MEDICAL UNIVERSITY
"PROF. DR. PARASKEV STOYANOV" - VARNA
FACULTY OF MEDICINE
DEPARTMENT OF PAEDIATRICS

MARGARITA IVANOVA NIKOLOVA, MD

GENOTYPE - PHENOTYPE CORRELATION IN PATIENTS WITH CYSTIC FIBROSIS

THESIS SUMMARY

Scientific specialty: **"Paediatrics"**

SCIENTIFIC SUPERVISORS:
PROF. DR. MIGLENA GEORGIEVA, DM, PhD
PROF. DR GERGANA PETROVA, DM, PhD

VARNA, 2024

The dissertation is presented in 137 pages and contains 16 tables, 32 figures and 2 annexes. 285 references are cited, of which 16 in Cyrillic and 268 in Latin. Publications related to the dissertation are 5.

The dissertation was discussed and directed for public defense at a meeting of the Departmental Council at the Department of Pediatrics, Medical University - Varna.

Scientific Jury

Prof. Dr. Dimitrichka Duceva Bliznakova, MD, PhD
Prof. Dr. Miroslava Nikolova Bosheva, MD, PhD
Assoc. Prof. Dr. Sirma Mileva Dimitrova, MD, PhD
Assoc. Dr. Ivan Ventsislavov Yankov, MD, PhD
Prof. Prof. Ruzha Zlatanova Pancheva - Dimitrova, MD, PhD

Additional Members
Assoc. Yordanka Georgieva Uzunova, MD, PhD
Assoc. Prof. Dr. Krasimira Ivanova Koleva, MD, PhD

The materials of the dissertacion are available in the scientific department of Medical University - Varna.

The public defense will take place on 11.06.2024. in the Department of Pediatrics, University Hospital "St.Marina, Varna, Hall 501 and in the Webex platform.

CONTENTS:

ABBREVIATIONS

BE - bronchiectasis

BOS - bronchobstructive syndrome

GIT - gastrointestinal tract

RF - respiratory failure

ECFPR - European Cystic Fibrosis Patient Registry

IRT - immunoreactive trypsin

CF - cystic fibrosis

MI - meconium ileus

NS - neonatal screening

USA - United States of America

CVI - cardiovascular accident

FEO1 - forced expiratory volume in 1 second

BMI - body mass index

CBAVD - Congenital bilateral absence of the vas deferens

CFLD - cystic fibrosis-associated liver disease

CFRD - cystic fibrosis-related diabetes

CFTR - cystic fibrosis transmembrane conductance regulator

DIOS - distal intestinal obstruction syndrome

PA - Pseudomonas Aeruginosa

SA - Staphylococcus Aureus

VUS - variant of uncertain significance

I. INTRODUCTION

Cystic fibrosis is an inherited autosomal recessive disease belonging to the group of rare diseases. It is caused by a defect in a gene encoding a protein - CFTR (cystic fibrosis transmembrane conductance regulator). CRTR is a chloride channel that regulates primarily water and electrolyte transport. Its dysfunction leads to the accumulation of viscous secretions in all ductal structures, and the narrowing and obstruction of their lumen damages the corresponding organ and maintains a process of chronic infectivity.

According to the type of defect in the CFTR, the mutations were divided into six classes, with a seventh class being later defined (Bell et al.). To date, the number of described mutations exceeds 2000 with more than 300 mutations associated with disease.

The phenotypic manifestation of CF ranges from involvement of a single system to multisystem and polyorgan involvement. The most significant changes are seen in the airways, gastrointestinal tract (GIT), sweat glands, and urogenital system.

Genotype-phenotype correlations in CF patients have been the subject of a number of studies and are relevant not only for the diagnosis but also for the prognosis of the disease. Phenotypic manifestations depend on the percentage of normally functioning CFTR and range from "classic" CF with exocrine pancreatic insufficiency, pulmonary involvement, high sweat test levels, and obstructive azoospermia, to involvement of a single system, e.g., obstructive azoospermia due to congenital bilateral absence of vas deferens (CBAVD).

Class I and II mutations generally predispose to a more severe clinical picture, with genotype-phenotype correlations being more demonstrative of symptoms on the GIT side. Despite the vast number of studies, a definitive relationship between genotype and pulmonary phenotype cannot be established. The severity of pulmonary involvement is a determinant of morbidity and mortality in CF patients. The variability of phenotypic manifestations is determined by the influence of genetic factors, environment, and adherence to therapy and is difficult to predict.

PART II. BACKGROUND, AIM AND TASKS

Modern treatment and care of CF patients by multidisciplinary teams largely depend on their diagnosis and timely referral to a specialized center. For countries without an established mass NS, such as Bulgaria, phenotypic manifestations are the main indicator to search for the underlying genotype and confirm the diagnosis. Furthermore, genotype-phenotype correlations are important predictors of disease course and timely prevention, diagnosis and treatment of complications. This is especially true for childhood CF patients because parents are primarily interested in their child's prognosis and what to do to improve the chances of a longer and more fulfilling life. This gave rise to the interest in the present work and defined its aim and objectives based on the following premises:

PREMISES:

1. Lack of developments on the topic in the territory of Northeast Bulgaria.
2. Lack of opportunity for early diagnosis due to lack of mass NS.
3. Concentration of multidisciplinary teams caring for CF patients in several regional cities of Bulgaria.
4. Importance of early diagnosis for timely initiation of appropriate therapy.
5. Importance of genotype for prognosis and timely prophylaxis and treatment.
6. Availability of a suitable group for the study - patients with CF from North-Eastern Bulgaria who are followed up and treated at the University Hospital "St. Marina".

WORKING HYPOTHESIS:

Patients with a genotype including class I or class II mutations according to the defect in CFTR have more severe phenotypic manifestations than patients with a genotype including mutations from the other classes. F508del is the most common mutation among CF patients and its presence in their genotype in homozygous or heterozygous carriers is associated with more severe phenotypic manifestations.

AIM

The aim of this dissertation is to evaluate the genotype-phenotype correlation in CF patients from Northeastern Bulgaria who are followed and treated at the St. Marina University Hospital.

TASKS

1. To include CF patients from the North East in the EPM.
2. To analyze CFTR mutations in CF patients from Northeast Bulgaria according to their frequency, distribution, ethnicity.
3. To analyze age and phenotypic manifestations in confirming the diagnosis in the setting of missing mass NS.
4. To look for genotype-phenotype correlations in CF patients from Northeast Bulgaria.
5. To look for a correlation between genotype and phenotypic manifestations on the respiratory side (FEO1, BE, bacterial colonization, nasal polyposis).
6. To look for a correlation between genotype and phenotypic manifestations on the GIT side (BMI, CFLD, MI, pancreatitis).
7. Look for correlation between genotype and other phenotypic manifestations (sweat test, CFRD, infertility).
8. To propose an algorithm based on alarming symptoms for early referral to an expert centre for CF diagnosis and subsequent treatment.

PART III. MATERIALS AND METHODS

1. PARTICIPANTS

The present study included 45 patients with proven CF who were enrolled in the ERPM and followed up at the St. Marina University Hospital.

Inclusion criteria:

1. Confirmed diagnosis of CF by detection of two pathological mutations in CFTR on genetic analysis and/or confirmed diagnosis of CF by two pathological sweat tests and presence of clinical manifestations.
2. Signed informed consent to be included in the ERPM and to use the entered data for this dissertation.
3. Availability of patient data for the last 5 years.

Exclusion Criteria:

1. Lack of sufficient criteria to definitively confirm a diagnosis of CF.
2. Other severe comorbidities (pulmonary, cardiovascular, neurological, etc.) due to overlapping phenotypic manifestations.
3. Lack of signed informed consent for inclusion of CF patients in the EPPM.
4. Patients with CF lost to follow-up in the last 5 years.

Patients or their parents/guardians (for children under 18 years of age) signed an informed consent to enter the data available about them in the system of St. Marina University Hospital in the EDPM. Protocol No. 75/07.06.2018 for the inclusion of the CF patients from the North-East was approved by the Research Ethics Committee (REC) at the Medical University-Varna, and the use of the data entered in the ERPM for this thesis was approved by an added change - protocol No. 87/24.10.2019.

1. SURVEY DESIGN

In this thesis a retrospective study was included using the information from the EPM for 45 patients from North-Eastern Bulgaria with a diagnosis of CF, who were followed up and treated in the VEC at St. Marina University Hospital. The period covered is 5 years - from January 2019 to December 2023.

2. METHODS

The retrospective study is an analysis of data from the EPM, which is based on the medical records of patients with CF at the St. Marina University Hospital for the period from January 2019 to December 2023. Data on 45 patients aged 5 months to 37 years were included.

1. The following information was systematized for patients from history and physical status:

- passport details - gender, age
- ethnicity
- pregnancy and delivery data - MI, NS
- data for the neonatal period
- clinical manifestation in the diagnosis of CF
- current clinical presentation by organ and system
- anthropometric data
- fertility
- complications
- life expectancy and cause of death

2. Data from the following studies conducted on CF patients were included:

- laboratory data
- sweat test results
- results of genetic analysis
- microbiological test results

- results of functional respiratory testing - spirometry
- imaging results

Description of methods

Anthropometric indicators

Patients with CF from the Northeast were followed up every 3 or 6 months at the VEC, and height (in cm), weight (in kg), and body mass index (BMI) (in kg/m^2), calculated using the standard formula: BMI = body mass (kg)/height2 (m^2), were recorded.

Microbiological tests

They are performed on the territory of the University Hospital "St. Marina" in the microbiology laboratory. They include a nasal swab, throat swab and sputum. The materials are plated on selective and differentiating culture media. After an incubation period of 24 hours, identification of the microorganism is made by manual biochemical or automated methods (MALDI, VITEK). In parallel with the identification process, antibiotic susceptibility testing of the isolated microorganism is performed. The antibiogram is produced by the Kirby-Bauer manual method on Mueller Hinton agar or by determination of minimum inhibitory concentration (VITEK). Interpretation of the results obtained for antibiotic susceptibility (sensitive, intermediate, resistant) is performed using the latest version of EUCAST.

Sweat test

The sweat test is performed in all patients with suspected CF, and in addition to diagnosis, it is also used to monitor the effect of CFTR-modulating therapy. The device used at St. Marina Hospital is a galvanostatic biopulser, which operates with galvanic current to perform iontophoresis. A photile diluted with distilled water is placed on the anode and fixed on the patient's arm. The cathode is fixed on the back by placing sodium carbonate on it. After the

iontophoresis is performed, a pre-measured filter paper is placed on the anode and left for one hour. Remove it, weigh it again and proceed to the laboratory determination of the quantity of chloride ions. This is done by the Sheils mercuriometric titrimetric method. The reagents used are diphenyl carbazone - 250 mg to 100 ml of 95% ethyl alcohol; mercuric nitrate solution - 800 mg of the substance dissolved in 10 ml of nitric acid followed by addition of 600 ml of distilled water; standard sodium chloride solution. Filtration is carried out using the following dilution factor: (5 ml + weight of sweat)/weight of sweat=A and the following formula: mmol/l sweat=A ml mercuric nitrate 4.4x2. Values below 40 mmol/l are considered normal, 40 to 60 mmol/l are considered borderline and above 60 mmol/l are considered pathological.

Genetic testing for identification of pathogenic mutations

The genetic studies that identified the pathogenic mutations in CF patients from the Northeast, followed up at the HPP, included the following panels:

- CFTR panel including the most common variants, which is suitable for patients with a clear diagnosis and is often used in screening.
- Sanger sequencing, which detects all exon and intron sequence changes but not large deletions and duplications.
- Multiplex ligation-dependent probe amplification (MLPA) analysis.
- Next-generation sequencing (NGS), which detects changes in the sequence of exons and introns, but requires specific additional laboratory analyses to detect large deletions and duplications.

In the presence of alarming symptoms and after performing a sweat test, patients consult a geneticist and the most appropriate type of genetic testing is recommended.

Spirometry

VEC has a MIR Spirolab spirometer, which is used to perform spirometry in CF patients over 6 years of age. The data required to calculate reference values are sex, age, height, weight and race. After explanation and demonstration by the examiner, the patient performs the

manoeuvre, and is monitored for correct performance, and subsequently the following three key points are recorded: maximal deep inspiration, maximal sharp expiration, sustained expiration to completion. The parameter we will use for functional assessment of breathing is FEO1 (forced expiratory volume in 1 second), which is calculated as a percentage (%). FEO1=100% means that the patient's lung function measurement is equal to the average lung function measurement of a healthy reference population with the same sex, age and anthropometric parameters. The FID is performed under strict infection control, additionally using a filter - Neumofilt Ergo. Absolute contraindications for the study are haemoptysis of unknown origin, pneumothorax, unstable state of the cardiovascular system, known aneurysm, immediately after ocular, abdominal or thoracic surgery.

Imaging studies

Chest radiography is the most commonly used imaging modality to assess structural lung injury in CF patients. The reason for this is the lower radiation exposure, greater accessibility and last but not least the lower cost. The disadvantage of conventional radiography is the lower informative value compared to other imaging studies, e.g. computed tomography (CT). Chest radiographs in patients with CF from Northeastern Bulgaria, followed up in the VPP, were performed on the territory of St. Marina University Hospital using Siemens Multix equipment. Routine radiographs to assess disease progression are performed once a year, and emergency radiographs - when pulmonary exacerbation, hemoptysis, pneumothorax or atelectasis is suspected.

CT (computer tomography**)** of the thorax is a more sensitive method for assessing structural lung involvement in CF patients. It detects early changes even in asymptomatic patients and allows more detailed monitoring of disease progression. It is suitable for searching for BE, atelectasis, consolidations and emphysematous changes. Disadvantages are higher radiation exposure, more difficult accessibility and higher cost. CT in patients with CF from the Northeast was performed using two devices, Siemens Somatom and Force Siemens, which have a standard low radiation dose (1-2 mSv).

Abdominal ultrasonography is a non-invasive and easily accessible imaging study suitable for the evaluation of anatomical features, dimensions and pathological changes of the abdominal organs. It is suitable for patients of all ages due to the absence of ionizing radiation. The ultrasound devices used in the territory of St. Marina Hospital are Hitachi Aloka F31 and Hitachi Aloka Arietta 70. They were used to assess the degree of structural liver involvement in CF patients from the Northeast of Bulgaria followed up in the HPP.

PART IV. RESULTS

The retrospective study included 45 patients from Northeastern Bulgaria with a diagnosis of CF who were followed and treated at St. Marina University Hospital. A period of 5 years was covered, from January 2019 to December 2023. During this period, after signing informed consent, patients were enrolled in the ECFPR and the data from the ECFPR was used for this thesis. Three patients died and two patients left the country and continued their treatment and follow-up abroad.

Of the 45 patients enrolled, 28 were male and 17 were female. There was a male predominance and the percentage is depicted graphically (Figure 1).

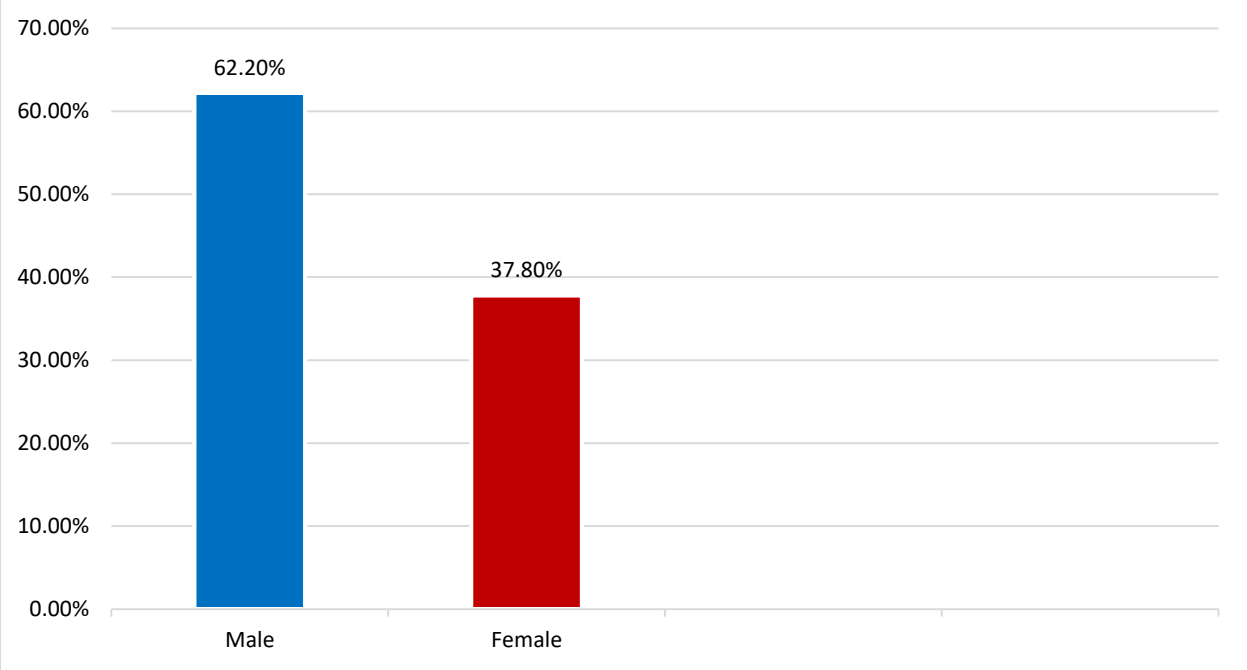


Figure 1. Distribution of patients by sex.

The age of CF patients ranges from 5 months to 37 years. Of the patients included, 31 were younger than 18 years and 14 were older than 18 years. The mean age of the population was 14.94±9.80 years. The percentage age distribution is presented in graphical form (Figure 2).

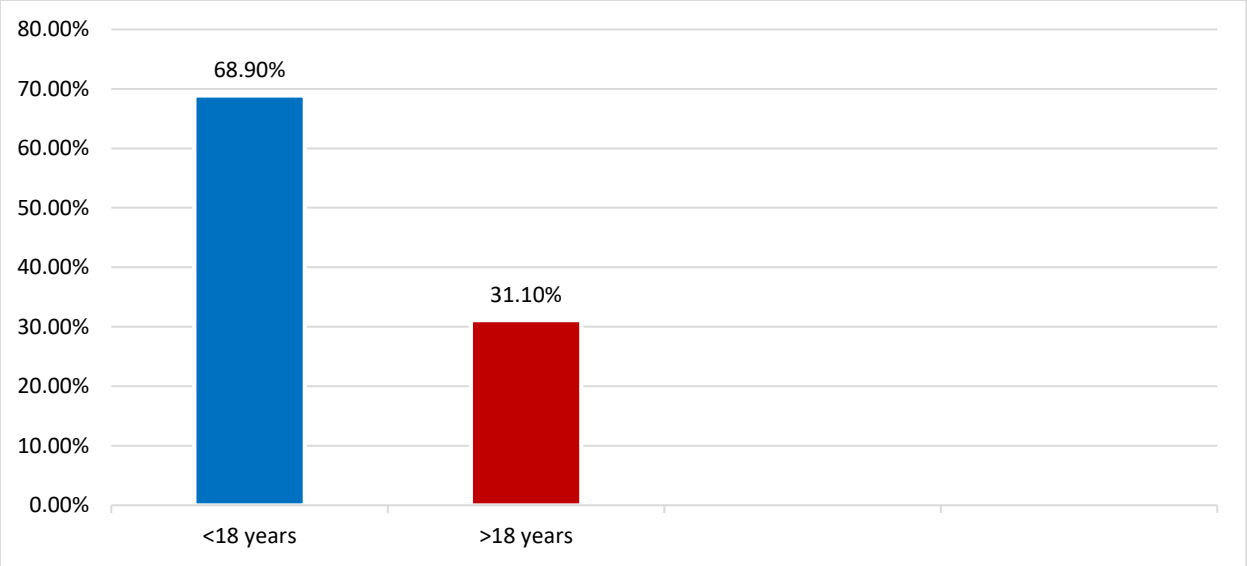


Figure 2. Distribution of CF patients by age.

According to ethnicity, 31 of the CF patients were Bulgarians, 12 - Bulgarian Turks and 2 Bulgarian Roma(Figure 3).

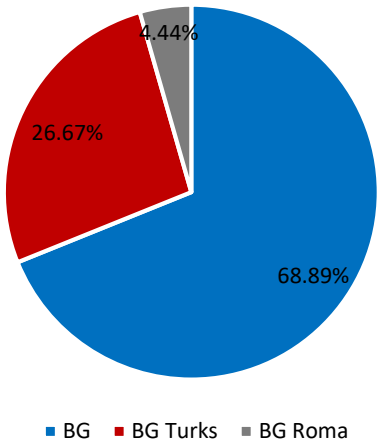


Figure 3. Distribution of CF patients according to their ethnicity.

Two CFTR mutations were identified in 44 CF patients by genetic analysis. In 1 patient, a second mutation was not detected, and the diagnosis of CF was confirmed by two pathological sweat tests and the presence of clinical manifestations (Table 2).

Table 2. Distribution of mutations in CF patients.

Mutations in CFTR	**Class**	**Number of patients**	**Frequency of mutation**
c.1521_1523delCTTp.(Phe508del) **F508del**	II	39	64.45%
c.1624G>T p.(Gly542*) **G542X**	I	4	4.45%
c.658C>T p.(Gln220X) **Q220X**	I	3	3.33%
c.442delA p.(Ile148Leufs*5) **574delA**	I	2	2.22%
c.3909C>G p.(Asn1303Lys) **N1303K**	II	2	2.22%
c.1040G>C p.(Arg347Pro) **R347P**	VI	1	2.22%
c.1712C>T p.(Leu571Ser) **L571S**	VUS	1	2.22%
c.3472C>T p.(Arg1158*) **R1158X**	I	1	1.11%
c.3846G>A p.(Trp1282X) **W1282X**	I	1	1.11%
c.2491G>T p.(Glu831) **E831X**	I	1	1.11%
c.2051_2052delAAinsG p.Lys684Serfs*38 **2183delAA>G**	I	1	1.11%
c.3731G>A p.(Gly1244Glu) **G1244E**	III	1	1.11%
c.254G>A p.Gly85Glu **G85E**	II	1	1.11%
c.274G>A p.(Glu92Lys) **E92K**	IV	1	1.11%
c.3160C>G p.(His1054Asp) **H1054D**	IV	1	1.11%
c.1766+1G>C **1898+1 G>T**	V	1	1.11%
c.3718-2477C>T **3849+10kbC>T**	V	1	1.11%
c.1585-8G>A **1717-8G>A**	V	1	1.11%
c.4242+1G>A **4374+1G>A**	V	1	1.11%
c.254G<A **p.Glu85Glu**	VUS	1	1.11%
c.3475T>C **pS1159P**	VUS	1	1.11%
c.3909>G **p.Asn1303Lys**	VUS	1	1.11%
L997F	VUS	1	1.11%
Unidentified		1	1.11%

The mutation with the highest frequency is class II F508del. Of all 89 mutations identified in CFTR, it occurs 58 times or in 64.45% in homozygous or heterozygous carriers. The next most frequent mutation was the class I G542X mutation, which was identified in heterozygous carriers in four of the CF patients and accounted for 4.45% of all mutations. In third place was the class I Q220X mutation, which occurred in heterozygous carriers in three patients or in 3.33% of cases. The N1303K (class II) and 574delA (class I) mutations were found in heterozygous carriers in two patients, or 2.22% of all mutations. The next most frequent mutations, R347P (VI) and L571S (VUS), had the same frequency of 2.22% but were found in homozygous carriers in one patient. The remaining 16 identified mutations were in heterozygous carriers in a single patient, 1.11% of all mutations.

The genotypic distribution of CFRT mutations in patients from the St. Marina University Hospital is presented in tabular and graphical form (Table 3, Figure 4).

Table 3. Distribution of patients by genotype.

Genotype	**Class**	**Number of patients**
F508del/F508del	II/ II	19
F508del/Q220X	II/ I	3
F508del/G542X	II/ I	2
F508del/574delA	II/ I	2
F508del/Druga	II/ -	13
Friend/friend	-/-	6

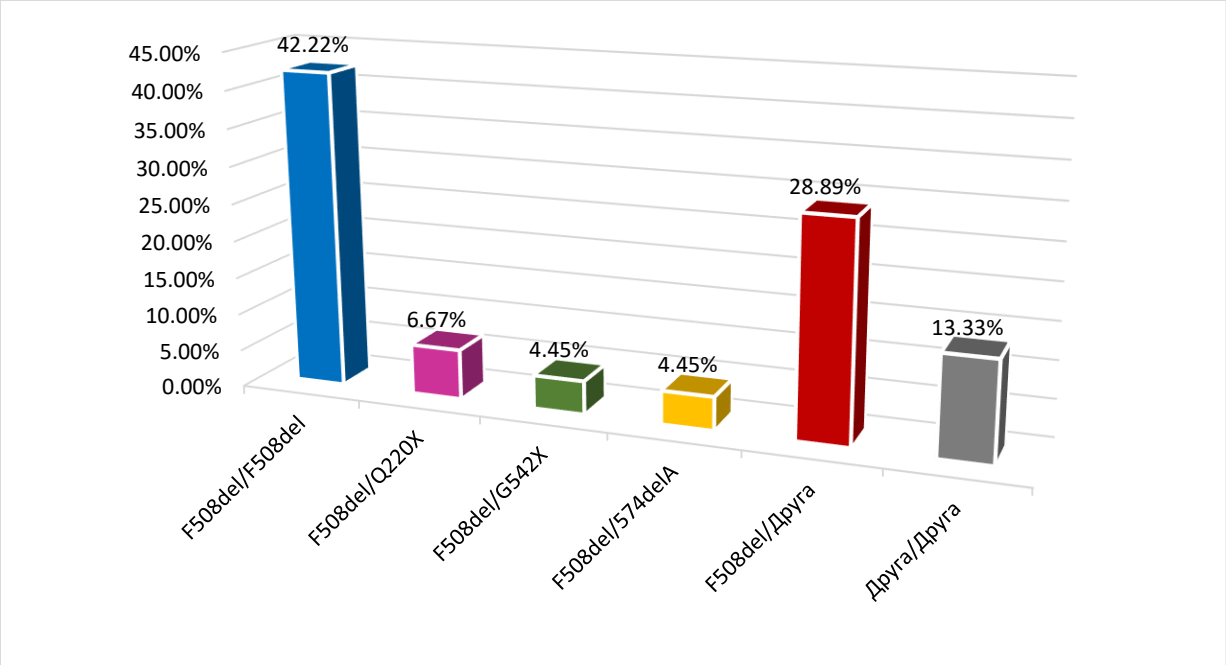


Figure 4. Percent genotype distribution.

The genotype with the highest frequency was F508del/F508del - in 19 patients or 42.22% of all genotypes included in the study. The next most frequent genotypes were F508del/1220X in three patients (6.67%), F508del/G542X in two patients (4.45%) and F508del/574delA also in two patients (4.45%). Outside of those listed, F508del in heterozygous carriers was present in 13 CF patients (28.89%). The percentage of patients with at least one F508del mutation was 84.45%. Genotypes with no F508del mutation were identified in six CF patients, or 13.33% of all genotypes. The genotypes that included F508del and occurred in one patient are presented in tabular form, and the mutation classes are noted (Table 4). Also presented are the six genotypes that do not have F508del and their respective combination of mutation classes (Table 5).

Table 4.F508del/ Druga.

Genotype F508del/Other	Mutation classes
F508del/S1159P	II/VUS
F508del/ E831X	II/ I
F508del/R1158X	II/ I
F508del/N1303K	II/ II
F508del/E92K	II/ IV
F508del/ 1898+1G>	II/V
F508del/ 2183AA+G	II/ I
F508del/ 1717-8G>A	II/V
F508del/H1054D	II/ IV
F508del/L997F	II/VUS
F508del/W1282X	II/ I
F508del/4374+1G>A	II/V

Table 5. Other/Fellow.

Genotype Other/Other	Mutation classes
R347P/R347P	VI/VI
L571S/L571S	VUS/VUS
G542X/3849+10kbC>T	I/V
c.3909C>G/c.254G<A	VUS/VUS
N1303K/G85E	II/II
G542X/G1244E	I/III

According to ethnicity, the distribution of these genotypes is as follows:

- Of the 31 patients of Bulgarians, 14 were F508del homozygotes, 11 were double heterozygotes with at least one F508del, and six were double heterozygotes without any F508del.
- Of the 12 patients of Bulgarian Turks, three were F508del homozygotes, 8 were double heterozygotes with at least one F508del, and one patient was a double heterozygote with no F508del.
- Of the two patients of Bulgarian Roma, one was homozygous for F508del and the other was a double heterozygote - F508del/Other.

Table 6: Percentage distribution of genotypes according to patients' ethnicity

Ethnicity	F508del/F508del	F508del/Other	Other/Other
English	**45.16%**	35.48%	19.36%
Turkish	25.00%	**66.67%**	8.33%
Roma	50%	50%	-

In summary, the F508del/F508del genotype is the most frequent among the Bulgarian ethnic group (45.16%), while among the Turkish ethnic group, double heterozygotes carrying one F508del and one other mutation predominate.

The mean age at confirmation of CF diagnosis of the 45 included patients was 2.14±5.32 years. The earliest diagnosed patient was 2 months old and the latest diagnosed patient was 18 years old.

Clinical manifestations at diagnosis include MI, failure to thrive, chronic diarrhea, OS, anemia, hypoproteinemia, recurrent pneumonia, liver involvement, vomiting with alkalosis, pancreatitis (Table 7, Figure 5).

Table 7. Leading clinical manifestations at diagnosis.

Leading clinical manifestation	Number of patients
Failure to thrive, diarrhea	12
Failure to thrive, hypoproteinaemia, anaemia	8
Failure to thrive, BOS	5
MI	5
Recurrent pneumonia	5
BOS	3
Vomiting, alkalosis	2
Other	5

One patient each with positive NS, pancreatitis, liver involvement, diarrhea, and BOS were included in "Other".

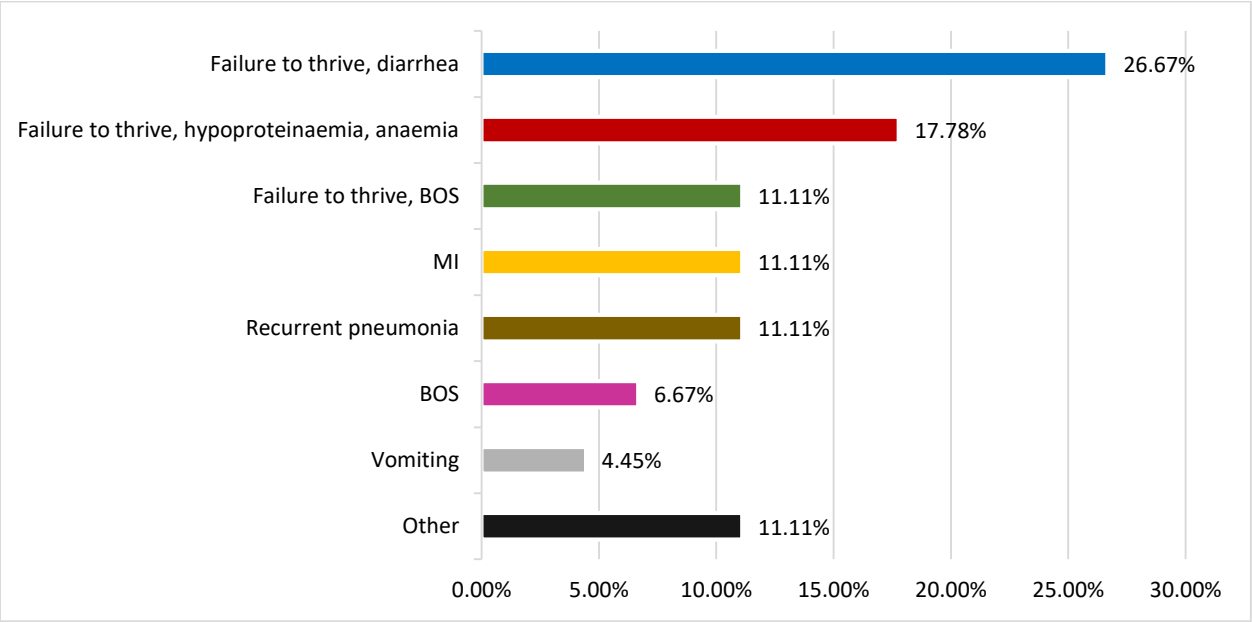


Figure 5. Percentage distribution of clinical manifestations at diagnosis.

The mean age at confirmation of diagnosis varies considerably between the leading clinical manifestations. Patients with MI, failure to thrive, hypoproteinemia, anemia, and vomiting with alkalosis are diagnosed in infancy, whereas those with recurrent pneumonia are diagnosed in adolescence (Table 8).

Table 8: Leading clinical manifestation and mean age at confirmatory diagnosis

Leading clinical manifestation	**Average age in years**
Failure to thrive, diarrhea	1.95±2.87
Failure to thrive, hypoproteinaemia, anaemia	0.35±0.15
Failure to thrive, BOS	1.2±0.66
MI	0.32±0.18
Recurrent pneumonia	15.14±6.82
BOS	2±0.61
Vomiting, alkalosis	0.35±0.07

There were three dead patients over the 5-year period covered. Two were male (M), aged 17 and 22 years, and one was female (F), aged 14 years. The mean life expectancy was 17.67±4.04 years. Two of the deceased were homozygous for the F508del/F508del genotype and the cause of death was end-stage lung disease RF. The third patient had the F508del/delA574 genotype and the cause of death was a cardiovascular event (CVI) (Table 7).

Table 9. Dead patients.

Patient	Gender	Age	Genotype	Reason
1	M	17 years.	F508del/F508del	RF
2	Ж	14 years.	F508del/F508del	RF
3	M	22 years old.	F508del/delA574	CVI

From 2020, CFTR-modulating therapy was initiated in 30 of the 42 CF patients enrolled, or 71.43%. The largest number of patients are on treatment with Elexacaftor/Tezacaftor/Ivacaftor, 20 patients or 47.62%. Ivacaftor/Lumikaftor treatment was conducted by 9 patients, 21.43%, and Ivacaftor treatment was conducted by only 1 patient, 2.38%. Twelve patients or 28.57% of the CF patients from the Northeast did not undergo gene-modification therapy, four of them - 9.52% did not meet the inclusion criteria according to the type of mutations in their genotype, and the remaining 8 of them, or 19.05%, were not included because of other criteria - not reaching the minimum age for initiation, lack of interest on the part of the parents and/or the patient, or a medical condition that did not allow starting therapy (Figures 6 and 7).

Figure 6: Percentage of CF patients receiving CFTR-modulating therapy.

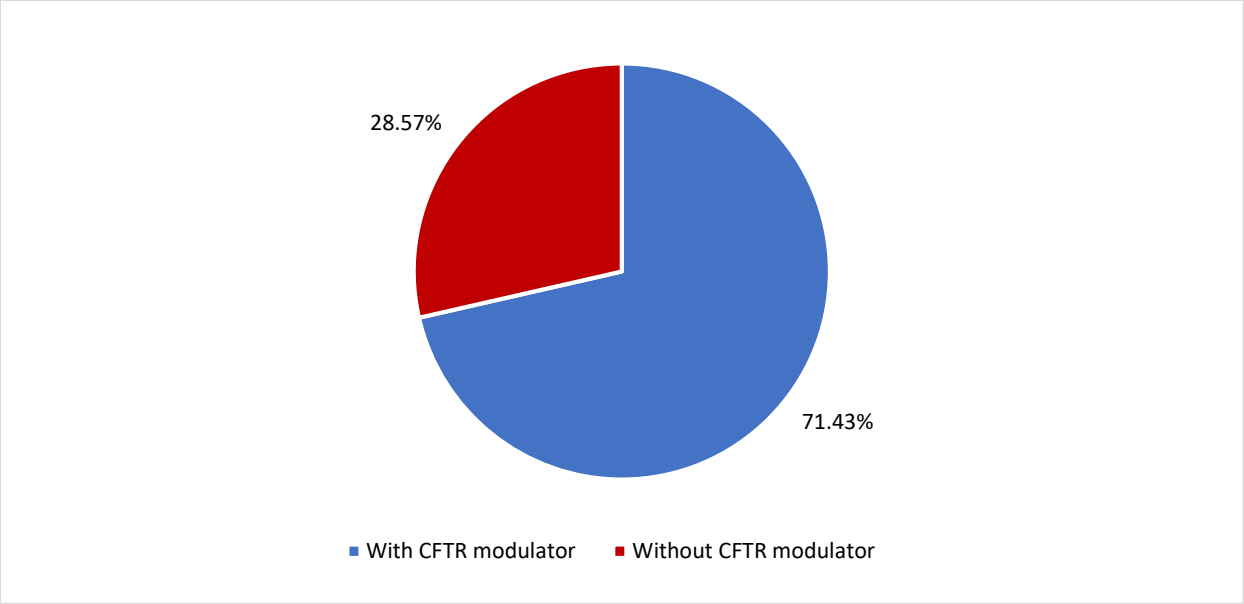
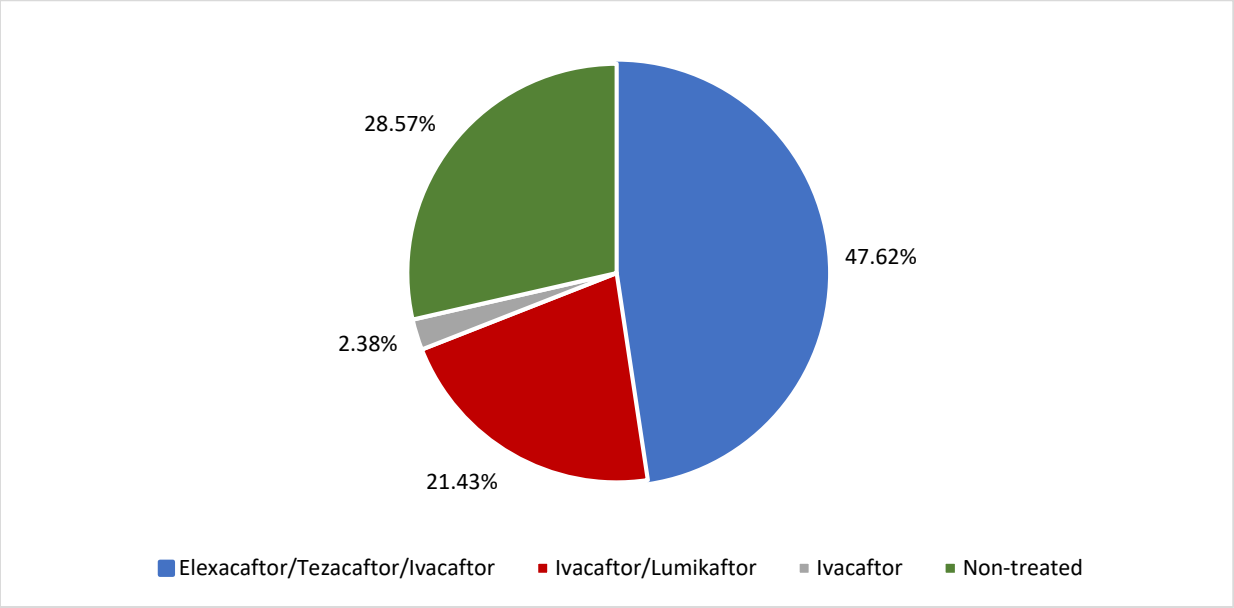


Figure 7. Percentage distribution of medications used.



Genotype-phenotype correlation – respiratory tract

To search for correlations between the genotypes described for CF patients from Northeastern Bulgaria and the phenotypic manifestations on the respiratory side, the following indices were used:

- Current value of FEO1 in % of predicted value for age, height, weight, race
- Presence of BE from the imaging studies performed
- Chronic infection with PA
- SA isolate (at least once)

The metrics listed were used to look for correlation between patients with the most frequent genotype, F508del/F508del, and the next most frequent genotypes, F508del/Q220X, F508del/G542X, and F508del/574delA, as well as between those with genotypes F508del/Other and Other/Other.

Genotype-phenotype correlation - FEO1

Patients with the F508del/F508del genotype had higher mean FEO1 values than those with the F508del/Q220X genotype and the F508del/delA574 genotype, but lower mean values than those with the F508del/G542X genotype. While the homozygotes included two class II mutations, the compared double heterozygotes had one class II mutation, F508del, and one class I mutation, Q220X, delA574, G543X (Figure 8).

With respect to FEO1, statistical hypothesis testing was performed by comparing two averages - in this case, the average value of FEO1 in %. Student's criterion was used, and the results are as follows:

- **There was no** statistically significant difference in FEO1 values between F508del homozygotes and F508del/Q220X double heterozygotes (temp. = 0.914 < tter. = 1.725).
- **There was no** statistically significant difference in FEO1 values between F508del homozygotes and F508del/delA574 double heterozygotes (temp. = 0.656 < tteror. = 1.729).
- **There was a** statistically significant difference in FEO1 values in favor of the **F508del/G542X.** double heterozygotes (temp. = 3.174 > tteor. = 1.729).

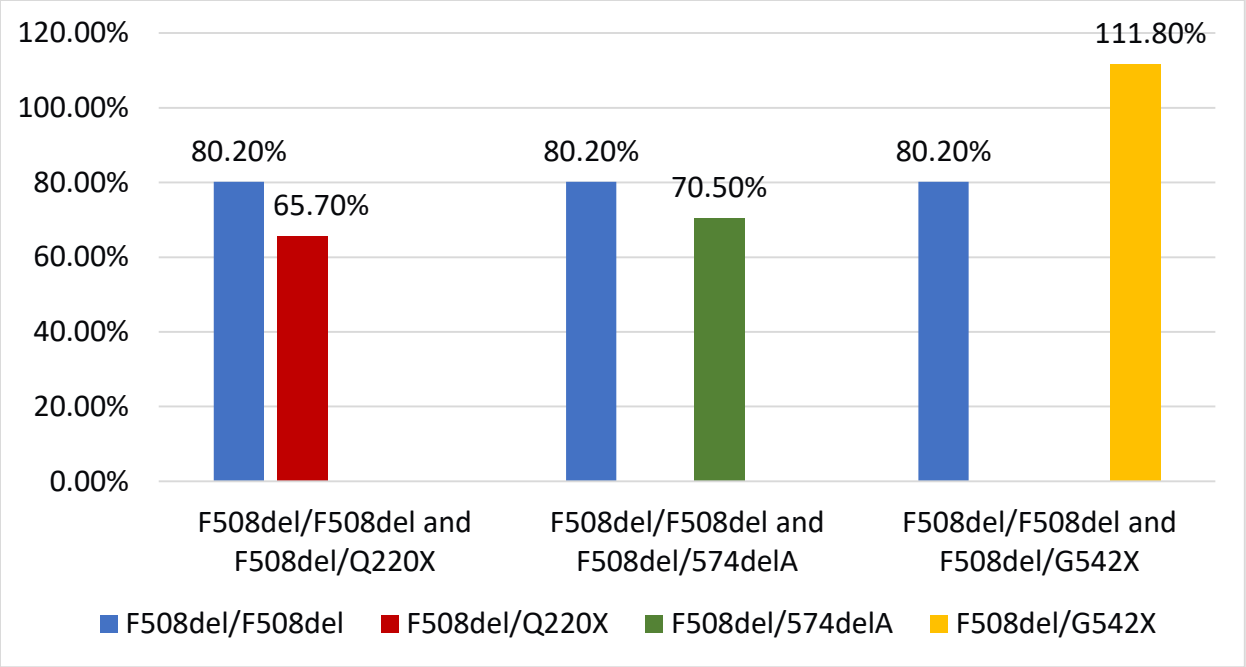


Figure 8: Comparison of FEO1 between patients with F508del/F508del genotype and the next most frequent genotypes.

It became clear that 29% of CF patients had one F508del and one other (non-F508del) mutation in their genotype. A comparison was again made by statistical hypothesis testing between two relative proportions (mean FEO1) between the two genotypes, F508del/Other and Other/Other, and the results showed a statistically significant difference in favor of patients with two mutations other than F508del - temp. = 3.796 > tteor. = 1.711. Their mean FEO1 was 95.90% and that of patients with the F508del/Druga genotype was 84.60% (Figure 9).

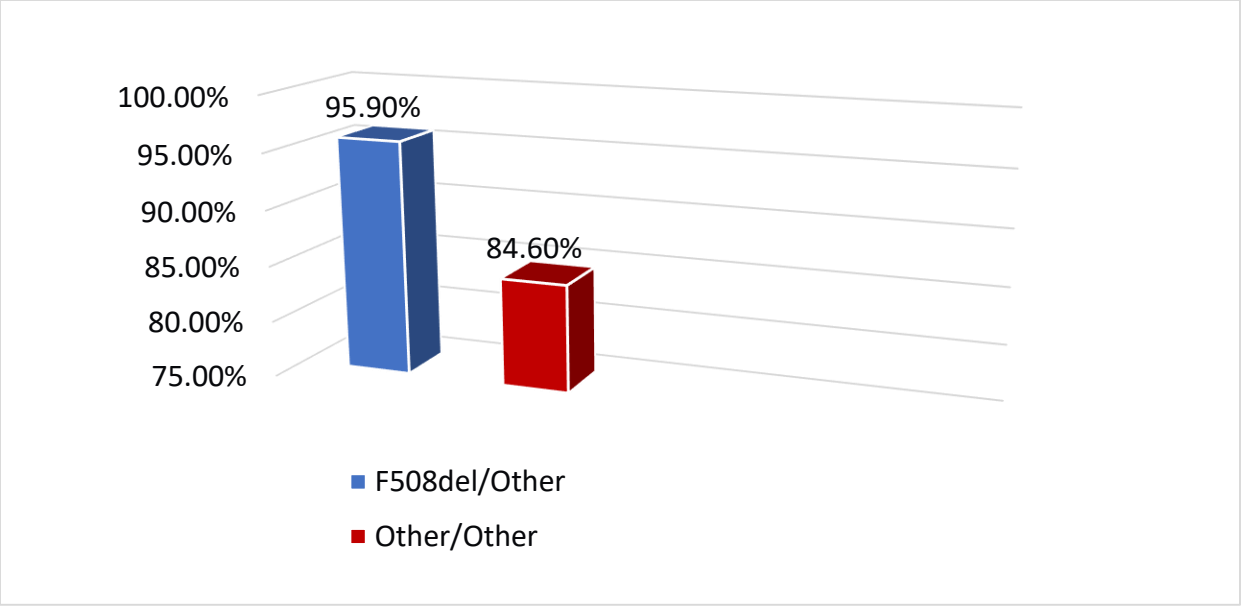


Figure 9. Comparison of FEO1 between patients with genotypes F508del/Other and Other/Other

Genotype-phenotype correlation - BE

BEs are proven in 69% of patients with CF from the North-East of Bulgaria, and it is assumed that their actual incidence is even higher (not all patients have undergone CT). All patients with genotype F508del/Q220X, F508del/G542X and F508del/574delA have formed BEs or this is 100%. In patients homozygous for F508del, the number of patients with BEs was 11 out of 19, which was 57.90%. By statistical hypothesis testing and comparing two relative proportions, again using Student's criterion, the difference between F508del and the other genotypes compared was shown to be statistically significant - temp. > tter (Figure 10).

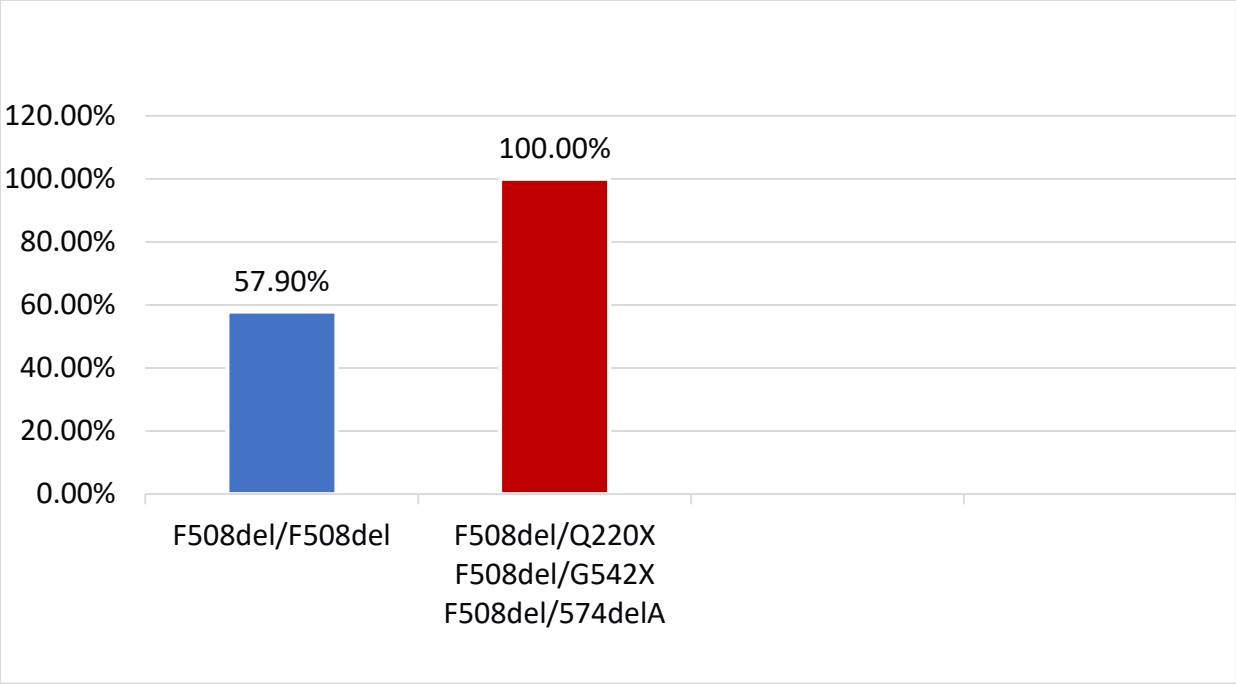


Figure 10. BE in the most frequent genotypes.

Patients with a genotype with at least one F508del (F508del/Other) have a significantly higher rate of BE than those with a genotype that does not include F508del (Other/Other). Of the 20 patients, 17 were found to have BE, while half of the six patients without F508del had none - 50%. The difference was again assessed by statistical hypothesis testing and comparing two relative proportions and was determined to be statistically significant - temp. = 3.796 > tteor. = 1.711 (Figure 11).

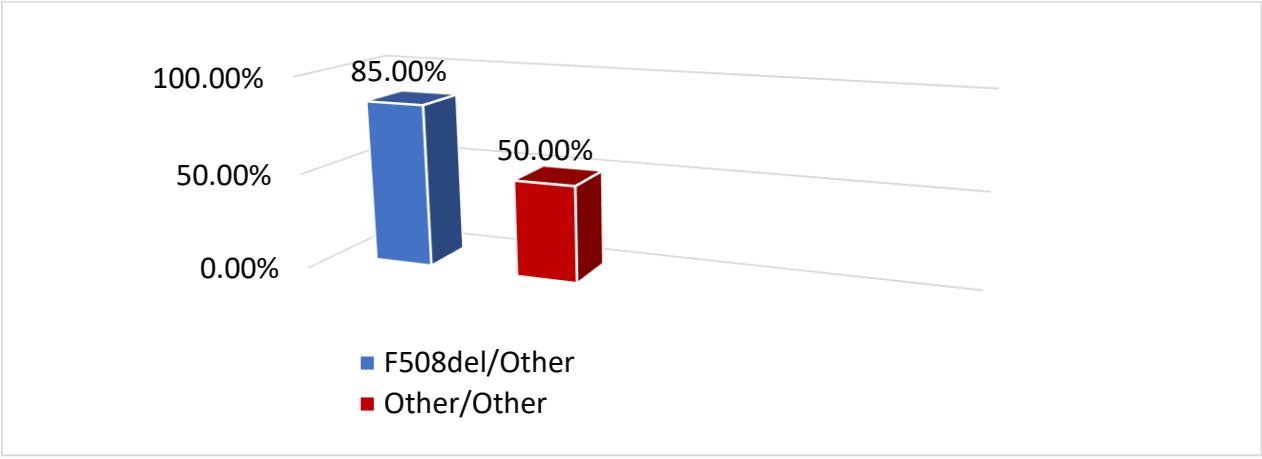


Figure 11. BE in patients with genotypes F508del/Other and Other/Other

Genotype-phenotype correlation - chronic PA infection

Chronic PA lung infection was present in 22 patients or 48.89% of all patients included in the study, with a rate of 32.26% for those under 18 years of age and 85.72% for those over 18 years of age.

Patients with the F508del/F508del genotype had colonization rates similar to those with the F508del/Q220X genotype and those with the F508del/574delA genotype and significantly higher than those with the F508del/G542X genotype. Again, the same statistical method was used - statistical hypothesis testing and comparison of two relative proportions, again using Student's criterion. The difference between F508del/F508del and F508del/G542X was statistically significant with respect to chronic PA infection (Figure 12).

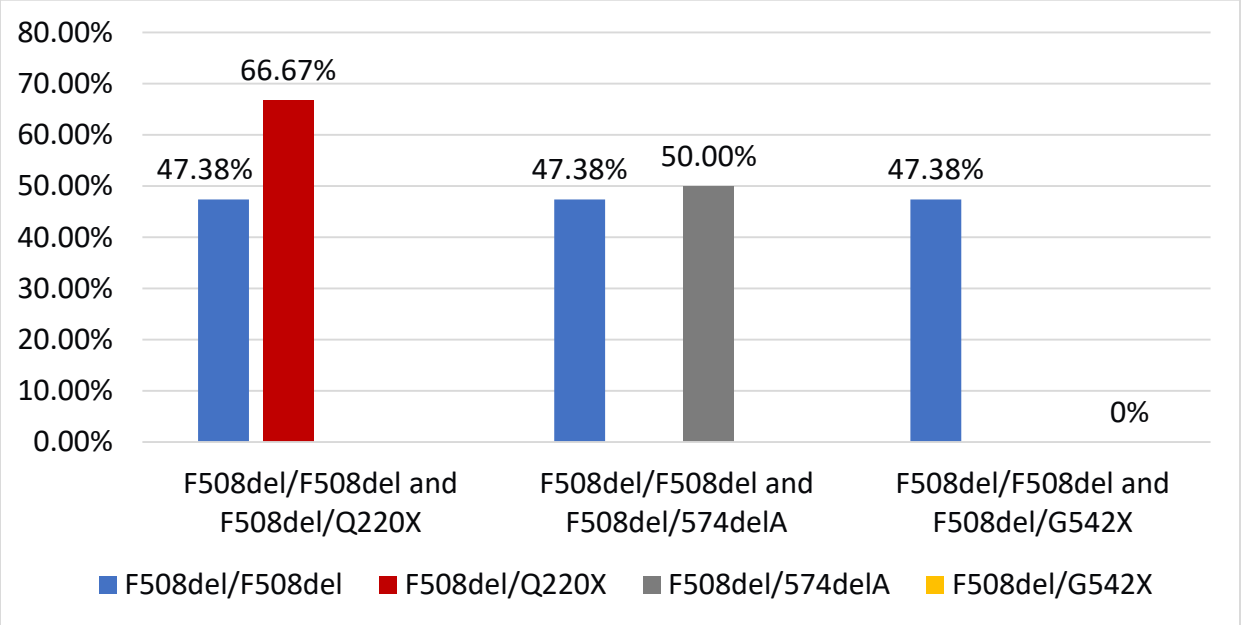


Figure 12. Chronic PA infection in the most common genotypes in CF patients.

Using the same statistical method, the rate of chronic PA infection was compared in patients with the F508del/other genotype and the Other/Other genotype. 11 of 20 patients with at least one F508del had chronic PA infection, and of the group with patients without F508del, two

of six were colonized. The difference was statistically significant - temp. = 2.523 > tteor. = 1.711 (Figure 13).

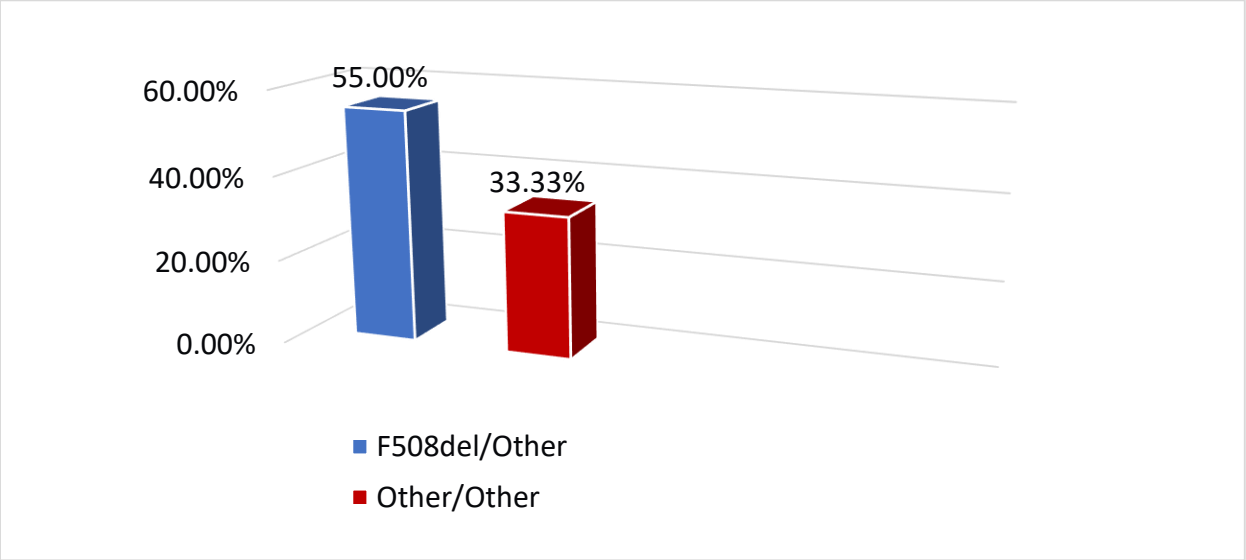


Figure 13. Chronic PA infection in patients with F508del/Other and Other/Other genotypes.

Genotype - phenotype correlation - SA

Intermittent, non-chronic infection with SA occurs in 33.33% of CF patients from Northeastern Bulgaria. A correlation was sought between the SA isolate and the most common genotypes. Double heterozygote patients showed a higher percentage of isolated SA compared to F508del homozygote patients. Using statistical hypothesis testing, there was a significant difference between F508del/F508del and F508del/Q220X - temp. = 1.976 > tteor. = 1.725 (Figure 14).

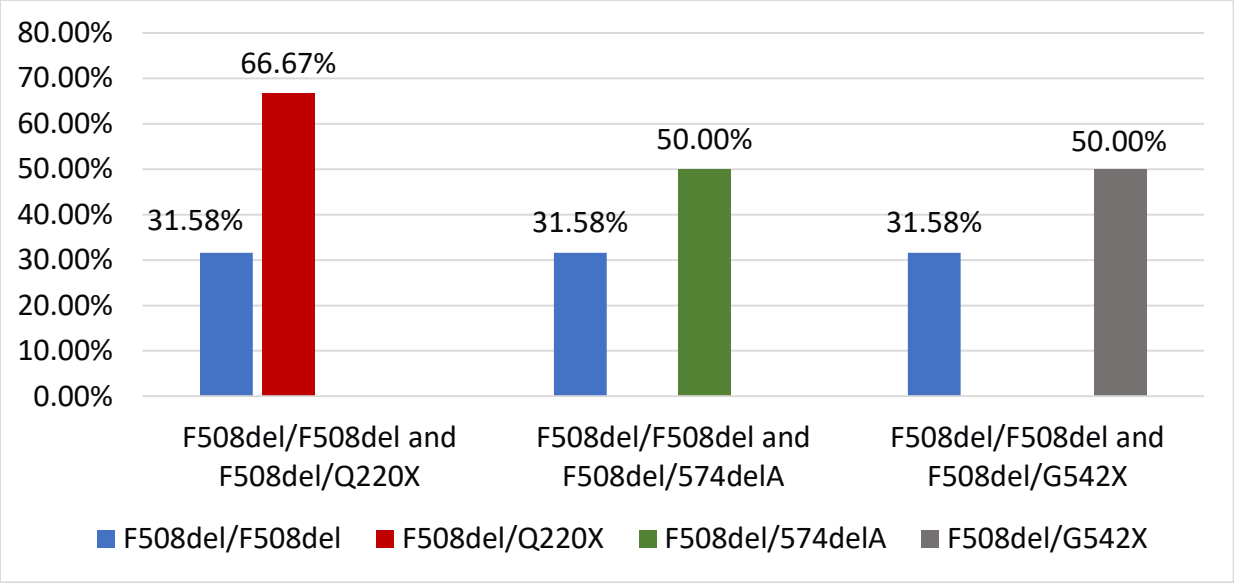


Figure 14. SA isolate in the most frequent genotypes.

There was also a statistically significant difference between the SA isolate in patients with at least one F508del and those with none - temp. = 2.032 > tteor. = 1.711 (Figure 15).

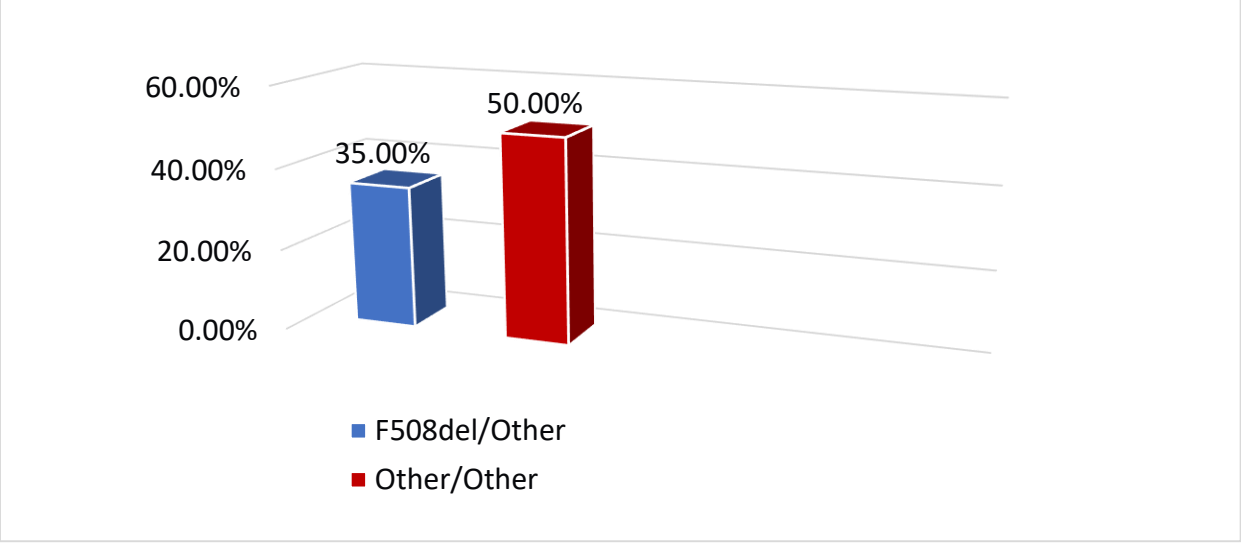


Figure 15. SA isolate in patients with F508del/Other and Other/Other genotypes.

From the studied parameters, it is noteworthy that the lung involvement of CF patients with F508del/G542X genotype is lower, although it is a combination of class II mutation and class I mutation, which are also F508del/Q220X and F508del/574delA genotypes.

Genotype-phenotype correlation in other respiratory manifestations

Nasal polyposis occurred in 20 out of 45 patients in the CF, or 44.45% of the patients included in the study. No statistically significant differences were demonstrated according to genotype, but there was an increase in the incidence of blue-nasal involvement with increasing age.

Allergic bronchopulmonary aspergillosis (ABPA) with IgE>1000 ng/mL was demonstrated in one female patient aged 13 years with F508del/F508del genotype. One male patient aged 22 years with genotype F508del/W1282X also tested *positive for* Aspergillus.

During the 5-year study period, there was also one successful **lung transplantation of** an adult female patient with end-stage lung disease of the F508del/F508del genotype.

Genotype-phenotype correlations - GIT

To search for correlations between the genotypes described for CF patients from Northeastern Bulgaria and the phenotypic manifestations of GIT, the following indices were used:

- BMI
- CFLD
- MI
- Pancreatitis

The metrics listed were used to look for correlation between patients with the most frequent genotype, F508del/F508del, and the next most frequent genotypes, F508del/Q220X, F508del/G542X, and F508del/574delA, as well as between those with the F508del/Other and Other/Other genotypes.

Genotype-phenotype correlations - BMI

A current BMI value was used. Patients with genotype F508del/F508del had lower mean BMI values compared to those with genotype F508del/Q220X and genotype F508del/delA574, but higher mean values than those with genotype F508del/G542X.

As with the FEO1 comparison, statistical hypothesis testing was used by comparing two means - in this case, mean BMI in kg/m2. Student's criterion was used, and the results are as follows:

- **There was a** statistically significant difference in BMI values between F508del homozygotes and F508del/Q220X double heterozygotes (temp. = 2.589 > temp. = 1.725).
- **There was a** statistically significant difference in BMI values between F508del homozygotes and F508del/delA574 double heterozygotes (temp. = 2.204 > tteor. = 1.729).
- **There was a** statistically significant difference in BMI values in favor of the **F508del/G542X.** double heterozygotes (temp. = 1.988 > tteor. = 1.729) (Figure 16).

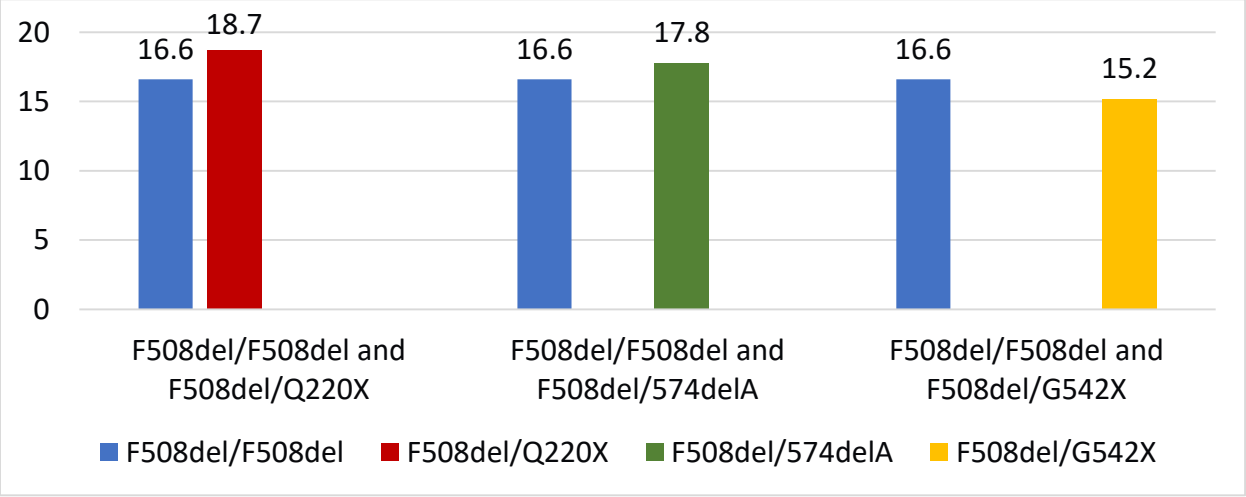


Figure 16. BMI comparison between patients with F508del/F508del genotype and the next most frequent genotypes.

According to the mean BMI values, F508del homozygous patients had moderate malnutrition, F508del/Q220X genotype had normal weight, F508del/574delA had mild malnutrition, and those with F508del genotype had severe malnutrition (Table 10). The results were also explained by the initiation of enzyme replacement therapy, hypercaloric diet and CFTR-modulating therapy.

Table 10. Nutritional status according to BMI value.

Nutritional status	**BMI *kg/m²***
Severe malnutrition	< 16,0
Average malnutrition	16 - 16,99
Mild malnutrition	17 - 18,49
Underweight	< 18,5
Normal weight	18,5 - 24,99

Patients with at least one F508del mutation had a mean MBI of 18.3 kg/m^2 , that is, underweight. The same is true for the group of patients with the Other/Other genotype who have a mean BMI of 18.2 kg/m^2 . Accordingly, no significant difference was found between the two study groups by statistical hypothesis testing and comparison of two means.

Genotype-phenotype correlation - CFLD

CFLD has been demonstrated in 20 of 45 CF patients, or this is a 44.45% incidence. In patients with the F508del/F508del genotype, it was present in 10 of 19 patients, an incidence of 52.63%. It was similar in frequency in patients with F508del/Q220X and F508/574delA genotype, 66.67% and 50.00% respectively. Again by testing hypotheses between two relative proportions and using Student's criterion, a significant difference was found between F508del homozygotes and patients with F508del/G542X genotype - rate. = 1.971 > tteor. = 1.729 (Figure 17).

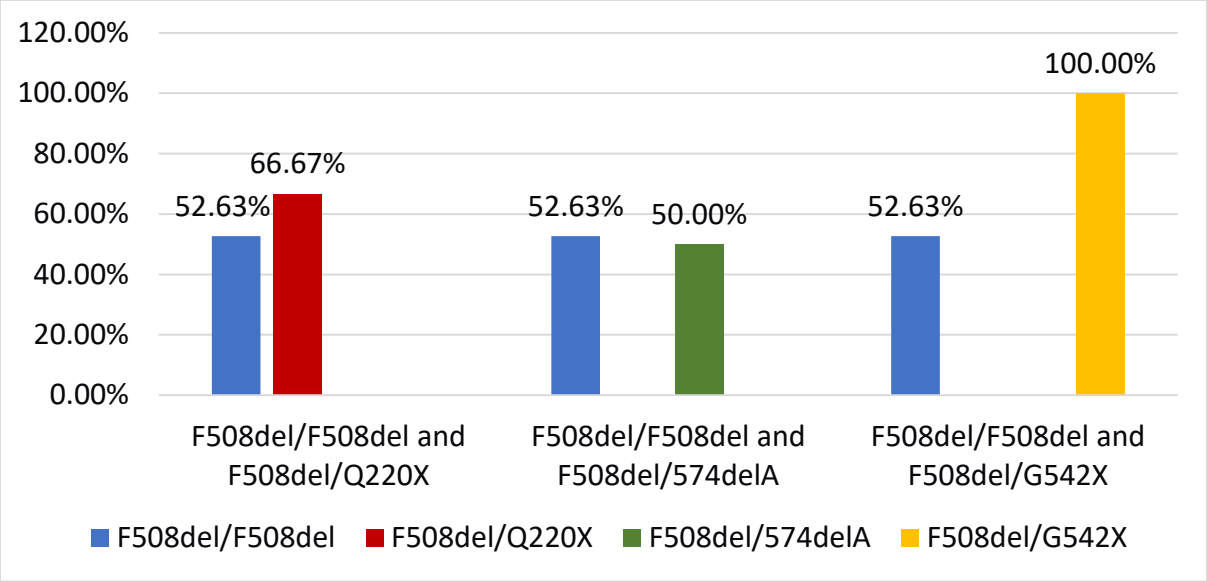


Figure 17. CFLD in the most common genotypes.

Patients with a genotype that includes at least one F508del (F508del/Other) have a significantly higher rate of CFLD than those with a genotype that does not include F508del (Other/Other). Of the 20 patients, 9 e demonstrated CFLD, whereas only one of the six patients without F508del had liver involvement. The difference was again assessed by statistical hypothesis testing and comparing two relative proportions and was determined to be statistically significant - temp. = 3.267> tteor. = 1.711 (Figure 18).

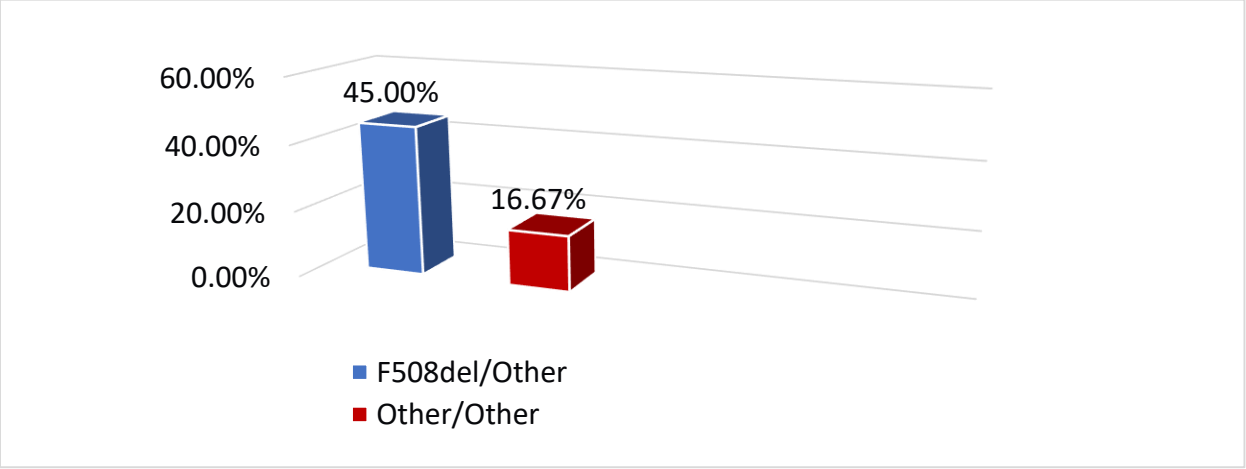


Figure 18. CFLD in patients with F508del/Other and Other/Other genotypes.

Genotype-phenotype correlation - MI

5 of the CF patients had undergone surgery for MI, respectively the prevalence among CF patients from Northeast Bulgaria was 11.11%. Due to the low incidence and different genotypes, no correlation can be sought on this indicator (Table 11).

Table 11. Genotype and class in MI patients.

Patient	Genotype	Class
1	F508del/F508del	II/II
2	F508del/F508del	II/II
3	F508del/G542X	II/I
4	F508del/574delA	II/I
5	c.3909C>G/c.254G<A	VUS/VUS

Genotype-phenotype correlation with respect to pancreatitis in CF

Acute (AP) or acute recurrent pancreatitis (ARP) was evidenced in 5 of the CF patients, and one patient had chronic pancreatitis (CP) (Table 12). The percentage incidence was 13.33% and was too small to look for statistically significant difference and correlations.

Table 12. Genotype and class in patients with pancreatitis.

Patient	Genotype	Class
1 - AP	R347P/R347P	VI/VI
2 - AP	F508del/ E831X	II/ I
3 - AP	F508del/F508del	II/II
4 - ARP	N1303K/G85E	II/II
5 - ARP	F508del/W1282X	II/ I
6 - CP	F508del/L997F	II/VUS

In the patient with the F508del/L997 genotype, acute recurrent pancreatitis was also the leading clinical manifestation to confirm the diagnosis of CF, and the pancreatitis subsequently became chronic.

During the 5-year follow-up period, there was one adult female patient with F508del/F508del genotype diagnosed with colon cancer.

Genotype - phenotype correlation to other phenotypic manifestations

Correlations were sought between the genotypes described for CF patients from Northeastern Bulgaria and the following other phenotypic manifestations:

- Value of sweat test
- CFRD
- Infertility

Genotype-phenotype correlation - sweat test values

Sweat test values were used to look for correlations in confirming the diagnosis of CF. Statistical hypothesis testing was performed by comparing two means using Student's criterion and the results showed that:

- **There was a** statistically significant difference in sweat test values between F508del homozygotes and F508del/Q220X double heterozygotes (temp. = 1.918 > tteor. = 1.725).
- **There was a** statistically significant difference in sweat test values between F508del homozygotes and F508del/delA574 double heterozygotes (temp. = 1.842 > tteor. = 1.729).
- **There was no** statistically significant difference in sweat test values in favor of the **F508del/G542X.** double heterozygotes (temp. = 0.928 < tteor. = 1.729) (Figure 19).

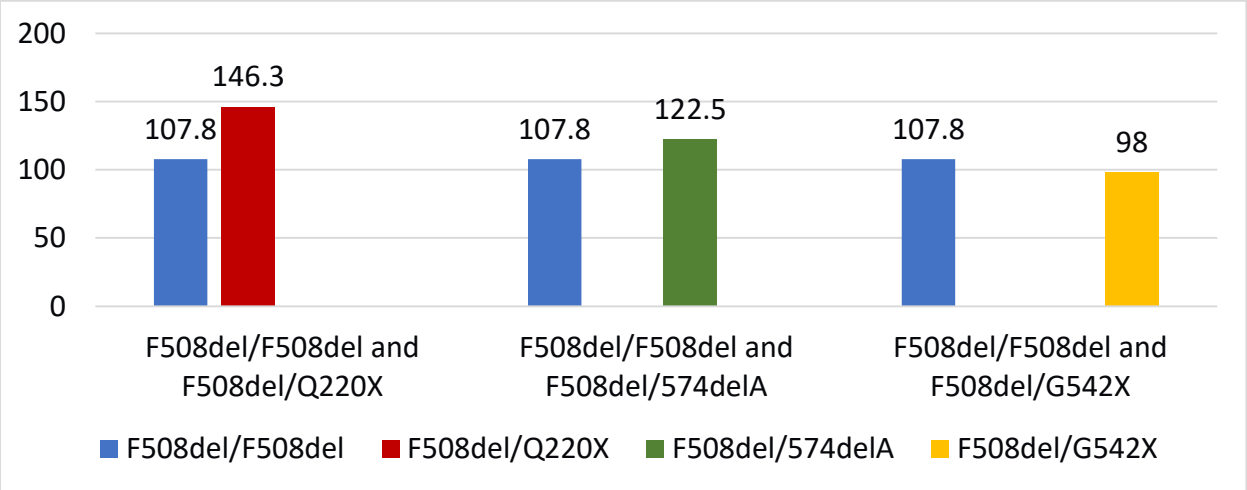


Figure 19. Mean sweat test value in mmol/l for the most frequent genotypes.

Patients with a genotype with at least one F508del (F508del/Other) had significantly higher mean sweat test values than those with a genotype that did not include F508del (Other/Other). The difference was again assessed by statistical hypothesis testing and comparing two means using Student's criterion and was determined to be statistically significant - temp. = 2.086 > tteor. = 1.711 (Figure 20).

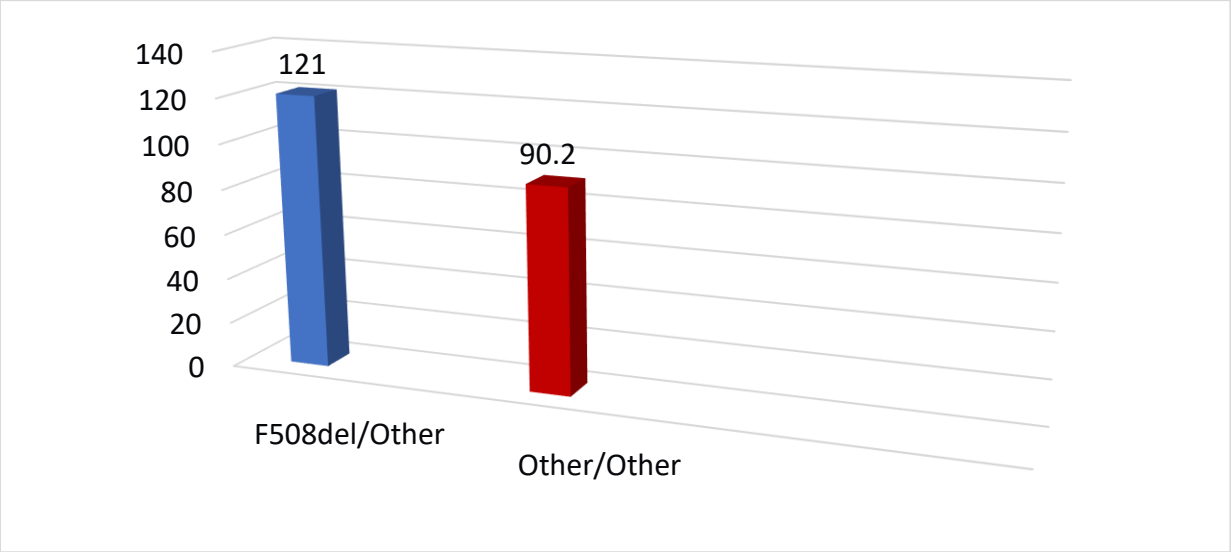


Figure 20. Mean sweat test values in mmol/L for genotype F508del/Other and Other/Other.

Genotype-phenotype correlation - CFRD

Five of the CF patients had CFRD, or 11.11% of all patients included in the study. All five had the F508del/F508del genotype, with a mean age of 21.60±7.27 years. Accordingly, the differences from each other genotype were significant and CFRD e with higher incidence in patients above 18 years of age (Figure 21).

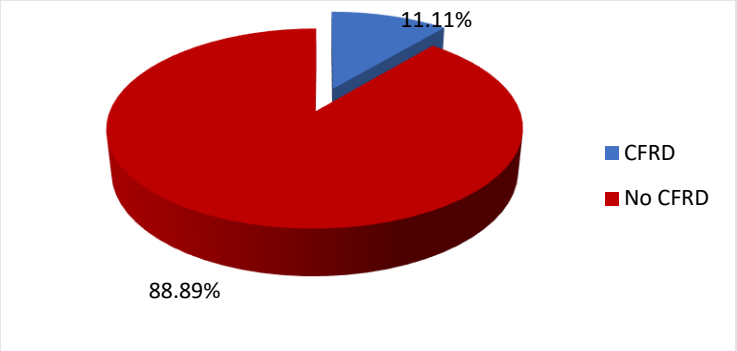


Figure 21. Incidence of CFRD among CF patients.

Genotype-phenotype correlation - infertility

Despite the lack of research conducted in male CF patients, it is assumed that 100% of them are infertile due to the lack of offspring. In female patients over the age of 18 years, there have been two successful deliveries and one pregnancy, in all three cases CFTR-modulating therapy was initiated. The number is small and does not allow searching for correlations.

Correlation between mutation class and phenotypic manifestation

To assess the correlation between mutation classes, CF patients from Northeastern Bulgaria were divided into three groups according to the combination of classes in their genotype (Table 13). One patient with one identified mutation and one unknown mutation was not included.

Table 13. Distribution of mutations in the genotype of CF patients by class.

Patient group	**Genotype according to the class of mutations involved**	**Number of patients**	**Percentage of patients**
First	I - II/ I - II	33	73.33%
Second	I - II/ III - VI or VUS	8	17.78%
Third	III - VI or VUS/ III - VI or VUS	3	6.67%

Correlation between mutation classes and lung involvement

Regarding lung involvement, the parameters of FEO1 value, presence of BE, PA colonization, SA isolate, and nasal polyposis were compared again.

Correlation between mutation classes and FEO1 value

Patients with two class I or class II mutations in their genotype (group 1) had lower mean FEO1 values than patients with a genotype including any of the following classes - I - II/ III - VI or VUS (group 2). Participating patients with genotype III - VI or VUS/ III - VI or VUS (third group) were not included in the comparison due to the low mean age of the sample - all were under 6 years of age and functional respiratory testing could not be performed. The mean age of patients in groups I and II was 14.24±8.36 and 19.25±13.20 years, respectively. Regarding FEO1, statistical hypothesis testing was performed between the first two groups of patients by comparing two means, in this case the mean value of FEO1 in %. The Student's criterion was used, and the results showed that temp. = 2.546 > tteor. = 1.684, therefore it is assumed that there is a statistically significant difference between the two means, indicating a correlation between more severe functional lung involvement and the presence of a class I or class II mutation in the patients' genotype (Figure 22).

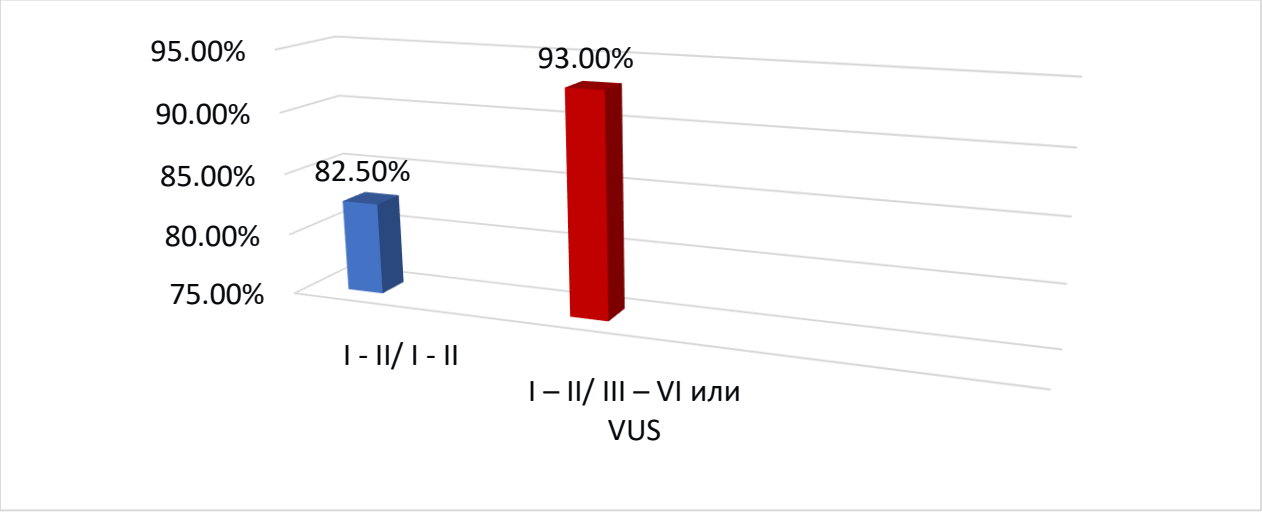


Figure 22. FEO1 in patients of the first and second group.

Correlation between mutation classes and the presence of BE

In terms of structural lung involvement - the presence of BE, the percentage was highest for the first group and lowest for the third, again taking into account the low average age of patients in the third group. Over 60% of patients who have at least one class I or class II mutation have proven BE. After using statistical hypothesis testing and comparing two relative proportions, again using Student's criterion, a statistically significant difference was found in the presence of BE between the three groups:

- Comparison between first and second group: temp. = 1.796 > tteor. = 1.684
- Comparison between first and third groups: temp. = 5.897 > tteor. = 1.697
- Comparison between second and third groups: temp. = 6.189 > tteor. = 1.833

These results show the correlation between more severe structural lung involvement in patients with at least one class I or II mutation compared with those with class III - VI or VUS mutations (Figure 23).

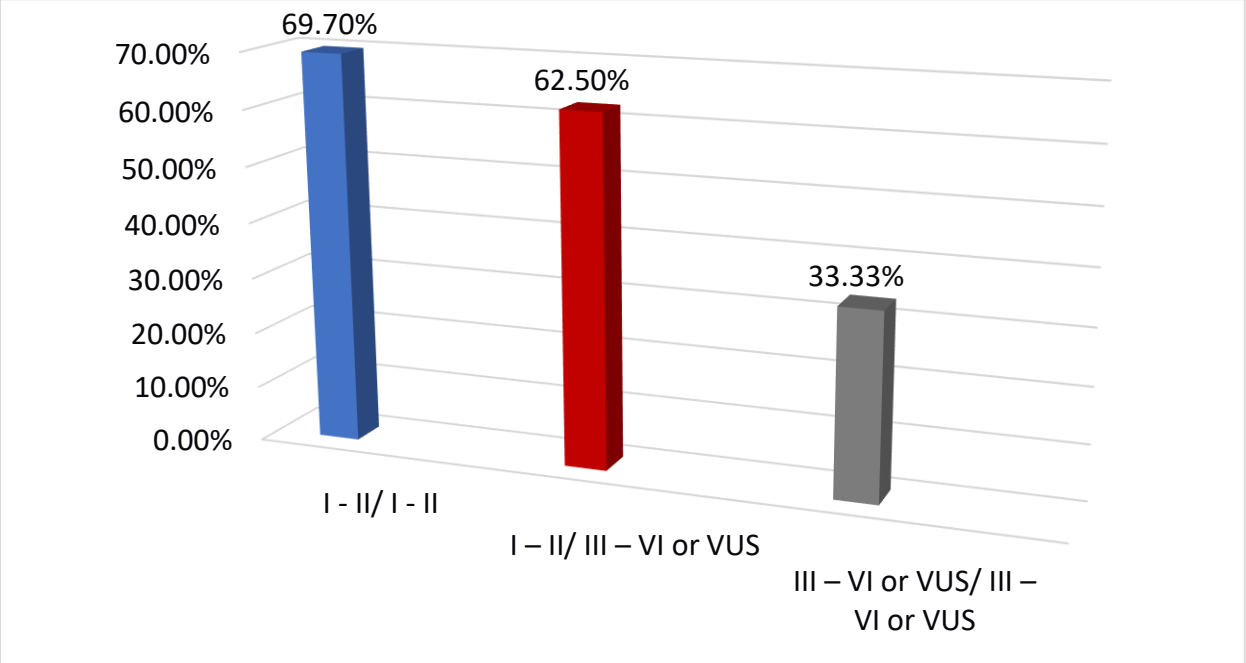


Figure 23. BE in patients of the three groups.

Correlation between mutation classes and chronic PA infection

The analysis of the data shows that the rate of chronic PA infection increases with increasing age. This explains why it was higher in the second group of patients who had a mean age of 19.25±13.20. There were no patients in the third group with microbiological examination positive for PA. After using statistical hypothesis testing and comparing two relative proportions, again using Student's criterion, there was a statistically significant difference in terms of chronic PA infection between the three groups:

- Comparison between first and second group: temp. = 2.123 > tteor. = 1.684
- Comparison between first and third groups: temp. = 10.063 > tteor. = 1.697
- Comparison between second and third groups: temp. = 12.321 > tteor. = 1.833

And with respect to chronic PA infection, the results show a correlation towards more severe involvement when a class I or class II mutation is present (Figure 24).

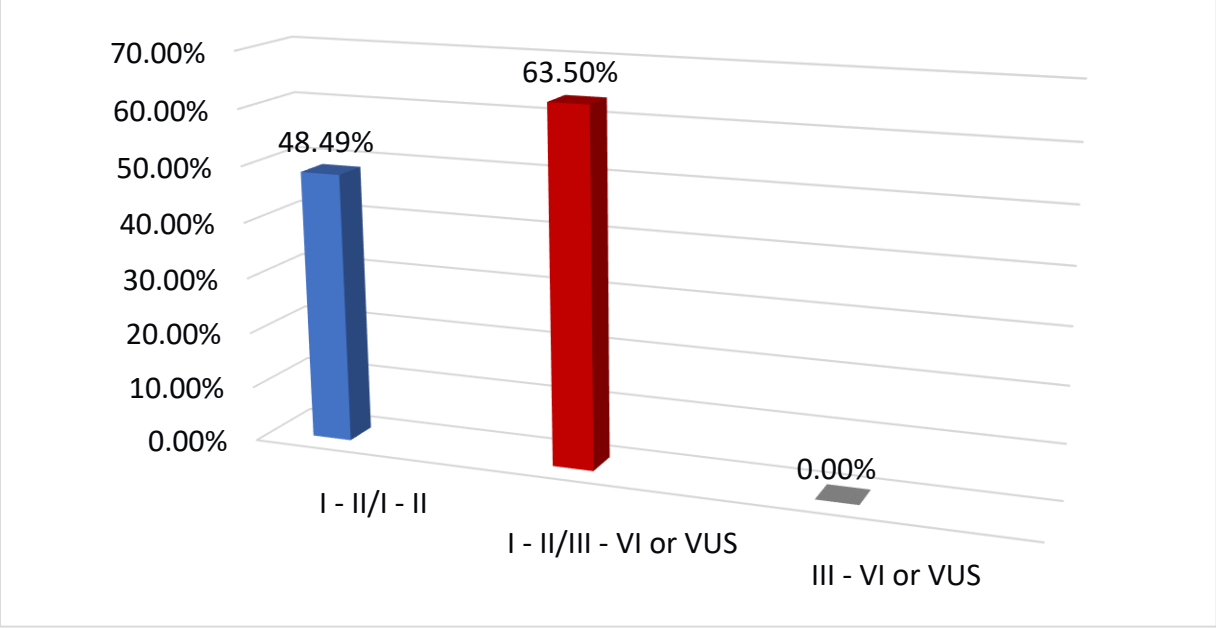


Figure 24. Chronic PA infection in the three groups.

Correlation between mutation classes and SA isolate

In all three groups of patients, SA was isolated in 30-40% on at least one sputum or CIS microbiological examination (Figure 25). No statistically significant difference was found between the rates in the three groups by statistical hypothesis testing and comparing two relative proportions, temp. < tteor. Therefore, no correlations were found in the patients included in the study with respect to the class of mutations in the SA genotype and isolate.

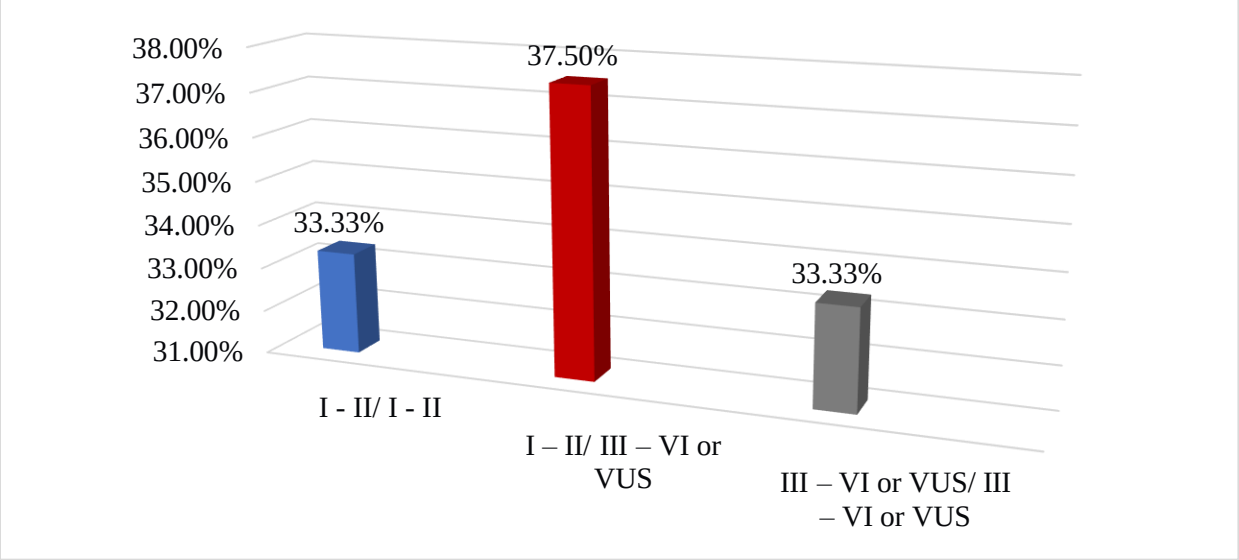


Figure 25. SA isolate in the three groups of patients.

Correlation between mutation classes and nasal polyposis

Nasal polyposis occurred at a higher rate in group one and group two patients, 48.49% and 50, respectively. 00%. For patients with a combination of mutation classes, the percentage was significantly lower at 33.33%, again drawing attention to the lower mean age of the sample (Figure 26). After using statistical hypothesis testing and comparing two relative proportions, again using Student's criterion, there was a statistically significant difference in terms of nasal polyposis between groups one and three and between groups two and three:

- Comparison between first and third groups: temp. = 1.823 > tteor. = 1.697

- Comparison between second and third groups: temp. = 2.852 > tteor. = 1.830
- Comparison between first and second group: temp. = 1.025 < tteor. = 1.833

The results of the analysis again showed that, in terms of nasal polyposis, patients with at least one class I or class II mutation had more pronounced involvement than those with none.

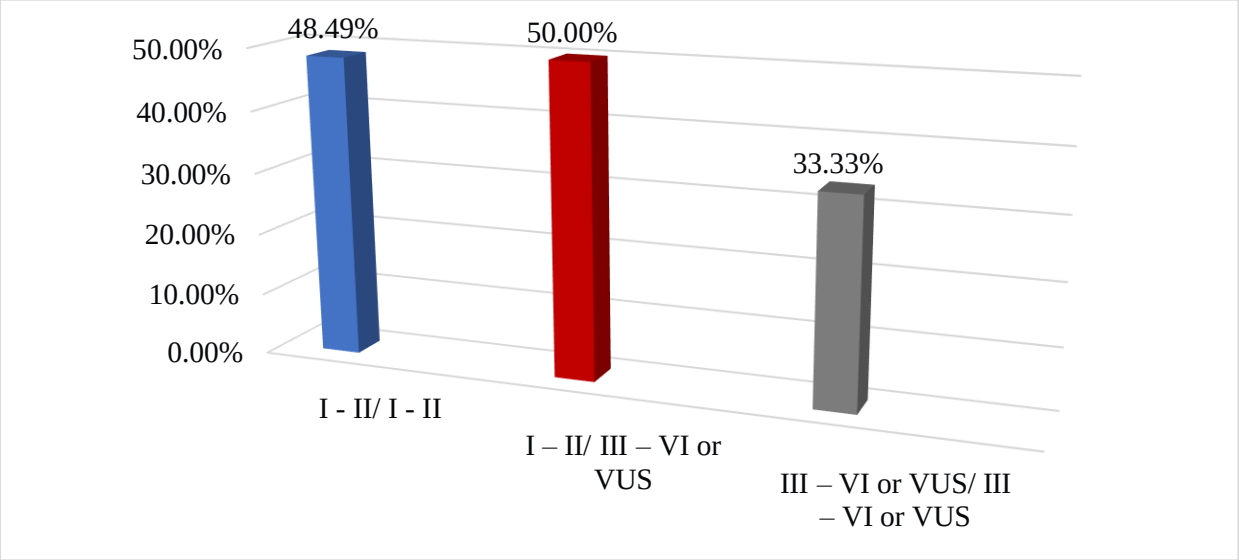


Figure 24. Nasal polyposis in patients of the three groups.

Correlation between mutation classes and GIT involvement

Regarding GIT involvement, the following parameters were again compared: BMI value, CFLD, MI and pancreatitis.

Correlation between mutation classes and BMI value

To look for correlations regarding BMI values, actual recent values were used against the background of enzyme complementation, hypercaloric diet and CFTR-modulating therapy in some of the patients. The mean BMI values of the three groups of patients showed a mean underweight for the first group (BMI - 16.6 kg/m^2) a normal mean weight for those in the second and third groups (BMI - 19.5 kg/m^2 and BMI 21.1 kg/m^2) (Figure 27). Statistical hypothesis testing was done by comparing two means using Student's criterion, and the results showed:

- Statistically significant difference between group one and group two - rate. = 2.875 > tteor. = 1.684
- Statistically significant difference between first and third group - rate. = 3.123 > tteor. = 1.697
- Lack of statistical difference between second and third group - temp. = 1.025 < tteor. = 1.833

Therefore, in the enrolled patients, there was a correlation between a genotype composed of class I and II mutations and lower BMI values.

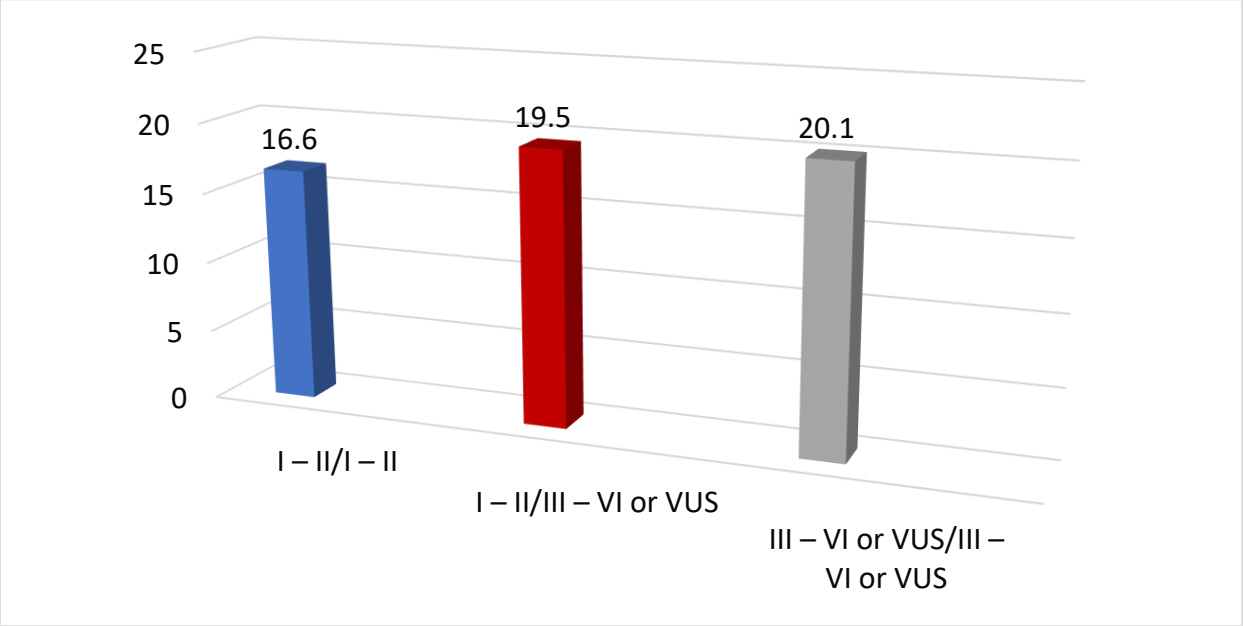


Figure 27. Mean BMI values in the three groups of patients.

Correlation between mutation classes and CFLD

CFLD occurs at a significantly higher rate in patients in the first group, with more than half of patients having varying degrees of liver involvement. There was only one patient with CFLD in the second group and none in the third group (Figure 28). After using statistical hypothesis testing and comparing two relative proportions, again using Student's criterion, there was a statistically significant difference in CFLD between groups one and two and between groups one and three:

- Comparison between first and second group: temp. = 8.267 > tteor. = 1.684
- Comparison between first and third groups: temp. = 14.123 > tteor. = 1.697
- Comparison between second and third groups: temp. = 2.468 > tteor. = 1.833

Accordingly, patients with a genotype composed of class I and class II mutations have a significant correlation with the presence of CFLD, in contrast to those who have only one class I or class II mutation or none. Although the difference between group II and group III was reported as significant both groups were classified as having minimal and no liver involvement.

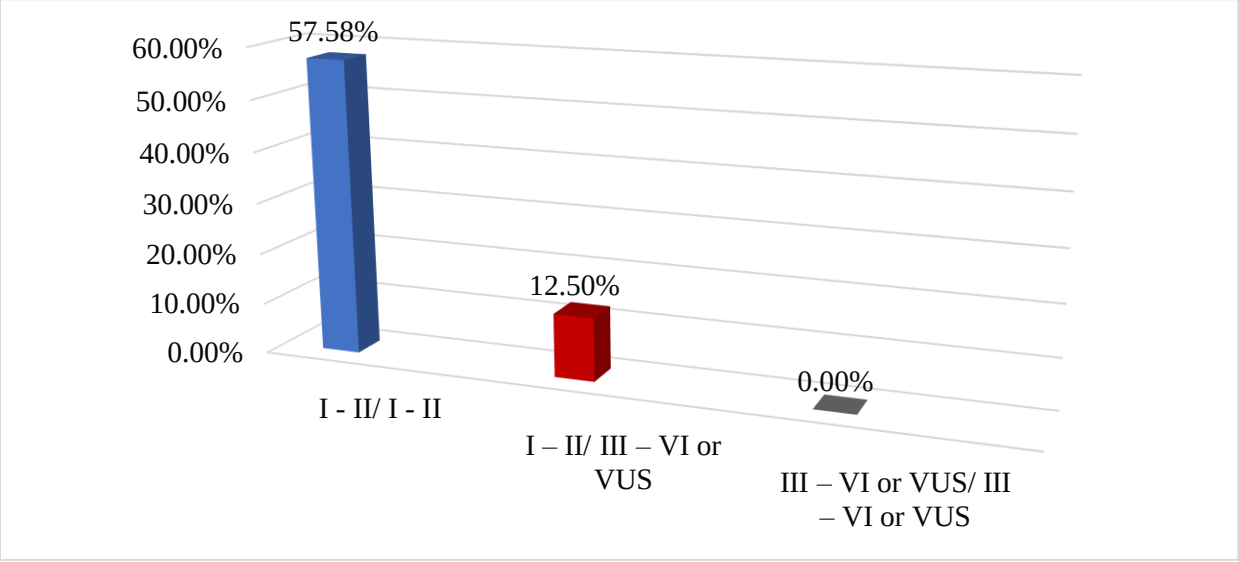


Figure 28. CFLD in the three groups of patients.

Correlation between mutation classes and MI

It became clear that 5 of the CF patients had undergone surgery for MI. According to their class distribution (Table 11), 4 of them had genotype of group 1 and one of group 3 (Figure 29). Therefore, based on the small sample size, it can be assumed that MI is significantly more common in those with a genotype composed of class I and II mutations.

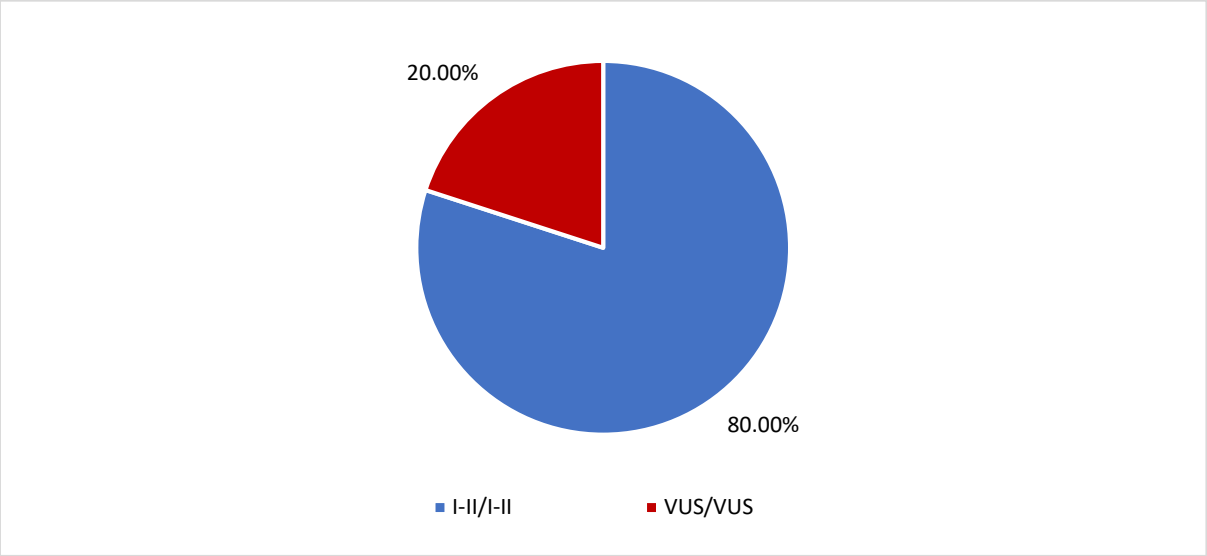


Figure 29. Distribution of MI rates between patient groups according to mutation class.

Correlation between mutation classes and pancreatitis in CF

Pancreatitis was also examined according to mutations in the genotype (Table 12), but relative to the classes of mutations included (Figure 30), when statistical hypothesis testing was performed by comparing two relative proportions and Student's criterion was used, a statistically significant difference was found between the third group of patients and the first and second groups:

- Comparison between first and third group: temp. = 3.462 > tteor. = 1.697
- Comparison between second and third groups: temp. = 3.855 > tteor. = 1.833
- Comparison between first and second group: temp. = 0.873 < tteor. = 1.684

Despite the small number of patients in the third group, due to the described statistically significant difference between the proportion of patients with pancreatitis among them and that of patients in the first and second groups, a correlation between genotype III - VI or VUS/ III - VI or VUS and pancreatitis can be assumed.

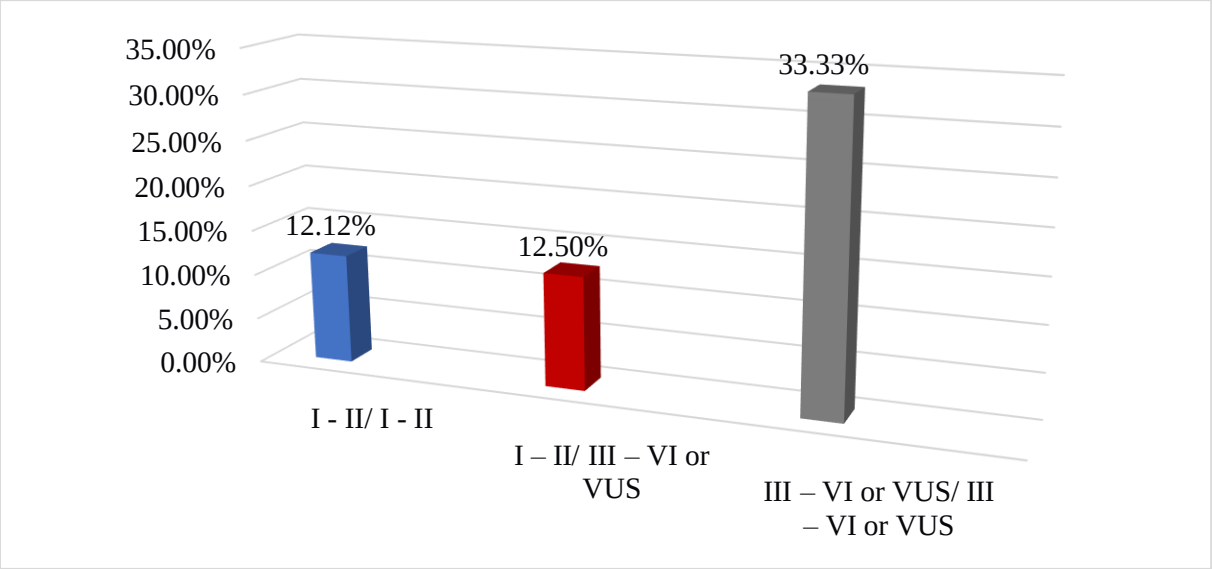


Figure 30. Pancreatitis in the three groups of patients.

Correlation between mutation classes and other phenotypic manifestations

From the "other" manifestations included in the search for genotype-phenotype correlations, it became clear that CFRD was described only in F508del homozygotes, i.e. in patients of the first group, and besides the mutation class, the correlation was also demonstrative with respect to the age of the patients.

For infertility, data are sparse and cannot be used to search for correlations by class of mutations in the genotype.

Correlation between mutation classes and sweat test value

Patients in group 1 had the highest mean sweat test values, followed by those in group 2 and finally those in group 3 (Figure 31). Statistical hypothesis testing was done by comparing two means between groups using Student's criterion, and the results showed:

- Statistically significant difference between group one and group two - rate. = 4.856 > tteor. = 1.684
- Statistically significant difference between first and third group - rate. = 7.873 > tteor. = 1.697
- Statistically significant difference between second and third group - rate. = 4.283 > tteor. = 1.833

Therefore, as the mutation class increases, the sweat test values decrease. Patients with class I and class II mutations have higher values than those with one class I or class II mutation, who in turn have higher values than those without either a class I or class II mutation.

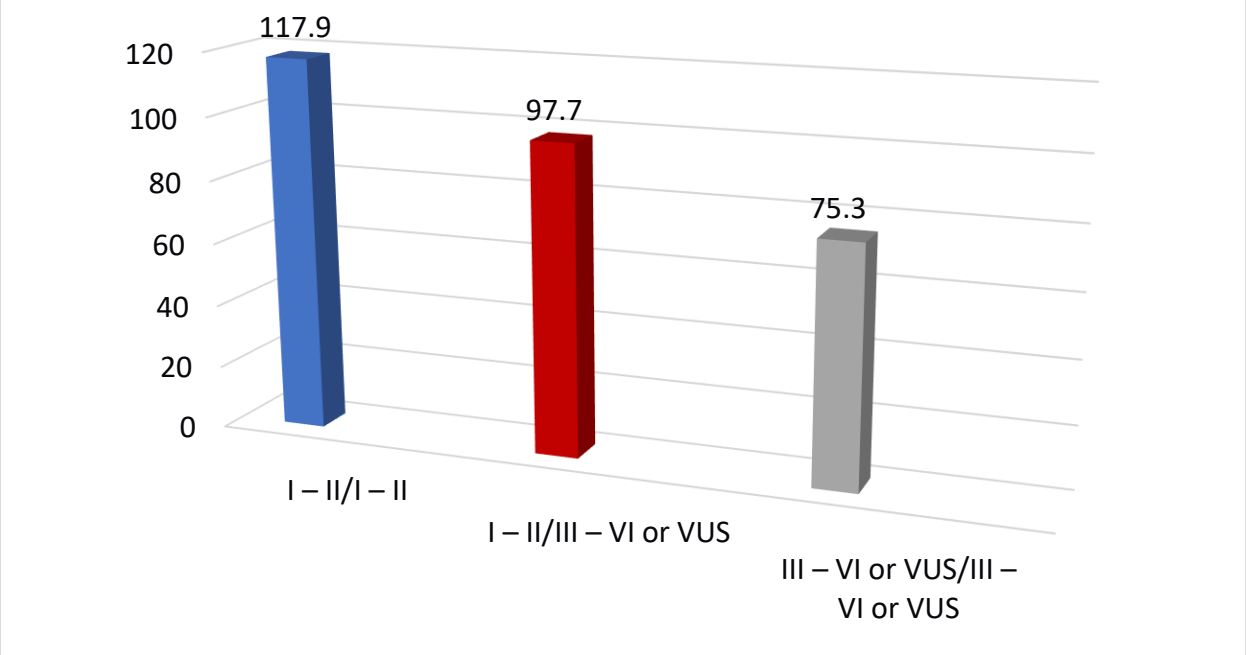


Figure 31. Mean sweat test values in mmol/L for the three groups.

Additional correlation analysis

For additional correlations between individual phenotypic manifestations, beyond genotype-phenotype, a correlation analysis was performed with the following results:

1. Correlations of patient's current age with other indicators:

- Significant positive correlation between patient's current age and BMI value by Pearson correlation coefficient - R = 0.649 (again the role of enzyme replacement therapy).
- Significant positive correlation between current age and presence of BE by biserial correlation coefficient - R = 0.545.
- Significant positive correlation between current age and CFRD by biserial correlation coefficient - R = 0.532.
- Significant negative correlation between patient's current age and FEO1 values by Pearson's correlation coefficient - R = - 0.533.

2. Correlations of the sweat test with other indicators:

- Moderate negative correlation between sweat test values and those of FEO1 by Pesrson's correlation coefficient - R = - 0.473.

3. Correlations of the FEO1 index with other indicators:

- Significant negative correlation between FEO1 values and presence of BE by biserial correlation coefficient: R = -0.532.
- Moderate negative correlation between FEO1 values and chronic PA infection by biserial correlation coefficient - R = - 0.487.

4. Correlations between the presence of BE with other indicators:

- Very strong positive correlation between the presence of BE and colonization with PA by Yule's association coefficient, R = 0.937.
- Strong correlation between the presence of BE and the S. Aureus isolate by Yule's association coefficient, R = 0.807.

By selecting the best fitting univariate regression model for regression analysis, the correlation between the indicators of current age, sweat test value, BMI value and FEO1 value was established (Tables 14, 15, 16).

Table 14. Influence of patient age on BMI values.

Model	b_0 , b_1	$T_{emp.}$	R^2 .100	F	σ	d
	13.8746 0.4443	17.851 5.6	42.17	31.362	2.756	1.76

Table 15. Effect of sweat test value on BMI value.

Model	b_0 , b_1	$t_{emp.}$	R^2 .100	F	σ	d
	97.7622 -2.1868	14.228 -2.299	28.41	5.287	4.372	1.68

Table 16. Influence of the sweat test value on the FEO1 value.

Model	b_0 , b_1	$T_{emp.}$	R^2 .100	F	σ	d
$Y_i = b_0 + b_1 X_i$	139.6125 -0.9408	6.605 -1.919	22.37	4.654	6.261	1.66

On the basis of the obtained results the following generalizations and conclusions can be drawn:

- As the age of CF patients increased by one year, the BMI value increased by 0.44 kg/m^2 (treatment effect).
- As the age of CF patients increased by one year, the FEO1 value decreased by 2.19%.
- As the sweat test value increased by one mmol/L, the FEO1 value decreased by 0.94%.

PART V. DISCUSSION OF RESULTS

1. Demographic features, CFTR mutation frequency and genotype distribution in CF patients

In this dissertation, data of 45 patients with proven CF who were previously enrolled is ECFPR were analyzed. The sample represents nearly ¼ of the patients with CF in Bulgaria. The ECFPR includes demographic and clinical data on more than 50 000 CF patients from 40 participating countries. After consent from patients to participate, information is submitted either from national registries or from local centres - in this case, the VEC. Participation of CF patients from the Northeast in the ERPM allows tracking of the course and treatment of the disease, as well as comparisons between participating countries, in order to increase knowledge about CF, promote new standards of care, and acquire epidemiological information to improve planning of public health interventions.

According to the ECFPR data on the gender distribution of patients, there is a slight male preponderance. For CF patients from Northeastern Bulgaria, the percentage of male patients was 62.20%. For Romania, for example, it is even higher. There has been a lot of work on the topic of why the disease progresses more in the female sex, with the main assumption being the influence of sex hormones on lung function by impairing mucociliary clearance and the immune response to infection.

Regarding age distribution, it is already known that from a purely paediatric disease with a poor prognosis, CF gradually becomes a disease of the elderly. In developed countries, a 70% increase in the number of adult patients is predicted between 2010 and 2025. According to ECFPR, 50% of CF patients are over 18 years of age. In contrast to most European countries, among our sample of patients, the percentage of elderly patients is low at 31.10%, and several reasons can be differentiated: lack of NS and therefore later diagnosis, introduction of CFTR-modulating therapy only in 2019, and last but not least, the socioeconomic conditions in the country and the structure of the health care system.

It is known that the average life expectancy of CF patients has increased significantly worldwide, but remains low at around 30 years for both men and women. The prognosis in countries of the European Union is better than that of those outside it and is highly dependent on

the quality of the healthcare system. Between 2012 and 2016, life expectancy was reported to be 43 years in the United States, 47 years in England and Germany, and 53 years in Canada. According to the 2021 EDPM, life expectancy for those born between 2018 and 2022 is projected to be 56 years. For the included patients from the Northeast, it is significantly lower, approximately 18 years, with the reasons again being late diagnosis, socioeconomic conditions, the functioning of the health care system, and the centralization of multidisciplinary teams.

The percentage distribution according to ethnicity of CF patients from Northeastern Bulgaria differs from the distribution for Bulgaria. In 2019, a large collaborative study of Bulgarian and Czech colleagues (Petrova et al), covering patients from all over the country, gave the following percentage distribution of CF according to patient ethnicity: 76.43% for Bulgarian ethnicity, 12.14% for Turkish ethnicity, and 11.43% for Roma ethnicity. For the patients included in the present study, the percentage of those of Bulgarian ethnicity was close to that for Bulgaria as a whole, but patients of Roma ethnicity accounted for only 4.44% at the expense of those of Turkish ethnicity - 26.67%.

Regarding the distribution of cystic fibrosis mutations, it is not surprising that the most frequent for the group of patients studied was F508del. Its frequency relative to all identified mutations was 64.45%, and the percentage of patients with at least one F508del was 84.45%. Worldwide, its frequency also reaches 70%, and in Europe 82.4% of CF patients have at least one F508del mutation, and it is more typical for Northern Europe - e.g. 83% in Denmark vs. 40-60% in Spain, Italy, Greece. For Bulgaria, again according to the publication of Petrova et al, F508del has an average frequency of 55% .

All other mutations in CF patients from Northeastern Bulgaria and Bulgaria as a whole, as well as in Europe, occur in only a few percent. The G542X and N1303K mutations are the next most frequent (after F508del) for the patients included in the study, for Bulgaria and for Europe. Even in Iceland, N1303K has a frequency of 46.4%. The Q220X mutation has an incidence of 3.37% in CF patients in Northeast Bulgaria and is among the ten most frequent mutations for Bulgaria, but not for Europe. The next most frequent for the participating patients was 574delA, which was identified in 2.25% of them. The frequency for Bulgaria is also 0.3%, but it is not among the most frequent mutations for Europe according to ECFPR.

The most common genotype is the homozygous F508del carrier. For the patients included in the thesis, its frequency was 42.22%. For Bulgaria it is 40% and is close to that in Europe. Its

frequency is lowest in Germany - less than 10%, and highest in Denmark and Albania - almost 70% of all genotypes.

The distribution of the F508del/Druga genotype, which has a similar mean frequency as F508del/F508del, ranging from 10% in Armenians to 50% in Ukrainians and Estonians, is also shown in the ERPM. For patients from Northeastern Bulgaria, its frequency is 28.98%, and for Bulgaria as a whole 40%. The Other/Other genotype has also been differentiated with an average frequency of nearly 20% for Europe as well as for Bulgaria, but with frequencies of up to 90% in Armenia and Germany. For patients with CF from northeastern Bulgaria, the frequency of genotypes that do not include any F508del mutation is 13.33% .

2. Impact of the absence of mass NS and the importance of early recognition of alarm symptoms on age of CF diagnosis

The reason for the late confirmation of the diagnosis in Bulgaria is the lack of mass NS. For the patients with CF included in the study from Northeast Bulgaria, the mean age of diagnosis was 2.14±5.32 years, varying according to the leading clinical manifestation. For countries such as the United States, Canada, Australia, Russia, Turkey, Brazil, and most European countries with NS present, diagnosis is confirmed early in infancy. For countries included in the EPPM, the median age of confirmation of CF diagnosis is 8 months. For Bulgaria it is 1.9 years. Serbia, Armenia, Belarus and Hungary have a similarly high mean age of diagnosis. Lithuania and Georgia have the highest such age at 2.7 and 2.6 years respectively. The ECFPR data also include the median age of diagnosis confirmation for patients who are now adults (over 18 years of age). The average for the participating countries is 7 years and is comparable to that for Bulgaria. For Spain, for example, there was a decline of 9.4 years in patients aged 18 years and over, at the expense of 8 months among all patients. This once again reflects the role of NS in timely diagnosis.

It has been reported that NS has a 98% disease detection rate, with false-positive results of about 5%. Neonates with positive NS in whom the diagnosis cannot be confirmed or excluded are referred to reference centres for follow-up, with 10% subsequently developing CF symptoms. Among the patients involved in the study, there was one patient with a cystic fibrosis mutation,

F508del, detected by NS abroad, who was referred to the VEC and subsequently diagnosed with the F508del/E831X genotype.

In the absence of mass NS, early recognition of alarming symptoms is relied upon to confirm a diagnosis of CF. The results obtained in this thesis show that the most frequent manifestations in patients from the Northeast are failure to thrive with diarrhea, failure to thrive with hypoproteinemia and anemia, failure to thrive and BOS, MI, and recurrent pneumonia. More indicative data for the Bulgarian population covering 167 patients, which is also a representative sample of patients in the country, are shown in the dissertation work of Baicheva, MD, with again failure to thrive and diarrhoea, failure to thrive and BOS and isolated BOS noted as the most frequent symptoms at confirmation of diagnosis.

Failure to thrive with or without diarrhoeal syndrome, in addition to being the most common reason for confirming a diagnosis of CF, is an early manifestation. It is well known that failure to thrive can be from an organic cause - a medical illness or from an inorganic cause - psychosocial factors. In this case, the cause was organic - malabsorption syndrome resulting from exocrine pancreatic insufficiency, and psychosocial factors were undoubtedly involved. The pancreas is one of the first organs affected in CF patients. Changes in it, as a result of accumulated viscous secretions, are already observed intrauterinally, with about 60% of newborns already having pancreatic insufficiency, with another 30% developing it in the next 36 months of life. It is important to differentiate these manifestations, especially in the absence of symptoms on the part of CF, from other diseases with such a clinic-most notably celiac disease, given the historical data. In patients from the Northeast, nonviraemia with diarrhea is the first manifestation of CF in 26.67% of patients, and for Bulgarian patients as a whole, it is about 50%.

In both developments, it is striking that patients with respiratory problems are diagnosed at a later age. Typical respiratory findings when confirming a diagnosis in older children, adolescents, and adults may include recurrent sinusitis, bronchitis, pneumonia, asthma that is not responsive to standard treatment, nasal polyposis, and BE of the lung on imaging studies.

For the patients from Northeastern Bulgaria, those with recurrent pneumonia were diagnosed the latest, with a mean age of diagnosis confirmation of 15.14±6.82 years. The reason for this is that children with CF are born with healthy lungs and the inflammation starts in the first weeks of life, often in the absence of symptoms and negative microbiological examinations. Another reason is that CF is a rare disease, and in the absence of other symptoms, more common

causes such as viral and bacterial lower respiratory tract infections and asthma come to the fore. Several of the patients included had been misdiagnosed and treated for asthma, which also caused a delay in confirming the diagnosis. CF is also characterized by symptoms such as coughing, difficulty breathing, and "wheezing" in the chest, with the difference being the presence of other alarming symptoms in addition to the lack of effect of step therapy. Cystic fibrosis-asthma overlap syndrome (CFAOS) is also a possibility in some cases.

Anaemia, hypoproteinaemia and oedema have long been known manifestations of CF, and in the past it was even classified as a separate form - the oedema-anaemia form of CF. A publication from the end of the last century with data from the Danish centre described that the age of diagnosis for this manifestation is low, between 1 and 5 months from birth. It is also one of the lowest for HEC patients, with an average age of about 4 months.

Vomiting and alkalosis with dysselectolemia or Pseudo-Barter syndrome are also important alarm manifestations in CF. Hypertrophic pyloric stenosis is known to be a more common cause of metabolic alkalosis in the first months after birth, with an incidence of 1-8 per 1000 live births in Caucasian children. Of the CF patients included in the present paper, there was one patient with vomiting and alkalosis as the leading clinical manifestations at diagnosis. In this patient, despite age-appropriate presentation, male gender, explosive vomiting "on fountain" and metabolic alkalosis with dyselectrolycemia, pyloric stenosis was ruled out by investigations, i.e., X-ray with contrast and ultrasound, and after a sweat test with a pathological value (115 mmol/L), the diagnosis of CF was confirmed by genetic analysis (p.R347P-homozygote).

Worldwide, MI occurs as the first manifestation of CF in 12.5% to 25.9% of newborns, for Bulgaria this percentage is between 15 and 20% (1,130). For CF patients from the Northeast, the rate is 11%. It is known that MI can be diagnosed ultrasonographically prenatally. When such a finding occurs, one proceeds according to algorithms developed to assess the risk for MI and CF and recommends delivery in a facility with a multidisciplinary pediatric team, intensive care, and pediatric surgery.

3. Genotype-phenotype correlations regarding respiratory system involvement, GIT and other manifestations such as sweat test values, CFRD and infertility

In addition to timely diagnosis of CF in the absence of mass NS, genotype-phenotype correlations are also important for disease prognosis. The results obtained from the analysis of the data of CF patients from Northeastern Bulgaria confirm that homozygous carriage of the F508del mutation is characterized by severe phenotypic manifestations. It has also been described in the literature as a "severe" mutation, as are other mutations without residual CFTR function, and phenotypically manifests with early exocrine pancreatic insufficiency, chronic obstructive pulmonary disease, high sweat test levels, and obstructive azoospermia.

For the patients included in this dissertation, in terms of lung involvement, F508del homozygotes had mean FEO1 values of approximately 80%. These values are similar to the mean FEO1 values for all Bulgarian patients included in the ECFPR, 85%, which is also related to the high frequency of the F508del/F508del genotype. While the next most frequent genotypes for patients from Northeast Bulgaria, F508del/Q220X and F508del/delA574, were associated with similar FEO1 values, patients with the F508del/G542X genotype had statistically significant higher mean values. Therefore, even mutations of the same class may be associated with different phenotypic manifestations. The mean FEO1 values of patients with the F508del/Druga genotype showed higher values compared to those of F508del homozygotes, 95.90%. Therefore, a comparison of this phenotypic manifestation according to the combination of mutation classes in CFTR was also performed and the results are consistent with previously published data of a more severe course of CF in mutation classes I and II.

In a Swedish study by Geborek et al. involving 266 CF patients with class I and II mutations, with respect to FEO1, results showed lower values for those patients carrying two class I mutations in their genotype.

Another prospective study by de Gracia et al. involving patients older than 18 years showed lower baseline FEO1 values and more pronounced loss of lung function at follow-up in patients with class I and II mutations .

For the CF patients included in the thesis from Northeastern Bulgaria, there was also a statistically significant difference in mean FEO1 values between patients with class I and II mutations and those with at least one mutation in a higher class. The mean FEO1 value for patients with genotype I - II/ I - II was the same as that for F505del homozygotes, again a result of the high frequency of the genotype.

A significant correlation was also found in terms of the decline in FEO1 and the increase in age of the patients, confirming the progression of the disease over time. Whereas end-stage lung disease used to include CF patients on the lung transplantation list, with the introduction of CFTR-modulating therapy this step is taken less frequently, not only because of the delay in progression but also because of the significant improvements in lung function.

Structural lung involvement in CF patients from the Northeast has been assessed by the presence of BE as the earliest predictor and found even in asymptomatic infants. In 69% of the included patients, BEs were proven by imaging studies, and it is assumed that their incidence is much higher due to the lack of CT performed in all patients, resulting in the lack of low-dose CT in the Northeast. According to the literature, BE are present in 30 to 40% of children aged 3 to 4 years and in up to 80% of children up to the age of 5 years. Regarding genotype-phenotype correlations, the results show a significantly higher involvement in patients with genotype I/II compared with II/II, and a higher percentage of BE in patients with at least one F508del compared with those with none in their genotype. Also confirming the results of de Gracia et al. - I and II mutations are associated with more severe lung involvement. The significant positive correlation between the presence of BE and increasing patient age was also confirmed by correlation analysis.

Regarding bacterial colonization in CF patients from the Northeast, half of the F508del homozygotes had chronic PA infection. The percentage is similar for patients with genotype I - II/ I - II. According to the ECFPR, 46% of patients in Bulgaria are colonized with PA, and among all participating countries the percentage is 23%. This shows on the one hand that F508del homozygotes can be considered a representative sample of CF patients in Bulgaria in terms of the percentage of colonization with PA and on the other hand that this percentage in Bulgarian patients in general and in those from the North-East of Bulgaria in particular is much higher than the European average. It should also be taken into account that some paediatric patients, especially those younger than 7 years of age, find it difficult to detach sputum and their colonisation is judged from examination of upper respiratory tract secretions, suggesting a lower than actual incidence (180, 213). When the correlations of this parameter were examined with respect to mutation class, the result obtained, a higher rate of chronic PA infection in patients with genotype I - II/ III - VI or VUS compared with those with genotype I - II/ I - II, was explained by the higher mean age of patients in the latter group. It is known that in childhood, 10 to 30% of CF patients are colonized with PA, whereas in adults this percentage is between 80 and 90% .For the included patients, in

addition to the influence of age, a correlation between FEO1 values and the presence of chronic PA infection was also described: its percentage was higher in patients with lower FEO1 values. The explanation here is twofold: on the one hand, colonization is the cause of CF disease progression, and on the other hand, the damaged lung is a favourable environment for colonization with pathogenic microorganisms.

Intermittent, non-chronic SA infection (sputum isolate at least once) in CF patients from Northeastern Bulgaria occurred in 33.33% of them. This percentage also occurs in patients homozygous for F508del and those with the F508/del genotype - again due to the high frequency of these genotypes. It was isolated in a higher percentage in the next most frequent genotypes F508del/Q220X, F508/574delA, F508del/G542X, as well as in the genotype Druga/Druga, which is explained by the younger age of the sample and their smaller number. According to literature data, colonization with SA begins at an early age, and is usually first isolated in up to 30% of infants in the first 6 months of life. Its presence in the lower airways is associated with pulmonary inflammation, even in the absence of clinical symptomatology. Its prevalence increases in the preschool years, and it is the most common pathogen during this period. According to the EPPM, 22.40% of patients under the age of 18 years had isolated it at least once, and in those over 18 years the percentage was 17.20%. These results were attributed to improved diagnosis and to the prevalence of isolates of other microorganisms with increasing age, such as PA, *Stenotrophomonas maltophilia*, and *Burcholderia cepacia complex.*

The prevalence of sinonasal involvement is also common and increases with age, with rates ranging from 6 to 48%. For patients from the Northeast, the rate was 44%, and there was no statistically significant difference in the prevalence of nasal polyposis among patients with the most common genotypes. With respect to the combination of mutation classes in the genotype of the CF patients included in the thesis, it was found that patients with at least one class I or class II mutation in their genotype had a higher percentage of sinonasal involvement than those whose genotype did not contain any mutation of the respective classes.

In 2000. Wang et al. published the results of a study of 147 patients with chronic rhinosinusitis and 127 healthy controls, demonstrating the higher frequency of at least one CFTR mutation in the former group of patients - 7%, compared to the latter - less than 1%. One patient was even diagnosed with CF by identification of two disease-defining mutations, which is another example of a phenotypic manifestation that may necessitate a search for an underlying genotype

(272). According to the frequency of blue-nasal involvement versus the class of mutations in the genotype, data by Berkhout et al. from a study in The Hague on 104 patients showed a higher rate in the presence of mutations from the first three classes.

Of the pulmonary impairment scores examined, the lesser pulmonary impairment of CF patients and the F508del/G542X genotype is notable, although it represents a combination of a class II mutation and a class I mutation, as are the F508del/Q220X and F508del/574delA genotypes. Also, patients homozygous for F508del at the same age may have varying degrees of lung damage, and despite a substantial number of studies, a definitive relationship between genotype and pulmonary phenotype has not been established. Variability, in addition to genotype, is determined by factors such as social status, adherence to therapy, health system status, and gene modifiers, and studies in CF patients with identical twins conducted after separation and living apart attributed their differences to the role of environmental and stochastic factors in 50% and to gene modifiers in the remaining 50%.

In contrast to pulmonary involvement, correlations in terms of GIT manifestations are more demonstrative. The pancreas is known to be the organ that is affected earliest in CF patients, with 60% already having pancreatic insufficiency at birth and 30% developing it in the following 36 months. It leads to malabsorption and therefore to assess the extent of damage for this thesis, actual BMI values were used to look for correlations. The results showed significantly lower BMI values in patients with a genotype including Class I or II mutations, with patients having a mean malnutrition BMI of 16.6 kg/m^2 . Patients with the F508del/F508del genotype had the same mean BMI values, again confirming the more severe phenotypic manifestations for the most common mutation. Prior to advances in CF patient care, BMI values decreased with increasing age due to disease progression. Data from the correlation analysis performed between BMI values and patient age in this thesis show an increase in BMI values with increasing age. The reason for these results is the implementation of enzyme replacement therapy, hypercaloric diet, and in some of the patients also the initiation of CFTR-modulating therapy.

CFLD in CF patients varies in severity and its percentage increases with increasing age (263). In the United States, its incidence reaches 47%. For Bulgarian CF patients, the incidence of CFLD is 51%. For CF patients from the Northeast included in the study, it is 44%, with significant genotype-phenotype correlations: patients with a genotype including class I or II mutations have a significantly higher rate of liver involvement (over 50%), as opposed to those with one mutation

in these classes or none. This is another example of the more severe impairment in mutations leading to the absence of CFTR, as well as the more compelling correlations regarding GIT.

MI in CF patients is the first manifestation in 12.5% to 25.9% of them. For patients from the Northeast, its incidence was 11.11%, with 80% of cases in patients with a genotype including class I or II mutations; accordingly, a demonstrable correlation was again observed for this manifestation on the GIT side.

According to the literature, pancreatitis is a rare manifestation of CF, with an incidence of 2%. In the patients included in the present study, varying degrees of pancreatic involvement occurred in 13%, and it is striking that it was significantly more common in patients with a genotype including mutations of classes III - VI or VUS. The L997F mutation, included in the F508del/L997F genotype, was identified in one patient from the Northeast of Bulgaria, with recurrent pancreatitis being the initial manifestation of CF, and subsequently chronic pancreatitis developed. This mutation was designated as VUS, a variant of uncertain significance.

In the literature, L997F is associated with idiopathic disseminated BE, recurrent pancreatitis, and hypertrypsinemia in infants. In a study by Castellani et al. including 32 patients with idiopathic pancreatitis, L997F was identified in 4 of them, and one even showed CF with L997F/F508del genotype .

In another study by Gomez et al. 49 infants with hypertrypsinemia were studied and the L997F mutation was identified in 4 of them.

Trininger et al. conducted studies of 14 adult patients with idiopathic chronic or acute relapsing pancreatitis, and L997F was identified in three of them .

A similar patient to the one included in this dissertation was described by Conklin et al. in 2008, a 4-year-old boy with recurrent pancreatitis, negative sweat test and pancreatic divisum. After measurement of differences in nasal transepithelial potentials and extended CFTR mutational analysis, two disease-defining mutations, F508del and L997F, were identified.

The sweat test values in the patient included in the thesis were borderline. It is known that mutations of classes after II are associated with lower sweat test values - they can be borderline or even in rare cases negative.

For the CF patients included in the study from Northeastern Bulgaria, those with genotypes including class I and II mutations had the highest rates, further evidence of the more severe and multi-organ involvement in these genotypes. It was also found by correlation analysis that higher

sweat test values, or initial classes in the genotype, were associated with lower FEO1 values - again an example of "severe" impairment.

The development of CFRD is also characteristic of patients with class I and II mutations and increases with age. In addition to the genotype, family history of diabetes, the presence of gene modifiers (e.g., TCF7L2), and the need to use corticosteroids as part of the therapeutic regimen in some CF patients are also important contributors to its occurrence. Approximately 11% of the patients included in this thesis have CFRD, with a mean age of nearly 22 years, and all have the F508del/F508del genotype. According to literature data, the prevalence is approximately 19% in patients under 18 years of age and 40-50% in adults. The reason for its low prevalence among patients from the Northeast is the significantly lower percentage of adult patients, as well as the low average age of the population.

While almost all male CF patients are infertile, 98%, women have reduced reproductive capacity, but with improved care and modern treatments, fertility rates have increased significantly. None of the CF patients from the Northeast of reproductive age had offspring, suggesting 100% infertility.

For female patients older than 18 years, there were only two successful deliveries on the background of initiated CFTR-modulating therapy, and the small sample of adult female patients included in the present work, 7 patients, should be taken into account. Worldwide, the number of pregnancies in CF patients is increasing, with many parallel studies on the safety of CFTR modulators.

4. Role of early confirmation of CF diagnosis for initiation of CFTR-modulating therapy in the setting of missing mass NS

The role of genotype-phenotype correlations in CF patients is important for predicting the course of the disease, as well as for searching for the underlying genotype in certain phenotypic manifestations and for timely diagnosis in the absence of mass NS. Even with its imminent introduction in Bulgaria, expected to start in 2025, there will be diagnosed cases of patients not covered by screening in the next 10 years or more. Furthermore, the screening introduced will not include inherited mutations in CFTR, but will be based on measurement of IRT only, which means

that genetic diagnosis may be influenced by a number of factors - financial, social, health system functionality, which in turn may lead to delays in the initiation of modern therapy.

From the present work, it became known that 71.34% of CF patients from Northeastern Bulgaria have already started their treatment with CFTR modulators, with an additional 19.05% suitable according to genotype. This percentage is also in line with the European average of 90% of CF patients having suitable mutations for CFRT modulating therapy.

The highest percentage of patients included in this dissertation were on triple therapy - Elexacaftor/Tezacaftor/Ivacaftor. The disadvantage is that its initiation is possible for patients over 6 years of age. During this time, especially in late diagnosed patients, severe involvement of a number of organs and systems is possible. Another disadvantage is the possible liver damage with the use of this drug. Given the high rate of CFLD, 44.62%, among CF patients in the Northeast, their regular follow-up after initiation of therapy to assess liver function is necessary.

Twice as many were on Ivacaftor/Lumikaftor therapy, and only one patient was on Ivacaftor therapy. For these medications, the appropriate age of initiation was over 2 years of age for Ivacaftor/Lumikaftor and over 4 months for Ivacaftor. This once again highlights the need for early diagnosis by NS and initiation of timely treatment before the development of permanent lesions - especially in the lung.

5. Algorithm based on alarming symptoms for early referral to an expert centre for CF diagnosis and subsequent treatment

On the basis of the results obtained and the problems discussed, an algorithm based on phenotypic manifestations was developed in order to confirm early the diagnosis of CF in the absence of mass neonatal screening, as well as in the period after its introduction, in which there will still be newly diagnosed, uncovered patients. The algorithm targets outpatient and inpatient practice and includes clinical guidelines for early recognition of alerting symptoms of CF and timely referral to a specialist centre for subsequent diagnosis, follow-up and treatment. It is important to note that CF does belong to the group of rare diseases and any symptom may be triggered by another more common disease (Figure 32).

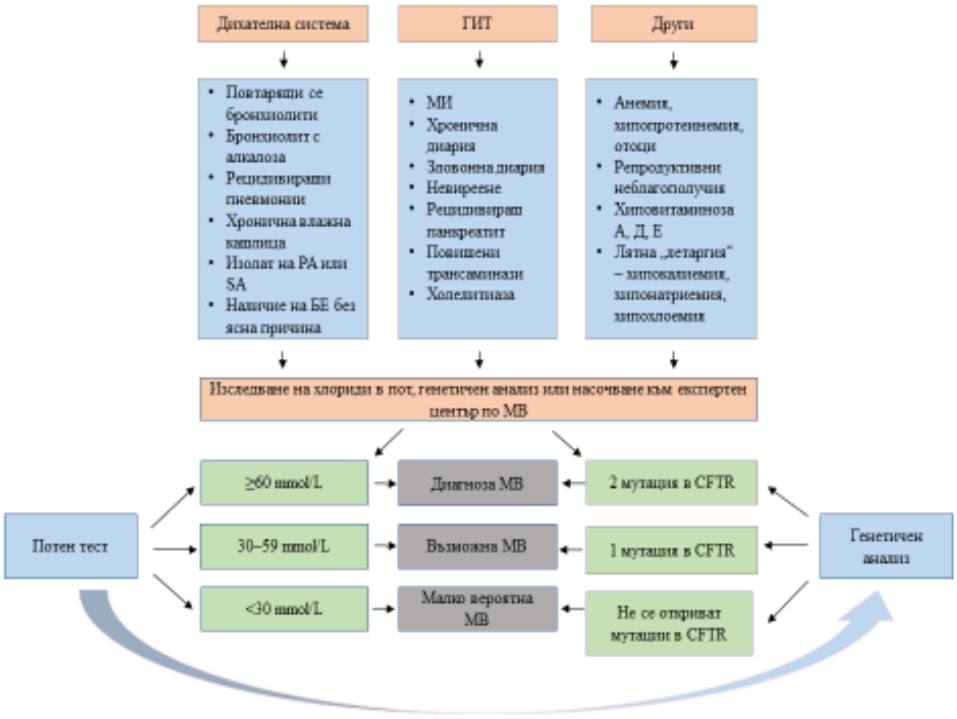


Figure 32. Algorithm based on alarming symptoms for early referral to an expert centre for CF diagnosis and subsequent treatment.

MAIN CONCLUSIONS

1. The inclusion of CF patients from Northeastern Bulgaria in the EPPM allows for the monitoring of the diagnostic, care and treatment processes, as well as comparison of the data entered with those for Europe and the other participating countries.

2. The most frequent mutation in CF patients from Northeast Bulgaria followed up in the VPP was F508del, followed by G542X, Q220X, 574delA and N13103K. The most frequent genotype in nearly 2/5 of patients was F508del/F508del, with the next most frequent genotypes being significantly rarer. At least one F508del mutation is found in 1/3 of CF patients. Among the Bulgarian ethnicity, homozygous F508del carriage prevails, and among the Turkish ethnicity, heterozygous F508del and other mutation carriage prevails.

3. The mean age at confirmation of CF diagnosis in patients from the Northeast is approximately 2 years, and the most common clinical manifestations are nonpneumonia and chronic diarrhea, followed by nonpneumonia, hypoproteinemia and anemia, nonpneumonia and BOS, MI, and recurrent pneumonia.

4. Genotype-phenotype correlations in respiratory system involvement, GIT, and other manifestations (sweat test values, CFRD) were found in CF patients from Northeastern Bulgaria followed up in the VPP.

   4.1 Patients with a genotype including class I or class II mutations have more severe respiratory phenotypic manifestations than patients with a genotype including class III - VI or VUS mutations. The G542X class I mutation has milder phenotypic manifestations in terms of pulmonary phenotype compared to the other mutations studied in the same class and to F508del.

   4.2 Patients with a genotype including class I or class II mutations have more severe phenotypic manifestations of GIT than patients with a genotype including class III - VI

or VUS mutations. Patients with the F508del/F508del genotype had intermediate malnutrition. The G542X class I mutation had more severe phenotypic manifestations in terms of BMI and CFLD compared to the other mutations studied in the same class and to F508del.

4.3 Patients with a genotype including class I or class II mutations had higher sweat test values than patients with a genotype including class III - VI or VUS mutations. The G542X class I mutation is characterized by lower sweat test values compared to the other mutations of the same class studied and to F508del. CFRD has only been reported in patients with the F508del/F508del genotype.

5. The algorithm with alerting symptoms for early diagnosis of CF, in the absence of mass NS, is applicable in Bulgarian outpatient and hospital practice.

CONTRIBUTIONS OF THE THESIS

1. Inclusion of CF patients from Northeast Bulgaria, followed up in the HPP, in the ERPM.
2. For the first time in Bulgaria, the distribution of CFTR mutations in CF patients from Northeastern Bulgaria followed up in the VPP was analyzed.
3. Confirmation of the more severe course of CF in patients with a genotype including class I and/or class II mutations compared to patients with a genotype including class III - VI or VUS mutations.
4. Detection of genotype-phenotype correlations in CF patients from Northeastern Bulgaria in terms of respiratory, GIT and other manifestations - sweat test, CFRD.
5. Creation of an actionable algorithm based on alerting symptoms for CF for early referral to a reference centre and subsequent diagnosis, treatment and follow-up.

CONCLUSION

Genotype-phenotype correlations in CF patients have been of interest since the end of the 20th century, but the present dissertation is the first work on the topic including CF patients from Northeastern Bulgaria.

In addition to the confirmatory nature of the study, namely that class I and class II mutations are associated with more severe phenotypic manifestations on the respiratory, GIT, and other sides - sweat test, CFRD, and worse prognosis, respectively, the diagnosis of CF based on alarming symptoms was also emphasized. Recognition of these is of particular importance in Bulgaria due to the lack of mass HC and although it is enshrined in the Maternal and Child Health programme, in the coming years there will still be reliance on these alerting symptoms for the unscreened patients.

The problem in diagnosis is determined by the fact that CF is a rare disease and there is experience with it only in a few of the regional cities of Bulgaria. Therefore, increasing awareness among outpatient and inpatient practice and knowledge of the alarming symptoms may prevent further delay in confirming the diagnosis.

The role of early detection of CF is continuously growing, especially with the therapeutic options available in developed countries and, as of 2019, in Bulgaria. The proven benefits of CFTR-modulating therapy are yet another reason to strive to confirm the diagnosis before the development of severe and irreversible changes, especially in the lungs of CF patients.

In conclusion, knowledge of genotype-phenotype correlations in CF is useful both for predicting the course of the disease and preventing complications, and for diagnosis by searching for the underlying genotype based on alarming phenotypic manifestations.

SCIENTIFIC PUBLICATIONS AND ANNOUNCEMENTS RELATED TO THE DISSERTATION

1. Nikolova M, Dyankova J, Hristov I, Georgieva M. Pulmonary colonization with Pseudomonas Aeruginosa in patients with cystic fibrosis. *Proceedings of the Union of Scientists - Varna*. Series medicine and ecology. 2019; 24 - 1: 54 - 60

2. Nikolova M, Koleva K, Rasheva N, Galabova M, Stoyanova M, Ivanov G, Dobrudjanska N, Georgieva M. L997F mutation phenotype - recurrent pancreatitis in a boy with Cystic Fibrosis. *Journal of Pediatrics*. 2020; 4: 49 - 52

3. Nikolova M, Galabova M, Dobrudjanska N, Rasheva N, Koleva K, Georgieva M. CFTR modulator therapy for cystic fibrosis. *Book of Proceedings - Symposium "Alumni Club and Friends"* - Medical University - Varna. 2020; 83-86

4. Nikolova M, Georgieva M, Petrova G. SARS-COV-2 and cystic fibrosis. *Journal of Pediatrics*. 2021; 4: 42 - 47

5. Nikolova M, Galabova M, Koleva KR, Rasheva N, Dobrudjanska N, Georgieva M. Management of adherence to therapy in cystic fibrosis. Comparison between two patients. *Journal of Pediatrics Plus*. September, issue 6/2022; 54-63

ACKNOWLEDGEMENTS

In connection with the development of this dissertation, I express my heartfelt Thanks:

- to my supervisors Prof. Dr. Miglena Georgieva and Prof. Dr. Gergana Petrova - for their understanding, patience and support;
- to all colleagues at the Pediatrics Clinic at St. Marina University Hospital - for their dedication to the care of CF patients;
- CF patients and their families - for their participation and support;
- to my family for their patience and support.