UNIVERSIDADE DE LISBOA
FACULDADE DE MEDICINA

Pneumococcal invasive disease in adults (2015-2017): epidemiological and molecular characterization of *Streptococcus pneumoniae* and genomic analysis of emerging clones

INÊS MARIA MARQUES TEODORO

Orientador: Prof. Doutora Ana Catarina da Silva-Costa

Dissertação especialmente elaborada para obtenção do grau de Mestre em Microbiologia Clínica e Doenças Infeciosas Emergentes

2020

UNIVERSIDADE DE LISBOA
FACULDADE DE MEDICINA

Pneumococcal invasive disease in adults (2015-2017): epidemiological and molecular characterization of *Streptococcus pneumoniae* and genomic analysis of emerging clones

INÊS MARIA MARQUES TEODORO

Orientador: Prof. Doutora Ana Catarina da Silva-Costa

Dissertação especialmente elaborada para obtenção do grau de Mestre em Microbiologia Clínica e Doenças Infeciosas Emergentes

2020

A impressão desta dissertação foi aprovada pelo Conselho Cientifico da Faculdade de Medicina de Lisboa em reunião de 30 Junho de 2020.

ACKNOWLEDGEMENTS

First, I would like to acknowledge my thesis supervisor, Professora Doutora Catarina Silva-Costa, for allowing me to develop this work and further my experience in the clinical microbiology field, for the support and shared knowledge during the research time and writing of this dissertation. I am incredibly grateful to Professor Doutor José Melo Cristino and Professor Doutor Mário Ramirez for the given opportunity to integrate their work team.

I would also like to thank all my colleagues in Mario Ramirez lab with whom I had the pleasure to work especially to my labmates Beatriz Esteves, Carolina Castro, Joana Silva and Soraia Guerreiro for helping me during the development of this thesis and for the companionship. I thank all bioinformatics members of Mario Ramirez lab, specially to Inês Mendes for the help in the bioinformatic analysis and the support in the writing process of this thesis.

I want to thank all my family, specially my parents and sister for the support in my academic career and life in general and to my friends, specially to Margarida Figueira for the support and friendship.

ABSTRACT

The introduction of the 13-valent pneumococcal conjugate vaccine (PCV13) for children in Portugal resulted in significant changes in the serotype distribution of the pneumococcal population responsible for invasive pneumococcal disease (IPD) not only in children, but also in adults, consistent with herd protection. A 23-valent pneumococcal polysaccharide vaccine (PPV23) is also available in Portugal since 1996, although adult uptake is believed to be low. The effect of broader PCV13 uptake in children in adult IPD is currently unknown. We aimed to characterize the pneumococcal population causing adult IPD (≥18 years old) in Portugal after the introduction of PCV13 in the Nacional Immunization Plan for children, in 2015.

A total of 1608 *Streptococcus pneumoniae* isolates were recovered from adult IPD in 61 hospitals in Portugal between 2015 and 2017. These strains were characterized phenotypically (serotyping and antimicrobial susceptibility testing) and by molecular methods (MLST). The results obtained were compared with those from previous studies to assess the effect of PCV13 introduction in Nacional Immunization Plan in the studied population.

Among the 1608 isolates, 56 different serotypes were found. The most common serotypes were serotype 8 (17.8%, n=287), 3 (14.7%, n=236), 22F (7.4%, n=119), 14 (6.9%, n=111), 19A (6.2%, n=100) and 9N (4.1%, n=66). Most of the cases corresponded to serotypes exclusively found in the PPV23 (42.4%, n=681). The PCV13 serotypes were responsible for 37.6% (n=605) of the cases, and 7-valent pneumococcal conjugate vaccine (PCV7) serotypes accounted for a small fraction of the cases (14.0%, n=225). Penicillin and erythromycin non-susceptibility were detected in 14.9% (n=239) and 16.4% (n=263) of the bacterial isolates, respectively. PCV7 and serotypes exclusively found in PCV13 were responsible for 51.0% (n=122) and 15.5% (n=37) of the penicillin non-susceptible isolates and 37.6% (n=99) and 19.0% (n=50) of the erythromycin resistant isolates, respectively. Regarding all collection, 31.2% (n=501) of the strains were non-susceptible to at least one of the tested antibiotics, where serotypes 14 and 19A were the most common.

MLST analysis was performed in 726 strains, where 56 different clonal complexes were found. PCV13 serotypes presented a high genetic diversity, while serotypes exclusively found in PPV23 and non-vaccine serotypes (NVT) presented a lower genetic

diversity. Thirteen major clonal complexes (CC) were defined in this study that accounted for 86.4% (n=627) of the isolates: CC156, CC180, CC433, CC378, CC97, CC235, CC439, CC199, CC260, CC315, CC994, CC30 and CC191. Eleven of the 43 PMEN clones were at least double-locus-variant of 33.6% (n=244) of the isolates, however no significant association was established between clonal complexes and penicillin and erythromycin non-susceptible isolates. Serotyping was also performed using *in silico* methods, where both approaches (SeroBa and PneumoCAT) were found to be good alternatives to conventional methods for the determination of *S. pneumoniae* serotypes.

Comparison with previous studies revealed an increase in the proportion of serotypes 8 and 12F, while serotypes 1, 12B and 7F decreased. Non-susceptible serotypes 8 and 22F isolates increased, whereas non-susceptible 19A isolates decreased. Although not significant, serotype 3 also presented an increase in the proportion of non-susceptible isolates to some of the antimicrobials tested. Regarding genetic lineages, sequence type (ST) 53 (serotype 8) increased, while ST191 (serotype 7F) and ST276 (serotype 19A) decreased.

Even with the introduction of PCV13 in the National Immunization Plan, serotype 3 is still one of the dominant serotypes in adult IPD in Portugal, along with serotypes 14 and 19A (all serotypes present in PCV13). The increase in serotypes that do not belong to PCV13, especially serotype 8, 22F and 9N is also of concern, particularly serotypes 8 and 22F where antimicrobial non-susceptible isolates increased. With these, our data suggests that, in a situation of higher vaccination coverage, PCV13 serotypes are still significant causes of adult IPD, especially serotypes 3. However, serotypes not present in PCV13 have been showing to be an important cause of adult IPD, particularly serotypes 8 and 22F, reinforcing the continuous need for surveillance studies.

Key-words: *Streptococcus pneumoniae*; adult invasive pneumococcal disease; phenotype; genotype; PCV13.

RESUMO

Streptococcus pneumoniae é um dos agentes patogénicos mais comuns do ser humano, responsável por elevadas taxas de mortalidade e morbilidade. Este microrganismo coloniza, tipicamente, a nasofaringe das crianças, podendo tornar-se patogénico, levando ao desenvolvimento de doenças pneumocócicas invasivas como a pneumonia, meningite e bacteriemia, ou não invasivas como otite média ou sinusite. Tem como principais grupos de risco crianças, indivíduos imunocomprometidos e idosos. Com a introdução das vacinas pneumocócicas conjugadas, observou-se um decréscimo na incidência da doença pneumocócica invasiva em crianças, assim como alterações na distribuição dos serotipos, também na população adulta. Este decréscimo deveu-se sobretudo à redução da incidência dos serotipos incluídos nestas vacinas. Contudo, observou-se um aumento na proporção de serotipos não vacinais, assim como a persistência de alguns serotipos vacinais (como foi o caso do serotipo 3 que se encontra presente na vacina pneumocócica conjugada 13-valente).

A vacina pneumocócica conjugada 7-valente (PCV7, inclui os serotipos 4, 6B, 9V, 14, 18C, 19F e 23F) foi introduzida em Portugal em 2001. Apesar de disponível apenas no sector privado, a cobertura desta vacina foi aumentado gradualmente ao longo dos anos, tendo-se estimado uma cobertura vacinal de 75% das crianças com ≤2 anos em 2008. Tal como nas crianças, nos adultos observou-se uma redução da proporção de doença pneumocócica invasiva causada por serotipos abrangidos pela vacina PCV7. No entanto, os serotipos 1, 3, 7F e 19A emergiram como causas importantes de doença pneumocócica invasiva nos adultos pós-PCV7 (estudo de 2006 a 2008). Em meados de 2009, a vacina pneumocócica conjugada 10-valente (PCV10, engloba todos os serotipos presentes na vacina PCV7, mais os serotipos 1, 5 e 7F) ficou disponível, seguindo-se em 2010 a disponibilização da vacina pneumocócica conjugada 13-valente (PCV13, abrange todos os serotipos presentes na vacina PCV10, mais os serotipos 3, 6A e 19A), ambas no sector privado. A introdução desta última vacina em Portugal resultou numa alteração da distribuição dos serotipos da população de *S. pneumoniae* responsáveis por doença pneumocócica invasiva em adultos, apesar de estes permanecerem como importantes causas de doença pneumocócica invasiva nos adultos, consistente com o efeito de proteção de grupo. A vacina pneumocócica polissacarídica 23-valente (PPV23, integra todos os serotipos presentes na vacina PCV13, à exceção do serotipo 6A, mais os serotipos 2, 8, 9N, 10A, 11A, 12F, 15B, 17F, 20, 22F e 33F) encontra-se disponível em

Portugal desde 1996, ainda assim estima-se que a cobertura desta vacina seja baixa. Em 2015, a vacina PCV13 foi introduzida no plano nacional de vacinação das crianças, esperando-se, deste modo, uma maior cobertura vacinal, o que pode levar a alterações na população deste microrganismo causador de doença pneumocócica invasiva em criança e, consequentemente, em adultos. Assim, este estudo teve como objetivo caracterizar a população de *Streptococcus pneumoniae* causadora de doença invasiva pneumocócica nos adultos (indivíduos com ≥18 anos) em Portugal, após a introdução da vacina PCV13 no plano nacional de vacinação para as crianças em 2015.

Um total de 1608 estirpes de *Streptococcus pneumoniae* responsável por doença invasiva pneumocócica em adultos foram recebidas por 61 hospitais de todo o país entre 2015 e 2017. Estas estripes foram caracterizadas fenotipicamente (através de serotipagem e teste de suscetibilidade aos antimicrobianos) e genotipicamente (através de métodos moleculares, obtendo-se o perfil de *multilocus sequence typing* por sequenciação total do genoma). Os resultados obtidos foram ainda comparados com os de estudos anteriores, de forma a avaliar o efeito da introdução da vacina PCV13 no plano nacional de vacinação na população em estudo.

De entre as 1608 estirpes de pneumococos, foram encontrados 56 serotipos diferentes, sendo que os mais frequentes foram os serotipos: 8 (17.8%, n=287), 3 (14,7%, n=236), 22F (7,4%, n=119), 14 (6,9%, n=111), 19A (6,2%, n=100) e 9N (4,1%, n=66). Grande parte dos casos pertenceram a serotipos exclusivos da vacina PPV23 (42,4%, n=681). Os serotipos da vacina PCV13 foram responsáveis por 37,6% (n=605) dos casos em estudo, sendo que serotipos da vacina PCV7 foram responsáveis por uma pequena fração dos casos em estudo (14,0%, n=225). A não suscetibilidade à penicilina e à eritromicina foi detetada em 14,9% (n=239) e 16,4% (n=263) das estirpes estudadas, respetivamente. Serotipos da vacina PCV7 e serotipos exclusivos da vacina PCV13 foram responsáveis por 51,0% (n=122) e 15,5% (n=37) dos casos de não suscetibilidade à penicilina e 37,6% (n=99) e 19,0% (n=50) dos casos de resistência à eritromicina, respetivamente. Em relação à coleção total de estirpes, 31,2% (n=501) das estirpes apresentaram não suscetibilidade a pelos menos um dos antimicrobianos testados, onde os serotipos 14 e 19 foram os mais comuns.

A analise dos perfis de *multilocus sequence typing* foi realizada em 50% das estipes que apresentaram mais do que 10 estirpes do mesmo serotipo, tendo-se obtido um total de 726 estirpes sequenciadas, onde foram detetados 56 complexos clonais diferentes. Serotipos da vacina PCV13 apresentaram uma elevada diversidade genética, enquanto

que os serotipos exclusivos da vacina PPV23 e serotipos não vacinais apresentaram uma diversidade genética baixa. Treze complexos clonais principais foram definidos neste estudo, englobando 86,4% (n=627) das estirpes sequenciadas: complexo clonal (CC) 156, CC180, CC433, CC378, CC97, CC235, CC439, CC199, CC260, CC315, CC994, CC30 e CC191. Onze dos 43 clones PMEN foram identificados como sendo, pelo menos, *double-locus-variants* de 33,6% (n=244) das estirpes sequenciadas. No entanto, não se estabeleceu nenhuma associação significativa, depois da correção para testes múltiplos, entre complexos clonais e não suscetibilidade à penicilina e à eritromicina. Determinaram-se, ainda, os serotipos através de métodos *in silico*, onde se observou que ambos os softwares SeroBa e PneumoCAT são boas alternativas aos métodos convencionais para a determinação dos serotipos de *S. pneumoniae*.

Em relação à analise comparativa com estudos anteriores, observou-se um aumento na proporção dos serotipos 8 e 12F, ao passo que os serotipo 1, 12B e 7F apresentaram um decréscimo significativo. A proporção de serotipos 8 e 22F não suscetíveis a antimicrobianos aumentou quando comparada com a proporção do estudo anterior, enquanto que a proporção de serotipo 19A não suscetível a antimicrobianos diminuiu significativamente. Apesar de não significativo após a correção para testes múltiplos, o serotipo 3 também apresentou um aumento na sua proporção de estirpes não suscetíveis a determinados antimicrobianos testados. Em relação à analise comparativa das linhagens genéticas, *sequence type* (ST) 53 (constituído maioritariamente por estirpes do serotipo 8) aumentou em proporção, ao passo que os ST191 (constituído sobretudo por estirpes do serotipo 7F) e ST276 (constituído principalmente por estirpes do serotipo 19A) diminuíram, todos de modo significativo.

Deste modo, observou-se que mesmo com a introdução da vacina PCV13 no plano nacional de vacinação, o serotipo 3 continua a ser um dos serotipos mais frequente como causa de doença invasiva pneumocócica em adultos em Portugal, juntamente com os serotipos 14 e 19A (todos estes serotipos estão presentes na vacina PCV13). O aumento da proporção de serotipos não abrangidos pela vacina PCV13, nomeadamente os serotipos 8, 22F e 9N, são de especial preocupação, principalmente o caso dos serotipos 8 e 22F onde se observou um aumento na proporção de estirpes não suscetíveis a antimicrobianos. Assim, podemos concluir com este estudo que numa situação de elevada cobertura vacinal, os serotipos da vacina PCV13 ainda são causas importantes de doença pneumocócica invasiva nos adultos, em particular o serotipo 3. Contudo, os serotipos não abrangidos por esta vacina têm-se mostrado, também, relevantes causas de doença

pneumocócica invasiva em adultos em Portugal, especialmente os serotipos 8 e 22F, reforçando desta forma a continua necessidade de realizar estudos de vigilância epidemiológica.

Palavras-chave: *Streptococcus pneumoniae*; doença invasiva pneumocócica em adultos; fenótipo; genótipo; PCV13.

ABBREVIATIONS

addPCV10 – serotypes present in the PCV10, but not in the PCV7

addPCV13 – Serotypes present in the PCV13, but not in the PCV7 and PCV10

AW – Adjusted Wallace coefficient

CBPs – Choline-Binding Proteins

C – Chloramphenicol

CC – Clonal Complex

CSF – Cerebrospinal Fluid

CLSI – Clinical and Laboratory Standards Institute

CPS – Capsular Polysaccharide

CT – Cefotaxime

DA – Clindamycin

DLV – Double-locus-variant

DNA – Deoxyribonucleic Acid

EPNSP –Erythromycin and Penicillin Non-Susceptible Pneumococci

ERSP –Erythromycin Resistant *S. pneumoniae*

ERY – Erythromycin

FDR – False Discovery Rate

IgA1 – Immunoglobulin A1

IPD – Invasive Pneumococcal Disease

LEV – Levofloxacin

LZ – Linezolid

M – Phenotype M

MDR – Multi-Drug Resistance

MLS_B – Macrolides, Lincosamides and Streptogramin B (resistance phenotype)

- $cMLS_B$ – Constitutive MLS_B resistance phenotype
- $iMLS_B$ – Inducible MLS_B resistance phenotype

MLST – Multilocus Sequence Typing

NIP – Nacional Immunization Plan

NOR – Norfloxacin

NVT – Non-Vaccine Serotypes

OR – Odds Ratio

PBPs – Penicillin-Binding Proteins

PCR – Polymerase Chain Reaction

PCV – Pneumococcal Conjugate Vaccine

PCV7 – 7-valent Pneumococcal Conjugate Vaccine

PCV10 – 10-valent Pneumococcal Conjugate Vaccine

PCV13 – 13-valent Pneumococcal Conjugate Vaccine

PEN – Penicillin

Ply – Pneumolysin

PMEN – Pneumococcal Molecular Epidemiology Network

PNSP – Penicillin Non-Susceptible Pneumococci

PPV23 – 23-valent Pneumococcal Polysaccharide Vaccine

rPCR – Real-time PCR

RNA – Ribonucleic Acid

rRNA – Ribosomal RNA

SID – Simpson's Index of Diversity

SLV – Single-Locus-Variant

ST – Sequence Types

SXT – Trimethoprim-Sulfamethoxazole

TEL – Telithromycin

TET – Tetracycline

VA – Vancomycin

WGS – Whole Genome Sequencing

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ..... **i**

ABSTRACT ..... **iii**

RESUMO ..... **v**

ABBREVIATIONS ..... **ix**

TABLE OF CONTENTS ..... **xi**

FIGURES AND TABLES: ..... **xiii**

1. INTRODUCTION **1**

1.1 Characterization of S. pneumoniae **3**

1.1.1 Identification 3

1.1.2 Virulence Factors 5

1.2 Host-pneumococcus interaction **10**

1.2.1 Host colonization 10

1.2.2 Pneumococcal Infections 13

1.2.3 Epidemiology 15

1.3 Vaccines **16**

1.3.1 Available pneumococcal vaccines 16

1.3.2 Serotype Replacement 18

1.4 Antimicrobial Resistance **19**

1.4.1 Definition of resistance 19

1.4.2 Mechanisms of Resistance 20

1.4.3 Epidemiology 24

1.4.4 PMEN Clones 25

1.5 Typing Methods **26**

1.5.1 Serotyping 26

1.5.2 Molecular typing 27

2. AIMS OF THE STUDY **29**

3. MATERIALS AND METHODS **30**

3.1 Collection of strains **30**

3.2 Serotyping 31

3.3 Antimicrobial susceptibility testing 32

3.4 DNA extraction and sequencing 33

3.5 Statistical Analysis 35

4. RESULTS 36

4.1 Isolate collection 36

4.2 Serotype distribution 36

4.3 Antimicrobial susceptibility 39

4.4 SeroBa 43

4.5 Molecular Characterization 44

5. DISCUSSION 53

6. CONCLUSION 62

7. FUTURE PERSPECTIVES 63

8. REFERENCES 64

9. SUPPLEMENTARY DATA 84

FIGURES AND TABLES:

Figure 1 – Illustration of *S. pneumoniae*. A- Gram stain of *S. pneumoniae*, reproduced from https://www.sciencephoto.com/media/13022/view/streptococcus-bacteria and B- electronic microscopy of *S. pneumoniae* non-encapsulated, reproduced from https://ppdictionary.com/bacteria/gpbac/pneumoniae.htm. 1

Figure 2 – Timeline of *Streptococcus pneumoniae* discoveries (Purple – discoveries related to *S. pneumoniae* and Orange – important discoveries related). Adapted from [3,6,124,125,129,189–193,7,48,73,79,80,90,91,123]..... 2

Figure 3 – Illustration of *S. pneumoniae* colonies on blood agar with alpha hemolysis. A- Mucoid colonies; B- Colonies not mucoid. 4

Figure 4 – Scheme of the principal pneumococcus virulence factors. PsaA, pneumococcal surface adhesin A; PspA, pneumococcal surface protein A; PspC, pneumococcal surface protein C; PiaA, pneumococcal iron acquisition A; PiuA, pneumococcal iron uptake A; PitA, pneumococcal iron transporter A. Reproduced from [18]..... 6

Figure 5 – Representation of the life cycle of Streptococcus pneumoniae and IPD. Reproduced from [70]. 11

Figure 6 – Global incidence rates of IPD in children with <5 years old in 2010. Reproduced from https://www.abcombibio.com/pneumococcal. 15

Figure 7 – Confirmed IPD cases per 100 000 population reported by Europeans countries in 2017. Countries: Austria, Bulgaria, Croatia, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom. Reproduced from [78]. 15

Figure 8 - Representation of serotypes included in pneumococcal vaccines. 17

Figure 9 – Representation of capsular switching. Reproduced from [81]. 19

Figure 10 – Proportion of S. pneumoniae combined non-susceptibility to penicillin and macrolides reported by Europeans countries in 2017. Countries: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, united kingdom. Reproduced from [126]..... 24

Figure 11 – Scheme of macrolide phenotypes identification. A- phenotype cMLS$_B$; B- phenotype iMLS$_B$ and C- phenotype M. “ERY”- Erythromycin; “DA”- Clindamycin. 33
Figure 12 – PCV vaccine serotypes distribution by age groups ([18-49]; [50-64] and ≥65 years old) in 2015, 2016 and 2017. Arrows represent the serotypes presented in PCV7, PCV10 and PCV13. "Others" refers to all serotypes presented in the study population, except for serotypes presented in PCV vaccines. 37
Figure 13 – Distribution of Non-PCV serotypes by age groups ([18-49]; [50-64] and >65 years old) in 2015, 2016 and 2017. A - Distribution of serotypes presented only in PPV23, where “Others” refers to all serotypes present in PPV23, which have one age group and year with <10 isolates. B – Distribution of NVT serotypes, where “Others” refers to all serotypes not included in a pneumococcal vaccine (PCVs and PPV23), which have one age group and year with <10 isolates. 38
Figure 14 – Resume of the antimicrobial non-susceptibility isolates included in this study. “PEN”- penicillin; “LEV”- levofloxacin; “NOR”- norfloxacin; “ERY”- erythromycin; “DA”- clindamycin; “TEL”- telithromycin; “TET”- tetracycline; “C”- chloramphenicol; “SXT”- trimethoprim-sulfamethoxazole; “CT”- cefotaxime. 39
Figure 15 – Distribution of penicillin and erythromycin susceptibility profiles for isolates of *S. pneumoniae* serotypes by study year (2015, 2016 and 2017). The arrows represent serotypes presented in PCV7, PCV10 and PCV13. "Others"- all serotypes presented in the study population, except for serotypes presented in PCV vaccines and serotypes 6C, 11A, 15A and 33F; “PNSP”- penicillin non-susceptible isolates; “ERSP”- erythromycin resistant isolates; and “EPNSP”- isolates presenting erythromycin resistance and penicillin non-susceptibility. 41
Figure 16 – Representation of all STs identified and indication of their CCs, highlighting the major CCs. STs linked by a line are SLVs and STs in grey are singletons. The scheme was done using PHYLOViZ [144]. 47
Figure 17 – Distribution of STs according to serotypes of the isolates causing adult IPD in 2015-2017 and expressing serotypes included in the conjugate vaccines. The respective SID values are indicated on top the bars and their respective confidence intervals are in parenthesis. “f” – represents the STS considered by goeBURST as founders of a CC; “p” – represents STs that matched with PMEN clones; “a”, “b”, “c” and “d” – represent STs that belong to the same CC in each serotype; “Others” – represent the isolates included in STs with <4 isolates..... 48

Figure 18 – Distribution of STs according to serotypes of the isolates causing adult IPD in 2015-2017 and expressing serotypes included exclusively in the PPV23 vaccine. The respective SID values are indicated on top the bars and their respective confidence intervals are in parenthesis. “f” – represents the STS considered by goeBURST as founders of a CC; “p” – represents STs that matched with PMEN clones; “a” – represent STs that belong to the same CC in each serotype; “Others” – represent the isolates included in STs with <4 isolates. 49
Figure 19 – Distribution of STs according to serotypes of the isolates causing adult IPD in 2015-2017 and expressing NVT serotypes. The respective SID values are indicated on top the bars and their respective confidence intervals are in parenthesis. “f” – represents the STS considered by goeBURST as founders of a CC; “p” – represents STs that matched with PMEN clones; “Others” – represent the isolates included in STs with <4 isolates; “NA” – Not applied. 49
Figure 20 – Representation of the STs and indication of PNSP, ERSP and EPNSP isolates. STs linked by a line are SLVs and STs with no CC associated are singletons. The scheme was done using PHYLOViZ [144]. 52

Table 1 – Risk factors associated with IPD. Adapted from [71,72]. 13
Table 2 – Former and current CLSI breakpoints for penicillin in S. pneumoniae. Adapted from [106]. 20
Table 3 – Resume of risk factors for resistance. Adapted from [105,107]. 20
Table 4 – Concentration and breakpoints for used antibiotics. 32
Table 5 – Simpson's index of diversity and respective 95% confidence intervals for serotypes distribution per study year. 39
Table 6 – Resume of the most frequent serotypes among non-susceptible isolates in adults between 2015 and 2017. 42
Table 7 – Serotypes distribution of the most frequent STs found in the 13 major CCs (n≥10 isolates) identified by goeBURST. 45
Table 8 – Difference between serotypes associated with STs found in the study and in the MLST database. 50

Table S1 – Results obtained when examined the association between age groups and serotypes. 84
Table S2 – Results obtained when examined the correlation of serotypes between 2012-2017. 84
Table S3 – Results obtained when examined the association between penicillin and erythromycin resistant profiles and serotypes. 84
Table S4 – Results obtained when examined the association between age groups and penicillin and erythromycin resistance profile. 85
Table S5 – Results obtained when examined the association between age groups and resistance to clindamycin. 85
Table S6 – Results obtained when examined the association between serotype and resistance to clindamycin, tetracycline and trimethoprim-sulfamethoxazole and MDR.85
Table S7 – Results obtained when examined the correlation between serotypes and resistance to some of the antibiotics analyzed during 2012-2017. 86
Table S8 – Serotype distribution of the STs found in CCs with <10 isolates by goeBURST..... 86
Table S9 – Results obtained when examined the correlation between STs and the study period (2015-2017). 88
Table S10 – Results obtained when examined the correlation between CCs and the study period (2015-2017). 88
Table S11 – STs that were associated with PNSP, ERSP and EPNSP. 88
Table S12 – Results obtained when examined the correlation of STs over time (2008-2011 and 215-2017). 89
Table S13 – Results obtained when analyzed the trend of serotype class between 2012-2017. 90
Table S14 – Results obtained when analyzed the trend of antimicrobials between 2012-2017. 90

1. INTRODUCTION

Streptococcus pneumoniae, also known as the pneumococcus, is responsible for severe infections such as bacteraemia, meningitis and pneumonia among children and adults worldwide [1]. With a high morbidity and mortality risk, this microorganism affects mostly young children, immunocompromised patients and the elderly [2,3]. It is capable of colonize asymptomatically the human nasopharynx, mostly in children [2,3] and, with the ability to adapt through genetic exchange and express diverse virulent factors, it is one of the *Streptococcus* with greater success as a human pathogen [4].

This pathogen is an encapsulated, Gram-positive, α-hemolytic and catalase-negative cocci that belongs to the viridans group of streptococci[1]. The cells are oval with 0,5 to 1,2 μm of diameter and arranged in pairs (diplococci) or short chains [2,5]. Colony morphology differs if the strain is encapsulated: generally large, round, and mucoid; or non-encapsulated: small and flat, and usually has a dimpled appearance in the central portion due to the expression of autolysin [2]. Pneumococcus is a facultative anaerobe that needs an enriched media supplemented with blood or serum and an atmosphere enhanced with 5-10% of CO_2 at 35-37ºC [5].

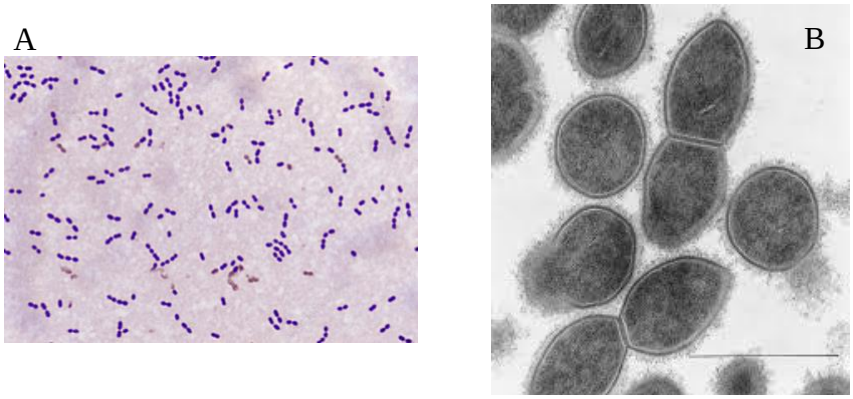

Figure 1 – Illustration of *S. pneumoniae*. A- Gram stain of *S. pneumoniae*, reproduced from https://www.sciencephoto.com/media/13022/view/streptococcus-bacteria and B- electronic microscopy of *S. pneumoniae* non-encapsulated, reproduced from https://ppdictionary.com/bacteria/gpbac/pneumoniae.htm.

[1] **Viridans group of streptococci –** Group of *Streptococcus* which forms a green pigment because of the partial hemolysis of blood agar. This group is subdivided into five distinct clinical groups: Anginous; Mitis (where *S. pneumoniae* is integrated); Mutans; Salivarius and Bovis [2].

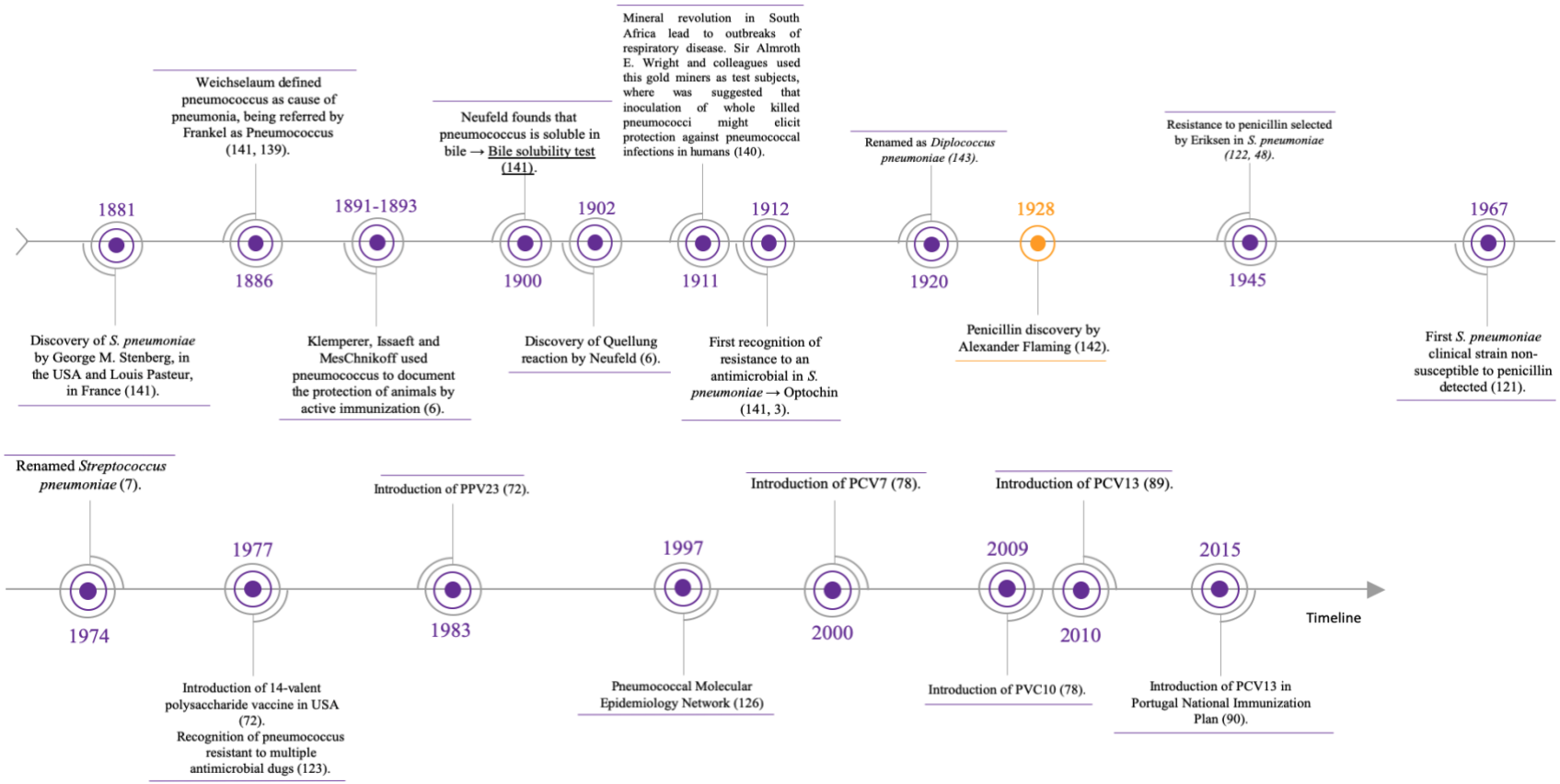

Figure 2 – Timeline of *Streptococcus pneumoniae* discoveries (Purple – discoveries related to *S. pneumoniae* and Orange – important discoveries related). Adapted from [3,6,124,125,129,189–193,7,48,73,79,80,90,91,123].

George M. Stenberg, in the United States of America (USA), and Louis Pasteur, in France, were the first to isolate, from carriers, what we call nowadays *Streptococcus pneumoniae*: "roughly lancet-shaped pairs of coccoid bacteria in human" (*Micrococcus pasteuri* and *Microbe septicemique du salive*, respectively) [6]. Over the years, this pathogen had many names, and based on its morphology when growing in liquid media, in 1974, was renamed *Streptococcus pneumoniae* [7] (Figure 2).

Since its discovery, pneumococcus has served as a cell model and led to a greater understanding of cellular biology, host-parasite relationship, antibiotic resistance and vaccine-related immunoprophylaxis [2,3] (Figure 2).

1.1 Characterization of *S. pneumoniae*

1.1.1 Identification

Despite some limitations regarding species identification [5], conventional microbiologic tests are useful to place isolates into the correct streptococcal groups, being still used in the identification of *S. pneumoniae*. However, these limitations make it challenging to establish the diagnosis of invasive pneumococcal disease (IPD), which refers to pneumonia, meningitis, bacteremia and infections of other normally sterile sites with *S. pneumoniae* [8]. In spite of the importance of obtaining accurate data on IPD burden, mainly to access the effectiveness of vaccination, there has not been significant development in laboratory diagnosis during the past few decades [8].

Laboratory identification of pneumococcus relies primarily on microscopy, culture and phenotypic testing. Microscopy allows the recognition of typical morphological characteristics of bacteria, in which *S. pneumoniae* appears as lancet-shaped, Gram-positive diplococci or short chains cocci (Figure 1). Nonetheless, this typical appearance can be altered by antimicrobial therapy [5]. In culture, overnight, at 35ºC with 5% CO2 in 5% sheep blood agar (optimal conditions), this pathogen appears to be small, greyish and mucoid, with a surrounded greenish zone of α-hemolysis. After 24-48h, the colonies present a central depression ("draughtsman colonies") (Figure 3) [8]. For the differentiation between pneumococcus and other viridans streptococci, two tests should be performed: optochin susceptibility test and bile solubility test.

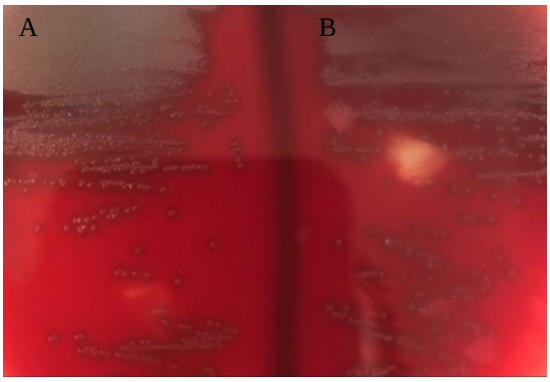

Figure 3 – Illustration of *S. pneumoniae* colonies on blood agar with alpha hemolysis. A- Mucoid colonies; B- Colonies not mucoid.

Pneumococci are usually susceptible to optochin (ethylhydrocupreine) even though this antibiotic is not used in therapy [8]. The optochin susceptibility test consists in inoculate a pure culture of *S. pneumoniae* in a sheep blood agar plate and add an optochin disk at optimal growth conditions. If a clear zone of inhibited growth around the disk is observed (with ≥14 mm), the isolate is susceptible, which means it is *S. pneumoniae* [5,9]. However, pneumococcal resistant strains have been reported, as well as optochin-susceptible *S. mitis* isolates. Isolates with these results should be subjected to a bile solubility test [5,10]. The bile solubility test consists of the lysis of the bacteria in the presence of sodium deoxycholate (bile) [8]. This test can be performed in a tube or directly applying the reagent on an agar plate, being the first method the most common [5]. *S. pneumoniae* expresses autolysins (enzymes that degrade the cell wall peptidoglycans) which are activated when exposed to bile salts, leading to this pathogen lysis (solution becomes transparent – positive result). Other α-hemolytic streptococci remain unchanged (solution remains turbid – negative result) [2,5]. Nonetheless, misidentification of other streptococci as pneumococcus has been reported, especially in case of *S. mitis*, *S. pseudopneumoniae* and *S. oralis* [11].

Alternative methods have been developed to overtake the limitations of conventional methods. The API 20 Strep System (bioMerieux, Inc., Hazelwood, Mo.) is a commercial, standardized test, composed of strips that contain biochemical tests, allowing the identification of *streptococci* including *S. pneumoniae* [5,12].

MALDI-TOF (Matrix-assisted laser desorption ionization time of flight mass spectrometry) is a routine tool vastly used for microbial identification, although differentiation between *S. pneumoniae* and other members of *S. mitis* group has been proven difficult to detect [5,13].

In the last decades, molecular methods have been actively developed. PCR (Polymerase Chain Reaction) detects minimal amounts of nucleic acid from potentially all pathogens and has been a useful tool to detect *S. pneumoniae* deoxyribonucleic acid (DNA) in all types of sources, especially in CSF (cerebrospinal fluid) and pleural fluid. This method presents advantages when compared to others methods as it does not depend on the viability of the target microbe, it is not affected by prior antimicrobial therapy and the results are obtainable in a short-time frame [8]. Multiple targets have increased the specificity of this method, offering advantages over other assays [14].

Despite the development of new methods and the advantages that they bring, especially the PCR, for the detection of *S. pneumoniae*, culture is still performed because it is the only technique that allows antimicrobial susceptibility testing.

The rapid immunochromatographic test (ICT, Now *S. pneumoniae* urinary antigen test, Binax) is also an alternative methods used in clinical detection of *S. pneumoniae*, that detects C polysaccharides cell wall antigens common to all strains of pneumococcus. This is a method with high sensitivity and specificity, however, it is only licensed for urine and LCR samples and in children results are frequently false-positives due to high rates of carriage in this age group, being recommended the used of this method only in adults [8,15].

1.1.2 Virulence Factors

For *S. pneumoniae* to be able to colonize and cause disease, it has to avoid the host's immune system activation and bacterial clearance by producing virulence factors [16]. There are many virulence factors, that contribute to pneumococcus persistence within a host (figure below). However, different strains may have different abilities to produce virulence factors, so understanding their role and distribution allows a better understanding of the pathogenesis of the infection and the ability to cause different types of disease [17].

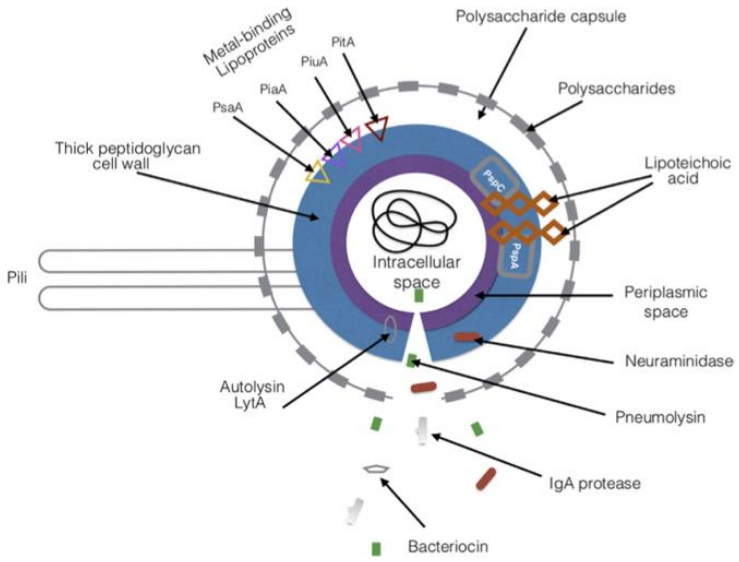

Figure 4 – Scheme of the principal pneumococcus virulence factors. PsaA, pneumococcal surface adhesin A; PspA, pneumococcal surface protein A; PspC, pneumococcal surface protein C; PiaA, pneumococcal iron acquisition A; PiuA, pneumococcal iron uptake A; PitA, pneumococcal iron transporter A. Reproduced from [18].

I. Polysaccharide Capsule

Polysaccharide capsule is one of the most studied virulence factor of *S. pneumoniae*, that consists of a 200-400 nm [19] thick layer of polysaccharides, which in the majority of cases are covalently linked to the cell-wall peptidoglycans [20,21]. Approximately 100 structurally and serologically distinct capsule polysaccharides (CPS) types have been recognized [22] characterizing pneumococcus serotypes, where some are more associated with disease than others.

In all pneumococcal serotypes, the *cps locus* is situated in the same chromosomal location, between genes *dexB* and *aliA*, except for serotype 37 (in which the *tts* gene is responsible for capsule production) [23]. The regulation of CPS synthesis helps to initiate infection by allowing the bacterium to adhere to host epithelium cells and cause inflammation, providing protection against the host's immune system [18,24]. The negative net charge present in the capsule (except for serotypes 7F, 7A, 15, 33F, 37 and 1, that have a neutral net charge [25]), given by the acidic polysaccharides and phosphates allows *S. pneumoniae* to avoid being trapped by mucus layers (first host barrier) due to electrostatic repulsion [18,26]. This virulence factor also prevents opsonophagocytosis by shielding the Fc region of IgG and complement system from interacting with their receptors on phagocytic cells [16,27].

II. Pneumolysin

Pneumolysin (Ply) is a wide-ranging virulence factor produced by more than 20 species of Gram-positive bacteria [16,17]. This protein belongs to the family of cholesterol-dependent cytolysins and is expressed during the late log phase of growth, being released upon *S. pneumoniae* lysis. By binding to membranes containing cholesterol, this virulent factor oligomerizes, forming a ring-shaped transmembrane pore with 260 Å of diameter, composed of approximately 50 monomers [16,17], leading to host cell lysis.

Ply also assists in the regulation of host-to-host transmission, by interfering in cytokine and chemokine production [28], and regulates the complement system, by reducing phagocytosis [17,29]. Recent studies have showed that Ply could cause DNA damage by creating ion channels that disrupt cell calcium levels, leading to overproduction of ROS (Reactive Oxygen Species) [30].

III. Cell wall components

S. pneumoniae cell wall, like other Gram-positive bacteria, is composed by peptidoglycans and teichoic acids (TAs) [31], that provide protection and shape to the cells. The TAs can be covalently attached to peptidoglycans, being exposed in the cell wall exterior (WTA) or anchored to the cytoplasmic membrane (lipoteichoic acids, LTA) [31]. These components have phosphocholine residues (PCho) that are important for the activation of autolysins [32] and to anchor choline-binding proteins (CBPs, family of cell surface proteins discussed below) [33]. Pneumococcus suffers secondary modifications of glycan chains [33] to resist lysozyme activity, avoiding host immune system.

WTA, LTA and peptidoglycans are pathogen-associated molecular patterns (PAMPs) [18] that activate the complement pathway, inducing the production of cytokines, necessary for the inflammation process [34].

IV. Pneumococcal surface proteins

A large group of proteins that aid in pathogenesis by acting as adhesins to host cells and hindering the host's immune system.

a. CBPs (Choline-Binding Proteins): proteins that bind to PCho presents in *S. pneumoniae* cell wall. These proteins interact with host immune system by: 1) helping the adhesion process [35]; 2) blocking the host complement system action, affecting its activity; and 3) by interacting with immunoglobulins, reducing their ability to eliminate pathogens [35,36]. Examples of these class of proteins in *S. pneumoniae* are pneumococcal surface protein A (PspA), choline-binding protein A (CbpA) and autolysin (LytA).

   PspA is a electronegative protein that blocks the complement system action, preventing opsonization [16,37]. This protein also possesses a lactoferrin-binding domain that binds to the N-lobe of human lactoferrin, blocking surface accessibility of this bactericidal peptide, preventing it from penetrating the bacterial membrane [38].

   CbpA, also known as PspC, SpsA and PbcA, is one of the most relevant pneumococcal adhesins [39], binds to the polymeric immunoglobulin receptor [18]. CbpA can also bind to the factor H, preventing the formation of C3b (a component of the complement system) facilitating the colonization process [18,40]. Some of these proteins may present a LPxTG motif, which incorporates them also in LPxTG-anchored proteins (referred below) [41].

   LytA is an autolysin that degrades peptidoglycans by cleaving the N-acetyl-muramoyl-L-alanine bond [42], inducing cell lysis. This leads to the release of toxins, such as Ply, peptidoglycans and TA, that are harmful to the host cells. *S. pneumoniae* has two more other lytic enzymes: LytB and LytC, but their virulence is not entirely understood [18].

b. Lipoproteins: metal-binding proteins that interact with ATP-binding cassette (ABC) transporter complexes.

   PsaA is an example of one of these proteins, that is involved in the transport of magnesium and zinc into the cell [35,43], which aids in the invasion of epithelial cells, as well as PiaA, PiuA and PitA which are involved in the regulation of iron [44].

c. LPxTG-anchored proteins: group of proteins that bound to peptidoglycans present in bacterial cell wall by a LPxTG-motif, in which x represents any amino acid [45]. Examples of these proteins include neuraminidase, hyaluronidase, pili and immunoglobulin A1 protease.

Neuraminidase is an exoglycosidase that cleaves siliac acid from glycoproteins, glycolipids and oligosaccharides present in cell-surface, exposing receptors for adherence [39,46].

Hyaluronidase is an enzyme encoded by the gene *hylA* [21] that depolymerizes hyaluronic acid from the mammalian connective tissue and extracellular matrix, allowing bacterial spreading and colonization [45].

Pili are hair-like structures, located on the cell wall surface of *S. pneumoniae* and other bacteria, that helps 1) in the process of colonization of nasopharynx and access to lung epithelial cells [47]; 2) to avoid phagocytosis by cells of the host immune system [48] and 3) to stimulate inflammatory responses of the host immune system [47]. There are two main types of pili found on pneumococcus: pilus-1and pilus-2. Pilus-1 is the most common pilus found in this pathogen (30% of clinical strains) [49], while pilus-2 is less common (16% of clinical strains) [50].

Immunoglobulin A1(IgA1) protease is an enzyme, that cleaves the IgA1 in the hinge region, interfering with the function of the antibody, by eliminating the Fc-mediating effector function, and consequently interfering with the host immune response [45,51].

d. Non-classical surface proteins (NCSPs): adhesins that can bind to host molecules, promoting a pneumococcal host cell invasion [52]. PavA, enolase and GAPDH are examples of these proteins present in *S. pneumoniae*.

PavA is a fibronectin-binding protein, similar to FBP54 present in *S. pyogenes* [53] that binds to fibronectin, mediating adherence to epithelial cells [45]. It also helps to control inflammation and inhibits recognition by dendritic cells [54].

Enolase is a plasminogen-binding protein, essential for proteolytic activity on the cell surface, a crucial process for adherence and colonization [55,56].

GAPDH, like enolase, is a plasmin(ogen)-binding protein with higher affinity to plasmin than to plasminogen [45]. It also plays a vital role in iron acquisition due to the ability to bind to haemoglobin and heme [45,57].

1.2 Host-pneumococcus interaction

1.2.1 Host colonization

S. pneumoniae is a pathogen which affects particularly children under the age of two, the immunocompromised and the elderly [58]. Virulence factors allow the establishment of a balance between the host immune system, the commensal flora and the pathogen, leading to colonization. This pathogen colonizes asymptomatically the nasopharynx, being part of the commensal microbiota [22,25]. From here, pneumococcus can successfully spread to other susceptible hosts [25], or disseminate into the lower airways and other organs and tissues becoming pathogenic [48].

Pneumococcus transmission occurs via direct contact with secretions of the respiratory system of a carrier or a contaminated surface [58,59]. There has been substantial debate regarding the age group that acts as the leading carrier/reservoir for *S. pneumoniae*. Although some researchers suggested infants as the primary carrier/reservoir and others suggested older children [18], children are commonly appointed as the leading carriers/reservoirs of this pathogen. Transmission is enhanced by the inflammatory response of the host, induced by Ply expression, suggesting that *S. pneumoniae* uses “the host’s inflammatory response as a signal for initiating its exit from the inhospitable host” [18]. When it finds a new host, *S. pneumoniae* has to avoid many defence barriers from the immune system until it can establish colonization.

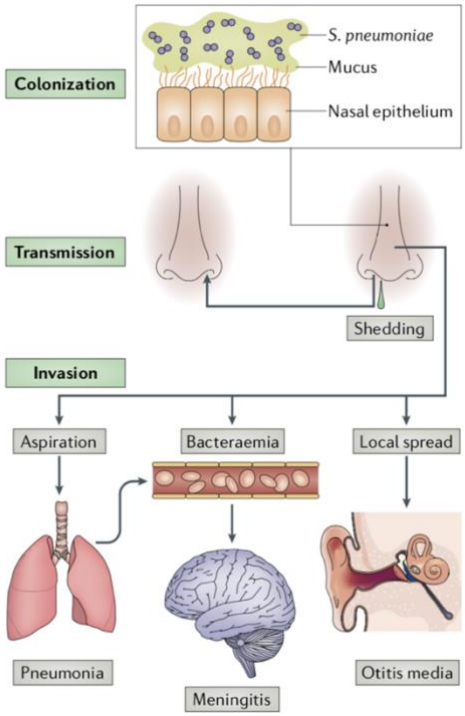

Figure 5 – Representation of the life cycle of *Streptococcus pneumoniae* and IPD. Reproduced from [70].

Mucus barrier is the first line of defence that pneumococcus encounters. This barrier consists of a viscous gel that contains antimicrobial compounds, that traps large particles (like bacteria) and transports them, through ciliary beating, towards the pharynx for removal via the oral-faecal route [60]. Mucins are glycoproteins covered with negative charge N-acetylneuraminic acid (sialic acid), present in this barrier. These glycoproteins can bind to positive charge particles, such as bacteria, with high avidity, preventing these particles from crossing the mucus barrier and facilitating their removal [25].

Pneumococcus overcomes this line of defence by expressing the different polysaccharide capsules. This virulence factor allows pneumococcus translocation across the negative charge mucous layer [26], facilitating pneumococcal adherence to epithelial cells, helping the establishment of colonization. However, the expression of a thick capsule inhibits adherence to epithelial cells, so this pathogen undergoes phase variation[2] between two forms of colony morphology: opaque and transparent [61]. Transparent isolates express a thinner capsule, which promotes the interaction of adhesins present in the pneumococcal surface and host receptors on epithelial cells. Neuraminidases (NanA, NanB and NanC) inactivate the adhesive properties of mucins, by cleaving the sialic acid, assisting polysaccharide capsules in neutralising mucus barrier function [62], enhancing

[2] **Phase variation –** molecular mechanism that leads to switching of the gene-expression state (on or off) [61].

S. pneumoniae capacity of establishing colonization. After establishing colonization, pneumococcus expresses a thicker capsule (opaque variant) to avoid the host immune system, leading to invasive infections [63].

Antibodies are also an essential barrier against this pathogen. IgA is an immunoglobulin present mainly in mucosal surfaces (gastrointestinal, respiratory and urogenital tracts) that inhibits bacterial adhesion, neutralises bacterial toxins (extra- and intracellularly) and eliminates pathogens via an IgA-mediated excretory pathway. There are two types of IgA in humans, IgA1 which constitutes more than 90% of IgA in the human airway, and IgA2, a secretory form of IgA [64]. *S. pneumoniae* developed an evasion strategy to overcome this barrier by expressing IgA1 protease. This enzyme cleaves the human IgA1, preventing inflammation from being initiated, which leads to antibody-mediated clearance to only occur after sufficient amounts of other classes of specific antibodies have been produced [16].

It is estimated that in a human pharynx more than 700 different microbial species can reside [65]. Pneumococcus, to successfully colonize the host, has to compete for resources (space and nutrients) with other colonizing pathogens and within the same species (different pneumococcal serotypes – co-colonization) [25]. By expressing pneumocins (bacteriocins[3]), peptides with antimicrobial activity and Ply, this pathogen is able to lyse other microorganisms, included other *S. pneumoniae*, becoming more competitive and at the same time originating a source of DNA [66]. *S. pneumoniae* can incorporate exogenous DNA from closely related oral streptococcal and co-colonizing pneumococci, increasing, in this way, its fitness [16]. This pathogen also secrets H_2O_2, which reduces the growth of bacteria that it may be competing with – bactericidal effect [67]. Pneumococcus produces H_2O_2 via pyruvate oxidase, which forms hydroxyl, radicals that are harmful to bacteria. However, pneumococcus developed the ability to reduce reactive OH before it comes in contact with its DNA, becoming resistant to H_2O_2 [68,69].

By overcoming all of these barriers, *S. pneumoniae* is able to colonize the nasopharynx and firmly attach to epithelial cells [25]. Translocation from the nasopharynx to deeper tissues leads to the progression to invasive disease [70].

[3] **Bacteriocins –** proteins or peptidic toxins produced by bacteria to inhibit the growth of similar or closely related bacteria [70].

1.2.2Pneumococcal Infections

S. pneumoniae is responsible for a variety of invasive (meningitis, bacteremic pneumonia, bacteremia and sepsis) and non-invasive infections (acute otitis media, sinusitis and non-bacteremic pneumonia). It presents a higher burden in children, immunocompromised and elderly individuals, being the risk factors summarised in the table below.

Table 1 – Risk factors associated with IPD. Adapted from [71,72].

Host factors	Age (children <2 years and the elderly) Ethnic distribution (black Americans, Alaska Natives and Australian Aborigines). Primary/secondary immunological deficiencies (HIV-positive, kidney disease, leukaemia, treatment with an immunosuppressive drug, and others) Spleen and liver failure Pulmonary and cardiac dysfunction Respiratory viral co-infection (e.g. influenza A)
Environmental factors	Smoking Crowding (day-care attendance, hospitalization and institutionalization) Seasonal distribution (winter months)
Pathogen factors	Nasopharyngeal acquisition of a new serotype.

I. Pneumonia

Community-acquired pneumonia is the fifth most common cause of death in people older than 65 years old, and it is the most common cause of infection-related mortality [73,74], with *S. pneumoniae* being one of the most common pathogens.

Pneumococcal pneumonia initiates with the aspiration of nasopharynx colonizing pneumococcus up to the lungs. Due to high organism densities and the proximity of blood vessels to the air spaces, pneumonia may complicate to bacteremia [73], leading to other diseases.

Treatment remains a challenge because of the difficulties in diagnosis and the prevalence of antimicrobial resistance [73].

II. Bacteremia and Sepsis

Bacteremia is defined as the presence of bacteria in the bloodstream. There are two types of bacteremia: Occult bacteremia – presence of bacteria in the blood of a well-appearing febrile individual without an identifiable focus of infection, and

Sepsis – bacteremia with clear signs of severe systemic illness caused by the activation of multiple overlapping cascades [73].

Twenty-five to 30% of patients with pneumococcal pneumonia and 80% of meningitis patients progress to bacteremia [2]. In the last years, the routine use of polysaccharide conjugate vaccines (PCVs) has reduced the incidence of occult bacteremia in developed countries [73].

III. Meningitis

Meningitis consists of an infection of the meninges (membranes that separate the brain and spinal cord). Pneumococcal meningitis usually occurs after bacteremia, infections of the ear or sinuses, or head trauma that causes communication between the subarachnoid space[4] and the nasopharynx [2].

Pneumococcal meningitis incidence varies with age, being the leading cause of disease in children (with ages between 1-23 months), along with *Neisseria meningitis*, and adults (with ≥60 years old) [73,75].

Since clinical characteristics cannot distinguish pneumococcal meningitis from other bacterial meningitis, microbiologic diagnosis is essential and requires a CSF sample [73].

Treatment requires the use of antimicrobial capable of achieving proper concentrations in the central nervous system (10-fold higher than the minimum inhibitory concentration (MIC) of the organism) [73].

IV. Sinusitis and Otitis Media

These infections usually precede viral infections of the URT, where polymorphonuclear neutrophils obstruct the sinuses and ear canal [2].

Otitis media, infection of the middle ear canal, occurs primarily in young children, while sinusitis occurs in patients of all ages [2,73]. Usually, otitis media and sinusitis are resolved without severe complications and antimicrobial therapy; however the same is not valid for those caused by *S. pneumoniae* [76].

[4] **Subarachnoid space –** space between the arachnoid and the pia mater (constituents of the meninges), which is filled with cerebrospinal fluid [186].

1.2.3 Epidemiology

Pneumococci infections are the primary cause of morbidity and mortality all over the world. WHO estimated that in 2010, approximately 1.3 million deaths annually were due to *S. pneumoniae* [77]. Despite the success that PCVs have achieved in developed countries, *S. pneumoniae* remains a major global healthcare challenge, especially in low-income countries.

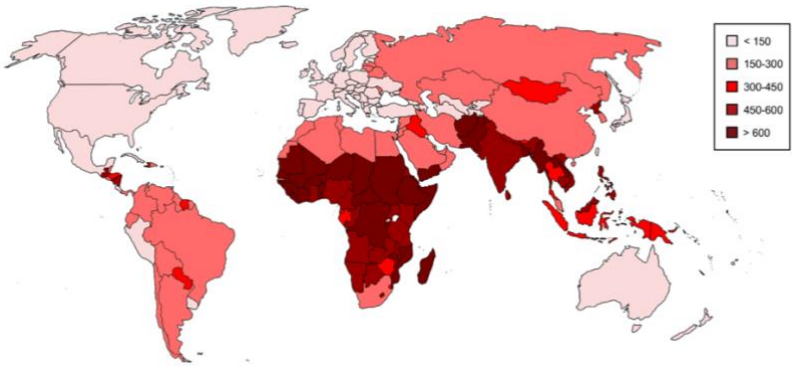

Figure 6 – Global incidence rates of IPD in children with <5 years old in 2010. Reproduced from https://www.abcombibio.com/pneumococcal.

In Europe, 23 886 confirmed cases of IPD were reported by 29 countries in 2017, with the highest number of confirmed cases being from the United Kingdom and France. The rates of IPD reported were higher in the elderly (18.9 confirmed cases per 100 000 population in adults ≥65 years old) and on infants (14.5 confirmed cases per 100 000 population in infants < 1 year old) [78].

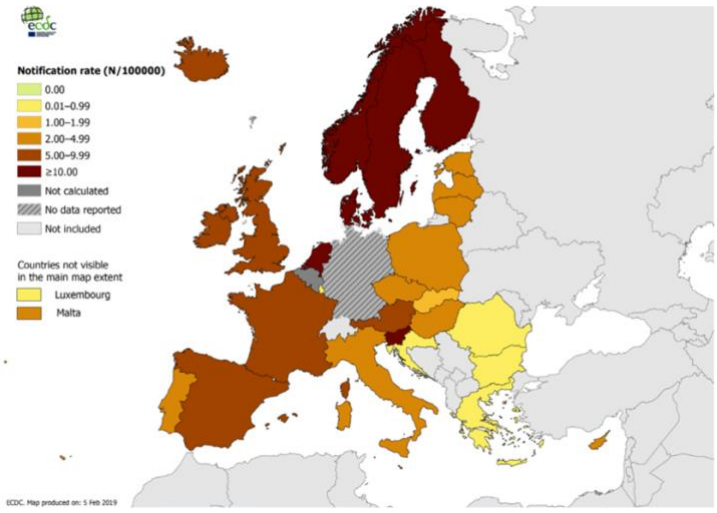

Figure 7 – Confirmed IPD cases per 100 000 population reported by Europeans countries in 2017. Countries: Austria, Bulgaria, Croatia, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom. Reproduced from [78].

In Portugal, the incidence of paediatric IPD in all age groups has been decreasing since 2008-2012 [79], presenting a rate of 2.9 confirmed cases per 100 000 population in 2017 [78].

1.3 Vaccines

1.3.1 Available pneumococcal vaccines

Due to the global impact and the high morbidity and mortality, *S. pneumoniae* was, since the beginning, a priority for the development of a strategy that protects against pneumococcal disease – vaccination.

Thirty years after the discovery of *S. pneumoniae*, Wright et al. conducted the first clinical trials of a whole-cell pneumococcal vaccination (Figure 2). However, it was only in 1977 that the first polysaccharide vaccine (vaccine against 14 pneumococcal serotypes) was licenced [80,81] (Figure 2). Nowadays, vaccination is the best strategy to prevent disease, by diminishing personal risks and allowing community protection – herd immunity[5] [81]. Polysaccharide vaccines confer protection against pneumococcal disease by opsonophagocytic antibodies directed to bacterial capsule polysaccharides [82].

Pneumococcal polysaccharide vaccine 23-valent (PPV23) was licenced in 1983 (Figure 2) and contains 23 purified capsular polysaccharides of *S. pneumoniae* (serotypes 1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19A, 19F, 20, 22F, 23F and 33F) that represent 85-90% of the serotypes that caused invasive pneumococcal infections among children and adults in the USA [83]. Although the efficacy of this vaccine against IPD in 65 years old and older individuals is 75% in USA [84], the response induced in younger children was low [81], since PPV23 induces a T-cell-independent response without immunological memory [85], being recommended its administration only in adults. To overcome this problem, in 2000 (Figure 2) a 7-valent pneumococcal conjugate vaccine (PCV7) against seven of the most frequent serotypes in the USA (4, 6B, 9V, 14, 18C, 19F and 23F) became available in the market [81]. This vaccine is composed of purified polysaccharides covalently attached to a protein (CRM197 [86]), which induces a T-cell-dependent response, producing IgG and memory

[5] **Herd Immunity –** indirect protection from infectious disease acquire when the majority of the population has become immune, protecting those who are not immune [187].

B-cells [87]. PCV7 demonstrated significant protection against IPD in all age groups, being recommended for use in children under five years old [81].

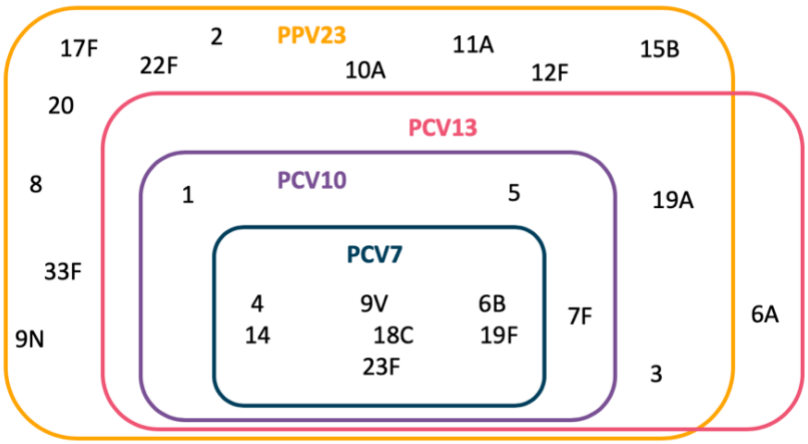

Figure 8 - Representation of serotypes included in pneumococcal vaccines.

Despite the reduction in PCV7 serotypes IPD and carriage [88] in USA, it was observed an increased rate of IPD caused by non-PCV7 serotypes (serotypes not present in the PCV7). This led to the licensure of 10-valent pneumococcal conjugate vaccine (PCV10) (including all PCV7 serotypes plus serotypes 1, 5 and 7F) in 2009 and the 13-valent pneumococcal conjugate vaccine (PCV13) (includes all PCV10 serotypes plus serotypes 3, 6A and 19A) in 2010 (Figures 2 and 8) [86]. Since its introduction, in 2001, in Portugal, the uptake of PCV7 in children was good, achieving 75% in 2008 [89], and with the replacement with PCV13, in 2012, the coverage was 63% [79]. PCV13 was introduced in the National Immunization Plan (NIP) in 2015 (Figure 2) with a 2+1 schedule for children and in particular risk groups followed by PPV23 [90,91]. PPV23 is available in Portugal since 1996, but its uptake among adults is estimated to be low (around 10%) [92].

The widespread use of PVCs led to a decrease in IPD incidence and carriage of these serotypes in vaccinated children and non-vaccinated population but allowed non-vaccine serotypes (NVT) to colonize the nasopharynx and possibly cause IPD – Serotype replacement [81,93], which led to the study of new vaccination strategies. New vaccines are based on pneumococcal proteins that contribute to virulence, and that are present in all serotypes, where proteins like Ply, PspA, CbpA, PsaA, NanA, PiuA and PiaA seem to be the most promising targets. [81]. These vaccines provide different degrees of protection, and there is clear evidence that immunization with certain combinations of

proteins provide additive or synergistic protection [94,95], being the combination Ply, PspA and CbpA the most effective [96].

1.3.2 Serotype Replacement

The decreased carriage of vaccine serotypes has as consequence the replacement of these serotypes for NVT in the nasopharynx, a phenomenon also known as serotype replacement. Several hypotheses have been proposed to explain this phenomenon. One hypothesis was the unmasking effect due to the limitations of the traditional serotyping methods [97,98]. Pneumococcus is able to co-colonize an individual, therefore, with reduction of vaccine serotypes, the increased detection of NVT would be attributed to the improvement of serotyping techniques and not because of a real increase in the acquisition of NVT. However, new molecular methods have shown an actual increase in non-vaccine serotypes together with a decrease in vaccine types, which presents a limitation to this hypothesis [81,99].

A broad consensus hypothesis that explains these observations is the co-colonization combined with capsular switching. Capsular switching, which can be triggered by vaccination and antibiotic use pressure, occurs when an isolate of *S. pneumoniae* replaces its capsular locus with that of another *S. pneumoniae* serotype present in the same environment [18,81]. Co-colonization is essential for this phenomenon since it is the DNA reservoir that will allow the horizontal gene transference [100] (Figure below). In this way, pneumococcus that expressed a vaccine serotype can now express a NVT. Recent studies have shown that capsular switching occurred regularly before PCVs introduction and widespread antibiotic use, suggesting that this phenomenon is not a consequence of vaccination and antibiotic use, but that its evolution and increase was influenced by these two selective pressures [101].

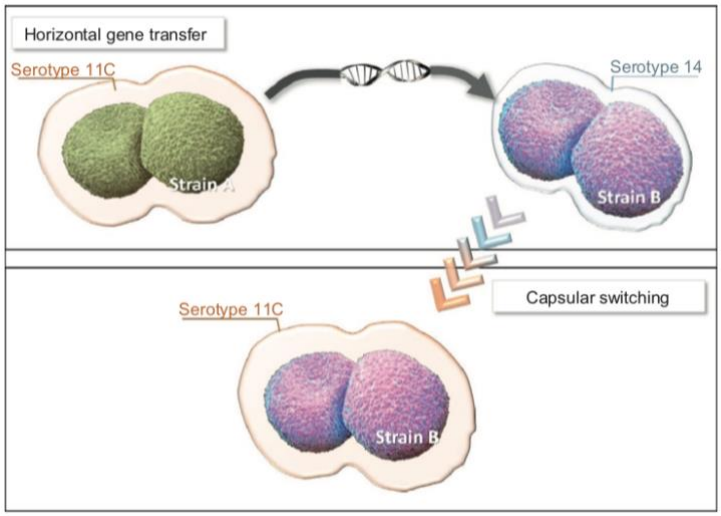

Figure 9 – Representation of capsular switching. Reproduced from [81].

1.4 Antimicrobial Resistance

1.4.1 Definition of resistance

Centers for Disease Control and Prevention (CDC) estimates that at least 2 million people acquired, in 2013, severe infections from pathogens that were antimicrobial resistant [102], being *S. pneumoniae* one of them.

S. pneumoniae antimicrobial resistance profiles are usually determined based on MIC using *Etest* strips or growth inhibition halos of disk diffusion susceptibility test. To assess antibiotic resistance, several professional bodies (the Clinical and Laboratory Standards Institute (CLSI); the British Society for Antimicrobial Chemotherapy (BSAC) and the European Committee on Antimicrobial Susceptibility Testing (EUCAST)) established rules regarding methods and their interpretations, to facilitate the comparison between studies [103]. The breakpoints, established by these institutions have in consideration the pharmacokinetic/pharmacodynamic properties of an antibiotic and the patient outcome, meaning that there can be changes in the definition of resistance[104]. An example of this situation was the breakpoints for penicillin, that were revised in January 2008 to redefine the susceptibility of meningeal and non-meningeal pneumococcal isolates (table below) [104,105].

Table 2 – Former and current CLSI breakpoints for penicillin in *S. pneumoniae*. Adapted from [106].

	MIC (μg/mL)		
	S	I	R
Before January 2008	≤0.06	0.12-1	≥2
After January 2008			
for meningitis	≤0.06	-	≥0.12
for non-meningitis			
via intravenous administration	≤2	4	≥8
via oral administration	≤0.06	0.12-1	≥2

"S" – susceptible; "I" – intermediate; "R" – resistant.

Many risk factors have been described to be associated with antibiotic-resistance in pneumococci. However, most of them can be linked to the selective pressure of antibiotics in eliminating susceptible strains [107].

Table 3 – Resume of risk factors for resistance. Adapted from [105,107].

Host factors	Age (children <2 years and elderly) Lack of PCV vaccination Exposure to antibiotics on an individual level HIV
Environmental factors	Crowding (hospitalization, urbanization and daycare attendance) Exposure to antibiotics at a national level
Pathogen factors	Bacterial clonal structure

1.4.2 Mechanisms of Resistance

I. β-lactam Resistance

(Penicillin, Cephalosporins and Carbapenems)

β-lactam antibiotics bound to penicillin-binding proteins (PBPs), that are responsible for cell wall synthesis, inhibiting the growth of pneumococci. *S. pneumoniae* has six different PBPs: 1a, 1b, 2a, 2b, 2x and 3 [107]. Resistance is achieved by alteration of target PBPs (mainly PBPs 2x, 2b and 1a), decreasing the antibiotic affinity [107]. *S. pneumoniae* acquires this altered PBPs by the acquisition of *pbp* genes from other streptococcal species, from the *S. mitis* group, by recombination creating mosaic genes [108]. Mutations in each PBP confer different degrees of nonsusceptibility to β-lactam: low-grade resistance is associated with mutations in *pbp2x*; mutation in *pbp2b* result in even higher levels of resistance and high levels of resistance required alteration in *pbp1a* [107]. Resistance to cephalosporins are associated with mutations in the *pbp1a* and *pbp2x*

genes and most of cephalosporin resistant strains are also penicillin resistant [107]. Carbapenem resistance, although the mechanism is not fully understood, it is known to be associated with decreasing affinity to PBP1a, PBP2x and PBP2b. In spite of the high prevalence of carbapenem resistance in Asia-Pacific region, this is not very common in European countries yet [109].

II. Macrolide and Lincosamides Resistance

(Erythromycin, Azithromycin, Clarithromycin and Clindamycin)

Macrolides are bacteriostatic agents[6] that act by inhibiting bacterial protein synthesis, by binding to the 23S ribosomal ribonucleic acid (rRNA) component of the 50S ribosomal subunit. The resistance to this antimicrobial class is mediated in *S. pneumoniae* by two primary mechanisms: target modifications and active efflux [105,107].

Target modification is a resistance mechanism that results from the expression of a 23S ribonucleic acid (RNA) methylase encoded in the *erm*(B) gene, usually present in the *Tn916* family transposons. This protein dimethylates the adenine residue at position 2058 on the 23S rRNA, reducing the affinity of the macrolides to the 23S binding site [107]. This process confers high-level resistance to macrolides, as well as lincosamides and streptogramins B[7] (MLS_B phenotype) [110]. Resistance can be expressed either constitutively ($cMLS_B$ phenotype – gene is always expressed, so isolates are always resistant to macrolides, lincosamides and streptogramins B [111]) or inducible ($iMLS_B$ phenotype – the presence of a macrolide induces resistance to lincosamides and streptogramins B [111]) [112]. An *erm*(A), subclass (TR) has also been shown to confer MLS_B resistance in *S. pneumoniae* [112]. Most macrolide resistant pneumococci are also resistant to tetracycline since *tet*(M) gene is typically carried by the *Tn916* family transposons [113].

Efflux pumps confer low-level resistance to macrolides, but not lincosamides or streptogramins B (M phenotype). Active efflux is encoded by *mef*-class genes

[6] **Bacteriostatic agents** – biological/chemical agents that stops bacteria from reproducing but do not kill them. Bacteriostatic antimicrobials, when used, the duration of the therapy must be sufficient to allow the host immune system to act [188].
[7] **Streptogramins:** Quinupristin and Pristinamycin.

that have three variants: the abundant *mef*(A), *mef*(E) and the rare variant *mef*(I) [107].

The dual phenotype resistance (presence of both the *erm* and the *mef* genes) in *S. pneumoniae* has been increasing, especially in Asia, but also in Europe, South Africa and United States [114].

III. Tetracycline Resistance

Tetracycline is a broad-spectrum bacteriostatic drug, active against pneumococci, which was vastly used in clinical practice [107]. This antimicrobial acts by binding to the 30S ribosome subunit, inhibiting the bacterial protein synthesis. By expressing TetM or TetO proteins, encoded in *tet*(M) and *tet*(O) genes, respectively, *S. pneumoniae* is able to block the tetracycline action, by binding these proteins to the 30S ribosome subunit [115]. Although tetracycline is no longer used in clinical practice, resistance to this antimicrobial still persists. This may be explained by the fact that the *tet*(M) gene (most common in *S. pneumoniae*) is usually present in *Tn916* transposons. This transposons often contain genes for resistance to other antibiotics, especially *erm*(B), which could explain the persistence of tetracycline resistance [105].

IV. Fluoroquinolone Resistance

(Levofloxacin, Ciprofloxacin and Moxifloxacin)

Fluoroquinolones are synthetic antibiotics used in the empirical treatment of respiratory tract infections and pneumonia in adults. This antibiotic binds to type II DNA topoisomerase enzymes (DNA gyrase and topoisomerase IV), vital for DNA supercoiling and chromosome segregation being, therefore, essential for DNA replication and cell division [107]. This antimicrobial binds to specific subunit sites of the enzymes (DNA gyrase was two subunits *gyrA* and *gyrB* that are homologous to the *parC* and *parE* of topoisomerase IV). Fluoroquinolones resistance in *S. pneumoniae* is mediated by two mechanisms: target alteration and active efflux.

Target alteration consists of changes in the conformation of the subunits of the fluoroquinolone-binding enzyme (primarily *gyrA* in the DNA gyrase and *parC* in topoisomerase IV) due to the stepwise accumulation of mutations in the quinolone

resistance determining the region (QRDR) [107]. Single mutations in one of the enzymes ("first-step mutation") often present a susceptible phenotype but confer an elevated risk of acquiring full resistance to fluoroquinolones by acquiring a second mutation in the other subunit of the enzymes [107,116].

The efflux pump mechanism is poorly characterized, being only known that it is mediated by the membrane ABC-transporter protein PmrA and other factors. It confers a low-level of resistance by lowering intracellular fluoroquinolones concentrations, promoting the occurrence of QRDR mutations that will lead to high levels of resistance [117].

V. Other Antimicrobial Resistance

a. Chloramphenicol: it is a bacteriostatic agent that binds to the 23S rRNA of the 50S ribosomal subunit, preventing peptide formation [118]. Resistance to this antimicrobial in *S. pneumoniae* is due to the expression of a chloramphenicol acetyltransferase enzyme (CAT), encoded by the *cat* gene, that inactivates this antibiotic [105,107].

b. Trimethoprim-sulfamethoxazole: it is a combined drug that inhibits successive steps in the folate synthesis pathway [119], essential for the DNA replication and cell division. Trimethoprim resistance in pneumococcus results from a single amino acid substitution (Ile-100 $\rightarrow$ Leu) in the dihydrofolate reductase (DHFR) protein. Resistance to sulphonamides is associated with chromosomal mutations within the gene encoding dihydropteroate synthase (DHPS) [107].

c. Telithromycin: it is a ketolide that binds to the domain II of the 23S rRNA of the 50S ribosomal subunit, inhibiting the protein synthesis [107]. Mutations in the domain II or V of the 23S rRNA ribosomal proteins L4 and L22 in *S. pneumoniae* confer resistance to this antimicrobial [120]. However, clinical telithromycin resistance in *S. pneumoniae* is rare [107].

d. Linezolid: it is an oxazolidinone that binds to the central loop segment of the domain V of the 23S rRNA of the 50S ribosomal subunit, inhibiting the protein synthesis. Resistance to this antimicrobial is rare in *S. pneumoniae*

and is due to the deletion of 6-bp in the gene encoding the riboprotein L4 [121].

1.4.3 Epidemiology

Antibiotic resistance is strongly associated with antibiotic consumption, varying based on geographic localization [122]. With the discovery of penicillin in 1928 [73], *S. pneumoniae* infections have been treated with this antibiotic [123]. However, 20 years later (in 1945) resistance to penicillin was selected by Eriksen et al. [124,125] and in 1967 the first *S. pneumoniae* clinical strain non-susceptible to penicillin was detected [123]. Since then, antibiotic resistance in *S. pneumoniae* emerged all over the globe. In 1977, the first multi-drug resistant (MDR) pneumococcus[8] was described [125] (Figure 2).

In Europe, *Streptococcus pneumoniae* combined non-susceptibility to penicillin and macrolides has been increasing over the years. In 2017, non-susceptibility to penicillin and macrolides varied from 0.1% to 45.5%, presenting lowest incidence in Belgium (0.1%), Netherlands (1.1%) and Estonia (1.6%) and the highest incidence in Cyprus (45.5%), Malta (26.3%) and Romania (24.05%) [126].

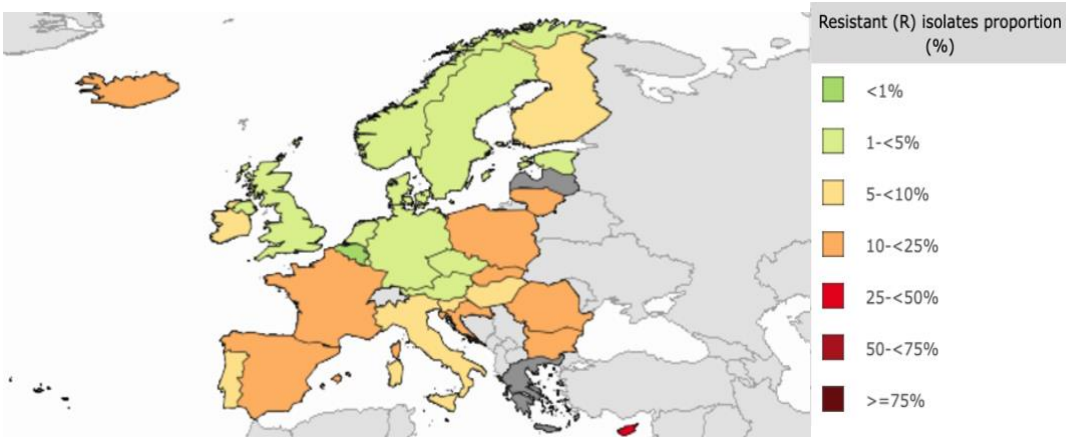

Figure 10 – Proportion of *S. pneumoniae* combined non-susceptibility to penicillin and macrolides reported by Europeans countries in 2017. Countries: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, united kingdom. Reproduced from [126].

[8] **MDR pneumococcus** – defined as having resistance to at least three different classes of antibiotics [122].

In Portugal, between 2012-2014, the proportion of pneumococcus with non-susceptibility to penicillin and macrolides has been decreasing, with a proportion of 15.4% and 17.0% respectively [127].

1.4.4 PMEN Clones

With the emergence of antimicrobial resistance pneumococci worldwide, the need for the identification of resistant clones and monitoring their spread became essential. Although there is considerable diversity among resistant strains within most serotypes, a small number of successful clones has emerged in various countries and spread massively [128]. Therefore, in 1997, the Pneumococcal Molecular Epidemiology Network (PMEN) was established, in order to standardize the nomenclature and classification of resistant pneumococcal clones in the world (Figure 2) [129]. To be included in the PMEN, a clone must fill in some criteria: 1) the clone should have wide geographic distribution (isolated at least in two continents); 2) should be resistant to one or more antibiotics that are in full clinical use or be a global susceptible clone known to be important to the disease; 3) data on the clone needs to be published or in press before ratification by the network; and 4) a representative isolate of the clone must be made available for confirmation by molecular typing methods [129]. Filled all these criteria, the clone is given a name that follows a particular nomenclature: Country$_{\text{first identified serotype}}$-sequential numbering in network-subsequent described serotype. In the case of capsular switch, it is indicated the new serotype variant after the sequential number [129].

The first PMEN identified was Spain$_{23F}$-1 [130], which is associated with resistance to penicillin, chloramphenicol and tetracycline, being, in some cases, also associated with resistance to erythromycin [128]. Currently there are 43 documented clones in the PMEN [129], of which 26 are MDR [107]. With the introduction of PCV7, the prevalence of MDR clones Spain$_{23F}$-ST81, Spain$_{6B}$-ST90 and Taiwan$_{19F}$-ST236 decreased, being replaced by serotype 19A strains mainly related to the expansion of two clonal complexes (CC) CC320 and CC230 [122]. Several studies also reported the emergence of new serotypes associated with MDR, such as 6C, 11, 15A, 23B, 33A and 35B [122].

1.5 Typing Methods

Typing methods are essential for the epidemiology study of bacterial diseases, providing valuable insight into the clonal spread analysis during epidemics, identification of route, transmission of infections and assessment of evolutionary changes following vaccination, in *S. pneumoniae* [131].

1.5.1 Serotyping

Serotyping methods in *S. pneumoniae* include the Quellung reaction, latex agglutination and PCR-based methods. These methods distinguish pneumococcus isolates based on the CPS. Although *S. pneumoniae* has a high CPS diversity, a restricted number of serotypes cause the majority of human infections [132].

Quellung reaction is the standard gold method for serotyping encapsulated bacteria such as pneumococcus [133]. It consists of the sequential test of commercially available antisera pools until a positive reaction – agglutination – is observed [134]. There are two different nomenclatures systems: the Danish system and the American system, however, the Danish system is more widely used, where serotypes are grouped by antigenic similarities, based on cross-reactions between different types. When types serological cross-react, they are assigned a common serogroup [135]. This method is very reliable and reproducible but it is also laborious (requires expertise and uses antisera with high cost [134]), having other methods been developed: latex agglutination and PCR-based serotyping. Latex agglutination test (Pneumotest-Latex) is fast and simple but does not detect multiple serotype carriage [131]. PCR-based serotyping methods reside on the amplification of specific serogroups or serotypes genes. However, this method is highly expensive and only detects a certain number of serotypes [131].

With the development of antimicrobial resistance in only a fraction of the isolates expressing a specific serotype, it became essential to distinguish isolates from the same serotype [128], and techniques like Pulsed-Field Gel Electrophoresis (PFGE) and later Multilocus Sequence Typing (MLST) were developed. Given serotype relevance in the evaluation of pneumococcal vaccine efficacy, serotyping will remain an essential method in epidemiological studies [131].

1.5.2 Molecular typing

PFGE is a typing method that compares profiles generated by the digestion of the total DNA of a strain using SmaI [132]. This method allows the identification of clones by visual comparison of the generated profiles using arbitrary define rules, or by defining cutoff values in software-generated dendrograms [136]. PFGE was the dominant method used to type pneumococci. However, it required specialized expertise, an appropriate software for the analysis of large-data studies, and the comparison of results between laboratories was not straightforward, being replaced by MLST, which overcame these limitations [132].

MLST is based on the sequence of internal fragments of seven housekeeping genes (approximately 450 bp nucleotides [137]) [132]. To each different locus sequence, is assigned an allele number, being defined as an allelic profile that will correspond to a sequence type (ST) [137]. In *S. pneumoniae, aroE* (shikimate dehydrogenase); *ddl* (D-alanine-D-alanine ligase); *gdh* (glucose-6-phosphate dehydrogenase); *gki* (glucose kinase); *recP* (transketolase); *spi* (signal peptidase I) and *xpt* (xanthine phosphoribosyltransferase) are the housekeeping genes used for the MLST analysis [138]. In December of 2004, it was established a public database (https://pubmlst.org/spneumoniae/) that allows a rapid comparison between isolates characterized by MLST and previously characterized isolates [139].

High throughput sequencing (HTS) is a method used to sequence multiple DNA molecules in parallel at the same time. Within this technology, different approaches have been developed that are used to create large data sets and allow a more comprehensive insight into cellular genomic and transcriptomic signatures [140]. With the development of relatively affordable techniques of HTS in the last decade, in particular of whole genome sequencing (WGS), the use of these techniques became common not only to determine the MSLT profiles, but also to obtain information regarding serotype identification, antibiotic resistance profiles, virulence factors [141], or even core genome MLST (cgMLST) [142]. cgMLST has the same principle as MLST, but instead of characterize genetic relationships between strains of the same species by the seven housekeeping genes, it types by doing a gene-by-gene allelic profiling of the core genome[9] of the isolates [142]. This presents a more accurate characterization, since strains

[9] **Core genome –** genes present in all isolates of a given organisms.

from the same MLST ST can be genetically distant regarding the rest of the genome [143]. However, this method is still limited since it is unable to analysed similarities/differences present in the accessory genome[10], which often integrate pathogenic components [142].

Software was developed to facilitate the analysis of broad MLST data. goeBURST is an algorithm that "constructs an unrooted tree representation of the relationship of the isolates analysed, based on the number of differences in the allelic profiles, assigning isolates to clonal complexes (CC)" [132]. It also includes a tie-break rule not present in its previous software (eBURST), where in the case of a draw of STs, the one with the lower ST IDs take precedence over the higher ST IDs as the CC ancestral [132]. This model allows the detection of emergent clones-defined as genotypes that increase in number in the population, due to a random genetic drift or fitness advantage [132]. PHILOViZ is a free software that allows the analysis of sequence-based typing methods and another kind of information, like phenotypic and epidemiological data, by interacting with many public databases [144].

With the accessibility to perform WGS, we are able to analyse additional genetic information. SeroBA is an *in silico* software that allows the prediction of *S. pneumoniae* serotypes from HTS data. It uses a k-mer and local-assembly approach to detect and identify the *cps* locus by comparing it to an included database, adapted from the previously published PneumoCAT tool [145,146], and infer the serotype of the strain. The original PneumoCAT database comprised typing information for 92 serotypes, including FASTA sequences, gene presence/absence and single-nucleotide polymorphisms (SNP) information [146]. SeroBA further extended it to include five additional serotypes and two NT references [145]. Both PneumoCAT and SeroBA run directly from the raw reads.

[10] **Accessory genome –** genes present in some of the strains of a given organism.

2. AIMS OF THE STUDY

The main objectives of this study were:

- Phenotypic characterization of *S. pneumoniae* responsible for IPD in adults, isolated from 2015 to 2017 in Portugal, by serotyping and antimicrobial susceptibility testing.
- Molecular characterization of the isolates by WGS to extract information regarding the MLST profiles of the isolates.
- Compare phenotypic serotyping data with data from genome sequencing to evaluate the efficiency of both methods.
- Compare data obtained with data from previous routine surveillance studies to evaluate the effect of vaccination in children in serotype distribution, levels of antimicrobial resistance and the circulating clones of *S. pneumoniae* in adults.

3. MATERIALS AND METHODS

3.1 Collection of strains

This study included *S. pneumoniae* isolates responsible for invasive infections in adults (≥18 years) during 2015 – 2017, provided by the Portuguese Group for the Study of Streptococcal Infections, involving microbiology laboratories of 61 hospitals throughout Portugal. Three different age groups were considered: young adults (18-49 years old), adults (50-64 years old) and the elderly (≥65 years old).

IPD was defined as the isolation of *S. pneumoniae* from a normally sterile body site (such as blood, CSF, and others) or pneumococcal DNA detected in sterile samples such as pleural fluid or CSF. Only one isolate was considered per patient.

Strains were identified as *S. pneumoniae* through conventional microbiology methods: colony morphology and haemolysis in blood agar, optochin susceptibility and bile solubility tests. When the identification was made using molecular methods, DNA was extracted from patient samples using the DNeasy Blood and Tissue kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. *S. pneumoniae* was identified by a singleplex real-time PCR (rPCR) of *lytA* and *wzg* genes performed on the Rotor-Gene 6000 (Corbett Research, Cambridge, United Kingdom), using the Platinum quantitative PCR SuperMix-UDG (Thermo Fisher Scientific, Massachusetts, USA). A conventional multiplex PCR was performed to evaluate the quality of the purified DNA, with the amplification of two human genes (human β-actin and RNaseP) [147].

Haemolysis in blood agar: strains were grown in optimal conditions and incubated overnight. After this period, the haemolysis was observed.

Optochin susceptibility: performed by the disk diffusion method, which consists of placing filter paper disks with 6 mm of diameter containing 5 μg of optochin in Tryptone Soya Agar (TSA) plates and incubated overnight at optimal growth conditions. The strains were considered resistant to optochin when the growth inhibition halos measured less than 14 mm [5]. When a strain presented resistance to optochin, bile solubility test was performed.

Bile solubility tests: this test was performed in tubes, where a control tube with the bacterial suspension plus 0.85% NaCl and other tube with the same bacterial suspension mixed with 2% sodium deoxycholate (DOC) were incubated at 35°C and

after 20-30 min turbidity was observed. If the tube containing the DOC changed from turbid to translucent, the strain is soluble in bile, i.e., it is considered to be *S. pneumoniae*.

All bacteria were stored at -80°C in Tryptone Soya Broth (TSB) (bioMérieux, Marcy-I'Étoile, France) supplemented with 15% glycerol.

3.2 Serotyping

Serotyping was performed through standard capsular reaction test, where specific antibodies reacted with their respective capsular polysaccharides. This method consists of blending approximately 1 μL of serum containing specific antibodies (Statens Serum Institute, Copenhagen, Denmark) into a glass slide, and add 3 μL of the bacterial suspension. When the capsular polysaccharides of the strain react with the serum, an agglutination occurs, usually macroscopically visible. The order of use of sera was based on the Chessboard system described by Sørensen [148], which intends to shorten the time of serotyping.

Serotypes were classified into vaccine serotypes included: in PCV7 (serotypes 4, 6B, 9V, 14, 18C, 19F, 23F), in PCV10 (all PCV7 serotypes plus 1, 5, and 7F), in PCV13 (all PCV10 serotypes plus 3, 6A and 19A), PPV23 (all PCV13 serotypes, except for serotype 6A, plus serotypes 2, 8, 9N, 10A, 11A, 12F, 15B, 17F, 20, 22F and 33F) and NVT that include all others serotypes and non-typable isolates.

Because serotypes 15B and 15C have a high frequency of capsular switching, we choose to group isolates with these serotypes into a single group, named 15B/C. Due to difficulties in phenotypically distinguishing isolates from serotype 25A and serotype 38, and serotype 29 and serotype 35B, it was decided to also group into 25A/38 and 29/35B, respectively.

In cases where the identification was performed by molecular methods, serotyping was realized by rPCR using 7 multiplex reactions that target three serotypes/serogroups: 3, 7F/7A and 19A; 1, 15B/C and 23F; 14, 18C and 19F; 4, 6 and 9V/9A; 5, 11A/11D and 16F; 8, 12F/12A/12B and 22F/22A; and 15A, 23A and 33F/33A/37 [147].

3.3 Antimicrobial susceptibility testing

Antimicrobial susceptibility was performed in Mueller-Hinton agar (Oxoid, Hampshire, United Kingdom) supplemented with 5% defibrillated sheep blood (Probiológica, Belas, Portugal) plates, inoculated with a 0.5 McFarland colony suspension and incubated overnight at 35°C with the antibiotics in an atmosphere enriched with 5% CO_2.

For levofloxacin, norfloxacin, erythromycin, clindamycin, telithromycin, tetracycline, chloramphenicol, trimethoprim-sulfamethoxazole, vancomycin and linezolid, the Kirby-Bauer disk diffusion method was used by commercial disks (Oxoid, Hampshire, United Kingdom), according to the CLSI recommendations, except for norfloxacin where it was used the EUCAST recommendations [149,150]. In the case of penicillin, cefotaxime and levofloxacin (when the isolates were resistant to norfloxacin, using the Kirby-Bauer disk diffusion method) were determined by the MIC, using Etest strips (bioMérieux, Marcy-I'Étoile, France). The reference strain used as a control was *S. pneumoniae* ATCC49619.

Table 4 – Concentration and breakpoints for used antibiotics.

	Antibiotics	Concentration (μg/mL)	S	Non-susceptible	
				I	R
Etest	Penicillin				
	Parenteral	0.002-32	≤2	4	≥8
	Meningitis		≤0.06	-	≥0.12
	Cefotaxime				
	Meningitis	0.002-32	≤0.5	1	≥2
	Nonmeningitis		≤1	2	≥4
Disk	Levofloxacin	5	≥17	14-16	≤13
	Norfloxacin		-	-	<11
	Erythromycin	15	≥21	16-20	≤15
	Clindamycin	2	≥19	16-18	≤15
	Tetracycline	30	≥28	25-27	≤24
	Telithromycin	15	≥19	16-18	≤15
	Chloramphenicol	30	≥21	-	≤20
	Trimethoprim-sulfamethoxazole	25	≥19	16-18	≤15
	Vancomycin	30	≥17	-	-
	Linezolid	30	≥21	-	-

"S"- susceptible; "I"- intermediate and "R"- resistant

Macrolide resistance phenotypes (M and MLS_B) were determined by the double disk method, in which erythromycin and clindamycin disks were placed 1 cm apart in a Mueller-Hinton agar plate inoculated. Simultaneous resistance to erythromycin and clindamycin characterizes the MLS_B phenotype (resistance to macrolides, lincosamides and streptogramins B), which can be constitutive ($cMLS_B$) or inducible ($iMLS_B$) (Figure 11 A

and B). Non-susceptibility to erythromycin only (Figure 11 C), indicates the M phenotype (resistance only to macrolides).

The antimicrobial susceptibility testing and serotyping results for the pneumococcal strains isolated in the years 2015 and 2016 were already available; for the purpose of this thesis, only the phenotypic characterization of isolates collected in 2017 was performed. As indicated below (DNA extraction and sequencing), the molecular characterization included isolates collected in all three years.

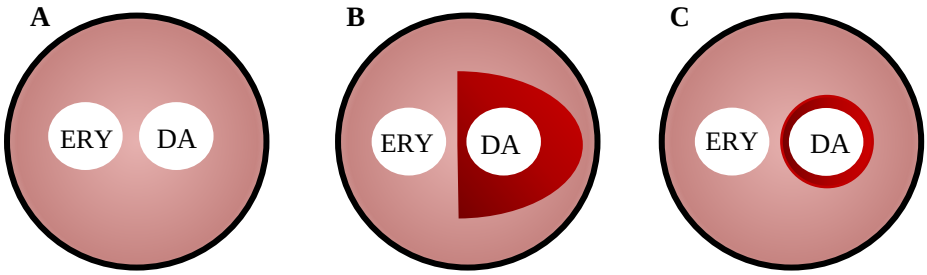

Figure 11 – Scheme of macrolide phenotypes identification. A- phenotype cMLS$_B$; B- phenotype iMLS$_B$ and C- phenotype M. "ERY"- Erythromycin; "DA"- Clindamycin.

3.4 DNA extraction and sequencing

Of the total strain collection, 50% of the strains of each serotype containing 10 or more isolates were selected for WGS. Whenever 50% of the strains of a given serotype represented less than 10 isolates, we sequenced 10 strains. For this, the bacteria were grown in Brain Heart Infusion (BHI) broth (Becton, Dickinson, New Jersey, USA) overnight at 37°C until an optical density between 0.7-0.9 was reached.

DNA was extracted using the PureLink Genomic DNA Mini Kit (ThermoFisher Scientific, Massachusetts, USA), according to the manufacturer's protocol, except for the addition of 20 µL of RNAse incubated for 5 min at room temperature before the addition of 200 µL of ethanol and a final elution volume of 5 µL. NanoDrop 2000 (ThermoFisher Scientific, Massachusetts, USA) was used to evaluate the purity of the DNA and a 1% agarose gel with 5 µL of sample with 2 µL of loading buffer and a 1 Kb plus DNA ladder run for approximately 60 minutes at 120 V was used to check the quality of the pneumococcal DNA. Qubit (Invitrogen by ThermoFisher Scientific, Massachusetts, USA) was used to determine the DNA concentration, and the dilutions were performed with Tris-HCl 10 mM pH 8 to obtain a final concentration of 10 ng/µL.

WGS libraries were prepared using paired-end Nextera XT DNA Library Prep Kit, Index Kit v2 (Illumina, San Diego, CA, USA) and sequenced on Illumina NextSeq 500 system (Illumina) using NextSeq 5007550 Mid-Output v2 Kit (300 cycles) at Instituto Gulbenkian de Ciências's Genomics Unit (Oeiras, Portugal).

The quality assessment, control and *de novo* assembly of paired-end reads obtained was performed with the INNUca pipeline (https://github.com/B-UMMI/INNUca). INNUca v4.0 was run using a Docker image, providing Nextera XT adapter sequences for adapter removal and using an expected genome size of 2.1 Mb.

Briefly, the quality of the reads was checked with FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and cleaned using Trimmomatic [151]. *De novo* assembly was performed using SPAdes [152] and subsequently polished using Pilon [153]. The coverage was evaluated based on genome size and on the genes included in the MLST schema of *S. pneumoniae*. Contamination is evaluated using Kraken [154] and the presence of multiple alleles in the genes included in the MLST schema of *S. pneumoniae*. MLST type was determined for the final draft assembly using the MLST software (https://github.com/tseemann/mlst), with alleles and STs being assigned using the pneumococcal database available at https://pubmlst.org/spneumoniae/. The draft assemblies were annotated using Prokka (https://github.com/tseemann/prokka). This analysis was performed by the bioinformatics group of MRamirez lab.

When new alleles were found, they were submitted to the pneumococcal database curator for number attribution. Genetic lineages were assigned using the goeBURST algorithm and all the STs available in the online pneumococcal MLST database (https://pubmlst.org/spneumoniae/) in December 2019. Results were visualized at PHYLOViZ [144]. CCs were defined at the single-locus-variant (SLV) level.

All sequenced strains were serotyped by SeroBa version 1.0.1 [145] with the provided database (updated at 7 August 2017) and the default parameters. When the serotyping and SeroBa results were not concordant, we used PneumoCAT, with default parameters, as it uses a strictly mapping based approach.

3.5 Statistical Analysis

To evaluate the diversity of the population, the Simpson's Index of Diversity (SID) was determined, which measures the discriminatory ability of typing systems (i.e., probability of two samples, randomly pick from a population, belonging to two different types) [155]. The Adjusted Wallace coefficient (AW) was calculated to determine how much new information was obtained from a typing method, given another method [155,156]. A high value of AW suggests that partitions given by one method could have been predicted by another, i.e., the use of both methods is redundant [157]. These two measures were obtained using the tool available at http://www.comparingpartitions.info.

Odds ratio (OR) was determined to assess the association between variables and the Cochran-Armitage test was used to calculate trends. The p-values obtained in both statistical methods were corrected using the False Discovery Rate (FDR) [158], and these two methods were estimated using the Epitools and Multtest packages for the R language. A $p<0.05$ was considered significant.

4. RESULTS

4.1 Isolate collection

A total of 1608 cases of adult invasive pneumococcal disease were reported between 2015 and 2017, of which 1607 were isolates and in one case, it was only possible to perform the diagnosis by PCR detection of pneumococcal DNA. Regarding the biological source of the isolates, 1499 (93.2%) were recovered from blood, 74 (4.6%) from CSF, 19 (1.2%) from pleural fluid, 9 (0.6%) from peritoneal fluid, 5 (0.3%) from synovial fluid and 2 (0.1%) from pericardial fluid. The case detected by molecular methods was recovered from CSF. The total number of cases per year remained constant, with 530 cases in 2015, 502 in 2016 and 576 in 2017.

The mean age of the patients was 66.9 and the mode was 79. Regarding the distribution of patients by age group, 299 (18.6%) of the cases were from young adults (between 18 and 49), 362 (22.5%) from adults with ages between 50 and 64 and lastly 947 (58.9%) from elderly (with 65 or more years old). Considering the gender of the patients, 59.8% (n=961) were male and 40.2% (n=647) were female.

4.2 Serotype distribution

Within this collection, 56 different capsular types were identified (SID=0.924; CI95%: 0.918-0.931), where serotypes 8 (n=287, 17.8%), 3 (n=236, 14.7%), 22F (n=119, 7.4%), 14 (n=111, 6.9%), 19A (n=100, 6.2%) and 9N (n=66, 4.1%) were the most frequent, being responsible for 57.1% (n=919) of invasive adult infections between 2015 and 2017.

The distribution of serotypes regarding age groups was highly diverse, but the overlapping of the confidence intervals implies that all age groups were similarly diverse. In young adults, 43 serotypes were detected (SID=0.894; CI95%: 0.868-0.921), in adults in between 50 and 64 years old, 44 serotypes were detected (SID=0.911; CI95%: 0.893-0.928) and in the elderly were detected 50 serotypes (SID=0.928; CI95%: 0.921-0.935). In young adults, the most frequent serotypes were: 8, 3, 22F, 9N, 19A and 20 (by decreasing order), accounting for 57.5% of the isolates. For the adults between 50 and 64 years old, the main serotypes were serotypes 8, 3, 14, 11A, and 19A (by decreasing order), together accounting for 54.2% of the isolates. For the elderly, serotypes 3, 8, 22F, 14, and 19A (by decreasing order) accounted for 53.5% of the isolates. The most important serotypes for all

age groups are represented in the Figures 12 and 13. In the first two age groups, young adults and adults aged between 50 and 64 years, serotype 8 was the most frequent serotype, responsible for 28.8% and 21.6% of the infections, respectively. However, in the last age group, the most frequent cause of infection was serotype 3 (15.8%). Nevertheless, analysis of the association between age groups and serotypes, revealed that serotype 8 had a positive association with adults 50-64 years (p=0.042) significant before FDR correction and with young adults (p<0.001) significant after FDR correction. Serotype 3 did not present an association with any of the age groups. Young adults also presented an association with serotypes 1, 4 and 7F significant after FDR correction (p=0.022, p=0.005 and p=0.035, respectively). Adults with ages between 50 and 64 were not associated with any other serotype and the elderly were associated with serotypes 31, 22F and 6C (p=0.003, p<0.001 and p=0.001, respectively, all significant after FDR correction) and serotype 19A significant before FDR correction (p=0.021) (Results resume in Table S1 of the Supplementary Data).

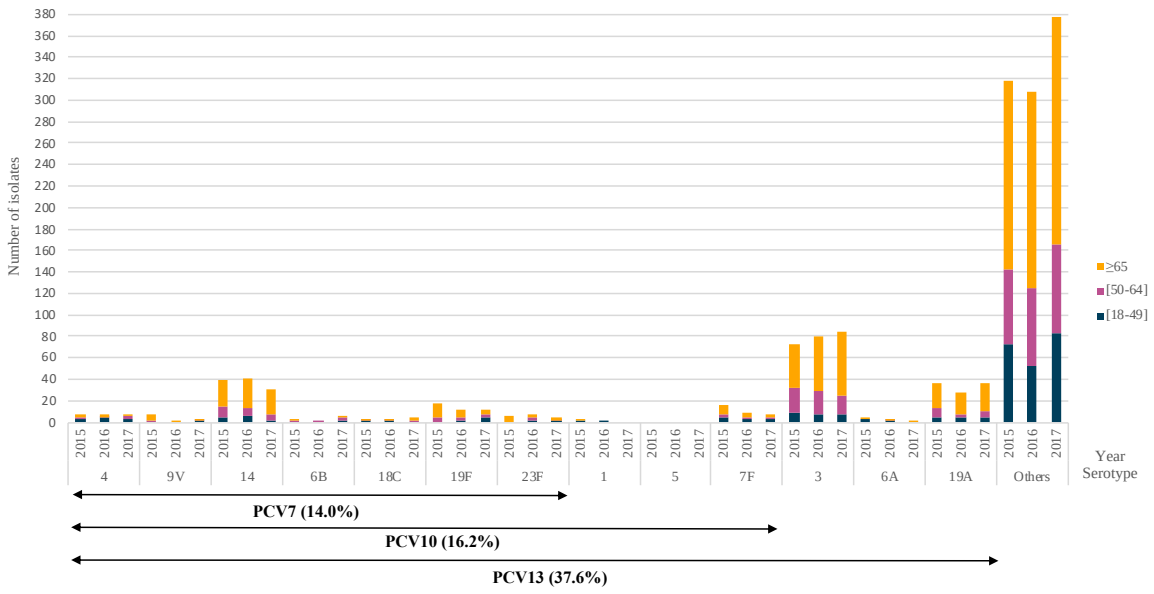

Figure 12 – PCV vaccine serotypes distribution by age groups ([18-49]; [50-64] and ≥65 years old) in 2015, 2016 and 2017. Arrows represent the serotypes presented in PCV7, PCV10 and PCV13. "Others" refers to all serotypes presented in the study population, except for serotypes presented in PCV vaccines.

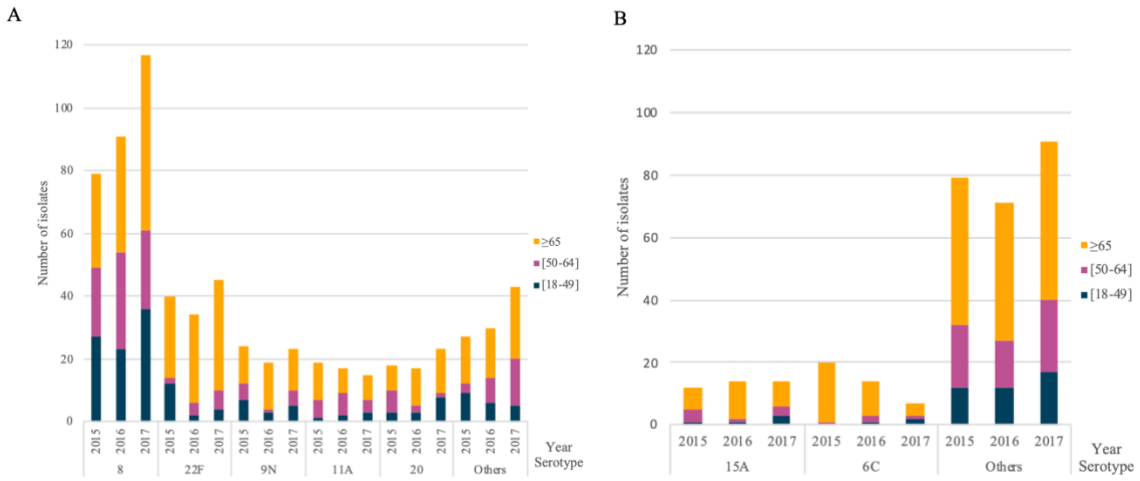

Figure 13 – Distribution of Non-PCV serotypes by age groups ([18-49]; [50-64] and >65 years old) in 2015, 2016 and 2017. A - Distribution of serotypes presented only in PPV23, where "Others" refers to all serotypes present in PPV23, which have one age group and year with <10 isolates. B – Distribution of NVT serotypes, where "Others" refers to all serotypes not included in a pneumococcal vaccine (PCVs and PPV23), which have one age group and year with <10 isolates.

PCV7 vaccine serotypes constituted 14.0% (n=225) of all isolates. AddPCV10 serotypes (serotypes present in PCV10, but not in the PCV7) were responsible for 2.2% (n=36) of the isolates and the addPCV13 serotypes (serotypes present in the PCV13, but not in the PCV7 and PCV10) were responsible for 21.4% (n=344) of the isolates. Serotype 5, one of the serotypes presented in PCV10 was not detected in the study population (Figure 12). Non-PCV serotypes held for 62.4% (n=1003) of the cases, of which 67.9% (n=681) were serotypes exclusively found in PPV23 (addPPV23) (Figure 13).

Serotype diversity was high in all study years and similar, due to the SID confidence intervals overlapping (Table below). The five most frequent serotypes (serotypes 8, 3, 22F, 14 and 19A) were the most frequent serotypes in all three study years, despite different distributions (Figures 12 and 13). Serotype 1 was isolated only in 2015 (n=2) and 2016 (n=2) (Figure 12). Some serotypes appeared only in one of the years, which was the case of serotypes: 37, 39, 24A, 28A, 33A and 35A in 2015; 43, 11D, 11F, 24B and 35B in 2016 and 12A, 15F, 18A, 6D and 9A in 2017.

Table 5 – Simpson's index of diversity and respective 95% confidence intervals for serotypes distribution per study year.

Year	SID (CI95%)
2015	0.934 (0.925-0.943)
2016	0.920 (0.908-0.932)
2017	0.917 (0.905-0.930)

Cochran-Armitage test was used to examined possible trends in serotypes distribution over the study period. Serotype 8 showed an increasing trend (p=0.019, significant before FDR correction) and serotypes 6C and 7F showed a decreasing trend (p=0.007 and p=0.033, respectively, significant before FDR correction). This test was also used to evaluate serotype variation between 2012-2017 [127], where serotypes 8 and 12F increased (both p<0.001) significantly after FDR correction and serotypes 1, 12B and 7F decrease significantly after FDR correction (all p<0.001) (results resume in Table S2 of the supplementary data).

4.3 Antimicrobial susceptibility

Antimicrobial susceptibility testing was performed for the 1607 isolates and it is resumed in the figure below. No resistance to vancomycin and linezolid was detected. There was no statistically significant change within the proportion of non-susceptibility to any antimicrobial class between 2015 and 2017.

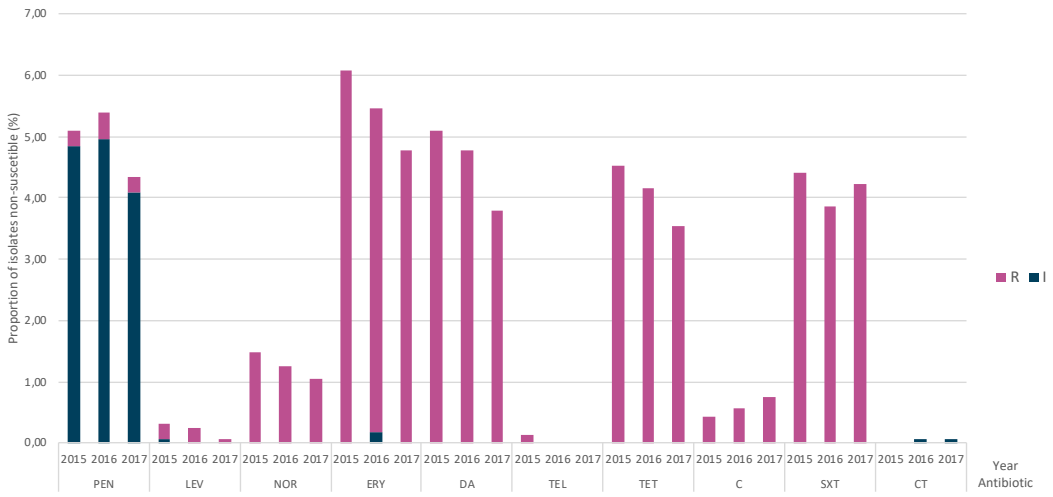

Figure 14 – Resume of the antimicrobial non-susceptibility isolates included in this study. “PEN”- penicillin; “LEV”- levofloxacin; “NOR”- norfloxacin; “ERY”- erythromycin; “DA”- clindamycin; “TEL”- telithromycin; “TET”- tetracycline; “C”- chloramphenicol; “SXT”- trimethoprim-sulfamethoxazole; “CT”- cefotaxime.

Overall, isolates Penicillin Non-Susceptible Pneumococci (PNSP) represented 14.9% (n=239) of the collection, of which 58.6% (n=140) were also non-susceptible to erythromycin (EPNSP, isolates non-susceptible to penicillin and erythromycin) (Figure 15). However, if the current CLSI breakpoints for parenteral penicillin [104] were considered, only 1.1% (n=17) of the isolates would have been considered resistant to penicillin. A total of 263 strains (16.4%) presented non-susceptibility to erythromycin (ERSP), of which 220 isolates (83.7%) expressed the cMLS$_B$ phenotype and 43 isolates (16.3%) expressed the M phenotype.

In most of the cases, the penicillin non-susceptible isolates were intermediately resistant (n=224, 93.7%). Moreover, only two cefotaxime non-susceptible isolates were detected, also intermediately resistant. Resistance to telithromycin (n=2, 0.1%) was observed only in 2015. Among the collection, 501 (31.2%) of the strains presented non-susceptibility to at least one class of antibiotics, and 226 strains (14.1%) were MDR, i.e., non-susceptible to at least three antibiotics.

The correlation between serotype and antimicrobial resistance was determined by the AW coefficient. The AW for serotype and PNSP was 0.649 (CI95%: 0.581-0.717), and the AW for serotype and ERSP was 0.5687 (CI95%: 0.495-0.640). Twenty-seven different serotypes were detected among the PNSP strains, where the most prevalent serotypes were 14, 19A, 15A, 6C and 11A (in decreasing order), responsible for 74.9% of these isolates. In the case of ERSP, 30 serotypes were found, where serotypes 14, 19A, 6C, 19F and 15A (in decreasing order) accounting for 68.1% of these isolates. Within the EPNSP, 16 serotypes were detected, in which serotypes 14, 19A, 15A, 6C and 19F (by decreasing order) were the most prevalent serotypes, contributing to 87.9% of the EPNSP (Figure 15).

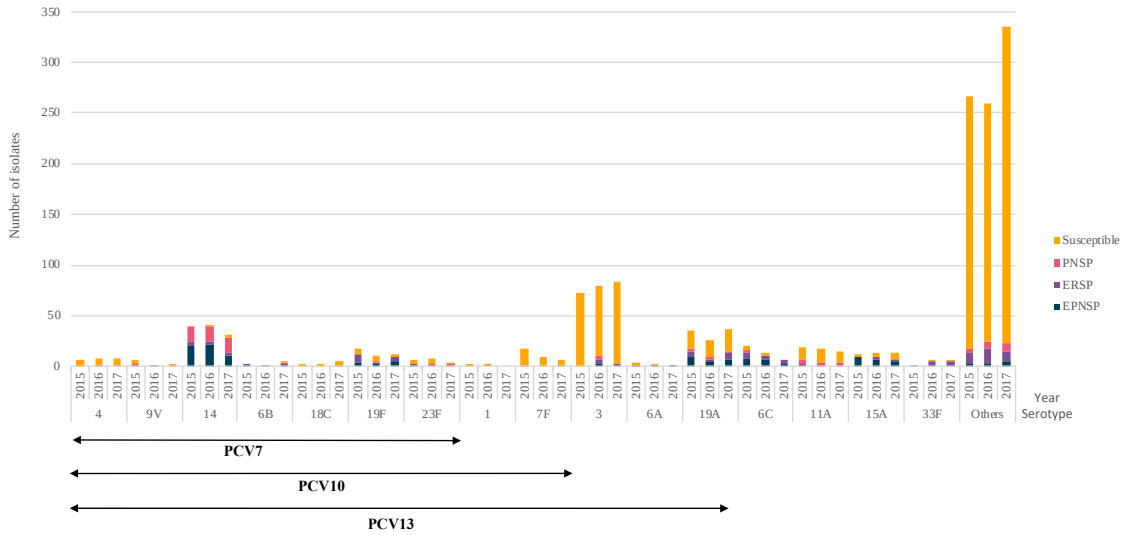

Figure 15 – Distribution of penicillin and erythromycin susceptibility profiles for isolates of *S. pneumoniae* serotypes by study year (2015, 2016 and 2017). The arrows represent serotypes presented in PCV7, PCV10 and PCV13. "Others"- all serotypes presented in the study population, except for serotypes presented in PCV vaccines and serotypes 6C, 11A, 15A and 33F; "PNSP"- penicillin non-susceptible isolates; "ERSP"- erythromycin resistant isolates; and "EPNSP"- isolates presenting erythromycin resistance and penicillin non-susceptibility.

PCV7 and addPCV13 vaccine serotypes represented 51.0% (n=122) and 15.5% (n=37) of the PNSP isolates, respectively, where there were no isolates non-susceptibility to penicillin among the addPCV10 vaccine serotypes. In ERSP isolates, PCV7 and addPCV13 vaccine serotypes represented 37.6% (n=99) and 19.0% (n=50) of the strains, respectively (no strain presented resistance to erythromycin among the addPCV10 vaccine serotypes).

When analyzing the association between antimicrobial non-susceptibility and serotype, we found that PNSP strains were significantly associated with serotype 14, 15A, 19A and 6C (all $p<0.001$), significant after FDR correction and serotypes 19F, 23F and 9V ($p=0.014$, $p=0.011$ and $p=0.022$, respectively), significant only before FDR correction. Regarding ERSP strains, serotypes 14, 15A, 19A, 19F, 33A, 33F, 6B and 6C (all $p<0.001$) presented a significant association after FDR correction. It was also observed an association between PNSP, ERSP and the elderly population, only significant before FDR correction ($p=0.032$ and $p=0.023$, respectively) and EPNSP and the same group of patients, significant after FDR correction ($p=0.015$) (results resume in the Tables S3 and S4 of the supplementary data).

Regarding other antibiotics, 13.7% (n=220) were clindamycin resistant, 12.3% (n=197) were tetracycline resistant and 12.5% (n=201) were trimethoprim-sulfamethoxazole resistant (Table 6). The most frequent serotypes presented by isolates resistant to these antibiotics are resume in the table presented below. Serotypes 14, 19A,

6C, 15A and 19F (by decreasing order) were responsible for 75.9% of clindamycin resistant isolates; serotypes 19A, 6C, 19F, 3 and 15A (by decreasing order) contributed to 60.9% of the tetracycline resistant isolates and serotypes 14, 19A, 24F, 11A and 19F (by decreasing order) were held for 56.2% of the trimethoprim-sulfamethoxazole resistant strains.

Table 6 – Resume of the most frequent serotypes among non-susceptible isolates in adults between 2015 and 2017.

Serotype	n	S	Non-susceptible to												
			PEN	LEV	NOR	ERY	DA	TEL	TET	SXT	C	CT	VA	LZ	MDR
3	236	200	6	4	18	11	10	0	18	6	3	0	0	0	12
6C	41	5	19	1	2	32	32	0	31	1	0	0	0	0	31
8	287	268	3	1	5	5	4	0	10	2	3	0	0	0	6
11A	51	33	11	0	5	3	0	0	2	11	0	0	0	0	2
14	111	3	99	1	3	61	50	0	13	54	2	2	0	0	57
15A	40	15	22	0	1	24	24	0	14	2	0	0	0	0	22
19A	100	54	28	2	4	36	36	0	38	23	3	0	0	0	37
19F	41	8	12	0	2	26	25	0	19	10	4	0	0	0	19
23F	18	9	7	0	0	2	0	2	0	2	0	0	0	0	1
24F	19	3	2	0	0	4	4	0	3	15	0	0	0	0	4
33F	14	3	0	0	0	11	11	0	8	0	0	0	0	0	8
Others	649	505	30	1	21	48	24	0	41	75	13	0	0	0	27
Total	1607	1106	239	10	61	263	220	2	197	201	28	2	0	0	226

"n"- total number of isolates; "S"- susceptible strains "PEN"- penicillin; "LEV"- levofloxacin; "NOR"- norfloxacin; "ERY"- erythromycin; "DA"- clindamycin; "TEL"- telithromycin; "TET"- tetracycline; "SXT"- trimethoprim-sulfamethoxazole; "C"- chloramphenicol; "CT"- cefotaxime; "VA"- vancomycin; "LZ"- linezolid; "MDR"- Multi-drug resistant; "Others"- other serotypes and non-typable strains detected. At bold is represented the PCV serotypes.

Serotype 19A was the serotype who contributed the most for tetracycline resistance (n=38, 19.3%), the second most frequent serotype among clindamycin resistant strains (n=35, 15.9%) and trimethoprim-sulfamethoxazole resistant isolates (n=23, 11.4%). Most strains of serotype 14 presented non-susceptibility to at least one antibiotic (n=108, 97.3%), being the ones who contributed the most for clindamycin resistance (n=51, 23.2%) and trimethoprim-sulfamethoxazole resistant isolates (n=53, 24.1%). However, among tetracycline resistant isolates, it was not one of the most frequent serotypes, being responsible for only 6.6% (n=13) of the cases. Concerning MDR strains, serotypes 14, 19A, 6C and 15A (in decreasing order) were the most frequent serotypes, responsible for 65.0% of these isolates.

To evaluate the association between the resistance to these antibiotics and their most frequent serotypes and age groups, the OR was calculated, being the results resume in Tables S5 and S6 of the supplementary data.

A diverse variety of serotypes was found in levofloxacin, norfloxacin and chloramphenicol resistant isolates, highlighting, in the case of norfloxacin, serotype 3, which was responsible for 29.5% (n=18) of these isolates. The two strains showing non-susceptibility to telithromycin were from serotype 23F, and both strains non-susceptible to cefotaxime were from serotype 14 (Table 6).

Cochran-Armitage test was used to evaluate antimicrobials non-susceptibility variation between serotypes from 2012-2017 [127], where it was observed an increased trend of serotypes 8 and 22F (both p=0.040) significant before FDR correction and a decrease trend of serotype 19A (p<0.001) significant after FDR correction among the PNSP isolates. In ERSP isolates, serotypes 3, 22F and 33F showed an increasing trend (p=0.014, p=0.026 and p=0.021, respectively), while serotypes 19A, 33A and 33B showed a decreasing trend (p=0.007, p=0.034 and p=0.031, respectively) before FDR correction (results resume in Table S7 of the supplementary data). Serotype 14 was only correlated with a decreasing trend (p=0.044) significant before FDR correction among the cefotaxime non-susceptible isolates between 2012-2014 and 2015-2017. In serotype 19A, resistance to penicillin, erythromycin, tetracycline, trimethoprim-sulfamethoxazole and clindamycin presented an overall decreasing trend, when comparing both periods (results resume in Table S7 of the supplementary data).

4.4 SeroBa

Of the 807 strains sequenced, 37 strains showed inconsistent serotyping results. Although most of the differences (n=27) were due to incorrect annotation, specially due to the annotation of serotype 12F as serotype 12B in the phenotypic serotyping assay, since the agglutination with 12b serum corresponds to serotype 12F, 10 strains (1.2%) still presented a discrepancy between phenotyping serotyping and SeroBa analysis. PneumoCAT results for these 10 strains were concordant with the SeroBa results. The strains were cultured and phenotypic serotyping was repeated having confirmed the initial results. Since the STs of these strains, as determined from the genomic information, showed a better association to the *in sillico* serotyping results, it was assumed that there was an error in the labelling during DNA extraction process of this strains and these strains were removed from the PHYLOViZ analysis.

4.5 Molecular Characterization

In addition to the 10 strains removed from the PHYLOViZ analyses due to inconsistencies between the phenotypic serotyping results and the genotypic serotyping results, 9 strains were also removed since they presented contaminations in their sequence results. Moreover, for 55 isolates initially selected, the sequencing results did not arrive in time and were excluded from the analysis. Given these limitations, the final collection of sequenced isolates comprised a total of 726 strains.

Among these, 172 different STs (SID=0.956, CI95%: 0.947-0.964) were found, grouped in 56 different CCs (SID=0.789, CI95%: 0.761-0.817) according to the goeBURST, when used all STs deposited in the MLST database (December 2019). The most frequent STs (grouping more than 10 isolates) were ST53 (mainly serotype 8; n=116, 16.0%), ST180 (serotype 3; n=62, 8.5%), ST433 (serotype 22F; n=51, 7.0%), ST66 (serotype 9N; n=30, 4.1%), ST143 (serotype 14; n=23, 3.2%), ST232 (serotype 3; n=19, 2.6%), ST156 (mainly serotype 14; n=17, 2.3%), ST408 (serotype 11A; n=14, 1.9%), ST386 (serotype 6C; n=12, 1.7%), ST235 (mainly serotype 20; n=10, 1.4%) and ST63 (mainly 15A; n=10, 1.4%), accounting for half of the isolates analyzed (50.1%).

Twenty-five new alleles were identified, of which 16 were designated: 505 (ST15150) for the *aroE*; 976 (ST15149), 977 (ST15157), 978 (ST15151) and 979 (ST 15152) for the *ddl*; 658 (ST15145) for the *gdh*, 693 (ST15156), 694 (ST15144), 695 (ST15153) and 696 (ST15147) for the *gki*; 459 (ST15148) for the *recP* and 643 (ST15155), 644 (ST15146) and 645 (ST15154) for the *spi*. The other 9 alleles were sent to the MLST database curator for allele number attribution. However, at the time of the writing of this thesis, the allele numbers have not been attributed yet, so a new allele number (absent from the MLST database) was ascribed to each of the alleles and profiles (STs: 30001, 30003, 30005, 30009, 30010, 30012, 30015, 30016 and 30017).

Twenty-nine new allelic combinations were identified as STs: 15144, 15145, 15146, 15147, 15148, 15149, 15150, 15151, 15152, 15153, 15154, 15155, 15156, 15157, 15064, 15065, 15066, 15067, 15068, 15069, 15070, 15071, 15081, 15082, 15083, 15084, 15085, 15086 and 15087.

There was a low correlation between CCs and serotypes (AW=0.308, CI95%: 0.232-0.384). The serotype distribution among the STs found in the 13 major CCs (n≥10 isolates) are presented in the Table below, accounting for 86.4% (n=627) of the sequenced isolates and Figure 16 represents the STs and the main CCs found among the collection.

CC156 (n=309) expressed mostly vaccine serotypes, especially PPV23 serotypes (61.8% of CC156 isolates), particularly serotypes 8, 9N, 11A and 17F. Among the major CCs, isolates presenting PPV23 serotypes isolates, were grouped also in other CCs, mainly serotype 22F in CC433, serotype 10A in CC97 and serotype 20 in CC235. PCV7 serotypes were mostly included in CC156 (25.9% of the isolates belonging to this CC), namely serotypes 14, 23F and 19F, and also in CC180 (serotype 18C).

AddPCV13 serotypes were included in diverse CCs, manly in CC180, CC260, CC378 (serotype 3), CC199 and CC994 (serotype 19A). Serotype 7F isolates, were included in CC191. The information regarding the STs found in CCs with less than 10 isolates are shown in Table S8 of the supplementary data.

Table 7 – Serotypes distribution of the most frequent STs found in the 13 major CCs (n≥10 isolates) identified by goeBURST.

CC (n)	ST	Total	Dominant serotype (n)	Other serotypes
156 (309)	53	116	8 (115)	3 (1)
	66	30	9N (30)	-
	143	22	14 (22)	-
	156	17	14 (16)	9V (1)
	408	14	11A (14)	-
	63	10	15A (9)	11A (1)
	162	6	24F (5)	9V (1)
	338	6	23F (6)	-
	392	6	17F (6)	-
	Others	82	19F (14)	8 (8), 14 (10), 11A (9), 12F (2), 15A (8), 15B/C (1), 17F (2), 19A (9), 22F (1), 23B (2), 23F (2), 24F (2), 33F (1), 6B (1), 6C (2), 9N (1), 9V (7)
180 (98)	180	62	3 (62)	-
	1766	9	31 (9)	-
	113	8	18C (7)	8 (1)
	Others	19	3 (15)	31 (3), 9N (1)
433 (55)	433	51	22F (51)	-
	10220	4	22F (4)	-
378 (22)	232	19	3 (19)	-
	1377	3	3 (3)	-
97 (21)	97	10	10A (10)	-
	Others	11	35F (9)	10A (2)

Table 7 (continued)

CC (n)	ST	Total	Dominant serotype (n)	Other serotypes
235 (21)	235	10	20 (9)	22F (1)
	10047	7	20 (7)	-
	Others	4	20 (4)	-
439 (21)	439	5	23B (5)	-
	Others	16	23A (12)	23B (3), 23F (1)
199 (15)	416	5	19A (5)	-
	Others	10	15B/C (5), 19A (5)	-
260 (15)	260	8	3 (8)	-
	1220	6	3 (6)	-
	15155	1	3 (1)	-
315 (15)	386	12	6C (12)	-
	Others	3	6C (3)	-
994 (13)	994	9	19A (9)	-
	4197	4	19A (4)	-
30 (11)	30	9	16F (9)	-
	Others	2	16F (1), 6C (1)	-
191 (11)	191	9	7F (9)	-
	Others	2	7F (2)	-

"Others": STs with <10 isolates.

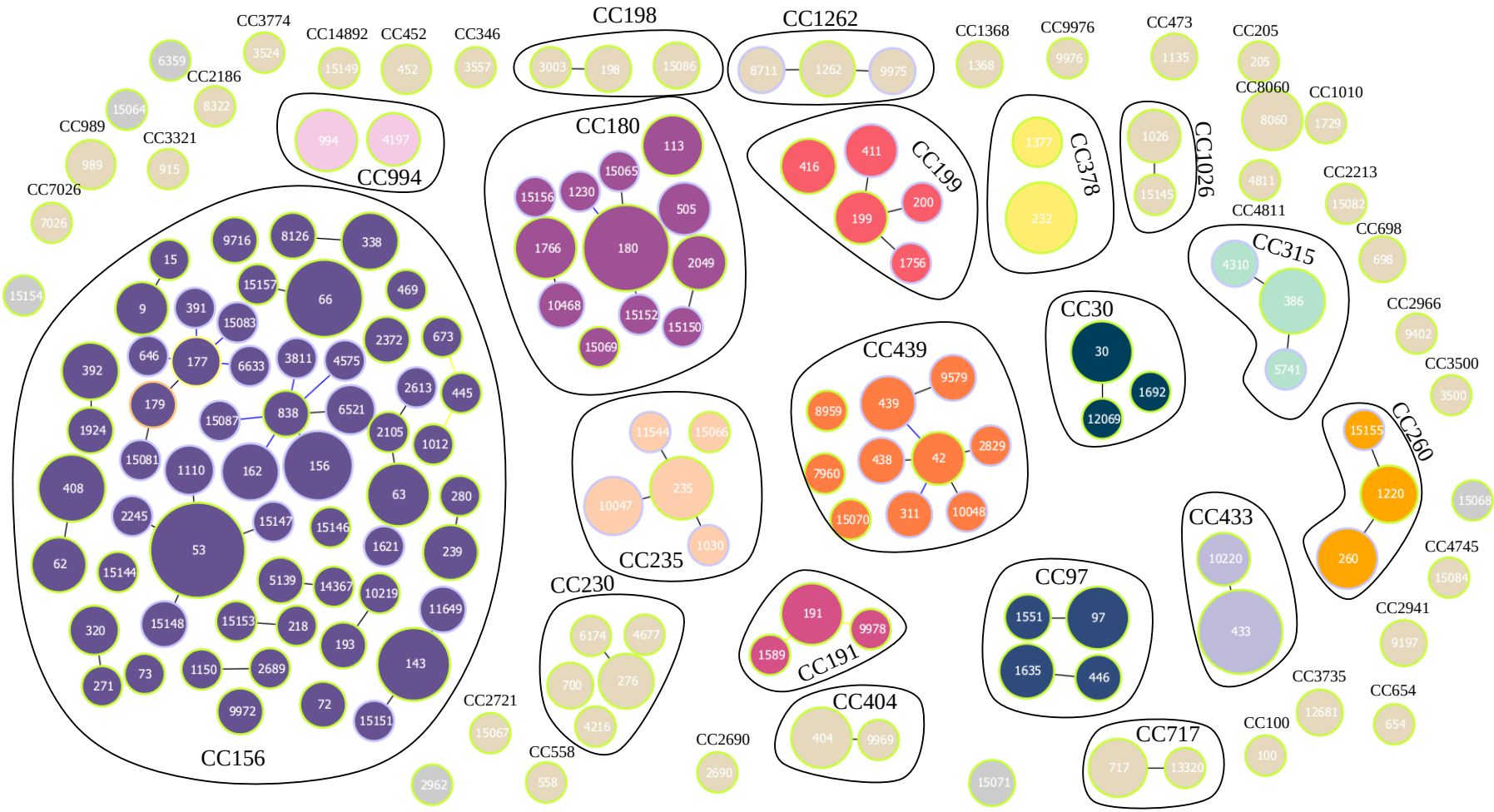

Figure 16 – Representation of all STs identified and indication of their CCs, highlighting the major CCs. STs linked by a line are SLVs and STs in grey are singletons. The scheme was done using PHYLOViZ [144].

There was a high correlation between STs and serotypes (AW=0.982, CI95%: 0.960-1.000), although some STs presented more than one serotype (Tables 7 and S8), which is consistent with the relatively low AW between serotypes and STs (AW=0.496, CI95%: 0.442-0.549). The genetic diversity varied with the serotypes, with some presenting a high diversity (SID>0.8), such as serotypes 19A, 19F, 9V, 20, 15B/C, 23A, 4 and 29/35B and others a low diversity (SID<0.3), such as serotypes 18C, 8, 22F and 9N (Figures 17, 18 and 19). Overall, there was a higher genetic diversity among PCV13 serotypes and a lower genetic diversity among PPV23 and NVT serotypes. However, for serotypes 19F, 9V and 23A, the wide variety of STs did not result in a high diversity of CCs, with a max of 3 CCs being detected.

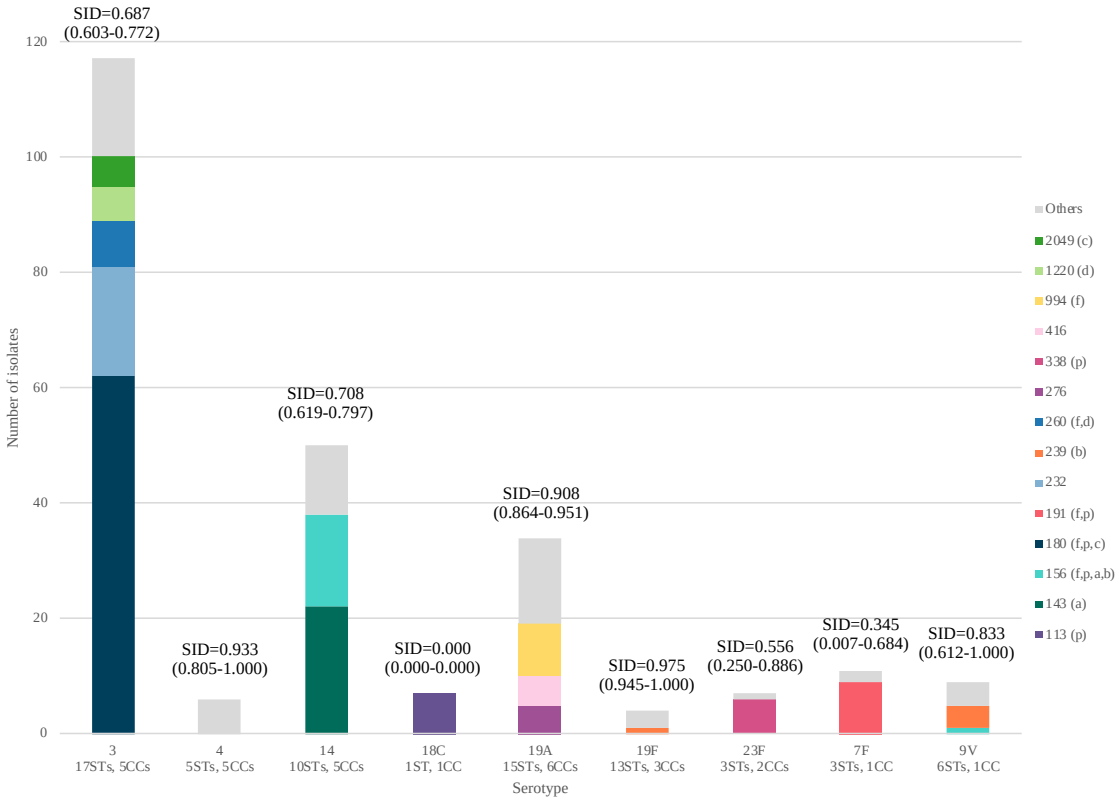

Figure 17 – Distribution of STs according to serotypes of the isolates causing adult IPD in 2015-2017 and expressing serotypes included in the conjugate vaccines. The respective SID values are indicated on top the bars and their respective confidence intervals are in parenthesis. "f" – represents the STS considered by goeBURST as founders of a CC; "p" – represents STs that matched with PMEN clones; "a", "b", "c" and "d" – represent STs that belong to the same CC in each serotype; "Others" – represent the isolates included in STs with <4 isolates.

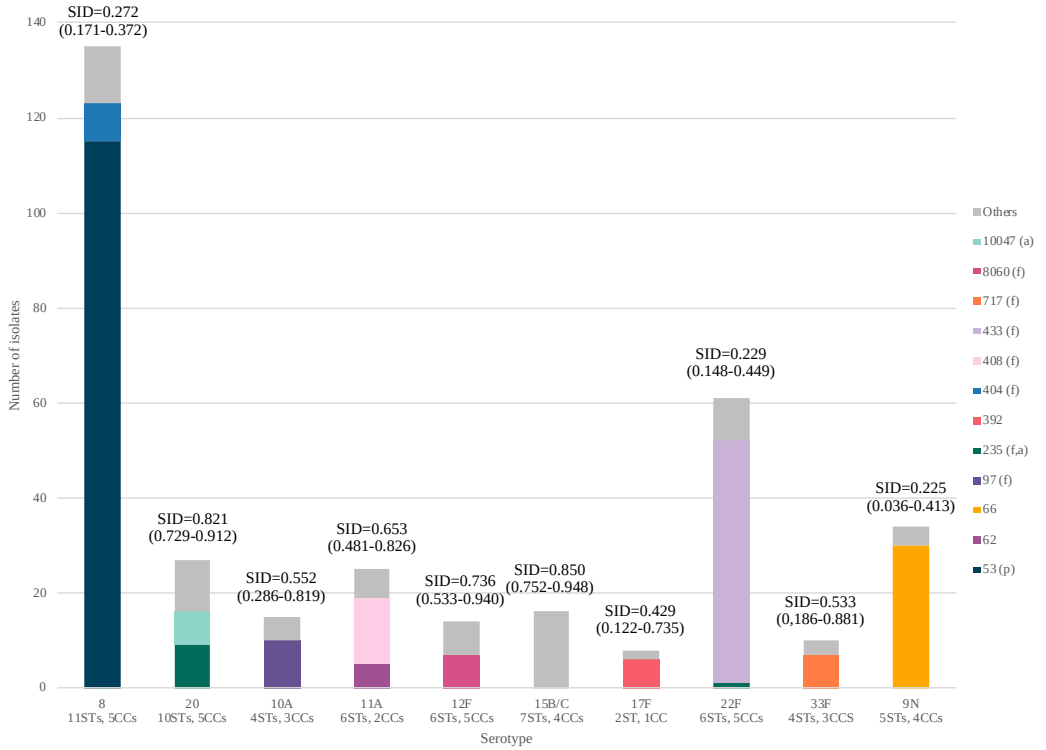

Figure 18 – Distribution of STs according to serotypes of the isolates causing adult IPD in 2015-2017 and expressing serotypes included exclusively in the PPV23 vaccine. The respective SID values are indicated on top the bars and their respective confidence intervals are in parenthesis. "f" – represents the STS considered by goeBURST as founders of a CC; "p" – represents STs that matched with PMEN clones; "a" – represent STs that belong to the same CC in each serotype; "Others" – represent the isolates included in STs with <4 isolates.

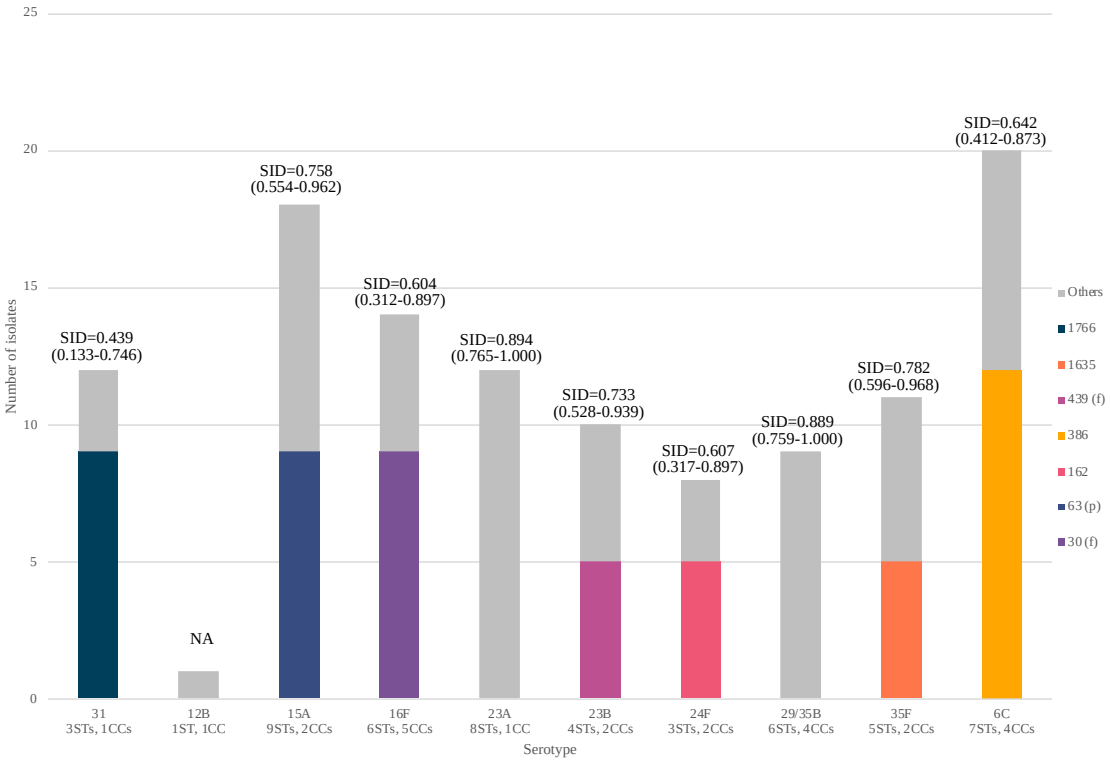

Figure 19 – Distribution of STs according to serotypes of the isolates causing adult IPD in 2015-2017 and expressing NVT serotypes. The respective SID values are indicated on top the bars and their respective confidence intervals are in parenthesis. "f" – represents the STS considered by goeBURST as founders of a CC; "p" – represents STs that matched with PMEN clones; "Others" – represent the isolates included in STs with <4 isolates; "NA" – Not applied.

A total of 216 isolates (29.8%) were related to 11 of the 43 PMEN clones (Colombia$_{23F}$-ST338, Denmark$_{12F}$-ST218, England$_{14}$-ST9, Netherlands$_{18C}$-ST113, Netherlands$_{3}$-ST180, Netherlands$_{7F}$-ST191, Netherlands$_{8}$-ST53, Portugal$_{19F}$-ST177, Sweden$_{15A}$-ST63, Sweden$_{4}$-ST205 and Spain$_{9V}$-ST156). Among these isolates, three strains have the same serotype and ST as the PMEN clones, but different antimicrobial resistant profiles. One of these cases was the PMEN clone Portugal$_{19F}$-ST177, where all the isolates identified (n=2) presented a susceptibility profile. Additionally, 9 SLVs and 19DLVs of PMEN clones were identified (Figures 17, 18 and 19). It was also observed that some STs were associated with serotypes that were not described in the *S. pneumoniae* MLST database (https://pubmlst.org/spneumoniae/), being the results resume in the Table below.

Table 8 – Difference between serotypes associated with STs found in the study and in the MLST database.

ST (n)	Serotype present in the collection (n)	Serotype present in the MLST database
53 (116)	3 (1)	8
63 (10)	11A (1)	15A, 19A, 23F, 19F, 14
113 (8)	8 (1)	18C, 24F, 7F, NT, 18B
193 (2)	15B/C (1)	19A, 21, 14, NT, 18C, 19F
235 (10)	22F (1)	20, 7C
445 (1)	22F (1)	33A, 33F
505 (4)	9N (1)	3, 15A, 6, 19A

"n"- number of isolates.

In the analysis of the evolution of ST between the study years (2015 to 2017), some changes were identified, although only significant before FDR correction. The STs for which there was a significant p-value before FDR correction in the Cochran-Armitage test for trend were: ST235 (serotype 20), ST1110 (serotype 8) and ST1262 (serotype 15B/C) (p=0.003, p=0.042 and p=0.036, respectively) showing an increasing trend, and ST408 (serotype 11A) and ST276 (serotype 19A) (p=0.009 and p=0.031, respectively) that decreased over the study period. Regarding changes in CCs over time, it was in agreement with the results obtained in STs, with a decline in CC230 significant before FDR correction (p=0.043) and an increase in CC235 significant before FDR correction (p=0.040) (results resume in Tables S9 and S10 present in the supplementary data).

PNSP and ERSP were more correlated with ST than with serotype. The AW for ST and PNSP and ERSP were 0.860 (CI95%: 0.761-0.960) and 0.787 (CI95%: 0.660-0.915), respectively, while for serotype were 0.667 (CI95%: 0.562-0.772) and 0.544 (CI95%: 0.421-0.667), respectively. Figure 20 shows the distribution in the CCs of the PNSP, ERSP and EPNSP isolates. When the association between PNSP and ERSP and STs was

analyzed, it was found that PNSP strains were associated with ST63, ST143, ST156, ST276 and ST386 (all $p<0.001$) significant after FDR correction (results resume in Table S11 of the supplementary data). Concerning ERSP, we found an association with ST9, ST63, ST143, ST276, ST386 and ST717 (all $p<0.001$) significant after FDR correction and (results resume in Table S11 of the supplementary data).

Although no association was observed between CCs and PNSP and ERSP after FDR correction, CC156, CC230 and CC315 presented an association with PNSP (all $p<0.001$) significant before FDR correction and CC156, CC230, CC315 and CC717 presented an association with ERSP (all $p<0.005$) significant before FDR correction.

Cochran-Armitage test was performed to evaluate ST variation between 2008-2011 (previous publish data [159]) and 2015-2017, where it was observed an increasing trend of ST53 (serotype 8) ($p<0.001$) and a decreasing trend of ST191 (serotype7F) and ST276 (serotype 19A) (both $p<0.001$) significant after FDR correction. ST392 (serotype 17F), ST505, ST2049 and ST1220 (all associated with serotype 3), ST1262 (serotype 15B/C), ST1635 (serotype 33F), ST8060 (serotype 12F), ST10047 (serotype 20), ST10220 and ST433 (both serotype 22F), ST113 (serotype 18C) and ST239 (serotype 9V) (all $p<0.05$) showed an increase trend significant before FDR correction, while ST156 (serotype 14) and ST179 (serotype 19F) ($p=0.011$ and $p=0.012$, respectively) decreased significantly before FDR correction (results resume in Table S12 of the supplementary data).

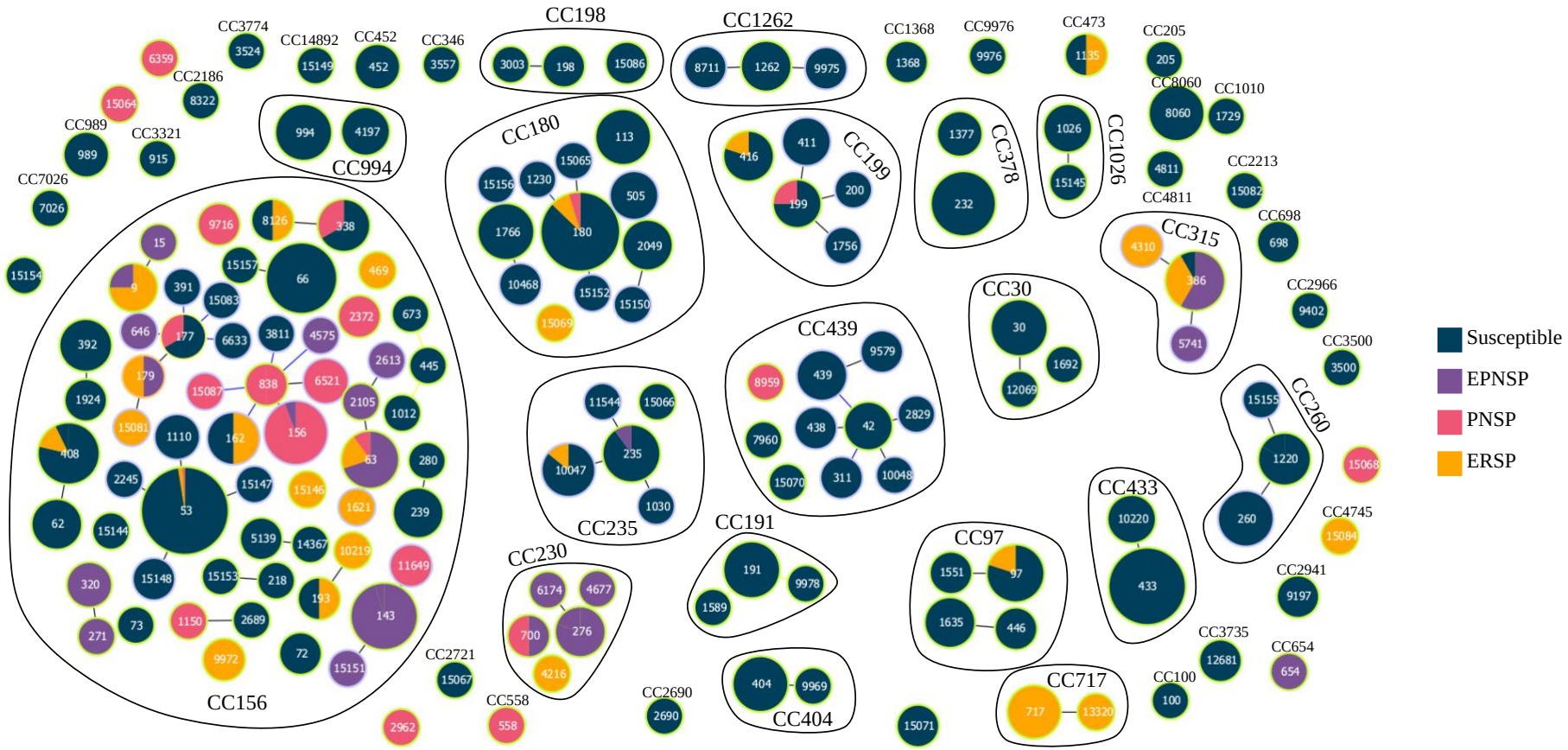

Figure 20 – Representation of the STs and indication of PNSP, ERSP and EPNSP isolates. STs linked by a line are SLVs and STs with no CC associated are singletons. The scheme was done using PHYLOViZ [144].

5. DISCUSSION

After the introduction of pneumococcal conjugate vaccines for children, a decrease in the number of cases of IPD in adults and IPD incidence in children has been observed, mainly due to the decrease in the number of infections caused by serotypes included in these vaccines [160,161]. At the same time, it has been observed, in some countries, an increased prevalence of non-PCV serotypes, which reinforces the importance of performing surveillance studies to determine vaccine impact on the epidemiology of IPD. Along with the increasing importance of non-PCV serotypes, it has also been observed that some VT (mainly serotypes associated with antibiotic resistance and serotype 3) persist as important causes of IPD both in children and in adults [147,161–164].

In Portugal, unlike other European countries, PCV13 was included in the NIP only in 2015. PCV7 was introduced in Portugal in 2001 and although available only through private market, the uptake of this vaccine increased gradually, reaching 75% of children with ≤2 years old in 2008 [79]. Similarly to the situation in the pediatric population, in adults, a reduction of the proportion of IPD caused by PCV7 serotypes after five years of this vaccine use was observed, while serotypes 3, 1, 7F and 19A emerged as significant causes of infection [92]. In mid-2009, PCV10 became available, followed by PCV13 in 2010, both used in the private sector. In adults, a study revealed a decrease in the proportion of IPD caused by the addPCV10 and addPCV13 serotypes shortly after PCV13 introduction (2012-2014), mainly due to the decrease in the proportion of serotypes 1, 5 and 19A, that was accompanied by an increase in the proportion of addPPV23 serotypes (mainly serotypes 8, 22F and 20) and NVT (mainly serotype 15A) [127]. However, PCV13 serotypes remained important, being responsible for 36.0% of the IPD cases between 2012-2014 [127]. PPV23 is available in Portugal since 1996, but the uptake is estimated to be low (~10%). However, with the introduction of PCV13 in the NIP, new guidelines recommend the vaccination of adults in particular risk groups (which include all adults with ≥65 years old) with PCV13 and PPV23 [90]. With the introduction of PCV13 in the NIP, a higher vaccinal coverage in children is expected, which may lead to changes in the *S. pneumoniae* population in children and consequently in adults. Having this in mind, this thesis intended to characterize the pneumococcus population responsible for adult IPD after PCV13 introduction in NIP.

In our study, most of the cases belonged to adults with ≥65 years old (58.9%) and more than half of the cases were from males (59.8%). Similar results were reported in

Italy and Spain [163,165]. The majority of the cases between 2015-2017 belonged to non-PCV serotypes (62.4%), of which, 67.9% were addPPV23 serotypes. However, PCV13 serotypes were still responsible for 37.6% of the cases, similarly to what was reported in other countries [161,164–166]. When comparing our data with the previous period (2012-2014) [127], a significant increase in the prevalence of NVT and addPPV23 serotypes and a decrease in the prevalence of addPCV10 serotypes were detected. Yet, no significant change was observed regarding the proportion of PCV7 and addPCV13 serotypes (Table S13).

The most frequent serotypes in the current study were serotypes: 8, 3, 22F, 14, 19A and 9N (accounting for 57.1% of the isolates). These serotypes were also among the most frequent serotypes observed in the previous study, but in a different rank order. This observation was similar to the most frequent serotypes found in other countries [161–164], except for serotype 14 that was only also present in Spain [163]. Serotype 7F decreased significantly during the study period and also when compared to the previous study. However, it still remains among the most prevalent serotypes in pediatric IPD (fourth most frequent serotype in 2012-2015) [167]. Among the most frequent serotypes, 7F was replaced by serotype 9N, a PPV23 serotype, and its proportion remained constant when compared to 2012-2014. Serotype 9N was not among the most prevalent serotypes in pediatric IPD [168] and was not found among the most frequent serotypes colonizing adults (study from 2010-2012) [169]. Nonetheless it was shown to have a high invasive disease potential [170]. Unlike the situation in Spain [163], in our work, this serotype belonged mainly to ST66, that remained constant when compared with the previous data published (2008-2011 – pre-PCV13 era), suggesting that serotype 9N is among the most frequent serotypes in 2015-2017 also because of the decrease in the proportion of serotype 7F. Nevertheless, it is essential to keep tracking of the evolution of this serotype in the post-PCV13 era, since a study report an increase in the incidence of serotype 9N three years after the introduction of PCV13 in the NIP in the United Kingdom [171].

Serotype 8, like serotype 9N, is a PPV23 serotype that was not among the most frequent serotypes in pediatric IPD in Portugal [167], nor is found among serotypes carried by adults [169], but it was shown to have a high invasive disease potential [170]. This serotype has been increasing in the post-PCV13 era (table S2), similar to what was reported in other countries [165,172,173]. It was the most frequent serotype in the current study and the second most frequent serotype in 2012-2014 [127]. It presented a significant association with young adults IPD and its increase in proportion when compared with

previous data (Table S2) may be due to the expansion of ST53 and ST1110 (both belonging to CC156) observed during the current study (Tables S9 and S12). Unlike what has been reported in other countries [163,174], in Portugal, most of the serotypes 8 isolates (85.2%) belonged to ST53 (Netherlands$_8$-ST53) and there were no ST63 among serotype 8 isolates, as it has been observed in Spain, where it was associated with levofloxacin resistance [162,163]. This serotype was not among the most frequent serotypes associated with non-susceptibility to any of the antimicrobials tested in our study. However, a slight increase in the proportion of non-susceptible serotype 8 isolates (primarily to penicillin and tetracycline) when compared to the previous study, was observed. This observation is of concern, considering that serotype 8 is not usually associated with antimicrobial non-susceptibility [161,165,166,175,176].

Serotype 22F is also a PPV23 serotype that was among the most frequent serotypes in 2015-2017 (third most frequent). Like serotype 8, this serotype was not among the most common serotypes in pediatric IPD [167], but was found among the most prevalent serotypes colonizing adults (≥60 years old) [169] and is becoming highly prevalent in carriage in children [177]. Associated with IPD in the elderly, the proportion of serotype 22F isolates remained constant in the post-PCV13 era (2012-2017), although an increase was observed in ST433 and ST10220 (both belonging to CC433) when compared with previous data (Table S12). This PPV23 serotype was not among the most frequent serotypes associated with non-susceptibility. In spite of this, like what was observed with serotype 8, an increase in antimicrobial non-susceptibility in serotype 22F isolates (especially to erythromycin) was detected, when compared both periods (Table S7). Similar to what was observed in other European countries [163,174,178], this serotype belonged mainly to ST433 (83.6%) in Portugal.

Serotype 3 is a PCV13 serotype usually associated with antimicrobial susceptibility, that was also among the most frequent serotypes in pediatric IPD in Portugal [167]. The proportion of this serotype was constant over the years (13.8% in 2012-2014 to 14.7% in 2015-2017), in agreement to what was observed in other countries [164,166,179], but contrarily to the situation, in Canada and Ireland, where an increase in the proportion of infections caused by serotype 3 isolates was reported [165,172]. Serotype 3 belonged mainly to ST180 (Netherlands$_3$-ST180), a lineage with a wide geographic dissemination [163,178,180]. Nevertheless, ST232 and ST260 were also among serotype 3 lineages in Portugal. A slight increase in the number of serotype 3 non-susceptible isolates (mainly to erythromycin, clindamycin and tetracycline, which may

indicate the presence of *Tn916* transposons) was observed when compared to the previous data (Table S7). Azarian et al. [181] observed that, although CC180 is the major CC associated with serotype 3, it can be divided in two different clades: clade I (that includes Iα and Iβ) and clade II, in which the last one emerged in the recent years (pre-PCV13 era). This study also showed that these two clades had distinct antigenic and antimicrobial resistance profiles, where clade II presented a higher level of antimicrobial resistance and a more competitive antigenic profile than clade Iα. In spite of the high genetic difference between serotype 3 isolates observed in this study (Figure 17), most of the strains (65.8%) belonged to the CC180, where an increase in antimicrobials non-susceptibility among these isolates was observed (Table S7), which may indicant that same of CC180 serotypes 3 in this study belong to clade II. However, although the emergence of clade II serotype 3 CC180 be of concern, mainly due to its association with high levels of antimicrobial resistance, this increase cannot explain the high proportion of serotype 3 isolates in post-PCV13 era, since both clades (clade Iα and clade II) need high antibody levels to occurred neutrophil-mediated opsonophagocytic killing by PCV13 [181]. A possible explanation for the low effect of PCV13 against serotype 3 could be the unusual polysaccharide synthesis pathway of serotype 3 isolates. CPS synthesis in *S. pneumoniae* has two different mechanisms: 1) *wzy*-dependent pathway, where the CPS is covalently linked to the peptidoglycans of the cell wall; and 2) synthase-dependent pathway, where the CPS is bound to the synthase or to phosphatidylglycerol on the cell wall, that can be ejected or dissociated. Although most serotypes produced their CPS by the *wzy*-dependent pathway, serotype 3 and 37 synthesize CPS by synthase-dependent pathway [23]. Choi et al [182] showed that CPS serotype 3 release interferes with antibody-mediated killing in opsonophagocytosis, either by absorbing free unbound antibody or by antibody bound to CPS that then are release. This reduces the protective efficacy of preexisting anti-CPS antibody, suggesting need for a higher antibody concentration to overcome this.

Serotype 19A was also a PCV13 serotype strongly associated with antimicrobial non-susceptibility (Tables 3, S3, S4, S5 and S6), similarly to what was observed in the USA [179] and other European countries [162,165,166]. This serotype was described as the most prevalent serotype colonizing adults [169] and a prevalent colonizer of non-vaccinated children in Portugal [177]. With the introduction of PCV13 in NIP, the number of non-vaccinated children was expected to decrease, which would contribute to a

reduction in children carrying this serotype, considering that vaccinated children carry fewer VTs [177]. This would lead to a reduction of IPD caused by this serotype in children and, consequently, due to the herd effect, a decrease in adult IPD. However, after three years of PCV13 used in the NIP in Portugal, there was no reduction observed in the proportion of serotype 19A in adult IPD. Nevertheless, a reduction of serotype 19A isolates associated with antimicrobials non-susceptibility when compared to previous data was observed (Table S7). This reduction was similar to what was reported in France, where serotype 19A decreased and remained constant among ERSP and PNSP isolates, respectively [176] and Germany, where serotype 19A was not even among the most prevalent serotypes in PNSP and ERSP isolates [175]. The reduction observed in this study may be associated with the decrease of ST276. In spite of the high genetic diversity of serotype 19A (CC156, CC230, CC199 and CC994), like what it was observed in Spain [163], ST276 (CC230) which was associated with non-susceptibility (Table S11), decreased during the current study (Table S9) and also when compared to the previous study (Table S12). Since genotypic data from 2012-2014 is not yet available, it is not possible to know if this observation was enhanced with the introduction of PCV13 in NIP or if the decrease in the proportion of ST276 was a consequence of the introduction of PCV13 in 2010. However, it was also observed a decrease in the proportion of serotype non-susceptible 19A isolates between 2012-2014 [127], which may indicate that this decrease in non-susceptible isolates was due to the introduction of PCV13 in 2010.

ST276 was also described previously as associated with serotype 24F in the pre-PCV13 era [159], which was of concern since this ST is an SLV of ST230 that was behind the expansion of serotype 19A as cause of IPD in Portugal in the post-PCV7 era [183]. However, in the current study, ST276 included only serotype 19A isolates. In fact, serotype 24F isolates in this study belonged mainly to ST162 (and also ST72 (2) and ST4677 (1)) (Tables 3, S3, S4, S5 and S6), which would explain why serotype 24F is not among the most frequent serotypes associated with antimicrobials non-susceptibility, as observed in France [176], Spain [162] and Canada [165].

Serotype 14 is the only PCV7 serotype among the most prevalent serotypes in adult IPD in Portugal. Strongly associated with antimicrobial non-susceptibility (Tables 3, S3, S4, S5, and S6), similarly to Spain and United Kingdom [166,184]. This serotype presented a significant decrease between 2012-2014 (mainly in 2014) [127], while this was not observed in the current study. One possible explanation for the continued high success of this serotype could be its association with clone England$_{14}$-ST9, associated

with antimicrobial non-susceptibility and the high antimicrobial consumption reported in our country as described before [127]. However, serotype 14 belonged mainly to ST156 and ST143 (both from CC156), and a decrease in ST156 when compared with previous data was detected. Serotype 14 is not currently among the most frequent serotypes colonizing both adults and children [169,177]. Contrarily to the situation reported here, a decrease in the prevalence of this PCV7 serotype, has been observed in other countries [165,175,176,179].

While serotype 12F was not among the most frequent serotypes in the current study, an increase in its proportion when compared to the previous study was observed. This increase can be associated with the increase of the genetic lineage ST8060 observed when compared with data from 2008-2011 (Table S12) or even with the increase of antimicrobial non-susceptible serotype 12F isolates that is observed when compared with data from 2012-2014 (Table S7). This observation is of concern since serotype 12F has been increasing in proportion, associated with antimicrobial non-susceptibility in the post-PCV13 era in France [176] and even reported among the most frequent serotypes in USA [161].

Serotypes 1 and 12B presented a significant decrease in their proportion when compared with data from 2012-2014 (Table S2). Regarding serotype 1, a PCV10 serotype, this trend was expected, but in the case of serotype 12B, an NVT, this decrease was not expected, reinforcing the importance of other factors besides vaccination that can cause changes in serotype prevalence as already described [79,127]. Serotype 6C, an NVT, also presented a significant decrease in proportion during the current study and this decrease was also observed in other countries, like Canada [165] and Norway [185], and it may be due to cross-protection of serotype 6A and 6B, present in PCV13 and PCV7, respectively. Serotype 6C, similar to what was reported in other countries [165,179,184], is strongly associated with antimicrobial non-susceptibility, being the third most frequent serotype among MDR isolates in the current study. Although not associated with any of the PMEN clones, serotype 6C belongs mainly to ST386, which is a genetic lineage associated with non-susceptibility in this study.

Overall, in this study, an increase in trimethoprim-sulfamethoxazole resistance and a decrease in cefotaxime resistance were detected, when compared with the previous study (Table S14). PCV7 and addPCV13 serotypes were responsible for a large proportion of PNSP and ERSP cases in this study (51.0% and 15.5% of PNSP and 37.6% and 19.0% of ERSP isolates, respectively), while resistance among addPPV23 and NVT

serotypes, although not significant, it has been decreasing in the post-PCV13 era (11.4% to 8.1% in the case of NVT and 8.2% to 6.1% in the case of addPPV23 serotypes between 2012-2014 and 2015-2017, respectively), unlike what has been observed in other countries [162,165,175,179]. Serotype 15A, like serotype 6C, is a NVT that was strongly associated with antimicrobial non-susceptibility, although its proportion has remained constant during the current study or when compared to the previous data. Serotype 15A was the fourth most frequent serotype among MDR isolates, like what has been observed in other countries [162,165,166,175,176,179]. This NVT belongs mainly to ST63 (Sweden$_{15A}$-ST63), associated with resistance to antimicrobials, which can lead to the expansion of this serotype, becoming one of the most frequent serotypes in adult IPD in Portugal, as already reported in France [164]. In Germany, serotypes 33A, 33F and 15B/C emerged as important causes of disease, associated with antimicrobial resistance which makes it important to maintain the surveillance of these serotypes [175]. Serotype 15B/C and serotype 33F are of particular importance since it was observed an increase in the proportion of STs associated with these serotypes (ST1262 and ST1635, respectively) during the current study and when compared to the previous data (Tables S9 and S12).

The major CCs present in this study were also among the major CCs of the previous study (2008-2011) [159], except for CC235 (serotype 20) and CC994 (serotype 19A). Serotype 20 was also associated with CC235 in 2008-2011, however this CC did not belong to the major CCs, and an increase in its proportion over the years was observed (Table S10). Serotype 19A, in the previous study, was associated with CC230, indicating expansion of distinct clonal lineages.

CC156 remains the most prevalent CC in the current study. This CC included mostly addPPV23 serotypes (8, 9N, 11A, and 17F, accounting for 61.8% of the CC156 isolates), which was not observed in 2008-2011 and children IPD in 2015-2017, where CC156 included mostly PCV7 serotypes [153 and unpublished data, respectively]. With the introduction of PCVs, CC156 did not lose its dominance as expected but remained as the most prevalent CC in both adults and children IPD. Nonetheless, in the current study, a shift was observed from PCV7 serotypes to addPPV23 serotypes. This shift was not observed in any other countries, where PCV7 remained the most prevalent serotypes express by this clonal lineage [162,176]. Two possible explanations can justify this observation: 1) there was an increase of genetic lineages associated with addPPV23 serotypes that already belong to CC156, due to the use of PCVs – selective pressure; or 2) the genetic lineages change, having emerged new STs associated with addPPV23

serotypes that belong to CC156. Since all genetic lineages associated with addPPV23 serotypes observed already belong to CC156 in 2008-2011 or belong to CCs that now belong to CC156 (since MLST database is constantly being updated), except for serotype 17F (belong mainly to CC113, that now is included in CC180 [159]), the first hypothesis seems to be the most likely explanations for the replacement of PCV7 serotypes by addPPV23 serotypes observed in this CC. However, serotype 8 in 2008-2011 belonged to the CC62, a lineage that in the present study was included in CC156. The regular update of the MLST database with new ST profiles can also be a possible explanation for this shift observed within the CC156, since serotype 8 represents a considerable proportion of adult IPD cases in both 2008-2011 and the current study. Yet, PCV7 serotypes are still mainly expressed in CC156 (79.2% of the PCV7 isolates), maybe due to the association of this CC with resistance to antimicrobials.

Among the major CCs, CC433, CC97, CC235, CC439, CC315, CC30 are of particular importance since they were associated with non-PCV serotypes (22F, 10A, 20, 23A, 6C and 16F, respectively) similar to what was reported in Spain [162], reinforcing the need to continue serotype surveillance.

PMEN clones recognized were not strongly represented in our collection, as already seen in the previous study [159], with 33.6% of the isolates being at the most DLVs of one of the 11 different PMEN clones identified [128]. These isolates identified belonged mainly to one of the 13 major CCs (essentially CC156, CC180 and CC191).

After the introduction of PCVs, the persistence of genetic lineages expressing a different serotype not covered by the PCVs was observed – a phenomenon described as capsular switching. Among our collection, possible capsular switching was observed (Table 5), reflecting mostly to occasional events, suggesting that these occurrences did not persist in the population. However, although rare, these events of capsular switching can be of importance since, in the future, they may proliferate if the conditions became favorable.

In this study, the serotype was also determined by *in silico* methods, where it was observed that both SeroBa and PneumoCAT were good approaches to determine *S. pneumoniae* serotypes since the results obtained were consistent with those obtains by traditional method. This analysis allowed also the detection of some errors in the annotation of serotypes determined by standard capsular reaction test and in the DNA extraction process, proving to be a relevant complementary analysis for this type of studies.

This study presented some limitations: 1) The information regarding the patients' health could be more detailed so it would be possible to detected more risk factors and their association with adult IPD; 2) The detection of genes associated with resistance to antimicrobials could be useful to this study since it would allow to conclude more about the increase in the proportion of non-susceptibility to trimethoprim-sulfamethoxazole over the years and the increase in the proportion of serotype 3, 8 and 22F non-susceptible to antimicrobials; 3) It would also be beneficial to study the presence of other genes coding for virulence factors, to better understand the invasive disease potential of same serotypes; 4) The characterization by cgMLST, instead of MLST, would provide a better insight into the genetic relationships of the isolates causing IPD and also to ascertain the differences between these two methods, since cgMLST allows a more detailed analysis; and 5) Although differences were observed regarding the allelic profiles when compared both periods (2008-2011 and 2015-2017), it is not possible to infer that the two periods are entirely different since this technique analyzed only part of the isolates in both periods (81.9% in 2008-2011 and 45.1% in 2015-2017).

6. CONCLUSION

This study intended to characterize phenotypically and genotypically the collection of pneumococcal strains responsible for adult IPD in Portugal during 2015-2017. A comparison was established between the obtained results and those of previous studies (2012-2014 for the antimicrobial non-susceptibility and serotyping results and 2008-2011 for the MLST analysis), so it would be possible to evaluate the impact of PCV13 introduction in NIP in adult IPD.

Three years after PCV13 introduction in NIP, changes were observed in the serotype distribution of *S. pneumoniae* population responsible for adult IPD when compared to previous data (pre-PCV13 introduction in NIP and post-PCV13 introduction in Portugal), where the most frequent serotypes were also present in 2012-2014. Regarding antimicrobials non-susceptibility, there were no major changes related to the main antimicrobials use in the treatment of these infections, although it was observed an increase in the proportion of trimethoprim-sulfamethoxazole resistant strains and a decrease in the proportion of cefotaxime resistant strains. PCV7 and addPCV13 serotypes still remained responsible for a significant portion of the PNSP and ERSP isolates in the current study, while addPPV23 and NVT PNSP and ERSP isolates decreased, which can be due to a decrease in genetic lineages associated with non-susceptibility. Nonetheless, an increase in the proportion of non-PCV serotypes was observed, whereas PCV7 and addPCV13 serotypes remained constant.

The MLST analysis was not always consistent with the analysis of serotype distribution and antimicrobial non-susceptibility profiles, possibly due to the gap of years when compared to the previous study (2012-2014 MLST information is not available, which corresponds to the period immediately before the current study and the introduction of PCV13 in Portugal). Anyhow, changes in serotype distribution were shown to be induced by the expansion of already circulating clones or due to declines of genetic lineages expressing a given serotype. In this study, capsular switching played a small role, yet these can be an important source of new variants that may emerge in the future.

In conclusion, we observed that in a situation of higher vaccination coverage, PCV13 serotypes are still an important cause of adult IPD in Portugal, especially serotype 3. However, non-PCV13 serotypes have emerge as important cause of adult IPD after the introduction of PCV13 in NIP, particularly serotypes 8 and 22F, reinforcing the continued need for surveillance studies.

7. FUTURE PERSPECTIVES

This study reinforced the need of routine surveillance studies in the future to continue to evaluate the effect of high coverage of PCV13 vaccination in children, in adult IPD, especially regarding serotypes 3, 14 and 19A that still represent important causes of adult IPD in Portugal. It is also important to see if changes in serotype distribution may lead to antimicrobial non-susceptibility alterations, particularly in serotypes 3, 8 and 22F where an increase in the proportion of non-susceptible strains was observed in the current study. Moreover, it is also crucial to follow the dynamics of *S. pneumoniae* population in order to monitor if the current CCs remain in circulation or new lineages emerge.

It would also be interesting to study the presence of other genes coding for virulence factors to better understand the genetic characteristics that potentially lead to a higher frequency of some serotypes in IPD when compared to the less frequent serotypes. A future characterization of the isolates by cgMLST will provide a better comprehension of the correlation between the isolates and the dynamics of the pneumococcal population causing adult IPD in Portugal.

8. REFERENCES

1. Nuorti JP, Whitney CG. Prevention of pneumococcal disease among infants and children - use of 13-valent pneumococcal conjugate vaccine and 23-valent pneumococcal polysaccharide vaccine - recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recomm reports Morb Mortal Wkly report Recomm reports. 2010 Dec;59(RR-11):1–18.
2. Murray PR, Rosenthal KS, Pfaller MA. Medical Microbiology. In: 8th ed. Philadelphia: Elsevier Inc.; 2016. p. 185–99.
3. Austrian R. Some observations on the pneumococcus and on the current status of pneumococcal disease and its prevention. Rev Infect Dis. 1981;3 Suppl:S1-17.
4. Weiser JN. The pneumococcus: why a commensal misbehaves. J Mol Med (Berl). 2010 Feb;88(2):97–102.
5. Jorgensen JH, Pfaller MA, Carroll KC, Funke G, Landry ML, Richter SS, et al., editors. Manual of Clinical Microbiology, Eleventh Edition. American Society of Microbiology; 2015.
6. Watson DA, Musher DM, Jacobson JW, Verhoef J. A brief history of the pneumococcus in biomedical research: a panoply of scientific discovery. Clin Infect Dis. 1993 Nov;17(5):913–24.
7. Deibel RH, Seeley Jr. HW. Family II. Streptococcacea. In: Buchanan RE, Gibbons N., editors. Bergey's Manual of Determinative Bacteriology. Eighth. Baltimore: The Willians & Wilkins Co; 1974. p. 490–515.
8. Lund E. Laboratory diagnosis of Pneumococcus infections. Bull World Health Organ. 1960;23(1):5–13.
9. Arbique JC, Poyart C, Trieu-Cuot P, Quesne G, Carvalho M da GS, Steigerwalt AG, et al. Accuracy of phenotypic and genotypic testing for identification of Streptococcus pneumoniae and description of Streptococcus pseudopneumoniae sp. nov. J Clin Microbiol. 2004 Oct;42(10):4686–96.
10. Richter SS, Heilmann KP, Dohrn CL, Riahi F, Beekmann SE, Doern G V. Accuracy of phenotypic methods for identification of Streptococcus pneumoniae isolates included in surveillance programs. J Clin Microbiol. 2008

Jul;46(7):2184–8.

11. Simoes AS, Tavares DA, Rolo D, Ardanuy C, Goossens H, Henriques-Normark B, et al. lytA-based identification methods can misidentify Streptococcus pneumoniae. Diagn Microbiol Infect Dis. 2016 Jun;85(2):141–8.

12. bioMérieux. API 20 Strep. Identification system for Streptococcacea and rekated organisms. 2010;Ref. 20600.

13. Marín M, Cercenado E, Sánchez-Carrillo C, Ruiz A, Gómez González Á, Rodríguez-Sánchez B, et al. Accurate Differentiation of Streptococcus pneumoniae from other Species within the Streptococcus mitis Group by Peak Analysis Using MALDI-TOF MS. Front Microbiol. 2017 Apr 25;8:698.

14. Carvalho M da GS, Tondella ML, McCaustland K, Weidlich L, McGee L, Mayer LW, et al. Evaluation and improvement of real-time PCR assays targeting lytA, ply, and psaA genes for detection of pneumococcal DNA. J Clin Microbiol. 2007 Aug;45(8):2460–6.

15. Roson B, Fernandez-Sabe N, Carratala J, Verdaguer R, Dorca J, Manresa F, et al. Contribution of a urinary antigen assay (Binax NOW) to the early diagnosis of pneumococcal pneumonia. Clin Infect Dis. 2004 Jan;38(2):222–6.

16. Kadioglu A, Weiser JN, Paton JC, Andrew PW. The role of Streptococcus pneumoniae virulence factors in host respiratory colonization and disease. Nat Rev Microbiol. 2008 Apr;6(4):288–301.

17. Mitchell AM, Mitchell TJ. Streptococcus pneumoniae: virulence factors and variation. Clin Microbiol Infect. 2010 May;16(5):411–8.

18. Brooks LRK, Mias GI. Streptococcus pneumoniae's Virulence and Host Immunity: Aging, Diagnostics, and Prevention. Front Immunol. 2018;9:1366.

19. Skov Sorensen UB, Blom J, Birch-Andersen A, Henrichsen J. Ultrastructural localization of capsules, cell wall polysaccharide, cell wall proteins, and F antigen in pneumococci. Infect Immun. 1988 Aug;56(8):1890–6.

20. Sorensen UB, Henrichsen J, Chen HC, Szu SC. Covalent linkage between the capsular polysaccharide and the cell wall peptidoglycan of Streptococcus pneumoniae revealed by immunochemical methods. Microb Pathog. 1990 May;8(5):325–34.

21. Nieto PA, Riquelme SA, Riedel CA, Kalergis AM, Bueno SM. Gene elements that regulate Streptococcus pneumoniae virulence and immunity evasion. Curr Gene Ther. 2013 Feb;13(1):51–64.

22. Bogaert D, De Groot R, Hermans PWM. Streptococcus pneumoniae colonisation: the key to pneumococcal disease. Lancet Infect Dis. 2004 Mar;4(3):144–54.

23. Standish AJ, Morona R. Capsule Structure, Synthesis, and Regulation. In: Brown JS, Hammerschmidt S, Orihuela CJ, editors. Streptococcus pneumoniae. Amesterdam: Academic Press; 2015. p. 169–79.

24. Hyams C, Camberlein E, Cohen JM, Bax K, Brown JS. The Streptococcus pneumoniae capsule inhibits complement activity and neutrophil phagocytosis by multiple mechanisms. Infect Immun. 2010 Feb;78(2):704–15.

25. Short KR, Diavotopoulos DA. Nasopharyngeal Colonization with Streptococcus pneumoniae. In: Brown J, Hammerschmidt S, Orihuela C, editors. Streptococcus pneumoniae. Amesterdam: Academic Press; 2015. p. 279–91.

26. Nelson AL, Roche AM, Gould JM, Chim K, Ratner AJ, Weiser JN. Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. Infect Immun. 2007 Jan;75(1):83–90.

27. Hostetter MK. Serotypic Variations among Virulent Pneumococci in Deposition and Degradation of Covalently Bound C3b: Implications for Phagocytosis and Antibody Production. J Infect Dis. 1986;153(4):682–93.

28. Zafar MA, Wang Y, Hamaguchi S, Weiser JN. Host-to-Host Transmission of Streptococcus pneumoniae Is Driven by Its Inflammatory Toxin, Pneumolysin. Cell Host Microbe. 2017 Jan;21(1):73–83.

29. Yuste J, Botto M, Paton JC, Holden DW, Brown JS. Additive inhibition of complement deposition by pneumolysin and PspA facilitates Streptococcus pneumoniae septicemia. J Immunol. 2005 Aug;175(3):1813–9.

30. Rai P, He F, Kwang J, Engelward BP, Chow VTK. Pneumococcal Pneumolysin Induces DNA Damage and Cell Cycle Arrest. Sci Rep. 2016 Mar;6:22972.

31. Gisch N, Peters K, Zahringer U, Vollmar W. The Pneumococcal Cell Wall. In: Brown J, Hammerschmidt S, Orihuela C, editors. Streptococcus pneumoniae. Amesterdam: Academic Press; 2015. p. 145–67.

32. Fischer W. Phosphocholine of pneumococcal teichoic acids: role in bacterial physiology and pneumococcal infection. Res Microbiol. 2000;151(6):421–7.

33. Bui NK, Eberhardt A, Vollmer D, Kern T, Bougault C, Tomasz A, et al. Isolation and analysis of cell wall components from Streptococcus pneumoniae. Anal Biochem. 2012 Feb;421(2):657–66.

34. AlonsoDeVelasco E, Verheul AF, Verhoef J, Snippe H. Streptococcus pneumoniae: virulence factors, pathogenesis, and vaccines. Microbiol Rev. 1995 Dec;59(4):591–603.

35. Jedrzejas MJ. Pneumococcal virulence factors: structure and function. Microbiol Mol Biol Rev. 2001 Jun;65(2):187-207 ; first page, table of contents.

36. Bergmann S, Hammerschmidt S. Versatility of pneumococcal surface proteins. Microbiology. 2006;152:295–303.

37. Tu AH, Fulgham RL, McCrory MA, Briles DE, Szalai AJ. Pneumococcal surface protein A inhibits complement activation by Streptococcus pneumoniae. Infect Immun. 1999 Sep;67(9):4720–4.

38. Senkovich O, Cook WJ, Mirza S, Hollingshead SK, Protasevich II, Briles DE, et al. Structure of a complex of human lactoferrin N-lobe with pneumococcal surface protein a provides insight into microbial defense mechanism. J Mol Biol. 2007 Jul;370(4):701–13.

39. Galán-Bartual S, Pérez-Dorado I, García P, Hermoso JA. Structure and Function of Choline Binding Proteins. In: Brown J, Hammerschmidt S, Orihuela C, editors. Streptococcus pneumoniae. Amesterdam: Academic Press; 2015. p. 207–30.

40. Dave S, Carmicle S, Hammerschmidt S, Pangburn MK, McDaniel LS. Dual roles of PspC, a surface protein of Streptococcus pneumoniae, in binding human secretory IgA and factor H. J Immunol. 2004 Jul;173(1):471–7.

41. Iannelli F, Oggioni MR, Pozzi G. Allelic variation in the highly polymorphic locus pspC of Streptococcus pneumoniae. Gene. 2002 Feb;284(1–2):63–71.

42. Bergmann S, Hammerschmidt S. Versatility of pneumococcal surface proteins. Microbiology. 2006 Feb;152(Pt 2):295–303.

43. Dintilhac A, Alloing G, Granadel C, Claverys JP. Competence and virulence of Streptococcus pneumoniae: Adc and PsaA mutants exhibit a requirement for Zn and Mn resulting from inactivation of putative ABC metal permeases. Mol Microbiol. 1997 Aug;25(4):727–39.

44. Whalan RH, Funnell SGP, Bowler LD, Hudson MJ, Robinson A, Dowson CG. Distribution and genetic diversity of the ABC transporter lipoproteins PiuA and PiaA within Streptococcus pneumoniae and related streptococci. J Bacteriol. 2006 Feb;188(3):1031–8.

45. Zomer A, Hermans PWM, Bootsma HJ. Non-Adhesive Surface Proteins of Streptococcus pneumoniae. In: Brown J, Hammerschmidt S, Orihuela C, editors. Streptococcus pneumoniae. Amesterdan: Academic Press; 2015. p. 231–44.

46. King SJ, Hippe KR, Gould JM, Bae D, Peterson S, Cline RT, et al. Phase variable desialylation of host proteins that bind to Streptococcus pneumoniae in vivo and protect the airway. Mol Microbiol. 2004 Oct;54(1):159–71.

47. Barocchi MA, Ries J, Zogaj X, Hemsley C, Albiger B, Kanth A, et al. A pneumococcal pilus influences virulence and host inflammatory responses. Proc Natl Acad Sci U S A. 2006 Feb;103(8):2857–62.

48. van der Poll T, Opal SM. Pathogenesis, treatment, and prevention of pneumococcal pneumonia. Lancet (London, England). 2009 Oct;374(9700):1543–56.

49. De Angelis G, Moschioni M, Muzzi A, Pezzicoli A, Censini S, Delany I, et al. The Streptococcus pneumoniae pilus-1 displays a biphasic expression pattern. PLoS One. 2011/06/22. 2011;6(6):e21269–e21269.

50. Bagnoli F, Moschioni M, Donati C, Dimitrovska V, Ferlenghi I, Facciotti C, et al. A second pilus type in Streptococcus pneumoniae is prevalent in emerging serotypes and mediates adhesion to host cells. J Bacteriol. 2008 Aug;190(15):5480–92.

51. Janoff EN, Rubins JB, Fasching C, Charboneau D, Rahkola JT, Plaut AG, et al. Pneumococcal IgA1 protease subverts specific protection by human IgA1. Mucosal Immunol. 2014 Mar;7(2):249–56.

52. Perez-Dorado I, Galan-Bartual S, Hermoso JA. Pneumococcal surface proteins:

when the whole is greater than the sum of its parts. Mol Oral Microbiol. 2012 Aug;27(4):221–45.

53. Holmes AR, McNab R, Millsap KW, Rohde M, Hammerschmidt S, Mawdsley JL, et al. The pavA gene of Streptococcus pneumoniae encodes a fibronectin-binding protein that is essential for virulence. Mol Microbiol. 2001 Sep;41(6):1395–408.

54. Noske N, Kammerer U, Rohde M, Hammerschmidt S. Pneumococcal interaction with human dendritic cells: phagocytosis, survival, and induced adaptive immune response are manipulated by PavA. J Immunol. 2009 Aug;183(3):1952–63.

55. Bergmann S, Wild D, Diekmann O, Frank R, Bracht D, Chhatwal GS, et al. Identification of a novel plasmin(ogen)-binding motif in surface displayed alpha-enolase of Streptococcus pneumoniae. Mol Microbiol. 2003 Jul;49(2):411–23.

56. Kolberg J, Aase A, Bergmann S, Herstad TK, Rodal G, Frank R, et al. Streptococcus pneumoniae enolase is important for plasminogen binding despite low abundance of enolase protein on the bacterial cell surface. Microbiology. 2006 May;152(Pt 5):1307–17.

57. Vazquez-Zamorano ZE, Gonzalez-Lopez MA, Romero-Espejel ME, Azuara-Liceaga EI, Lopez-Casamichana M, Olivares-Trejo J de J. Streptococcus pneumoniae secretes a glyceraldehyde-3-phosphate dehydrogenase, which binds haemoglobin and haem. Biometals. 2014 Aug;27(4):683–93.

58. Infante AJ, McCullers JA, Orihuela CJ. Mechanisms of Predisposition to Pneumonia: Infants, the Elderly and Viral Infections. In: Brown J, Hammerschmidt S, Orihuela C, editors. Streptococcus pneumoniae. Amesterdam: Academic Press; 2015. p. 363–82.

59. Marks LR, Reddinger RM, Hakansson AP. Biofilm formation enhances fomite survival of Streptococcus pneumoniae and Streptococcus pyogenes. Infect Immun. 2014 Mar;82(3):1141–6.

60. Dahl, Mygind. Anatomy, physiology and function of the nasal cavities in health and disease. Adv Drug Deliv Rev. 1998 Jan;29(1–2):3–12.

61. Weiser JN, Austrian R, Sreenivasan PK, Masure HR. Phase variation in

pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization. Infect Immun. 1994 Jun;62(6):2582–9.

62. Xu G, Kiefel MJ, Wilson JC, Andrew PW, Oggioni MR, Taylor GL. Three Streptococcus pneumoniae Sialidases: Three Different Products. J Am Chem Soc. 2011 Feb 16;133(6):1718–21.

63. Oliver MB, Swords WE. Pneumococcal biofilms and bacterial persistence during otitis media infections. In: Brown J, Hammerschmidt S, Orihuela C, editors. Streptococcus pneumoniae. Amesterdam: Academic Press; 2015. p. 293–308.

64. Woof JM, Russell MW. Structure and function relationships in IgA. Mucosal Immunol. 2011;4(6):590–7.

65. Aas JA, Paster BJ, Stokes LN, Olsen I, Dewhirst FE. Defining the normal bacterial flora of the oral cavity. J Clin Microbiol. 2005 Nov;43(11):5721–32.

66. Bogaardt C, van Tonder AJ, Brueggemann AB. Genomic analyses of pneumococci reveal a wide diversity of bacteriocins - including pneumocyclicin, a novel circular bacteriocin. BMC Genomics. 2015 Jul 28;16(1):554.

67. Pericone CD, Overweg K, Hermans PWM, Weiser JN. Inhibitory and Bactericidal Effects of Hydrogen Peroxide Production by Streptococcus pneumoniae on Other Inhabitants of the Upper Respiratory Tract. Infect Immun. 2000 Jul 1;68(7):3990 LP – 3997.

68. Pesakhov S, Benisty R, Sikron N, Cohen Z, Gomelsky P, Khozin-Goldberg I, et al. Effect of hydrogen peroxide production and the Fenton reaction on membrane composition of Streptococcus pneumoniae. Biochim Biophys Acta - Biomembr. 2007;1768(3):590–7.

69. Yesilkaya H, Andisi VF, Andrew PW, Bijlsma JJE. Streptococcus pneumoniae and reactive oxygen species: an unusual approach to living with radicals. Trends Microbiol. 2013;21(4):187–95.

70. Weiser JN, Ferreira DM, Paton JC. Streptococcus pneumoniae: transmission, colonization and invasion. Nat Rev Microbiol. 2018;16(6):355–67.

71. Mehr S, Wood N. Streptococcus pneumoniae-a review of carriage, infection, serotype replacement and vaccination. Paediatr Respir Rev. 2012 Dec;13(4):258–64.

72. Schoenmakers MCJ, Hament J-M, Fleer A, Aerts PC, van Dijk H, Kimpen JLL, et al. Risk factors for invasive pneumococcal disease. Rev Med Microbiol. 2002;13(1).

73. Bridy-Pappas AE, Margolis MB, Center KJ, Isaacman DJ. Streptococcus pneumoniae: description of the pathogen, disease epidemiology, treatment, and prevention. Pharmacotherapy. 2005 Sep;25(9):1193–212.

74. Mandell LA, Bartlett JG, Dowell SF, File TMJ, Musher DM, Whitney C. Update of practice guidelines for the management of community-acquired pneumonia in immunocompetent adults. Clin Infect Dis. 2003 Dec;37(11):1405–33.

75. Schuchat A, Robinson K, Wenger JD, Harrison LH, Farley M, Reingold AL, et al. Bacterial meningitis in the United States in 1995. Active Surveillance Team. N Engl J Med. 1997 Oct;337(14):970–6.

76. Klein JO. Microbiologic efficacy of antibacterial drugs for acute otitis media. Pediatr Infect Dis J. 1993 Dec;12(12):973–5.

77. Walker CLF, Rudan I, Liu L, Nair H, Theodoratou E, Bhutta ZA, et al. Global burden of childhood pneumonia and diarrhoea. Lancet (London, England). 2013 Apr;381(9875):1405–16.

78. European Center for Disease Prevention and Control (ECDC). Invasive pneumococcal disease. Inc: ECDC. Annual epidemiological report for 2017. Stockhalm: ECDC. 2019.

79. Aguiar SI, Brito MJ, Horacio AN, Lopes JP, Ramirez M, Melo-Cristino J. Decreasing incidence and changes in serotype distribution of invasive pneumococcal disease in persons aged under 18 years since introduction of 10-valent and 13-valent conjugate vaccines in Portugal, July 2008 to June 2012. Euro Surveill. 2014 Mar;19(12):20750.

80. Regev-Yochay G, Paran Y, Bishara J, Oren I, Chowers M, Tziba Y, et al. Early impact of PCV7/PCV13 sequential introduction to the national pediatric immunization plan, on adult invasive pneumococcal disease: A nationwide surveillance study. Vaccine. 2015 Feb;33(9):1135–42.

81. de Paz HD, Selva L, Muñoz-Almagro C. Pneumococcal Vaccination and Consequences. In: Brown J, Hammerschmidt S, Orihuela C, editors.

Streptococcus pneumoniae. Amesterdam: Academic Press; 2015. p. 41–57.

82. Kim JO, Romero-Steiner S, Sorensen UB, Blom J, Carvalho M, Barnard S, et al. Relationship between cell surface carbohydrates and intrastrain variation on opsonophagocytosis of Streptococcus pneumoniae. Infect Immun. 1999 May;67(5):2327–33.

83. Butler JC, Shapiro ED, Carlone GM. Pneumococcal vaccines: history, current status, and future directions. Am J Med. 1999 Jul;107(1A):69S-76S.

84. Butler JC, Breiman RF, Campbell JF, Lipman HB, Broome C V, Facklam RR. Pneumococcal Polysaccharide Vaccine Efficacy: An Evaluation of Current Recommendations. JAMA. 1993 Oct 20;270(15):1826–31.

85. de Roux A, Schmole-Thoma B, Siber GR, Hackell JG, Kuhnke A, Ahlers N, et al. Comparison of pneumococcal conjugate polysaccharide and free polysaccharide vaccines in elderly adults: conjugate vaccine elicits improved antibacterial immune responses and immunological memory. Clin Infect Dis. 2008 Apr;46(7):1015–23.

86. Food and Drugs Administration (FDA). Prevenar 13 [Internet]. 2018 [cited 2019 Oct 9]. Available from: https://www.fda.gov/media/107657/download

87. Paradiso PR. Pneumococcal conjugate vaccine for adults: a new paradigm. Clin Infect Dis. 2012 Jul;55(2):259–64.

88. Pilishvili T, Lexau C, Farley MM, Hadler J, Harrison LH, Bennett NM, et al. Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. J Infect Dis. 2010 Jan;201(1):32–41.

89. Aguiar SI, Serrano I, Pinto FR, Melo-Cristino J, Ramirez M. Changes in Streptococcus pneumoniae serotypes causing invasive disease with non-universal vaccination coverage of the seven-valent conjugate vaccine. Clin Microbiol Infect. 2008 Sep 1;14(9):835–43.

90. Direção-Geral da Saúde (DGS). Normas da Direcção-Geral da Saúde. Norma nº 011/2015 - Vacinação contra infeções por Streptococcus pneumoniae de grupos com risco acrescido para doença invasiva pneumocócica (DIP).Adultos (≥18 anos de idade). 2015.

91. Direcção-Geral da Saúde (DGS). Normas da Direcção-Geral da Saúde. Norma nº

008/2015 - Programa Nacional de Vacinação. Introdução da vacina conjugada de 13 valências contra infeções por Streptococcus pneumoniae (Pn13). 2015.

92. Horácio AN, Diamantino-Miranda J, Aguiar SI, Ramirez M, Melo-Cristino J. Serotype changes in adult invasive pneumococcal infections in Portugal did not reduce the high fraction of potentially vaccine preventable infections. Vaccine. 2012;30(2):218–24.

93. Weinberger DM, Malley R, Lipsitch M. Serotype replacement in disease after pneumococcal vaccination. Lancet (London, England). 2011 Dec;378(9807):1962–73.

94. Briles DE, Ades E, Paton JC, Sampson JS, Carlone GM, Huebner RC, et al. Intranasal immunization of mice with a mixture of the pneumococcal proteins PsaA and PspA is highly protective against nasopharyngeal carriage of Streptococcus pneumoniae. Infect Immun. 2000 Feb;68(2):796–800.

95. Ogunniyi AD, Folland RL, Briles DE, Hollingshead SK, Paton JC. Immunization of mice with combinations of pneumococcal virulence proteins elicits enhanced protection against challenge with Streptococcus pneumoniae. Infect Immun. 2000 May;68(5):3028–33.

96. Ogunniyi AD, Grabowicz M, Briles DE, Cook J, Paton JC. Development of a vaccine against invasive pneumococcal disease based on combinations of virulence proteins of Streptococcus pneumoniae. Infect Immun. 2007 Jan;75(1):350–7.

97. Lipsitch M. Bacterial vaccines and serotype replacement: lessons from Haemophilus influenzae and prospects for Streptococcus pneumoniae. Emerg Infect Dis. 1999;5(3):336–45.

98. Lipsitch M. Vaccination against colonizing bacteria with multiple serotypes. Proc Natl Acad Sci U S A. 1997 Jun 10;94(12):6571–6.

99. Rivera-Olivero IA, Blommaart M, Bogaert D, Hermans PWM, de Waard JH. Multiplex PCR reveals a high rate of nasopharyngeal pneumococcal 7-valent conjugate vaccine serotypes co-colonizing indigenous Warao children in Venezuela. J Med Microbiol. 2009;58(5):584–7.

100. Brugger SD, Frey P, Aebi S, Hinds J, Muhlemann K. Multiple colonization with

S. pneumoniae before and after introduction of the seven-valent conjugated pneumococcal polysaccharide vaccine. PLoS One. 2010 Jul;5(7):e11638.

101. Wyres KL, Lambertsen LM, Croucher NJ, McGee L, von Gottberg A, Linares J, et al. Pneumococcal capsular switching: a historical perspective. J Infect Dis. 2013 Feb;207(3):439–49.

102. Centers for Disease Control and Prevention (CDC). Antibiotic resistance threats in the United States, 2013. Centers For Disease Control and Prevention, Athanta, GA; 2014.

103. Edson DC, Glick T, Massey LD. Susceptibility testing practices for Streptococcus pneumoniae: results of a proficiency testing survey of clinical laboratories. Diagn Microbiol Infect Dis. 2006;55(3):225–30.

104. CLSI. Performance Standards for Antimicrobial Susceptibility Testing. 28th ed. CLSI supplemented M100. 2018.

105. Kim L, McGee L, Tomczyk S, Beall B. Biological and Epidemiological Features of Antibiotic-Resistant Streptococcus pneumoniae in Pre- and Post-Conjugate Vaccine Eras: a United States Perspective. Clin Microbiol Rev. 2016 Jul;29(3):525–52.

106. Weinstein MP, Klugman KP, Jones RN. Rationale for revised penicillin susceptibility breakpoints versus Streptococcus pneumoniae: coping with antimicrobial susceptibility in an era of resistance. Clin Infect Dis. 2009 Jun;48(11):1596–600.

107. McGee L, Pletz MW, Fobiwe JP, Klugman KP. Antibiotic Resistance of Pneumococci. In: Brown J, Hammerschmidt S, Orihuela C, editors. Streptococcus pneumoniae. Amesterdam: Academic Press; 2015. p. 21–40.

108. Hakenbeck R, Bruckner R, Denapaite D, Maurer P. Molecular mechanisms of beta-lactam resistance in Streptococcus pneumoniae. Future Microbiol. 2012 Mar;7(3):395–410.

109. Doi A, Iwata K, Takegawa H, Miki K, Sono Y, Nishioka H, et al. Community-acquired pneumonia caused by carbapenem-resistant Streptococcus pneumoniae: re-examining its prevention and treatment. Vol. 7, International journal of general medicine. New Zealand; 2014. p. 253–7.

110. Weisblum B. Erythromycin resistance by ribosome modification. Antimicrob Agents Chemother. 1995 Mar 1;39(3):577 LP – 585.

111. Schroeder MR, Stephens DS. Macrolide Resistance in Streptococcus pneumoniae. Front Cell Infect Microbiol. 2016;6:98.

112. Leclercq R, Courvalin P. Resistance to macrolides and related antibiotics in Streptococcus pneumoniae. Antimicrob Agents Chemother. 2002 Sep;46(9):2727–34.

113. Varaldo PE, Montanari MP, Giovanetti E. Genetic elements responsible for erythromycin resistance in streptococci. Antimicrob Agents Chemother. 2008/11/10. 2009 Feb;53(2):343–53.

114. Felmingham D, Canton R, Jenkins SG. Regional trends in beta-lactam, macrolide, fluoroquinolone and telithromycin resistance among Streptococcus pneumoniae isolates 2001-2004. J Infect. 2007 Aug;55(2):111–8.

115. Widdowson CA, Klugman KP, Hanslo D. Identification of the tetracycline resistance gene, tet(O), in Streptococcus pneumoniae. Antimicrob Agents Chemother. 1996 Dec;40(12):2891–3.

116. Lim S, Bast D, McGeer A, de Azavedo J, Low DE. Antimicrobial susceptibility breakpoints and first-step parC mutations in Streptococcus pneumoniae: redefining fluoroquinolone resistance. Emerg Infect Dis. 2003 Jul;9(7):833–7.

117. Pletz MW, Michaylov N, Schumacher U, van der Linden M, Duesberg CB, Fuehner T, et al. Antihypertensives suppress the emergence of fluoroquinolone-resistant mutants in pneumococci: an in vitro study. Int J Med Microbiol. 2013 May;303(4):176–81.

118. Matthew E, Levison MD. Chloramphenicol [Internet]. MERCK. 2009 [cited 2019 Oct 19]. Available from: https://web.archive.org/web/20100310105845/http://www.merck.com/mmpe/sec14/ch170/ch170d.html

119. Wormser GP, Keusch GT, Heel RC. Co-trimoxazole (trimethoprim-sulfamethoxazole): an updated review of its antibacterial activity and clinical efficacy. Drugs. 1982 Dec;24(6):459–518.

120. Walsh F, Willcock J, Amyes S. High-level telithromycin resistance in laboratory-

generated mutants of Streptococcus pneumoniae. J Antimicrob Chemother. 2003 Oct 1;52:345–53.

121. Wolter N, Smith AM, Farrell DJ, Schaffner W, Moore M, Whitney CG, et al. Novel Mechanism of Resistance to Oxazolidinones, Macrolides, and Chloramphenicol in Ribosomal Protein L4 of the Pneumococcus. Antimicrob Agents Chemother. 2005 Aug 1;49(8):3554 LP – 3557.

122. Cilloniz C, Ardanuy C, Vila J, Torres A. What is the clinical relevance of drug-resistant pneumococcus? Curr Opin Pulm Med. 2016 May;22(3):227–34.

123. Cornick JE, Bentley SD. Streptococcus pneumoniae: the evolution of antimicrobial resistance to beta-lactams, fluoroquinolones and macrolides. Microbes Infect. 2012 Jul;14(7–8):573–83.

124. Austrian R. Pneumococcus: the first one hundred years. Rev Infect Dis. 1981;3(2):183–9.

125. Jacobs MR, Koornhof HJ, Robins-Browne RM, Stevenson CM, Vermaak ZA, Freiman I, et al. Emergence of multiply resistant pneumococci. N Engl J Med. 1978 Oct;299(14):735–40.

126. European Center for Disease Prevention and Control (ECDC). Surveillance Atlas of Infectious Diseases: S. pneumoniae combined non-susceptibility to penicillins and macrolides (2017) [Internet]. [cited 2019 Oct 17]. Available from: https://atlas.ecdc.europa.eu/public/index.aspx?Dataset=27&HealthTopic=4

127. Horacio AN, Silva-Costa C, Lopes JP, Ramirez M, Melo-Cristino J. Serotype 3 Remains the Leading Cause of Invasive Pneumococcal Disease in Adults in Portugal (2012-2014) Despite Continued Reductions in Other 13-Valent Conjugate Vaccine Serotypes. Front Microbiol. 2016;7:1616.

128. McGee L, McDougal L, Zhou J, Spratt BG, Tenover FC, George R, et al. Nomenclature of major antimicrobial-resistant clones of Streptococcus pneumoniae defined by the pneumococcal molecular epidemiology network. J Clin Microbiol. 2001 Jul;39(7):2565–71.

129. Pneumococcal Molecular Epidemiology Network (PMEN). Pneumococcal Molecular Epidemiology Network. 2014.

130. Munoz R, Coffey TJ, Daniels M, Dowson CG, Laible G, Casal J, et al.

Intercontinental spread of a multiresistant clone of serotype 23F Streptococcus pneumoniae. J Infect Dis. 1991 Aug;164(2):302–6.

131. Donkor ES. Molecular typing of the pneumococcus and its application in epidemiology in sub-Saharan Africa. Front Cell Infect Microbiol. 2013 Mar 14;3:12.

132. Ramirez M, Carriço JA, van der Linden M, Melo-Cristino J. Molecular Epidemiology of Streptococcus pneumoniae. In: Brown JS, Hammerschmidt S, Orihuela C, editors. Streptococcus pneumoniae. Amesterdam: Academic Press; 2015. p. 3–19.

133. Baek JY, Ko KS, Kim SH, Kang C-I, Chung DR, Peck KR, et al. Comparison of methods for detection of Streptococcus pneumoniae serotype 6C. Diagn Microbiol Infect Dis. 2011;69(4):457–9.

134. Habib M, Porter BD, Satzke C. Capsular serotyping of Streptococcus pneumoniae using the Quellung reaction. J Vis Exp. 2014 Feb;(84):e51208.

135. Henrichsen J. The pneumococcal typing system and pneumococcal surveillance. J Infect. 1979;1:31–7.

136. Carrico JA, Pinto FR, Simas C, Nunes S, Sousa NG, Frazao N, et al. Assessment of band-based similarity coefficients for automatic type and subtype classification of microbial isolates analyzed by pulsed-field gel electrophoresis. J Clin Microbiol. 2005 Nov;43(11):5483–90.

137. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J Bacteriol. 2004 Mar;186(5):1518–30.

138. Enright MC, Spratt BG. A multilocus sequence typing scheme for Streptococcus pneumoniae: identification of clones associated with serious invasive disease. Microbiology. 1998 Nov;144 (Pt 1:3049–60.

139. Feil EJ, Spratt BG. Recombination and the population structures of bacterial pathogens. Annu Rev Microbiol. 2001;55:561–90.

140. Churko JM, Mantalas GL, Snyder MP, Wu JC. Overview of high throughput sequencing technologies to elucidate molecular pathways in cardiovascular diseases. Circ Res. 2013 Jun;112(12):1613–23.

141. Jolley KA, Maiden MC. Automated extraction of typing information for bacterial pathogens from whole genome sequence data: Neisseria meningitidis as an exemplar. Euro Surveill. 2013 Jan;18(4):20379.

142. Maiden MCJ, Jansen van Rensburg MJ, Bray JE, Earle SG, Ford SA, Jolley KA, et al. MLST revisited: the gene-by-gene approach to bacterial genomics. Vol. 11, Nature reviews. Microbiology. England; 2013. p. 728–36.

143. Medini D, Donati C, Tettelin H, Masignani V, Rappuoli R. The microbial pan-genome. Curr Opin Genet Dev. 2005 Dec;15(6):589–94.

144. Francisco AP, Vaz C, Monteiro PT, Melo-Cristino J, Ramirez M, Carrico JA. PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods. BMC Bioinformatics. 2012 May;13:87.

145. Epping L, van Tonder AJ, Gladstone RA, Consortium TGPS, Bentley SD, Page AJ, et al. SeroBA: rapid high-throughput serotyping of Streptococcus pneumoniae from whole genome sequence data. Microb Genomics. 2018;4(7).

146. Kapatai G, Sheppard CL, Al-Shahib A, Litt DJ, Underwood AP, Harrison TG, et al. Whole genome sequencing of Streptococcus pneumoniae: development, evaluation and verification of targets for serogroup and serotype prediction using an automated pipeline. PeerJ. 2016;4:e2477.

147. Silva-Costa C, Brito MJ, Pinho MD, Friaes A, Aguiar SI, Ramirez M, et al. Pediatric Complicated Pneumonia Caused by Streptococcus pneumoniae Serotype 3 in 13-Valent Pneumococcal Conjugate Vaccinees, Portugal, 2010-2015. Emerg Infect Dis. 2018 Jul;24(7):1307–14.

148. Sorensen UB. Typing of pneumococci by using 12 pooled antisera. J Clin Microbiol. 1993 Aug;31(8):2097–100.

149. CLSI. Performance Standars for Antimicrobial Susceptibility Testing; Seventeen Informational Supplement. CLSI document M100-S17. 2007.

150. The European on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 9.0. 2019.

151. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014 Aug;30(15):2114–20.

152. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012 May;19(5):455–77.

153. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One. 2014;9(11):e112963.

154. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019 Nov 28;20(1):257.

155. Carrico JA, Silva-Costa C, Melo-Cristino J, Pinto FR, de Lencastre H, Almeida JS, et al. Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant Streptococcus pyogenes. J Clin Microbiol. 2006 Jul;44(7):2524–32.

156. Pinto FR, Melo-Cristino J, Ramirez M. A confidence interval for the wallace coefficient of concordance and its application to microbial typing methods. PLoS One. 2008;3(11):e3696.

157. Severiano A, Pinto FR, Ramirez M, Carriço JA. Adjusted Wallace Coefficient as a Measure of Congruence between Typing Methods. J Clin Microbiol. 2011 Nov 1;49(11):3997 LP – 4000.

158. Benjamini Y, Hochberg Y. Controlling The False Discovery Rate - A Practical And Powerful Approach To Multiple Testing. J R Stat Soc, Ser B. 1995 Nov 30;57:289–300.

159. Horacio AN, Silva-Costa C, Diamantino-Miranda J, Lopes JP, Ramirez M, Melo-Cristino J. Population Structure of Streptococcus pneumoniae Causing Invasive Disease in Adults in Portugal before PCV13 Availability for Adults: 2008-2011. PLoS One. 2016;11(5):e0153602.

160. Horacio AN, Diamantino-Miranda J, Aguiar SI, Ramirez M, Melo-Cristino J. The majority of adult pneumococcal invasive infections in Portugal are still potentially vaccine preventable in spite of significant declines of serotypes 1 and 5. PLoS One. 2013;8(9):e73704.

161. Moore MR, Link-Gelles R, Schaffner W, Lynfield R, Lexau C, Bennett NM, et al. Effect of use of 13-valent pneumococcal conjugate vaccine in children on

invasive pneumococcal disease in children and adults in the USA: analysis of multisite, population-based surveillance. Lancet Infect Dis. 2015 Mar;15(3):301–9.

162. Fenoll A, Ardanuy C, Linares J, Cercenado E, Marco F, Fleites A, et al. Serotypes and genotypes of S. pneumoniae isolates from adult invasive disease in Spain: A 5-year prospective surveillance after pediatric PCV13 licensure. The ODIN study. Vaccine. 2018 Dec;36(52):7993–8000.

163. Del Amo E, Esteva C, Hernandez-Bou S, Galles C, Navarro M, Sauca G, et al. Serotypes and Clonal Diversity of Streptococcus pneumoniae Causing Invasive Disease in the Era of PCV13 in Catalonia, Spain. PLoS One. 2016;11(3):e0151125.

164. Hays C, Vermee Q, Agathine A, Dupuis A, Varon E, Poyart C, et al. Demonstration of the herd effect in adults after the implementation of pneumococcal vaccination with PCV13 in children. Eur J Clin Microbiol Infect Dis. 2017 May;36(5):831–8.

165. Demczuk WHB, Martin I, Desai S, Griffith A, Caron-Poulin L, Lefebvre B, et al. Serotype distribution of invasive Streptococcus pneumoniae in adults 65years of age and over after the introduction of childhood 13-valent pneumococcal conjugate vaccination programs in Canada, 2010-2016. Vaccine. 2018 Jul;36(31):4701–7.

166. Sheppard C, Fry NK, Mushtaq S, Woodford N, Reynolds R, Janes R, et al. Rise of multidrug-resistant non-vaccine serotype 15A Streptococcus pneumoniae in the United Kingdom, 2001 to 2014. Euro Surveill. 2016 Dec;21(50).

167. Silva-Costa C, Brito MJ, Aguiar SI, Lopes JP, Ramirez M, Melo-Cristino J. Dominance of vaccine serotypes in pediatric invasive pneumococcal infections in Portugal (2012-2015). Sci Rep. 2019 Jan;9(1):6.

168. Echaniz-Aviles G, Guerreiro SI, Silva-Costa C, Mendes CI, Carrico JA, Carnalla-Barajas MN, et al. Streptococcus pneumoniae Serotype 3 in Mexico (1994 to 2017): Decrease of the Unusual Clonal Complex 4909 Lineage following PCV13 Introduction. J Clin Microbiol. 2019 Jan;57(1).

169. Almeida ST, Nunes S, Santos Paulo AC, Valadares I, Martins S, Breia F, et al.

Low prevalence of pneumococcal carriage and high serotype and genotype diversity among adults over 60 years of age living in Portugal. PLoS One. 2014;9(3):e90974.

170. Sá-Leão R, Pinto F, Aguiar S, Nunes S, Carrico JA, Frazao N, et al. Analysis of invasiveness of pneumococcal serotypes and clones circulating in Portugal before widespread use of conjugate vaccines reveals heterogeneous behavior of clones expressing the same serotype. J Clin Microbiol. 2011 Apr;49(4):1369–75.

171. Houseman C, Hughes GJ, Chapman KE, Wilson D, Gorton R. Increased Invasive Pneumococcal Disease, North East England, UK. Emerg Infect Dis. 2017 Jan;23(1):122–6.

172. Corcoran M, Vickers I, Mereckiene J, Murchan S, Cotter S, Fitzgerald M, et al. The epidemiology of invasive pneumococcal disease in older adults in the post-PCV era. Has there been a herd effect? Epidemiol Infect. 2017 Aug;145(11):2390–9.

173. Waight PA, Andrews NJ, Ladhani SN, Sheppard CL, Slack MPE, Miller E. Effect of the 13-valent pneumococcal conjugate vaccine on invasive pneumococcal disease in England and Wales 4 years after its introduction: an observational cohort study. Lancet Infect Dis. 2015 May;15(5):535–43.

174. Duvvuri VR, Deng X, Teatero S, Memari N, Athey T, Fittipaldi N, et al. Population structure and drug resistance patterns of emerging non-PCV-13 Streptococcus pneumoniae serotypes 22F, 15A, and 8 isolated from adults in Ontario, Canada. Infect Genet Evol. 2016 Aug;42:1–8.

175. Imohl M, Reinert RR, van der Linden M. Antibiotic susceptibility rates of invasive pneumococci before and after the introduction of pneumococcal conjugate vaccination in Germany. Int J Med Microbiol. 2015 Oct;305(7):776–83.

176. Janoir C, Lepoutre A, Gutmann L, Varon E. Insight Into Resistance Phenotypes of Emergent Non 13-valent Pneumococcal Conjugate Vaccine Type Pneumococci Isolated From Invasive Disease After 13-valent Pneumococcal Conjugate Vaccine Implementation in France. Open forum Infect Dis. 2016 Jan;3(1):ofw020.

177. Valente C, Hinds J, Gould KA, Pinto FR, de Lencastre H, Sá-Leão R. Impact of the 13-valent pneumococcal conjugate vaccine on Streptococcus pneumoniae multiple serotype carriage. Vaccine. 2016 Jul;34(34):4072–8.

178. Galanis I, Lindstrand A, Darenberg J, Browall S, Nannapaneni P, Sjöström K, et al. Effects of PCV7 and PCV13 on invasive pneumococcal disease and carriage in Stockholm, Sweden. Eur Respir J. 2016 Apr 1;47(4):1208 LP – 1218.

179. Richter SS, Diekema DJ, Heilmann KP, Dohrn CL, Riahi F, Doern G V. Changes in pneumococcal serotypes and antimicrobial resistance after introduction of the 13-valent conjugate vaccine in the United States. Antimicrob Agents Chemother. 2014 Nov;58(11):6484–9.

180. Groves N, Sheppard CL, Litt D, Rose S, Silva A, Njoku N, et al. Evolution of Streptococcus pneumoniae Serotype 3 in England and Wales: A Major Vaccine Evader. Genes (Basel). 2019 Oct;10(11).

181. Azarian T, Mitchell PK, Georgieva M, Thompson CM, Ghouila A, Pollard AJ, et al. Global emergence and population dynamics of divergent serotype 3 CC180 pneumococci. PLOS Pathog. 2018 Nov 26;14(11):e1007438.

182. Choi EH, Zhang F, Lu Y-J, Malley R. Capsular Polysaccharide (CPS) Release by Serotype 3 Pneumococcal Strains Reduces the Protective Effect of Anti-Type 3 CPS Antibodies. Clin Vaccine Immunol. 2016 Feb;23(2):162–7.

183. Aguiar SI, Pinto FR, Nunes S, Serrano I, Melo-Cristino J, Sa-Leao R, et al. Denmark14-230 clone as an increasing cause of pneumococcal infection in Portugal within a background of diverse serotype 19A lineages. J Clin Microbiol. 2010 Jan;48(1):101–8.

184. Camara J, Marimon JM, Cercenado E, Larrosa N, Quesada MD, Fontanals D, et al. Decrease of invasive pneumococcal disease (IPD) in adults after introduction of pneumococcal 13-valent conjugate vaccine in Spain. PLoS One. 2017;12(4):e0175224.

185. Steens A, Bergsaker MAR, Aaberge IS, Ronning K, Vestrheim DF. Prompt effect of replacing the 7-valent pneumococcal conjugate vaccine with the 13-valent vaccine on the epidemiology of invasive pneumococcal disease in Norway. Vaccine. 2013 Dec;31(52):6232–8.

186. Graaf V De. Nervous Tissue and the Central Nervous System. In: Human Anatomy. The McGraw-Hill Companies; 2001. p. 343–99.

187. Gordis L. The dynamics of Disease Transmission. In: Epidemiology. Elsevier Health Science; 2013. p. 19–37.

188. Pankey GA, Sabath LD. Clinical relevance of bacteriostatic versus bactericidal mechanisms of action in the treatment of Gram-positive bacterial infections. Clin Infect Dis. 2004 Mar;38(6):864–70.

189. Fränkel A. Weitere Beiträge zur Lehre von den Mikrococcen der genuinen fibrinosen Pneumonic. In: Zeitschrift fur Klinische Medicin. 1886. p. 437–58.

190. Wright AE, Margan WP, Colebrook L, Dodgson RW. Observations on prophylactic inoculation against pneumococcus infections and on the results which have beeb achieved by it. In: Lancet 1. 1914. p. 1–10, 87–95.

191. Winslow CE, Broadhurst J, Buchanan RE, Krumwiede C, Rogers LA, Smith GH. The Families and Genera of the Bacteria: Final Report of the Committee of the Society of American Bacteriologists on Characterization and Classification of Bacterial Types. J Bacteriol. 1920 May;5(3):191–229.

192. Tan SY, Tatsumura Y. Alexander Fleming (1881-1955): Discoverer of penicillin. Singapore Med J. 2015 Jul;56(7):366–7.

193. Dawson MH. The transformation of pneumococcal types : II. The interconvertibility of type-specific s pneumococci. J Exp Med. 1930 Jan;51(1):123–47.

9. SUPPLEMENTARY DATA

Table S1 – Results obtained when examined the association between age groups and serotypes.

Age group	Serotype	OR (CI95%)	p-value	
			Before FDR	After FDR
[18,49]	1	13,23 (1.06-693.48)	0.022	0.308
	4	3.45 (1.34-8.60)	0.005	0.092
	7F	2.23 (1.01-5.15)	0.040	0.365
	8	2.22 (1.64-3.00)	<0.001	<0.001
[50,64]	8	1.36 (1.003-1.837)	0.042	0.451
≥65	31	5.21 (1.56-27.31)	0.003	0.045
	6C	3.48 (1.51-9.35)	0,001	0.022
	22F	2.18 (1.41-3.46)	<0.001	0.006
	19A	1.68 (1.07-2.70)	0.021	0.201

"OR"- Odds Ratio

Table S2 – Results obtained when examined the correlation of serotypes between 2012-2017.

Serotype	Study period		p-value	
	2012-2014	2015-2017	Before FDR	After FDR
1	26	4	<0.001	<0.001
8	123	287	<0.001	<0.001
12B	18	3	<0.001	<0.001
12F	0	17	<0.001	0.006
33A	9	4	0.017	0.165
33B	2	0	0.024	0.198
33F	3	14	0.010	0.107
7F	61	32	<0.001	<0.001

Table S3 – Results obtained when examined the association between penicillin and erythromycin resistant profiles and serotypes.

	Serotype	OR (CI95%)	p-value	
			Before FDR	After FDR
PNSP	14	79.94 (42.50-164.80)	<0.001	<0.001
	15A	7.62 (3.83-15.34)	<0.001	<0.001
	19A	2.40 (1.45-3.86)	<0.001	0.003
	6C	5.30 (2.66-10.44)	<0.001	<0.001
	19F	2.45 (1.12-5.03)	0.014	0.065
	23F	3.73 (1.21-10.66)	0.011	0.055
	9A	Inf (1.08-Inf)	0.022	0.096
ERSP	14	7.88 (5.17-12.05)	<0.001	<0.001
	15A	8.39 (4.20-17.17)	<0.001	<0.001
	19A	3.20 (2.01-5.02)	<0.001	<0.001
	19F	9,78 (4.91-20.18)	<0.001	<0.001
	33A	Inf (3.42-Inf)	<0.001	<0.003
	33F	19.60 (5.13-110.47)	<0.001	<0.001
	6B	21.15 (4.19-205.08)	<0.001	<0.001
	6C	20.67 (9.48-49.94)	<0.001	<0.001

"PNSP" – penicillin non-susceptibility; "ERSP" – erythromycin resistance; "OR" – Odds Ratio; "Inf" – infinite.

Table S4 – Results obtained when examined the association between age groups and penicillin and erythromycin resistance profile.

	Age group	OR (CI95%)	p-value	
			Before FDR	After FDR
PNSP	≥65	1.37 (1.02-1.85)	0.032	0.070
ERSP	≥65	1.38 (1.04-1.84)	0.023	0.070
EPNSP	≥65	1.59 (1.08-2.36)	0.015	0.023

"PNSP" – penicillin non-susceptibility; "ERSP" – erythromycin resistance; "EPNSP" – penicillin and erythromycin non-susceptibility; "OR" – Odds Ratio.

Table S5 – Results obtained when examined the association between age groups and resistance to clindamycin.

Antibiotic	Age group	OR (CI95%)	p-value	
			Before FDR	After FDR
DA	≥65	1.45 (1.07-1.99)	0.015	0.045

"DA"- clindamycin; "OR" – Odds Ratio

Table S6 – Results obtained when examined the association between serotype and resistance to clindamycin, tetracycline and trimethoprim-sulfamethoxazole and MDR.

Antibiotic	Serotype	OR (CI95%)	p-value	
			Before FDR	After FDR
DA	14	6.35 (4.14-9.73)	<0.001	<0.001
	15A	10.42 (5.21-21.40)	<0.001	<0.001
	19A	4.02 (2.52-6.34)	<0.001	<0.001
	19F	10.42 (5.21-21.40)	<0.001	<0.001
	33A	Inf (4.19-Inf)	<0.001	0.002
	33F	24.10 (6.30-135.68)	<0.001	<0.001
	6C	25.86 (11.84-62.59)	<0.001	<0.001
TET	15A	4.07 (1.93-8.26)	<0.001	0.001
	19A	5.20 (3.26-8.19)	<0.001	<0.001
	19F	6.34 (3.14-12.65)	<0.001	<0.001
	33A	Inf (4.78-Inf)	<0.001	0.002
	33F	9.89 (2.97-35.01)	<0.001	<0.001
	6C	26.07 (12.19-60.73)	<0.001	<0.001
SXT	14	8.69 (5.65-13.36)	<0.001	<0.001
	34	8.93 (1.90-45.39)	0.002	0.014
	15B/C	5.45 (1.43-7.80)	0.003	0.015
	18C	32,81 (6.72-313.76)	<0.001	<0.001
	19A	2.23 (1.30-3.71)	0.003	0.014
	24F	28.21 (8.86-118.26)	<0.001	<0.001
	9V	32.81 (6.72-313.76)	<0.001	<0.001
	19F	2.40 (1.03-5.15)	0.026	0.101
	35A	Inf (1.32-Inf)	0.016	0.066
	9A	Inf (1.32-Inf)	0.016	0.066
MDR	14	8.29 (5.54-12.70)	<0.001	<0.001
	15A	8.17 (4.10-16.46)	<0.001	<0.001
	19A	4.10 (2.58-6.45)	<0.001	<0.001
	19F	5.95 (2.97-11.84)	<0.001	<0.001
	33A	Inf (4.08-Inf)	<0.001	0.002
	33F	8.41 (2.53-29.68)	<0.001	0.001
	6C	21.77 (10.20-50,66)	<0.001	<0.001

"DA"- clindamycin; "TET"- tetracycline; "SXT"- Trimethoprim-sulfamethoxazole; "MDR"- Multi-Drug Resistance; "OR"- Odds Ratio; "Inf"- Infinite

Table S7 – Results obtained when examined the correlation between serotypes and resistance to some of the antibiotics analyzed during 2012-2017.

Antibiotic	Serotype	Study period		p-value	
		2012-2014	2015-2017	Before FDR	After FDR
PEN	8	0	3	0.040	0.303
	19A	38	28	<0.001	0.014
	22F	0	3	0.040	0.303
ERY	3	3	11	0.014	0.210
	19A	42	36	0.007	0.210
	22F	0	4	0.026	0.210
	33A	9	4	0.034	0.210
	33B	2	0	0.031	0.210
	33F	3	11	0.021	0.210
DA	3	3	10	0.031	0.158
	19A	42	36	0.005	0.140
	22F	0	3	0.036	0.158
	33A	8	4	0.035	0.158
	33B	2	0	0.031	0.128
	33F	3	11	0.021	0.158
TET	3	4	18	0.009	0.141
	8	1	10	0.031	0.222
	12B	3	0	0.016	0.141
	12F	0	5	0.014	0.141
	19A	40	38	0.004	0.141
	22F	0	3	0.049	0.291
C	12B	3	0	0.029	0.243
	12F	0	5	0.021	0.243
SXT	12B	7	1	0.016	0.233
	12F	0	4	0.013	0.233
	19A	35	23	<0.001	0.022
CT	14	4	2	0.044	0.222

"PEN"- penicillin; "ERY"- erythromycin; "DA"- clindamycin; "TET"- tetracycline; "C"- chloramphenicol; "SXT"- trimethoprim-sulfamethoxazole; "CT"- cefotaxime

Table S8 – Serotype distribution of the STs found in CCs with <10 isolates by goeBURST.

CC (n)	ST (n)	Serotype (n)
230 (9)	276 (5)	19A (5)
	700 (2)	3 (2)
	4677 (1)	24F (1)
	6174 (1)	19A (1)
404 (9)	404 (8)	8 (8)
	9969 (1)	8 (1)
1262 (9)	1262 (5)	15B/C (5)
	8711 (2)	15B/C (2)
	9975 (2)	15B/C (2)

Table S8 (continued)

CC (n)	ST (n)	Serotype (n)
717 (8)	717 (7)	33F (7)
	13320 (1)	33F (1)
8060 (8)	8060 (8)	12F (7), 12B (1)
1026 (5)	1026 (4)	20 (4)
	15145 (1)	20 (1)
198 (4)	198 (2)	29/35B (2)
	3003 (1)	29/35B (1)
	30001 (1)	29/35B (1)
452 (3)	452 (3)	29/35B (3)
989 (3)	989(3)	12F (3)
473 (2)	1135 (2)	6C (2)
698 (2)	698 (2)	22F (2)
1151 (2)	15086 (2)	12F (1), 19A (1)
1368 (2)	1368 (2)	35F (2)
2941 (2)	9197 (2)	4 (2)
3735 (2)	12681 (2)	10A (2)
100 (1)	100 (1)	33F (1)
205 (1)	205 (1)	4 (1)
346 (1)	3557 (1)	15B/C (1)
558 (1)	558 (1)	29/35B (1)
654 (1)	654 (1)	19F (1)
1010 (1)	1729 (1)	4 (1)
2186 (1)	8322(1)	15A (1)
2213 (1)	15082 (1)	4 (1)
2721 (1)	15067 (1)	16F (1)
2762 (1)	4216 (1)	8 (1)
2966 (1)	9402 (1)	16F (1)
3321 (1)	915 (1)	14 (1)
3500 (1)	3500 (1)	8 (1)
3774 (1)	3524 (1)	12F (1)
4745 (1)	15084 (1)	20 (1)
4811 (1)	4811 (1)	9N (1)
7026 (1)	7026 (1)	4 (1)
9976 (1)	9976 (1)	16F (1)
14892 (1)	15149 (1)	20 (1)

Table S9 – Results obtained when examined the correlation between STs and the study period (2015-2017).

ST (n)	Study year			p-values	
	2015	2016	2017	Before FDR	After FDR
235 (10)	0	2	8	0.003	0.470
276 (5)	4	1	0	0.031	0.470
408 (14)	9	4	1	0.009	0.470
1110 (3)	0	0	3	0.042	0.470
1262 (5)	0	1	4	0.036	0.470

Table S10 – Results obtained when examined the correlation between CCs and the study period (2015-2017).

CC (n)	Study year			p-values	
	2015	2016	2017	Before FDR	After FDR
230 (9)	4	4	0	0.043	0.479
235 (21)	4	5	12	0.040	0.479

Table S11 – STs that were associated with PNSP, ERSP and EPNSP.

Antimicrobial Resistance	ST (n)	OR (CI95%)	p-value	
			Before FDR	After FDR
PNSP	63 (8)	24.73 (4.84-241.41)	<0.001	<0.001
	143 (22)	Inf (38.56-Inf)	<0.001	<0.001
	156 (17)	Inf (27.56-Inf)	<0.001	<0.001
	276 (5)	Inf (5.44-Inf)	<0.001	0.002
	320 (3)	Inf (2.42-Inf)	0.003	0.060
	386 (7)	8.55 (2.29-34.88)	<0.001	0.013
	700 (2)	Inf (1.09-Inf)	0.022	0.246
	838 (2)	Inf (1.09-Inf)	0.022	0.246
	2372 (2)	Inf (1.09-Inf)	0.022	0.246
	6521 (3)	Inf (2.42-Inf)	0.003	0.060
	9716 (2)	Inf (1.09-Inf)	0.022	0.246
	11649 (2)	Inf (1.09-Inf)	0.022	0.246
ERSP	9 (4)	Inf (3.73-Inf)	<0.001	0.011
	63 (9)	53.82 (7.32-2349.80)	<0.001	<0.001
	143 (22)	150.44 (23.73-6036.54)	<0.001	<0.001
	179 (2)	Inf (1.05-Inf)	0.023	0.308
	276 (5)	Inf (5.22-Inf)	<0.001	0.002
	320 (3)	Inf (2.32-Inf)	0.003	0.067
	386 (11)	67.06 (9.56-2867.50)	<0.001	<0.001
	717 (7)	Inf (8.34-Inf)	<0.001	<0.001
	4310 (2)	Inf (1.05-Inf)	0.023	0.308
	9972 (2)	Inf (1.05-Inf)	0.023	0.308

Table S11 (continued)

Antimicrobial Resistance	ST (n)	OR (CI95%)	p-value	
			Before FDR	After FDR
EPNSP	63 (7)	28.24 (6.24-174.43)	<0.001	<0.001
	143 (22)	Inf (87.63-Inf)	<0.001	<0.001
	276 (5)	Inf (10.48-Inf)	<0.001	<0.001
	320 (3)	Inf (4.60-Inf)	<0.001	0.016
	386 (7)	16.95 (4.46-69.98)	<0.001	<0.001

"PNSP" – penicillin non-susceptibility; "ERSP" – erythromycin resistance; "EPNSP" – penicillin and erythromycin non-susceptibility; "OR" – Odds Ratio; "Inf" – infinite.

Table S12 – Results obtained when examined the correlation of STs over time (2008-2011 and 215-2017).

ST	Study period		p-value	
	2008-2011	2015-2017	Before FDR	After FDR
53	39	116	<0.001	<0.001
113	0	8	0.005	0.134
156	35	17	0.011	0.194
179	11	2	0.012	0.194
191	86	9	<0.001	<0.001
239	0	5	0.025	0.217
276	31	5	<0.001	<0.001
392	0	6	0.014	0.194
433	28	51	0.008	0.156
505	0	4	0.045	0.331
1220	0	6	0.0141	0.194
1262	0	5	0.025	0.217
1635	0	5	0.025	0.217
2049	0	5	0.025	0.217
8060	0	8	0.005	0.134
10047	1	7	0.033	0.279
10220	0	4	0.045	0.331

Table S13 – Results obtained when analyzed the trend of serotype class between 2012-2017.

Serotypes	Study period		p-value	
	2012-2014	2015-2017	Before FDR	After FDR
addPCV10	88	36	<0.001	<0.001
addPPV23	381	681	<0.001	<0.001
NVT	275	322	0.022	0.037

Table S14 – Results obtained when analyzed the trend of antimicrobials between 2012-2017.

Antimicrobial	Study period		p-value	
	2012-2014	2015-2017	Before FDR	After FDR
SXT	162	201	0.045	0.178
CT	13	2	<0.001	0.002

"SXT" – Trimethoprim-sulfamethoxazole and "CT" – Cefotaxime.