements in the evaluation phase are control groups and patient samples. It is very important to check that the run in a completed PCR test is working properly and according to the protocol. For this first, the positive and negative controls in the 96-well plate are examined. These controls give us information about whether there is a problem in the COVID-19 diagnostic kit prepared and in the work done with this kit. The radiation obtained from the positive control (PC) well is emitted in both channels and sigmoidal graphs are obtained in both channels (FAM, HEX). In the radiation obtained from the negative control(NC) well, sigmoidal curves cannot be obtained in both channels (FAM, HEX). Straight lines are obtained within both channels. Positive and negative controls graphs of the Biospeedy SARS-CoV-2 Double Gene RT-qPCR diagnostic kit are shown in Figure 1.2 and Figure 1.3[38].

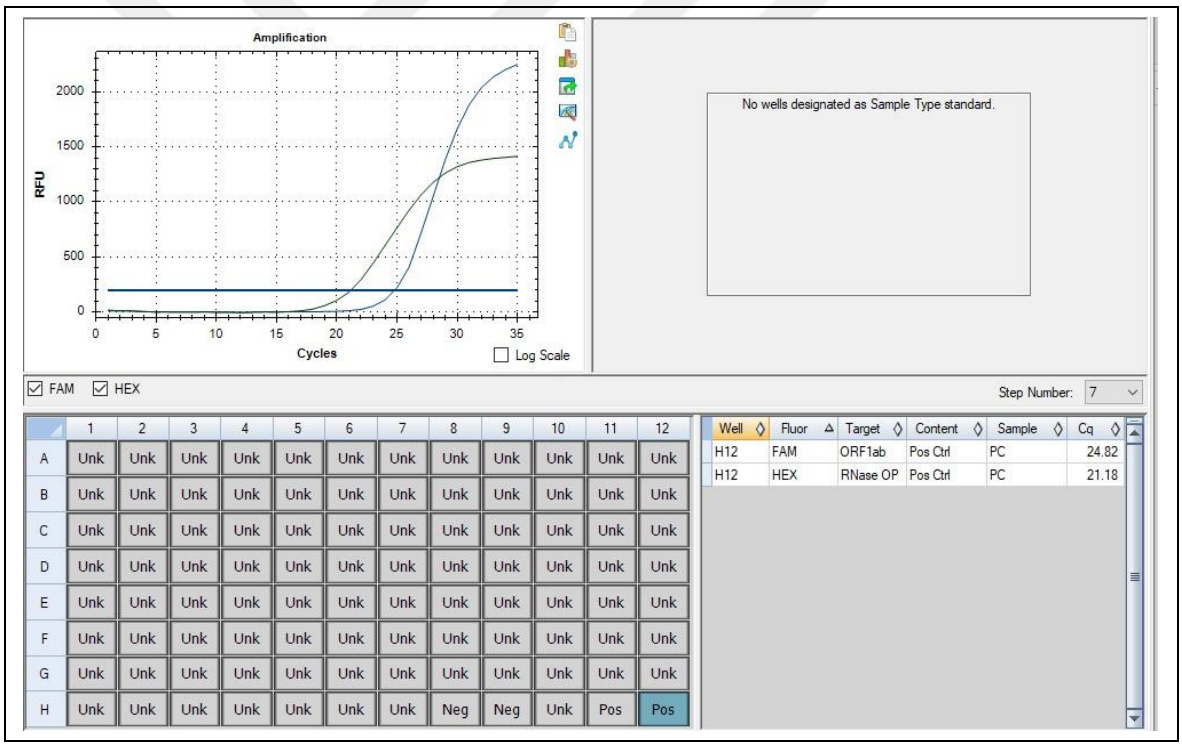

Figure 1.2. Positive controls graph of the Biospeedy SARS-CoV-2 Double Gene RT-qPCR diagnostic kit

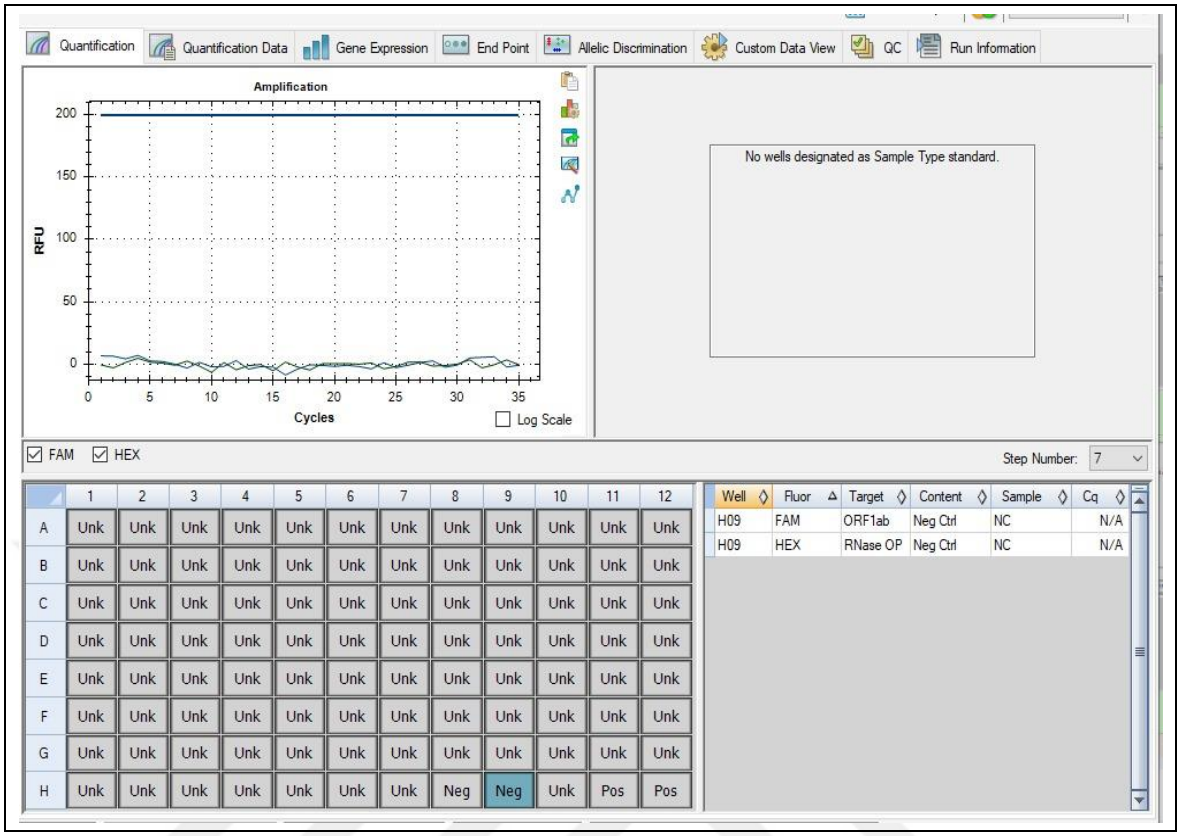

Figure 1.3. Negative control graph of the Biospeedy SARS-CoV-2 Double Gene RT-qPCR diagnostic kit

If there is an opposite condition in each control well (positive and negative), this indicates a problem in the run and the run should be carefully repeated with the caveats noted. After the examination of the positive and negative control wells is completed, the examination of the samples in each well is started. The HEX channels of the samples are checked first. In this way, it is checked whether the appropriate sample is taken or not. Sigmoidal radiations from the HEX channel indicate that the samples are properly taken and working. Then the FAM channels of the samples are examined. Sigmoidal graphs obtained from the FAM channel indicate that the sample is positive. If no sigmoidal graph is obtained from the sample, it indicates that the sample is negative. Samples that do not have any HEX radiation in the examinations and analyzes are reported as an inappropriate samples. Graphs of positive sample according to the Biospeedy SARS-CoV-2 Double Gene RT-qPCR diagnostic kit are shown in Figure 1.4.

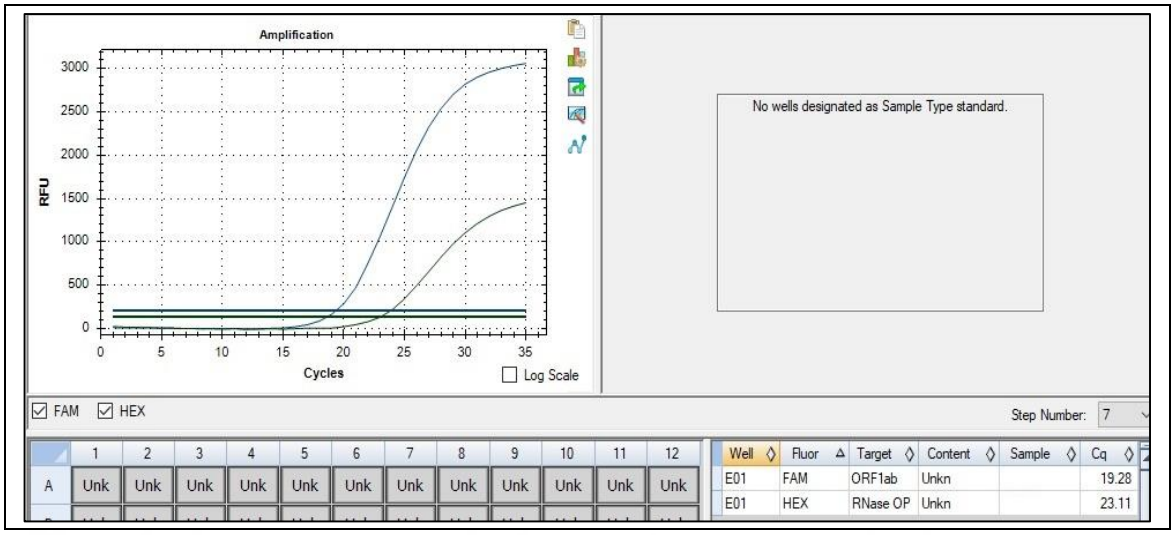

Figure 1.4. Positive sample FAM and HEX graphs according to the Biospeedy SARS-CoV-2 Double Gene RT-qPCR diagnostic kit

To be expressed in detail; If genetic materials found in samples collected from individuals match the primers defined in the COVID-19 diagnostic kit, millions of copies are obtained and the result is interpreted as positive. If the genetic materials of the samples collected from individuals do not match the primers defined in the COVID-19 diagnostic kit, no amplification is obtained and the result is interpreted as negative. In addition to these, there are also false-positive and false-negative results. If a sample does not contain genetic material, it is unlikely that the result will be positive. If such a situation occurs (contamination etc.), these results are called false positives. Likewise, if a sample contains genetic material, the result is unlikely to be negative. If such a situation occurs, these results are called false negatives. Graphs of the false-negative and false-positive sample according to the Biospeedy SARS-CoV-2 Double Gene RT-qPCR diagnostic kit are shown in Figure 1.5, Figure 1.6, Figure 1.7 and Figure 1.8. Figure 1.5 shows the FAM graph of the first test result of a sample with a false positive result, while Figure 1.6 shows the FAM graph of the second and final test result of this sample. While the sample was positive according to Figure 1.5, it was seen to be negative in Figure 1.6. Similarly, Figure 1.7 shows the FAM graph of the first test result of a sample with a false negative result, while Figure 1.8 shows the FAM graph of the second and final test result of this sample. While the sample was negative according to Figure 1.7, it was seen to be positive in Figure 1.8.

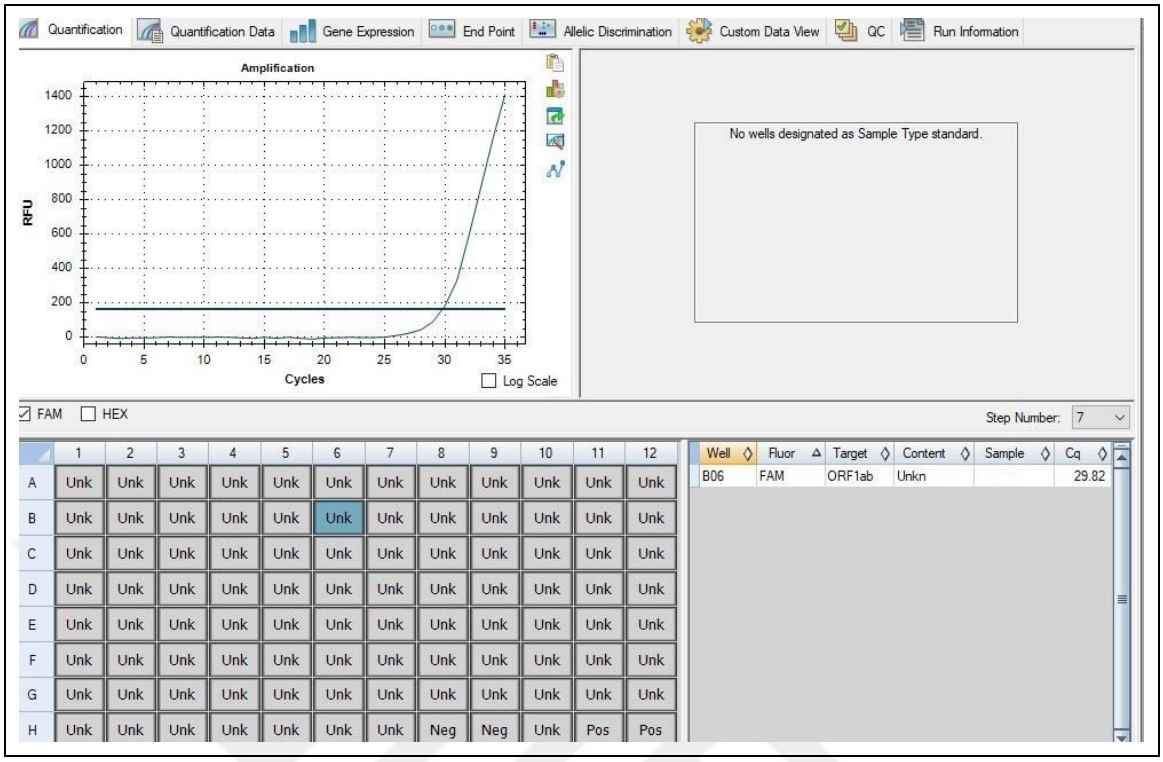

Figure 1.5. The FAM channel graph of the first test result of a false positive sample

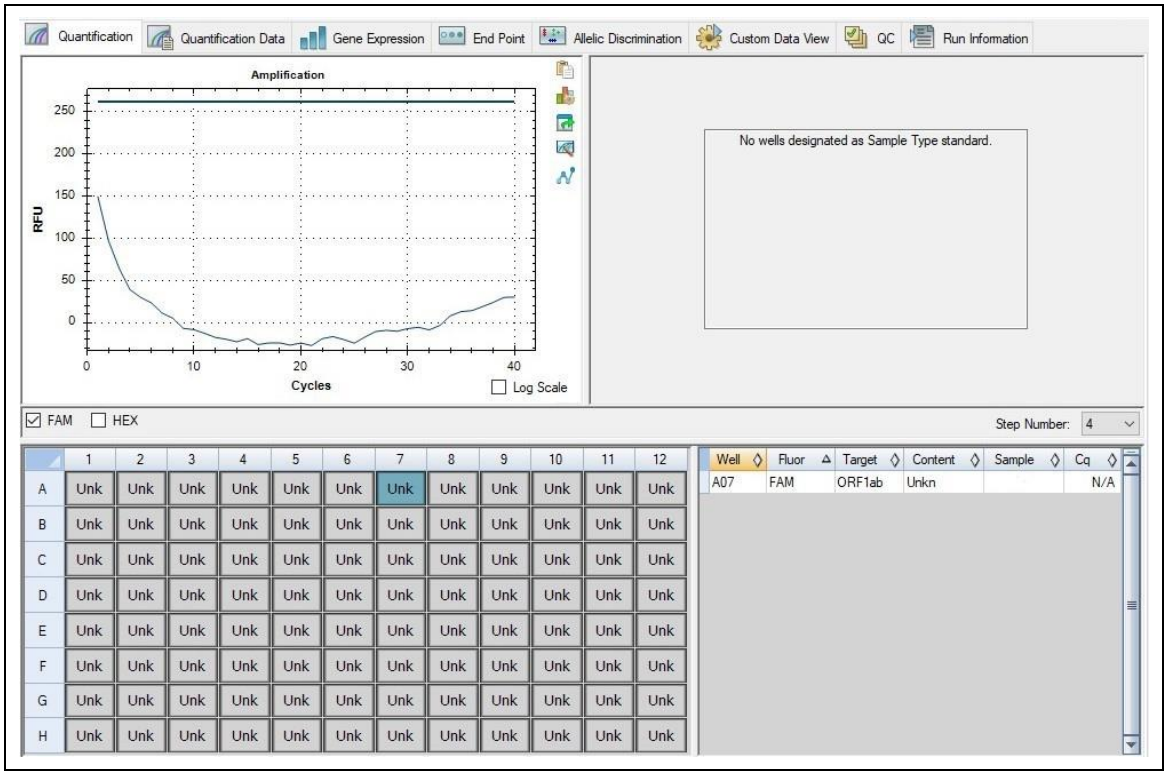

Figure 1.6. The FAM channel graph of the second test result of a false positive sample

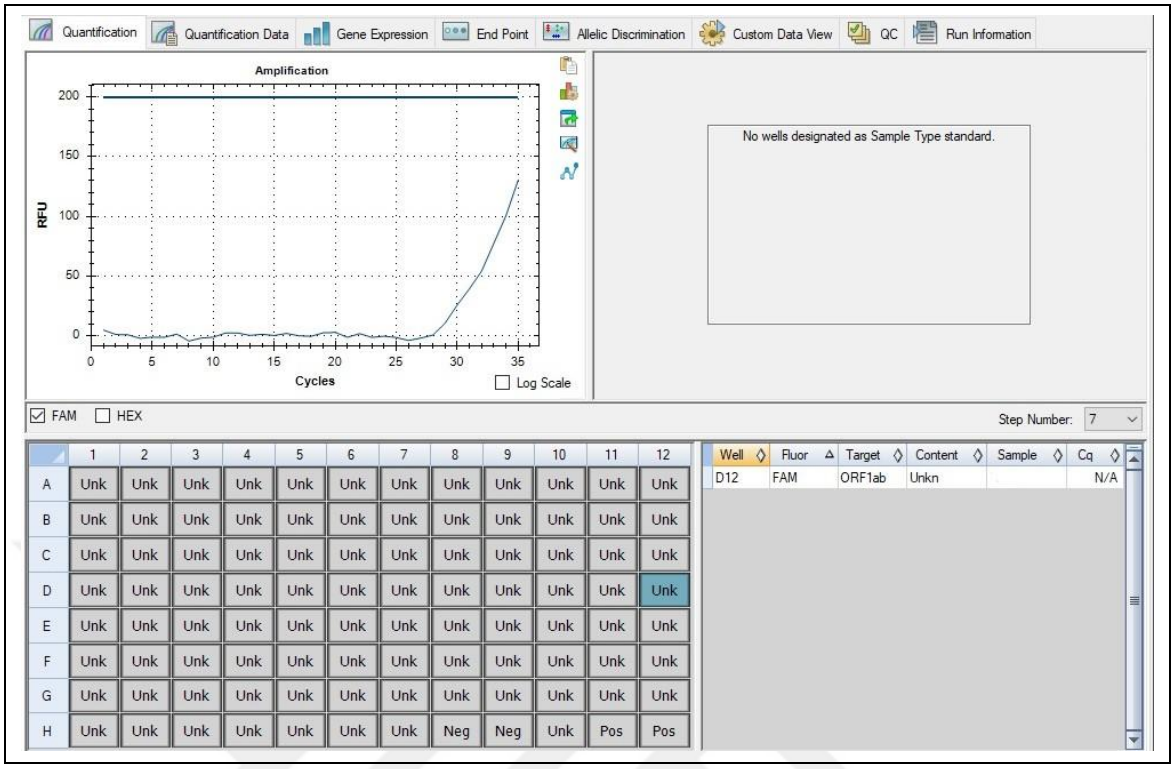

Figure 1.7. The FAM channel graph of the first test result of a false negative sample

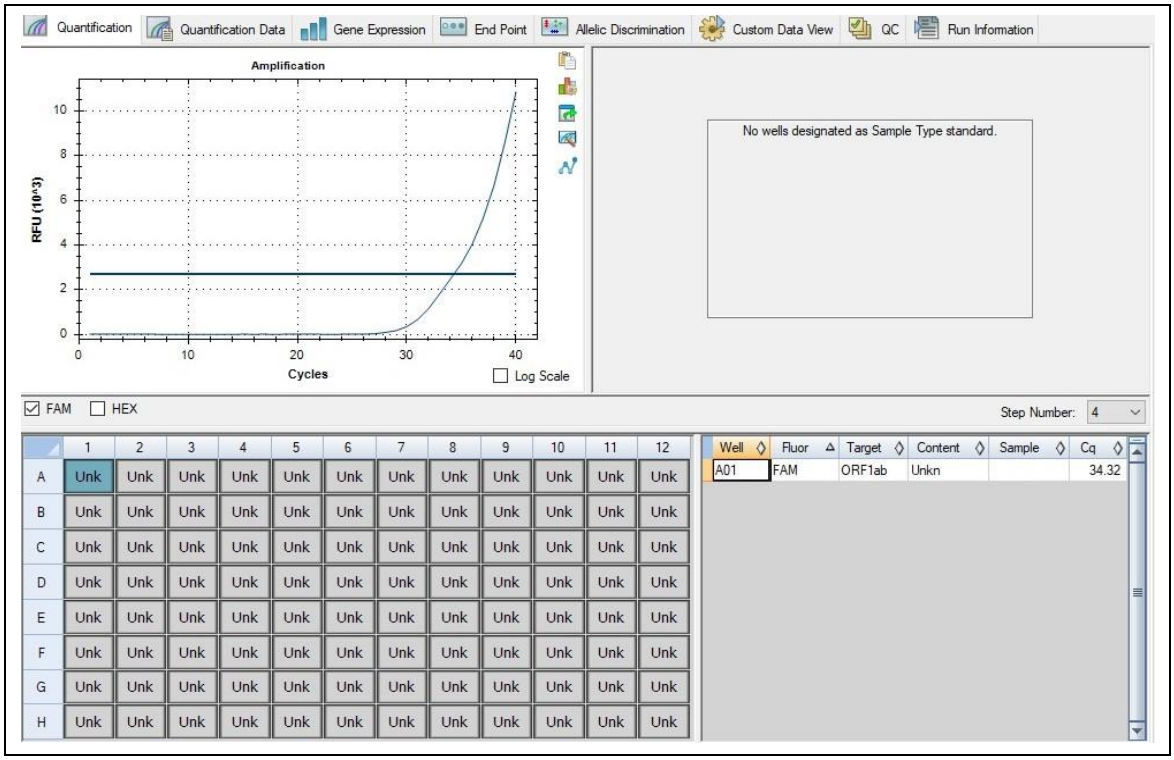

Figure 1.8. The FAM channel graph of the second test result of a false negative sample

A false-negative result is usually decided when patient is infected but does not have sufficient viral load to diagnose the disease. False-negative results are more likely to occur than false positives [39]. The reasons for false-negative results may be improper sample collection, insufficient viral load, the PCR COVID-19 detection kit not working properly, etc[40].

During PCR amplification, signals are acquired simultaneously and plotted along the axes of these signals and amplifying cycles. All kits include cycle threshold (Ct) information. This Ct value is used to report detectable fluorescence during PCR cycles. Within the scope of this information, a low Ct value indicates a higher viral load. The earlier the sample starts to radiate at the Ct value and to plot sigmoidal graphs during the operation of the kit, the higher the viral load will be. According to the general standards determined by the manufacturers, analysis with a Ct value less than 40 in the diagnosis of COVID is considered positive. These Ct values vary according to the manufacturers. Analyzes with Ct values greater than 40 are considered negative. If the Ct value is between 37-40 as a result of the analysis of the samples, the result is to repeat the test [41]. These Ct values vary according to the manufacturers. In some kits, if the Ct value is less than 38, it is considered positive, and greater than 38 is considered negative.

1.7. HYPOTHESIS TESTING

Various methods are used in performing statistical analysis. One of them is hypothesis testing. Hypothesis is expressed as a claim about the value of a population parameter based on sampling. Population parameters, not sample parameters, are used in hypothesis formation. This is an important point of confusion in hypothesis development. Hypothesis testing, on the other hand, consists of a series of systematic procedures based on sampling to decide the population through the sample drawn from the determined population. In this context, hypothesis testing is investigated which of the hypotheses and its opposite (H0 and Ha) are better compatible with the result obtained in the sample. It starts with determining a research hypothesis and determining the alternative hypothesis (H1 and Ha) and null hypothesis (H0) under the concept of statistical hypothesis related to them and determination of the level of significance (α). After determining these two steps, four possibilities emerge. These are Type I error, Type II error, and two right decisions. The significance level (α) is

the maximum probability of error that we accept to process while rejecting the null hypothesis. In other words, it is the probability of rejecting the H0 hypothesis as a result of the test when the H0 hypothesis is true. Because while the H0 hypothesis is true, it is undesirable to reject it. The determined significance level is an indicator of the quality of the hypothesis test. This value is determined by the researcher. In statistical analyzes, the significance level is generally determined as 0.001, 0.01, and 0.05. A null hypothesis is a hypothesis that is generally put forward in a test and is desired to be tested. In other words, the current belief about the population parameter is expressed as the status quo. In the null hypothesis, there is no distinction between the value obtained from the sample and the known value of the population. On the other hand, the alternative hypothesis is defined as the opposite hypothesis formed against the null hypothesis, which contradicts the null hypothesis. In the alternative/opposite hypothesis, there is a distinction between the value obtained from the sample and the known value of the population. There are major errors that can occur during the performance of hypothesis testing. These are Type I and Type II errors[42]. Type I errors are symbolized with α and Type II errors with β. Type I error (α) is the mistaken rejection of the null hypothesis (H0) when it is true. Type II error (β) is the false acceptance of this hypothesis when the null hypothesis (H0) is false. General information about hypothesis testing is given in Table 1.1.

Table 1.1. Hypothesis tests summary evaluation table

	The decision as a result of hypothesis testing		
True Scenario		**H0 is rejected (H1/Ha is not rejected)**	**H0 is not rejected (H1/Ha is rejected)**
	H0 (True)	Type I Error (α) (Significance Level)	(1-α)
	H0 (False)	(1-β)	Type II Error (β) (Power of Tests)

When the situations and conditions are evaluated, 4 results are possible. The first conclusion, H0 is true and is not rejected. The second result, H0 is true and not rejected (Type I). The third result, H0 is false but not rejected (Type II). The fourth result, H0 is false and was rejected. There are major factors affecting the sample size. These are Type I error (α), Power factor (1-β), and effect size(δ). Type I error is the probability of accidentally rejecting a true null hypothesis. (1- α) is defined as the reliability level of the test. Power (1- β) is the probability of rejecting the H0 hypothesis when the H0 hypothesis is false. In other words,

when there is a difference between the groups, in reality, H0 is the difference to be rejected as a result of the test. β is a Type II error. The effect size (δ) indicates how much the estimated value obtained as a result of the study may deviate from the population value. This value is determined by the researcher.

1.7.1. Stages of Hypothesis Testing

There are certain steps involved in successfully performing hypothesis testing. These steps are listed below respectively.

i. Generating Hypotheses: In this step, null and alternative hypotheses are created for the research and analysis to be carried out.
ii. Determination of significance level (α): It is the value that shows the quality of the hypothesis test and expresses the critical region(rejection region). It is determined by the researcher.
iii. Determination of sampling distribution: It is the step of deciding on the statistical tables to be used in the analysis. (Z table, T table, X^2 table, F table)
iv. Determination of the rejection region: The critical value is obtained from the sample distribution according to the table values determined in the previous step, and the rejection region is determined.
v. Calculation and interpretation of test statistics: Calculations are made with the formulations of the table used. By comparing these calculations (test values) with the critical values attached to the tables, the rejection and non-rejection of the hypotheses are interpreted.

1.7.2. Hypothesis Establishment

There are 3 different conditions for the hypothesis formation phase, which is the first step for analysis in statistics. These conditions are formed on the null hypothesis and the alternative hypothesis. These hypotheses are divided into two double-sided and one-sided. The tests are double or single-sided, depending on whether the opposing hypothesis is "different" or "large" or "small" than the null hypothesis. The conditions are listed below :

i. Double- Sided Hypothesis : H0 : $\mu=\mu_0$

H1(Ha): $\mu \neq \mu_0$

ii. Single – Sided Hypothesis: : H0 : $\mu=\mu_0$

H1(Ha): $\mu < \mu_0$

iii. Single – Sided Hypothesis: : H0 : $\mu=\mu_0$

H1(Ha): $\mu > \mu_0$

2. MATERIALS AND METHODS

The data used in this research study were obtained through a comprehensive literature study on the subject. The research was mainly carried out on PCR tests used in the diagnosis of SARS-COV-2, errors related to these tests, comparisons, and specifically false negative and false positive results, and the factors causing these results.

The formulas used in the calculation of the data obtained from the articles as a result of the research will be given under this heading.

2.1. FALSE NEGATIVE

In the literature study on false negative (FN) errors, nineteen articles were used, from which we could access numerical data. According to these articles, the average total number of patients included in the studies was 109,680. Approximately 104,210 of these patients have positive results and 4,692 negative results. 1,025 patients were accepted as false negative. In some of the studies, the calculations were made approximately because the patients were initially included in the study and were not included in the study as a result of the later evaluations, and some data were not given clearly. At this point, the calculations of the data were made carefully and the most accurate and clear data was tried to be obtained. Articles and relevant information about false negatives are given in Table 2.1.

Table 2.1. Data from articles about false negative

FALSE NEGATIVE				
Reference	**Total Patient**	**Positive**	**Negative**	**False Negative**
[43]	1014	601	413	15
[44]	128	-	-	11
[45]	64	58	-	6
[46]	1502	64	1438	6
[47]	102	-	-	12
[48]	87	36	41	6
[49]	82	34	48	7
[50]	568	87	481	19
[51]	51	-	-	15
[52]	80	-	-	39
[53]	1324	1324	-	601

[54]	167	-	-	5
[55]	1890	145	1745	4
[56]	610	226	384	48
[57]	95,919	-	49	5
[58]	70	-	-	15
[59]	301	-	70	21
[60]	21	14	-	7
[61]	5700	5517	-	183
Total	≈ 109,680	≈ 104,210	≈ 4,692	≈ 1,025

2.2. FALSE POSITIVE

In the literature study on false positive (FP) errors, twelve articles were used, from which we could access numerical data. In the articles where we did not reach the total number of patients within the calculations and information, the average total number of patients was tried to be reached, at least over the existing patients, by using the negative/positive patient numbers given.

According to these articles, the average total number of patients included in the studies was 78,521. Approximately 3,485 of these patients have positive results and 75,036 negative results, and 447 patients were accepted as false positives. In some of the studies, the calculations were made approximately because the patients were initially included in the study and were not included in the study as a result of the later evaluations, and some data were not given clearly. At this point, the calculations of the data were made carefully and the most accurate and clear data was tried to be obtained. Articles and relevant information about false positives are given in Table 2.2.

Table 2.2. Data from articles about false positive

FALSE POSITIVE				
Reference	**Total Patient**	**Positive**	**Negative**	**False Positive**
[62]	5110	31	5079	26
[63]	261	45	216	-
[64]	24,717	228	-	20
[65]	141	130	-	11
[66]	1202	4	1198	-
[67]	105	14	91	-
[68]	493	133	360	-

[69]	-	-	10,538	336
[70]	42	39	3	3
[71]	34,383	2861	-	40
Total	≈ 78,521	≈ 3,485	≈ 75,036	≈ 447

2.3. CALCULATION OF FALSE NEGATIVE AND FALSE POSITIVE RATES

In the study, it is necessary to calculate certain values in line with the purpose of the study. Two of these are false negative and false positive rates, which form the basis of our research. In addition to these concepts, there are specificity (true negative rate (TNR)) and sensitivity (true positive rate (TPR)) values. In our study, false negative and false positive rates were examined. It is given as information in the calculation of other concepts.

Equation (2.1) and Equation (2.2) were used to calculate the false negative rate (FNR) and the false positive rate (FPR), respectively.

$$False\ Negative\ Rate(FNR) = \frac{False\ Negative\ (FN)}{Positive(P)} = \frac{False\ Negative(FN)}{False\ Negative + True\ Positive(TP)} \quad (2.1)$$

$$False\ Positive\ Rate(FPR) = \frac{False\ Positives(FP)}{Negatives(N)} = \frac{False\ Positive(FP)}{False\ Positive(FP) + True\ Negative(TN)} \quad (2.2)$$

Equation (2.3) and Equation (2.4) are used to calculate sensitivity (true positive rate (TPR)).

$$True\ Positive\ Rate(TPR) = \frac{True\ Positive(TP)}{Positives(P)} = \frac{True\ Positive(TP)}{True\ Positive(TP) + False\ Negatives\ (FN)} \quad (2.3)$$

$$Sensitivity = 1 - Type\ II\ Error(\square) = 1 - FNR \quad (2.4)$$

Equation (2.5) and Equation (2.6) are used to calculate specificity (true negative rate (TNR)) values.

$$True\ Negative\ Rate(TNR) = \frac{True\ Negatives(TN)}{Negatives\ (N)} = \frac{True\ Negatives(TN)}{True\ Negatives(TN) + False\ Positives(FP)} \quad (2.5)$$

$$Specificity = 1 - Type\ I\ Error(\alpha) = 1 - FPR \quad (2.6)$$

3. RESULT

There are nineteen false negative and twelve false positive articles obtained through the literature review. Each of these articles is independent from the other. When 31 articles are evaluated within a single study, the total number of patients is approximately 188,201. Of these, approximately 107,695 were positive, 79,728 were negative, and 1,472 were discordant (false positive and false negative). The positivity rate calculated within these data was found to be 57.2% and is shown in Table 3.1.

Table 3.1. Approximate total data of thirty-one articles

Total	Positive	Negative	Discordant	Positivity rate
188,201	107,695	79,728	1472	0.5722

In our study conducted through the literature review, 447(30,37%) of 1,472 discordant results were recorded as false positives and 1,025(69,63%) as false negatives, and are shown in Table 3.2.

Table 3.2. Data belongs to false positive and false negative results

Discordant	False Positive	False Negative	True Positive	True Negative
1,472	447	1,025	106,670	79,281

As a result of the calculations made using Equation (2.1) and Equation (2.2), false negative rate and false positive rate values were calculated, respectively. Accordingly, false negative rate values of the articles scanned from the literature are given in Table 3.3, and false positive rate values are given in Table 3.4.

Table 3.3. Data used in calculations of false negative rate

FALSE NEGATIVE					
Reference	**Total Patient**	**Positive**	**Negative**	**False Negative**	**False Negative Rate**
[43]	1014	601	413	15	2.5%
[44]	128	-	-	11	8.6%-28.1%
[45]	64	58	-	6	9.4%
[46]	1502	64	1438	6	9.4%
[47]	102	-	-	12	11.8%
[48]	87	36	41	6	16.7%
[49]	82	34	48	7	20.6%
[50]	568	87	481	19	21.8%
[51]	51	-	-	15	29.4%
[52]	80	-	-	39	48.8%
[53]	1324	1324	-	601	52.2%
[54]	167	-	-	5	3%
[55]	1890	145	1745	4	3.5%
[56]	610	226	384	48	17.52%
[57]	95,919	-	49	5	9.3%
[58]	70	-	-	15	21.4%
[59]	301	-	70	21	30%
[60]	21	14	-	7	3.4%
[61]	5700	5517	-	183	3.2%
Total	≈ 109,680	≈ 104,210	≈ 4,692	≈ 1,025	0.98%

Table 3.4. Data used in calculations of false positive rate

FALSE POSITIVE					
Reference	**Total Patient**	**Positive**	**Negative**	**False Positive**	**False Positive Rate**
[62]	5110	31	5079	26	0.5%
[63]					0.8%-4.0%
[64]	261	45	216	-	20.8%
[65]	24,717	228	-	20	0.3%-6.9%
[66]	141	130	-	11	7%
[67]	1202	4	1198	-	0.3%
[68]	105	14	91	-	15.4%
[69]	493	133	360	-	36.9%
[70]	-	-	10,538	336	3.2%
[71]	42	39	3	3	7.1%

[62]	34,383	2861	-	40	0.1%
[63]	-	-	1,529	11	0.7%
Total	≈ 78,521	≈ 3,485	≈ 75,036	≈ 447	0.57%

The false negative rate (FNR) calculated according to the table of only false negative studies (Table 3.3) was found to be 0.98%. Likewise, the false positive rate calculated according to the table of only false positive studies (Table 3.4) was 0.57%.

Calculations were made for the general study by using the data in Table 3.1 and Table 3.2 belonging to a total study of 31 articles. As a result of this study, the false positive Type I error (α) value of the study was found to be 0.56% in line with the analyzed data. Based on this value, the specificity value calculated in Equation (2.3) was confirmed with Equation(2.4) and found to be 0.9944 (99.44%). Similarly, the false-negative Type II error (β) value of the study was found to be 0.95%. Based on this value, the sensitivity value calculated in Equation (2.5) was confirmed by the Equation (2.6) formula and was found to be 0.9905 (99.05%). These data are listed in Table 3.5.

Table 3.5. Data obtained as a result of calculations

	Positive (+)	**Negative (-)**
Positive (+)	Sensitivity (1-α) 0.9905 (99.05%)	Type I Error (α) 0.0056 (0.56%)
Negative (-)	Type II Error (β) 0,0095 (0.95)	Specificity 0.9944 (99.44%)

4. DISCUSSION

The research was carried out with a large sample size of thirty-one articles. The reasons for retests or false negative/false positive results may differ according to each article. Discordant results(false positive/false negative) were found in 0.78% of the tested patients. The FNR and sensitivity values of the study were 0.95% and 99.05%, respectively, and the FPR and specificity values were 0.56% and 99.44%, respectively. In studies conducted during the pandemic, it has been stated that the false-negative rate can vary between approximately 2% and 29% [72]. In addition to this statement, the FNR found in similar studies were found to be 9,3% [57], 13% [73], 21,4% [58], 30% [74], 38% [75] and 3,45% [55]. Likewise, for false-positive rate estimations, the false-positive rate (FPR) was expressed as 0,8%-4,0% according to the generalizations made in the United States [76]. In another study, this rate was limited to 0.2%-0.9% [77]. In different studies, this rate was reported as 0.3% [64] and 0.5% [62]. The 1025 false-negative results are thought to be due to sample enrollment without adequate waiting time, sampling site, insufficient viral load of the sample due to poor quality or empty swab, improper handling and transportation of specimens, poor RNA isolation, and inactivation [78,79]. It is thought that the results of 447 samples, which were evaluated as false positive, are caused by splashing between the wells in the cabinet, contamination from the equipment used, contamination that may occur during sample collection, contamination that may occur due to the carelessness of the experts working in the laboratory, sample entanglement, materials incompatible with the kits (transparent/white well plate), software problems and cross-reaction [76,69,80]. False positive results may lead to potential consequences such as delaying treatments unnecessarily, exposure to infection in case of hospitalization for treatment, suspension of current social life due to quarantine, not making planned travels, fear of infecting others psychologically due to the current result [81]. This study have some limitations. Since the study consisted of a compilation of 31 articles and the numbers to be used in the calculations were not the same in all the articles, the total number of patients, the number of positive patients and the number of negative patients were calculated approximately. Using these approximate values, false negative, false positive, specificity and sensitivity values of the study were calculated. These approximate calculations are the reason why some values do not fall within the ranges specified in other literature studies.

5. CONCLUSIONS

In this study, false negative (FN) and false positive (FP) PCR test results in the literature were evaluated and the reasons causing these results were investigated. Although it differs according to the articles, false positive and false negative results were obtained by resampling from symptomatic and asymptomatic individuals after the first sample collection or by performing control studies. Taking samples from people incorrectly, not storing the samples in a suitable environment and for a sufficient time, and taking samples in early stages of the disease are among the causes of false negative results. On the other hand, expert laboratory personnel, laboratory consumables, contamination during sampling, pipetting errors, spatter between sleepers, cross-reactions, and the use of inappropriate materials have led to false positive results.

This study will be supportive for future studies by providing information about the errors that may occur The Type I Error (False Positive Rate) and Type II Error (False Negative Rate) values calculated with the approximate data collected from the literature will provide great benefits in the evaluation and analysis of the error margins of the diagnostic kits that have been developed and will be developed.

REFERENCES

1. Asrani P, Eapen SE, Chia C, Haug G, Weber HC, Hassan I, Sohal SS. Diagnostic approaches in COVID-19: clinical updates. *Expert Review of Respiratory Medicine*. 2020;15(2):197-212.
2. Koonin E, Senkevich T, Dolja V. The ancient Virus World and evaluation of cells. *Biology Direct*. 2006; 1(29).
3. Yin J, Redovich J. Kinetic Modeling of Virus Growth in Cells. *American Society of Microbiology*. 2018;82(2).
4. Breitbart M, Rohwer F. Here a virus, there a virus,everywhere the same virus?. *Trends in Microbiology*. 2005;13(6):278-84.
5. Caspar D L, Klug A. Physical principles in the construction of regular viruses. *Cold Spring Harb Symp Quant Biol.* 1962;27:1-24.
6. Crick F H, Watson J D. Structure of small viruses. *Nature*. 1956; 177(4506):473-5.
7. Rybicki E. The classification of organisms at the edge of life, or problems with virus systematics. *South African Journal of Science*. 1990;86.
8. Koonin E, Starokadomskyy P. Are viruses alive? The replicator paradigm sheds decisive light on an old but misguided question. *Studies in History and Philosophy of Biological and Biomedical Sciences*. 2016;59:125-134.
9. Nguyen M, Haenni A, Expression strategies of ambisense viruses. *Virus Research*. 2003;93(2):141-50.
10. Gelderblom H R. Structure and Classification of Viruses. *Medical Microbiology*.1996.
11. Uzunoğulları N, Gümüş M. VİRÜS TAKSONOMİSİNİN TARİHSEL GELİŞİMİ VE SON DURUMU. Bahçe. 2017;46(2):51-58.
12. ICTV Code; [cited 2022 9 October]. Available from: https://ictv.global/about/code.
13. Baltimore Classification of Viruses ; [cited 2022 16 October]. Available from: http://www.web-books.com/MoBio/Free/Ch1E2.htm.

14. Sinsheimer R. A single-stranded deoxyribonucleic acid from bacteriophage φX174†. Journal of Molecular Biology. 1959;1(1):43-53.

15. Tyler K L. *Segmented Double-Stranded RNA Viruses: Structure and Molecular Biology*. Caister Academic Press; 2008.

16. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in china: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* 2020; 323(13), 1239–1242.

17. Prasad N, Gopalakrishnan N, Sahay M, Gupta A, Agarwal SK. Epidemiology, Genomic Structure, the Molecular Mechanism of Injury, Diagnosis, and ClinicalManifestations of Coronavirus Infection: An Overview. *Indian J Nephrol.* 2020; 30(3):143-154.

18. Eurosurveillance Editorial Team. Note from the editors: World Health Organization declares novel coronavirus (2019-nCoV) sixth public health emergency of international concern. *Eurosurveillance*.2020; 25(5).

19. WHO Coronavirus (COVID-19) Dashboard. [cited 2022 16 October]. Available from: https://covid19.who.int/.

20. World Health Organization. Coronavirus disease 2019 (COVID-19). Situation report – 51. Available from: https://www.who.int/docs/default- source/coronaviruse/ situation-reports/20200311-sitrep-51-covid-19. pdf?sfvrsn=1ba62e57_10 (2020).

21. Touma M. COVID-19: molecular diagnostics overview. *Journal of Molecular Medicine.* 2020; 98:947-954.

22. Hu B, Guo H, Zhou P, Li Shi Z. Characteristics of SARS- CoV-2 and COVID-19. *Nature Reviews Microbiology*. 2020; 19:141-154.

23. Udristioiu A. Molecular mechanisms and methods used to detect COVID-19 infection. 2020.

24. SARS-CoV-2 Mechanism of Cell Entry and Replication. [cited 2022 10 October]. Available from: https://www.tocris.com/research-area/covid-19-research.

25. Flower TG, Buffalo CZ, Hooy RM, Hurley JH. Structure of SARS-CoV-2 ORF8, a rapidly evolving immune evasion protein. *PNAS.* 2020; 118(2).

26. Zhou P, Lou Yang X, Li Shi Z. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020; 579, 270–273.
27. Zhao W, Song S, Chen M. The 2019 novel coronavirus resource. *Yi Chuan.* 2020; 42, 212–221.
28. Li X, Geng M, Peng Y, Meng L, Lu S. Molecular immune pathogenesis and diagnosis of COVID-19. *Journal of Pharmaceutical Analysis.* 2020; 10: 102-108.
29. Rahman MT, Uddin MS, Sultana R, Moue A, Setu M. Polymerase Chain Reaction (PCR): A Short Review. *Anwer Khan Modern medical College Journal.* 2013; 4(1): 30-36.
30. Rajalakshmi S. DIFFERENT TYPES OF PCR TECHNIQUES AND ITS APPLICATIONS. *International Journal of Pharmaceutical, Chemical and Biological Sciences (IJPCBS)*. 2017;7(3):285-292.
31. COVID-19 (SARS-CoV-2 INFECTION) GUIDE; [cited 2022 10 October]. Available from: https://hsgm.saglik.gov.tr/depo/birimler/goc_sagligi/covid19/rehber/COVID-19_Rehberi20200414_eng_v4_002_14.05.2020.pdf.
32. Cleveland Clinic's homepage 2021[cited 2021 24 August]. Available from: https://my.clevelandclinic.org/health/diagnostics/21462-covid-19-and-pcr-testing.
33. Kilic T, Weissleder R, Lee H. Molecular and Immunological Diagnostic Tests of COVID-19: Current Status and Challenges. *iScience.* 2020. 23(8).
34. Depalma A. Product Focus: Biological Safety Cabinets. *Lab Manager*. 2014;9(6).
35. Interim Laboratory Biosafety Guidelines for Handling and Processing Specimens Associated with Coronavirus Disease 2019 (COVID-19); [cited 2022 15 October]. Available from: https://www.cdc.gov/coronavirus/2019-ncov/lab/lab-biosafety-guidelines.html.
36. Delidow B C, Lynch J P, Peluse J J, White B A. Polymerase chain reaction : basic protocols. *Methods in Molecular Biology*. 1993;(15):1-29.
37. Oliveira BA, Oliveira LC, Sabino EC, Okay TS. SARS-CoV-2 andthe COVID-19 disease: a mini review on diagnostic methods. *RevInst Med Trop Sao Paulo*. 2020;62:e44.
38. Bioeksen AR-GE Teknolojileri. SARS-CoV-2 Double Gene RT-qPCR Kit. 2021.
39. Ozilgen M. Handbook of Food Process Modeling and Statistical Quality Control. *CRC Press*. 2011: 59-136.

40. National Human Genome Research Institute's home page 2022 [cited 2012 18 January]. Available from: https://www.genome.gov/about-genomics/fact-sheets/Understanding-COVID-19- PCR Testing.
41. Bohn MK, Lippi G, Horvath A. Molecular, serological, and biochemical diagnosis and monitoring of COVID-19: IFCC taskforce evaluation of the latest evidence. *Clin Chem Lab Med.*2020. 58(7):1037-1052.
42. Afzal A. Molecular diagnostic technologies for COVID-19: Limitations and challenges. *Journal of Advances Research.*2020.26:149-159.
43. Ai T, Yang Z, Hou H. Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology*. 2020;296(2):E32-E40.
44. Luo L, Liu D,Liao X. Modes of contact and risk of transmission in COVID-19 among close contacts. *MedRxiv*. 2020.
45. Wong HY, Lam HY, Fong AH, et al. Frequency and distribution of chest radiographic findings in COVID-19 positive patients. *Radiol*. 2020;296(2):E72-E78.
46. Kim H, Hong H. Yoon S, Diagnostic Performance of CT and Reverse Transcriptase Polymerase Chain Reaction for Coronavirus Disease 2019: A Meta-Analysis. *Radiology*. 2020;296(3):E145-E155.
47. Bernheim A, Mei X, Huang M. Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection. *Radiology*. 2020;295(3):200463.
48. Long C, Xu H, Shen Q. Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT?. *European Journal of Radiology*. 2020;126:108961.
49. He J, Luo L, Luo Z. Diagnostic performance between CT and initial real-time RT-PCR for clinically suspected 2019 coronavirus disease (COVID-19) patients outside Wuhan, China. *Respiratory Medicine*. 2020;168:105980.
50. Ren X, Liu Y, Chen H. Application and optimization of RT-PCR in diagnosis of SARS-CoV-2 infection. *medRxiv*. 2020.
51. Fang Y, Zhang H, Xie J. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. Radiology. 2020;296(2):E115-E117.
52. Wu J, Liu J, Zhao X. Clinical characteristics of imported cases of Coronavirus Disease 2019 (COVID- 19) in Jiangsu Province: a multicenter descriptive study. *Clin Infect Dis*. 2020;71(15):706-712.

53. Xu H, Yan L, Qiu C. Analysis and Prediction of COVID-19 Patients' False Negative Results for SARS-CoV-2 Detection with Pharyngeal Swab Specimen: A Retrospective Study. medRxiv. 2020.
54. Xie X, Zhong Z, Zhao W. Chest CT for Typical Coronavirus Disease 2019 (COVID-19) Pneumonia: Relationship to Negative RT-PCR Testing. *Radiology*. 2020;296(2):E41-E45.
55. Takashi H, Ichinose N, Okada Y. False-negative rate of SARS-CoV-2 RT-PCR tests and its relationship to test timing and illness severity: A case series. *IDCases.* 2022; 28:e01496.
56. Li Y, Yao L, Li J. Stability issues of RT-PCR testing of SARS-CoV-2 for Hospitalized Patients Clinically Diagnosed With COVID-19. *Journal Of Medical Virology*.2020;92(7):903-908.
57. Kanji JN, Zelyas N, MacDonald C, Pabbaraju K. False-negative rate of COVID-19 PCR testing: a discordant testing analysis. *Virology Journal.* 2021. 18:13.
58. Xiao AT, Tong YX, Zhang S. False-negative of RT-PCR and prolonged nucleic acid conversion in COVID-19: Rather than recurrence. *J Med Virol.* 2020. 92(10) 1755-1756.
59. Xiao A, Tong Y, Gao C. Dynamic profile of RT-PCR findings from 301 COVID-19 patients in Wuhan, China: A descriptive study. *J Clin Virol*. 2020;127:104346.
60. Chen D, Jiang X, Hong Y. Can Chest CT Features Distinguish Patients With Negative From Those With Positive Initial RT-PCR Results for Coronavirus Disease (COVID-19)?. *AJR Am J Roentgenol*. 2021;216(1):66-70.
61. Richardson S, Hirsch J, Narasimhan M. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *JAMA.* 2020;323(20):2052-2059.
62. Healy B, Khan A, Metezai H, Blyth I, Asad H. The impact of false-positive COVID-19 results in an area of low prevalence. *Clin Med (Lond).* 2021. 21(1): e54-e56
63. Supplementary Materials for False positives in reverse transcription PCR testing for SARS-CoV-2. [cited 2020 6 October]. Available from: https://www.researchgate.net/publication/342397948_Supplementary_Materials_for_False_positives_in_reverse_transcription_PCR_testing_for_SARS-CoV-2.
64. Layfield LJ, Camp S, Bowers K, Miller DC. SARS-CoV-2 detection by reverse transcriptase polymerase chain reaction testing: Analysis of false-positive results and

recommendations for quality control measures. *Pathology-Research and Practice.* 2021. 225.

65. Wang P. Combination of serological total antibody and RT-PCR test for detection of SARS-COV-2 infections. *Journal of Virological Methods*. 2020;283:113919.
66. Corman V, Landt O, Kaiser M. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill*.2020;25(3):2000045.
67. Etievant S, Bal A, Escuret V. Performance Assessment of SARS-CoV-2 PCR Assays Developed by WHO Referral Laboratories. *J Clin Med*. 2020;9(6):1871.
68. Barra G, Rita T, Mesquita P. Analytical sensibility and specificity of two RT-qPCR protocols for SARS-CoV-2 detection performed in an automated workflow. *Genes(Basel)*.2020;11(10):1183.
69. Cohen A, Kessel B, Milgroom M. Diagnosing SARS-CoV-2 infection: the danger of over-reliance on positive test results. *MedRxiv*. 2020.
70. Katz A, Civantos F, Sargi Z. False-positive reverse transcriptase polymerase chain reaction screening for SARS-CoV-2 in the setting of urgent head and neck surgery and otolaryngologic emergencies during the pandemic: Clinical implications. *Head Neck*. 2020;42(7):1621-1628.
71. Lin L, Carlquist J, Sinclair W. Experience With False-Positive Test Results on the TaqPath Real-Time Reverse Transcription-Polymerase Chain Reaction Coronavirus Disease 2019 (COVID-19) Testing Platform. *Arch Pathol Lab Med*. 2021;145(3):259-261.
72. Woloshin S, Patel N, Kesselheim A. False Negative Tests for SARS-CoV-2 Infection — Challenges and Implications. *The New England Journal of Medicine*.2020;383:E38.
73. Arevalo-Rodriguez I, Buitrago-Garcia D, Simancas-Racines D. False-negative results of initial RT-PCR assays for COVID-19: A systematic review. *PLOS ONE*. 2020. 15(12).
74. Yang H, Wang C, Poon L C. Novel coronavirus infection and pregnancy. *Ultrasound Obstet Gynecol*.2020;55(4):434-437.
75. Kucirka L, Lauer S, Laeyendecker O. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction–Based SARS-CoV-2 Tests by Time Since Exposure. *Ann Intern Med*. 2020;173(4):262-267.
76. Surkova E, Nikolasyevskyy V, Drobniewski F. False-positive COVID-19 results: hidden problems and costs. *Lancet Respir Med.* 2020. 8(12):1167-1168.

77. Cohen AN, ICD10 Monitor's homepage 2020 [cited 2020 9 November]. Available from: https://icd10monitor.com/false-positives-in-pcr-tests-for-covid-19/.
78. Fevraleva I, Glinshchikoca O, Makarik T. How to Avoid False-Negative and False-Positive COVID-19 PCR Testing. *Int. J. Transl. Med*. 2022;2(2):204-209.
79. Çubukçu H, Coşkun A. False-negative results and tolerance limits of SARS-CoV-2 laboratory tests. *Pathog Glob Health*. 2021;115(3): 137-138.
80. Impact of false-positives and false-negatives in the UK's COVID-19 RT-PCR testing programme; [cited 2011 1 October]. Available from: https://www.gov.uk/government/publications/gos-impact-of-false-positives-and-negatives-3-june-2020/impact-of-false-positives-and-false-negatives-in-the-uks-covid-19-rt-pcr-testing-programme-3-june-2020.
81. Levinson S. False Positive Results in Real-time Reverse Transcription-Polymerase Chain Reaction (rRT-PCR) for SARS-CoV-2?. The American Association for Clinical Chemistry (AACC). 2022.