The impact of pneumococcal 7- and 13- valent conjugate vaccines on the molecular epidemiology of invasive *Streptococcus pneumoniae* in South Africa

Kedibone Maria Ndlangisa

A thesis submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Doctor of Philosophy.

Johannesburg, November 2018

Declaration

I, Kedibone Maria Ndlangisa, declare that this thesis is my own work. It is being submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other University.

Kedibone Ndlangisa

_____day of November 2018

Presentations arising from this study

1. Kedibone M. Ndlangisa, Mignon du Plessis, Linda de Gouveia, Mushal Allam, Nicole Wolter, Keith P. Klugman, Cheryl Cohen, Susan Mering, Rebecca A. Gladstone, Stephen D. Bentley, Lesley McGee, Robert F. Breiman, and Anne von Gottberg, for GERMS-SA. Non-vaccine *Streptococcus pneumoniae* serotypes causing invasive disease in children <5 years pre and post conjugate vaccine introduction in South Africa. 10th International Symposium on Pneumococci and Pneumococcal Diseases. 26-30 June 2016. Glasgow, United Kingdom.

2. Kedibone M. Ndlangisa, Mignon du Plessis, Mushal Allam, Linda de Gouveia, Keith P. Klugman and Anne von Gottberg for GERMS-SA (Group for Enteric, Respiratory and Meningeal Disease Surveillance). Molecular characterization of co-infecting, invasive *Streptococcus pneumoniae* in South Africa, 2005-2013. Federation of Infectious Diseases Societies of South Africa, 5-8 Nov 2015, Champagne Sports Resort, Drakensberg, South Africa.

3. Kedibone M. Ndlangisa, Mignon du Plessis, Nicole Wolter, Linda de Gouveia, Keith P. Klugman and Anne von Gottberg for GERMS-SA (Group for Enteric, Respiratory and Meningeal Disease Surveillance). Population snapshot of *Streptococcus pneumoniae* causing invasive disease in South Africa prior to introduction of pneumococcal conjugate vaccines. Wits Health Sciences Research Day and Postgraduate Expo, 17 Sept 2014.

4. Kedibone M. Mothibeli, Mignon du Plessis M., Phumza V.Tikili, Akhona Tshangela, Keith Klugman, Anne von Gottberg, and for GERMSSA. A rare genotype (ST2062) is a major cause of invasive pneumococcal serotype 19A disease among children in the pre-conjugate vaccine era, South Africa, 2007. 10th EUROPEAN MEETING ON THE MOLECULAR EPIDEMIOLOGY OF THE PNEUMOCOCCUS. June 2011. Amsterdam, Netherlands.

Publications arising from this study

1. Ndlangisa KM, du Plessis M, Wolter N, de Gouveia L, Klugman KP, von Gottberg A; GERMS-SA. Population snapshot of *Streptococcus pneumoniae* causing invasive disease in South Africa prior to introduction of pneumococcal conjugate vaccines. PLoS One. 2014 Sep 18;9(9):e107666.

2. Kedibone Ndlangisa, du Plessis M, Allam M, Wolter N, de Gouveia L, Klugman KP, Cohen C, Gladstone RA and von Gottberg A. Invasive disease caused simultaneously by dual serotypes of *Streptococcus pneumoniae*. J Clin Microbiol. 2017 Dec 26,56(1): pii:e01149-17

3. Ndlangisa KM, du Plessis M, Allam M, Wolter N, Mohale T, de Gouveia L, Birkhead M, Klugman KP, von Gottberg A. Two cases of serotypeable and non-serotypeable

variants of *Streptococcus pneumoniae* detected simultaneously during invasive disease. BMC Microbiol. 2016 ;16(1):126

Additional manuscript in preparation

1. Kedibone M. Ndlangisa, Mignon du Plessis, Linda de Gouveia, Mushal Allam, Nicole Wolter, Keith P. Klugman, Rebecca A. Gladstone, Stephen D. Bentley, Lesley McGee, Robert F. Breiman, and Anne von Gottberg, for GERMS-SA. Population snapshot of *Streptococcus pneumoniae* isolates causing invasive disease before and after introduction of pneumococcal conjugate vaccines in South Africa.

Abstract

Introduction

There are at least 98 known pneumococcal serotypes; however, the majority of disease is caused by approximately 20 serotypes. The pneumococcal conjugate vaccine (PCV) significantly decreased pneumococcal disease however following its introduction some countries reported increases in non-PCV serotypes. Cases of capsular switching have been reported however, in most cases increases were driven by pre-existing clones that were circulating prior to PCV introduction. The 7-valent pneumococcal conjugate vaccine (PCV7) was introduced in South Africa in 2009 and was replaced with the higher valency PCV13 in 2011. Pneumococcal disease is usually caused by a single serotype and disease by more than one serotype is rarely reported. Some pneumococci do not react with commercially available antisera used for serotyping and are thus regarded as non-serotypeable (NT). NT pneumococci are commonly isolated during carriage studies and very rarely cause invasive disease.

Materials and methods

Patient information and bacterial isolates were obtained through the national, laboratory-based surveillance for invasive pneumococcal disease (IPD) in South Africa, coordinated by the Group for Enteric, Respiratory and Meningeal Disease Surveillance in South Africa (GERMS-SA). Genotypes (sequence types) of isolates that caused IPD in 2005 through 2013 prior and during routine use of PCV7 and PCV13 were compared. Sequence type diversity by serotype was assessed using Simpson's index of diversity. For dual serotype IPD (co-infection with two serotypes) cases identified in 2005 through 2014, factors associated with dual serotype IPD

were assess and genomes of isolate pairs from co-infected individuals were sequenced. In two cases of IPD where a non-serotypeable and a serotypeable isolate were co-detected during routine serotyping, genomes of isolate pairs were compared.

Results

From 2005 through 2013, 36,267 IPD cases were reported to GERMS-SA of which 70% (25,508/36,267) had viable isolates available for further characterization. The average incidence of PCV serotypes declined 49% (95% confidence interval (CI), -44% to -54%) from 18.6/100,000 to 4.0/100,000 population during the pre-PCV (2005-2008) and PCV13 period (2011-3013), among infants and young children (<5 years old). For this age group, while no change occurred among all non-PCV serotypes combined, average 35B incidence increased 76% (95% CI, 26% to 94%) from 0.078/100,000 population in the pre-PCV period to 0.32/100,000 population in the PCV13 period. Among older children and adults (≥5 years old), declines occurred in average incidences of PCV serotypes from 7.2/100,000 to 2.1/100,000 population and in non-PCV serotypes from 3.0/100,000 to 1.5/100,000 population between the pre-PCV and PCV13 period. Similar genotypes circulated among isolates from infants and young children and among older children and adults. For most serotypes, sequence type diversity and predominant genotypes did not change between the pre-PCV and PCV13 period. While globally disseminated clones were common among some serotypes (e.g., serotype 1 [clonal complex (CC) 217] and 14 [CC230]), some were represented mainly by clonal complexes rarely reported elsewhere (e.g., serotype 3 [CC458, 60%] and 19A [CC2062, 83%]). Increase in non-PCV serotypes appeared to be due to expansion of pre-existing clones. For example, CC172, a genotype previously common among multiresistant serotypes 6A, 19A, 19F

and 23F, was represented by 13% (1/8) and 56% (9/16) of 35B isolates from infants and young children during the pre-PCV and PCV13 period, respectively.

IPD co-infection with two serotypes was identified in 33 IPD patients in 2005-2014. For 30 (91%) of the 33 patients, one or both isolates was a PCV13 serotype. Dual serotype IPD was associated with infants and young children <5 years of age [adjusted odds ratio (aOR), 4.7, 95% CI, 1.8-11.7], underlying illness (other than HIV) (aOR, 2.8; CI, 1.1-6.6) and death (aOR, 2.5; CI, 1.08-6.09). For each co-infecting pair, isolates were genotypically unrelated and their genotypes were common among isolates of the same serotype in South Africa. Of 701 accessory genes identified among dual serotype IPD isolates, four were common between isolate pairs.

A serotype 1 (isolate not available for further characterization) and 18C isolate were each co-detected with a non-serotypeable (NT) isolate in 2009 (case A) and 2010 (case B). Comparison of a non-serotypeable case A isolate with a serotype 1 genome (isolate of the same genotype as the NT case A isolate) revealed only the presence of four genes (*rmlA*, *B*, *C* and *D*) in the capsular locus, all other capsular genes were absent. Nonetheless it had a ST associated with serotype 1 (ST217 and ribosomal ST3462) and its core genome clustered with other ST217 isolates. The case B non-serotypeable isolate had all serotype 18C capsular genes except for variation in two of the capsular genes (*wchA* and *wze*), compared to the 18C isolate. Both case B isolates were ST9817 and their core genomes were identical.

Conclusions

In this study, genotypes circulating in South Africa before and during PCV use did not differ. Genotypes rarely reported in other parts of the world but common among some of our serotypes highlight the importance of these data. Genome comparison of 35B isolates belonging to CC172

and CC172 isolates expressing other capsular types may provide some information on the origin and evolution of the 35B/CC172 clone, which appears to have increased during the PCV era in South Africa. The association of dual serotypes with death warrants increased awareness of IPD co-infection caused by two or more serotypes. Although IPD due to non-serotypeable pneumococci is currently rare and there is no reported evidence of adaption from PCV serotypes to non-serotypeable pneumococci thus far, the ability of pneumococci to alter capsule production is a concern and therefore non-serotypeable pneumococci should be monitored.

Acknowledgements

I would like to express my deepest gratitude to:

- My supervisors: Keith Klugman, Anne von Gottberg and Mignon du Plessis for their support, encouragements, patience and many opportunities they provided during this journey
- CRDM bacteriology staff for their technical support
- GERMS-SA for coordinating the surveillance study and all patients for participating
- NICD/National Health Laboratory Service; PEPFAR (President's Emergency Plan for AIDS Relief) for funding the surveillance study
- Pfizer South Africa and Bill and Melinda Gates Foundation for funding molecular characterization of isolates

Table of contents

List of figures

List of tables

List of abbreviations

aOR	Adjusted odds ratio
BIGSdb	Bacterial Isolate Genome Sequence database
BRIG	BLAST Ring Image Generator
bp	base pair
CLSI	Clinical and Laboratory Standards Institute
CC	clonal complex
cps	capsular polysaccharide
CRDM	Centre for Respiratory Diseases and Meningitis
CSF	cerebrospinal fluid
D	Simpson's index of diversity
DLV	double-locus variants
DMP	Diagnostic Media Products
DNA	Deoxyribonucleic acid
HIV	human immunodeficiency virus
IPD	invasive pneumococcal disease
MDR	multidrug resistant
MIC	minimum inhibitory concentration
MLST	multilocus sequence typing
NHLS	National Health Laboratory Service
NICD	National Institute for Communicable Diseases
NT	Non-serotypeable
PCV	pneumococcal conjugate vaccine

PCV7	7-valent pneumococcal conjugate vaccine
PCV10	10-valent pneumococcal conjugate vaccine
PCV13	13-valent pneumococcal conjugate vaccine
PCR	polymerase chain reaction
PFGE	pulsed-field gel electrophoresis
PMEN	pneumococcal molecular epidemiology network
rMLST	ribosomal multilocus sequence typing
SLV	single-locus variant
SNV	single nucleotide variation
ST	sequence type
TEM	Transmission electron microscopy
UK	United Kingdom
USA	United States of America

1 Chapter One

Introduction

1.1 Microbiology of the pneumococcus

Streptococcus pneumoniae (the pneumococcus) is a Gram-positive, alpha-hemolytic diplococcus bacterium. In almost every clinical isolate, the pneumococcal cell wall is surrounded by a polysaccharide capsule which plays an important role in the pathogenesis of the pneumococcus as it provides protection against complement-mediated phagocytosis. It is also immunogenic and can be detected by specific antisera, and hence forms the basis for serotyping. Ninety-eight pneumococcal serotypes has been identified to date (1).

Some pneumococci do not react with commercially available sera used for serotyping and are thus regarded as non-serotypeable (NT). With the exception of serotype 37, genes responsible for production of the polysaccharide capsule are located in the capsular polysaccharide (*cps*) region in a single locus flanked by the *dexB* and *aliA* genes (2). The first four *cps* genes (*cpsA, cpsB, cpsC and cpsD* – also known as *wzg, wzh, wzd and wze*, respectively) are conserved among all serotypes and are involved in the regulation and processing of the capsule (2). Non-serotypability is due to partial or complete loss of the *cps* gene cluster (3), replacement of *cps* genes with other genes (3-5), sequence duplication (6) or single point mutations, commonly within the *wchA* (also known as *cpsE*) gene (7;8).

1.2 Epidemiology of the pneumococcus

The pneumococcus is commonly carried asymptomatically in the human nasopharynx. Occasionally it evades the immune system, crosses the mucous membranes and is an important cause of invasive disease such as meningitis and bacteremia and non-invasive disease such as acute otitis media (9). The burden of the disease is higher among young children and elderly adults (10;11). In South Africa, a country with a high prevalence of HIV infection, the burden of the disease is also higher among adults 25-45 years old (11;12). In 2012, the rate of invasive pneumococcal disease (IPD) among South African children <2 years of age and adults 25-45 years, was estimated to be 17/100,000 cases and 7.8/100,000 cases, respectively (11).

The majority of pneumococcal disease is caused by approximately 20 serotypes which are incorporated into different vaccine formulations (13-15). Disease is usually caused by a single serotype and only a few cases of disease caused by more than one serotype have been reported so far (16-18). The frequency of carriage and disease caused by individual serotypes varies by age, season and geographic area. For example, seven serotypes that are included in PCV7 caused approximately 80% of IPD in children less than 5 years of age living in the USA prior to the routine use of PCV7 (19). These serotypes are also associated with carriage and antimicrobial resistance (20).

1.3 Pneumococcal polysaccharide vaccines

Due to its immunologic effect, the polysaccharide capsule forms the basis for current vaccines against pneumococcal disease. The 23-valent polysaccharide vaccine was developed in the 1960's. It is composed of purified proteins of 23 pneumococcal serotypes (1, 2, 3, 4, 5, 6B, 7F,

8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19F, 19A, 20, 22F, 23F, 33F) that commonly cause disease. Due to the poor immunogenicity of purified polysaccharide in infants, this vaccine is used in adults and immuno-compromised children over the age of two years. Newer generations of vaccines are conjugated to a carrier protein which renders them the ability to induce T-cell-dependent immune response and are thus immunogenic in children younger than two years. Conjugated vaccines also reduce nasopharyngeal carriage of serotypes included in vaccines. Four formulations of pneumococcal conjugate vaccines (PCV) have been developed. PCV7 contains capsular polysaccharide from serotypes 4, 6B, 9V, 14, 18C, 19F and 23F that commonly caused IPD in young children prior to its introduction. Serotypes included in the remaining three vaccines are as follows: PCV10 (PCV7 plus 1, 5 and 7F), PCV13 (PCV10 plus 3, 6A and 19A) and PCV15 (PCV13 plus 22F and 33F). PCV7 was first introduced into a routine childhood immunization program by the USA in 2000 and was replaced by PCV13 in February 2010 (21). In South Africa, PCV7 was registered in 2005 but was initially only available in the private health-care sector (22). It was introduced into the childhood immunization program in April 2009 administered in a novel 2 + 1 schedule at 6, 14 and 40 weeks of age and was replaced by PCV13 in 2011 (22).

PCV significantly reduced vaccine-serotype IPD among vaccinated children in countries where it has been introduced (23-25). Reduction of disease in unvaccinated children and adults because of herd effect was also reported in those countries (24). Despite the benefits, increase in disease due to non-vaccine serotypes was reported in some countries where the vaccine was used routinely (23-27). Most countries reported serotype 19A as the dominant non-PCV serotype that increased post PCV7 introduction (28;29). The 19A increase was driven mainly by expansion

of the pre-existing susceptible clonal complex (CC) 199 (30;31) however, emergence of clones (e.g., CC172, CC81, CC236) associated with vaccine (PCV7) serotypes and antimicrobial resistance was reported (32;33). PCV7 replacement with PCV13, which include serotype 19A, resulted in reduction in 19A disease and more reports on other increasing non-PCV serotypes (12F, 35B, 15A, 22F, 24F) (28;33;34).While these increases were mainly due to expansion of pre-existing clones (e.g. CC63) (33;34) emergence of clones (e.g. CC156) previously associated with PCV serotypes were also reported (32;33).

Changes not associated with vaccine pressure have been reported in some countries. In Korea and Israel, serotype 19A increased prior to PCV introduction due to high antimicrobial use (35;36). Demark and Scotland, both of which had limited antibiotic and PCV use at the time of their studies, reported natural fluctuations in some pneumococcal serotypes indicating that changes may occur in the absence of any known pressures (37;38). An increase in the proportion of serotype 19A disease was reported in Denmark while an increase in serotype 1, associated with sequence type (ST) 306, was reported in Scotland (37;38).

1.4 Molecular characterization of the pneumococcus

Multilocus sequence typing (MLST) is a sequence-based method commonly used for analyzing genetic population structures of pneumococci and other infectious microorganisms (39). Isolates are characterized based on DNA sequences of seven housekeeping genes. With the exception of *ddl*, which is in close proximity to *pbp2b*, these genes are relatively resistant to mutations as they are not under any selective pressure and these characteristics make MLST appropriate for global and long-term epidemiology studies of isolates collected from different

time periods and geographic regions. Sequence types are assigned based on the combination of seven alleles derived from sequences of the seven genes. The eBURST algorithm is used for analysis of MLST data and groups related MLST genotypes (STs) together into clonal complexes and can also predict the founder (ancestor) of each group (40).

Capsular switching is a process whereby a recipient pneumococcal strain acquires capsular genes from a donor strain (with a different capsular type) resulting in the expression of the donor serotype by the recipient strain (41-43). Although not always the case, isolates with the same ST or CC are usually of the same serotype (40). An unusual serotype-ST/CC combination is therefore often an indication of capsular switching. Although genes can be acquired naturally, pressures such as high use of antibiotics and the use of serotype-specific vaccines such as PCV can drive pneumococci to switch genes. Using MLST, some studies have identified potential capsular switch events through the detection of unusual serotype/ST combinations (44). Through analysis of the capsular locus, Brueggemann and colleagues demonstrate that serotype 19A ST695 isolates, detected in the USA only after PCV introduction, resulted from capsular switch between serotype 19A and serotype 4 (41).

Whole genome sequencing (WGS) and the continuous development of bioinformatics software for analysis of such data enables an in-depth analysis of genomes. Since the entire genome is sequenced, more information can be extracted from WGS data than from MLST. These data include identification of resistance genes and changes within virulence genes such as the capsular loci, which could result from recombination events. The significance of WGS in characterizing pneumococci has been demonstrated by several studies including one by

Croucher and colleagues who analyzed 240 pneumococcal strains collected between 1984 and 2008 in 22 countries, including South Africa (45). All strains belonged to the Spain^{23F} ST81 multidrug-resistant clone and significant genetic diversity was found amongst them. A number of recombination events were detected between strains, including serotype switching. More recently Metcalf and colleagues explored the utility of WGS data in predicting pneumococcal antimicrobial resistance phenotypes (46).

1.5 Aim of this study

The main aim of this study was to determine the molecular epidemiology of *S. pneumoniae* causing invasive disease before and after PCV introduction in South Africa. In addition, isolate pairs collected from IPD patients co-infected with more than one pneumococcal serotype (dual-serotype IPD) were compared and characteristics of these patients and factors associated with dual-serotype IPD were defined. Genomes of isolate pairs from cases of IPD where a non-serotypeable and a serotypeable isolate were co-detected were also compared.

2 Chapter Two

Material and methods

2.1 Bacterial strains

Isolates and patient data were collected as part of the GERMS-SA (Group for Enteric, Respiratory and Meningeal Disease Surveillance) national, laboratory-based surveillance for IPD in South Africa, initiated in 1999 (47). In 2003, national surveillance was enhanced to include additional data such as patient outcome and HIV serological status from approximately 30 sentinel sites in all nine provinces. Surveillance reporting increased from 2003 onwards, and stabilized by 2005 (48). Isolates and patient data were submitted to the National Institute for Communicable Diseases (NICD) from approximately 200 participating laboratories across all nine provinces. A case of IPD was defined as the isolation of *S. pneumoniae* from a normally sterile-site specimen (e.g., blood, cerebrospinal fluid [CSF], pleural fluid, joint fluid). Isolates were transported on Dorset media (Diagnostic Media Products [DMP], Johannesburg, South Africa) and sub-cultured on arrival at the NICD, on 5% horse blood agar plates (DMP) in the presence of an optochin disc (Mast Diagnostics, Virginia, USA). Pure cultures were stored in 10% skim milk (DMP) at -70^{o}C. In addition, cases included patients with normally sterile-site specimens testing positive by PCR, or bacterial latex antigen supported by Gram stain microscopy (49). Laboratory-confirmed but unreported cases identified through audits of databases in participating laboratories were also included.

2.2 Serotyping and antimicrobial susceptibility testing

Serotypes were determined by the Quellung method using serotype-specific antisera (Statens Serum Institute, Copenhagen, Denmark) (50). Antimicrobial susceptibility testing was performed by agar dilution (for penicillin and ceftriaxone) or Etest (amoxicillin, erythromycin, clindamycin, chloramphenicol, tetracycline, rifampicin, cotrimoxazole, ofloxacin, linezolid and vancomycin) (AB Biodisk, Solna, Sweden) from 2005 through 2008 and by the broth microdilution method using commercially prepared Sensititre-SASP2 panels (Trek Diagnostics Inc., Cleveland, OH) from 2009 onwards. Results were interpreted according to the Clinical and Laboratory Standards Institute guidelines and breakpoints (51). Isolates were considered non-susceptible to penicillin at MICs ≥0.12mg/L using the oral penicillin meningitis breakpoints. For other antimicrobials, isolates were defined as non-susceptible if they were intermediately or fully resistant to the agent tested. Isolates were considered to be multidrug resistant if they were non-susceptible to beta-lactams and at least two other classes of antimicrobials.

2.3 DNA extraction [conventional multilocus sequence typing (MLST)]

For conventional MLST, DNA was extracted directly from 100µl of stored culture in 10% skim milk using a MagNA Pure LC 2.0 instrument with the MagNA Pure LC DNA Isolation kit III (Roche, Mannheim, Germany) according to the manufacturer's instructions. Extracted DNA was eluted into 100µl of elution buffer and stored at -20ºC. If poor quality sequencing results were obtained, DNA was re-extracted by boiling pure overnight cultures suspended in 100µl of sterile deionized water for 10 minutes at 95ºC.

2.4 Conventional MLST

Conventional MLST was performed as previously described (40), using primer pairs described by the Centers for Diseases Control and Prevention, Atlanta, USA (http://www.cdc.gov/ncidod/biotech/strep/alt-MLST-primers.htm) or primers listed on the MLST website (http://spneumoniae.mlst.net/). MLST genes were sequenced in one direction using the Big Dye Terminator v3.1 cycle sequencing kit (Applied Biosystems [ABI], Foster City, CA, USA), and products were analyzed using the ABI 3500 genetic analyzer. Sequences were aligned and analyzed using Lasergene software v9 (DNASTAR, Wisconsin, USA) and compared with sequences on the global MLST database (http://spneumoniae.mlst.net/) to assign allele numbers and ST. For new alleles, amplicons were sequenced in both directions and trace files were submitted to the MLST database curator. New profiles arising as a result of new alleles were also submitted to the curator for assignment of a sequence type.

2.5 DNA extraction [whole genome sequencing (WGS)]

Overnight fresh cultures were inoculated into brain heart infusion broth (DMP) and incubated overnight at 37°C in 5% CO_2. Cells were pre-lysed at 37°C for 1 hour in 10 mg/ml of lysozyme (Sigma-Aldrich, St. Louis, MO) and DNA was extracted using the QIAamp DNA mini kit (Qiagen, Venlo, Netherlands), following the manufacturer's instructions. DNA extracts were quantified using the Qubit instrument and dsDNA BR Assay kit (Life Technologies, Carlsbad, CA, USA).

2.6 Whole genome sequencing and MLST

Multiplexed paired-end libraries were prepared using the Nextera XT DNA sample preparation kit (Illumina, San Diego, CA, USA). Genome sequencing was carried out on an Illumina MiSeq platform at the NICD core sequencing facility or Sanger Institute in the UK. Paired-end reads were checked for quality, trimmed and *de novo* assembled using the CLC Genomics Workbench version 8 (CLC Inc., Aarhus, Denmark). The resultant contiguous sequences were ordered using Mauve (52) and *S. pneumoniae* ATCC 700669 as a reference [GenBank: FM211187], and annotated using Prokka version 1.11 (53).

Sequences of the 7 MLST genes were extracted from the assembled genomes and allele numbers and sequence types were assigned using scripts available at http://search.cpan.org/dist/Bio-MLST-Check/lib/Bio/MLST/Check.pm or https://github.com/sanger-pathogen/mlst_check.

2.7 MLST data analysis

The eBURST v3 algorithm available on the MLST website was used to determine relationships between sequence types (ST) and to generate population snapshots (54). The minimum number of identical loci was set at 0 for group definition. Single- (SLV) and double-locus (DLV) variants were defined as ST's differing from each other by one and two alleles, respectively. The default bootstrap (1 000) was used to assign the group founder ST. A clonal complex was defined as a group of ST's sharing at least 6 of 7 identical alleles with another ST in the group. Clonal complexes were named after the group founder ST as predicted by eBURST using data from the current study.

Sequence type diversity was assessed using Simpson's index of diversity (D). The value of D ranges from 0 to 1 and the greater the value the greater the sample diversity. Analysis was performed through an online tool for diversity and partition congruency calculations (http://darwin.phyloviz.net/ComparingPartitions/index.php?link=Tool).

2.8 Ethics consideration

The surveillance study and molecular characterization of isolates were approved by the Human Research Ethics Committee, University of the Witwatersrand, South Africa (protocol numbers: M081117 and M111008, respectively). Ethics for GERMS-SA surveillance is approved annually.

3 Chapter Three

Population snapshot of *Streptococcus pneumo*niae causing invasive disease in South Africa prior to introduction of pneumococcal conjugate vaccines

3.1 Introduction

The incidence of pneumococcal disease due to PCV7 serotypes decreased significantly among South African children less than 2 years of age in 2012 compared to the pre-PCV period (11). However, a non-significant increase in the incidence of disease caused by non-vaccine serotypes 8 and 15B was noted in this age group in 2012 compared to the pre-PCV period. Among South African adults aged 25 to 45 years, non-significant increases in some non-PCV serotype (7C, 9N, 10A, 12F, 22F and 34) was observed in 2012 compared to the pre-PCV period (11).

Natural fluctuations in circulating genotypes in the absence of any known interventions have been shown to occur (55) and factors other than vaccine pressure can result in proliferation of pneumococcal strains. This was demonstrated by an increase in antibiotic-resistant serotype 19A in Spain, Korea and Israel before routine use of PCV7 (56-58). Baseline data describing the genetic structure of pneumococci prior to the introduction of new interventions such as conjugate vaccines are therefore required to detect genotype changes potentially resulting from such interventions.

In this study multilocus sequence typing was used to investigate the clonal composition of invasive pneumococci prior to routine use of pneumococcal conjugate vaccines in South Africa.

3.2 Materials and Methods

3.2.1 Strain selection

Isolates were selected from IPD cases reported in 2007, a stable, pre-vaccine year. PCV serotypes (1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F and 23F) and 6C were grouped by patient age category (infants and young children <5 years [n=927] and older individuals ≥5 years [n=1,510]). Sampling was conducted as follows: for serotypes represented by ≤100 isolates per serotype, per age category, all isolates were included for characterization. Due to the large numbers of isolates available for certain serotypes (4, 6A, 6B, 14, 19F, 23F) sampling was rationalized by selecting 50% where numbers exceeded 100 (per serotype and per age category). Isolates were sorted by collection date and every second isolate was selected. In addition, all non-PCV serotype isolates that caused disease in 2007 among children <2 years of age were analyzed (n=134).

3.2.2 Multilocus sequence typing

DNA extraction and conventional MLST was performed as described in chapter 2. Relationships between sequence types and diversity by serotype and age category (infants and young children <5 years of age and individuals ≥5 years) were respectively assessed using the eBURST algorithm and Simpson's index of diversity as described in chapter 2.

3.3 Results

3.3.1 National surveillance

A total of 4,733 cases of IPD were reported from January through December 2007, of which, 70% (3,329/4,733) had viable isolates available for further characterization. The majority of cases with viable isolates were diagnosed from positive blood cultures (1,879/3,329, 56%), followed by CSF (1,005/3,329, 30%). In addition, isolates were recovered from both blood and CSF (247/3,329, 7%), pleural fluid (144/3,329, 4%) or other normally sterile-site specimens (54/3,329, 2%).

Among cases with known gender (97% [3,241/3,329]), females and males were equally represented (1,638/3,241, 51% and 1,603/3,241, 49%, respectively). Age was known for 96% (3,194/3,329) of cases with viable isolates, and of these, 34% (1,084/3,194) were infants and young children <5 years, 10% (333/3,194) were older children 5-17 years and 56% (1,777/3,194) were adults ≥18 years. Overall, PCV serotypes were responsible for 86% (927/1,084), 82% (273/333) and 69% (1,235/1,777) of IPD in infants and young children, older children 5-17 years and adults ≥18 years, respectively. Serotypes 14, 6B and 6A were the most prevalent serotypes among infants and young children accounting for 41% (446/1,084), while 1, 14 and 6B were the most common serotypes among older children 5–17 years (135/333, 41%) and 1, 19A and 4 were prevalent among individuals ≥18 years (505/1,777, 28%) (Figure 3.1). Among infants and young children <2 years, serotype 15B/C and 8 were the most common serotypes among non-PCV serotypes, and represented 16% (22/134) and 15% (20/134) of the isolates, respectively. The remaining isolates (69%, 92/134) belonged to 32 serotypes. Non-susceptibility to penicillin was detected in 43% (1,429/3,329), of which 39% (562/1,429) were multidrug resistant. HIV status was known for 32%

(1,064/3,329) of cases with viable isolates, the majority of whom were HIV infected (816/1,064, 77%). Overall, PCV serotypes were responsible for 77% (633/816) and 81% (186/231) of IPD in HIV infected and uninfected individuals, respectively (p=0.002).

3.3.2 Multilocus sequence typing

A total of 1,461 isolates representing PCV serotypes and 6C (1,461/2 437, 60%) were selected for MLST, of which complete results were available for 73% (1,064/1 461). Poor sequence data or PCR amplification failure for some of the genes resulted in incomplete MLST results for some isolates, 43% (169/397) of which were from infants and young children <5 years. MLST results were available for 33% (269/816) of isolates from HIV infected and 37% (85/231) of isolates from HIV uninfected individuals. All 134 isolates representing non-PCV serotypes, from children <2 years of age were selected for characterization and complete results were obtained for 55% (74/134) of these isolates.

3.3.3 MLST genotypes

MLST data are summarized in Table 3.1. For PCV serotypes, similar clonal complexes were found among isolates from infants and young children and among isolates from older individuals and there were no differences in sequence type diversities between the two age groups (data not shown). Simpson's diversity indices were therefore calculated for each serotype using combined sequence type data from isolates of children and adults (Figure 3.2). The majority of serotypes were heterogeneous, however, serotypes 1 and 5 were the least diverse (D=0.38 and 0.55, respectively) while 6A, 6B, 18C and 23F were the most diverse serotypes, with D ranging from

0.97 for 18C and 23F to 0.98 for 6A and 6B. The majority of serotype 1 isolates were CC217 (143/146, 98%). CC458, which was the largest clonal complex among serotype 3 isolates, represented 60% (56/93) of the serotype 3 isolates. The main clonal complex among serotype 4 isolates was CC32 (31/81, 38%). All serotype 5 isolates were CC289. For serotype 6A, three common clonal complexes were identified, of which the most prevalent, CC473, represented 39% (37/95) of the isolates. Several complexes were identified among serotype 6B isolates, of which CC2421 (26/113, 23%) and CC156 (26/113, 23%) were the most common. CC2185 represented majority of serotype 6C isolates (9/11, 82%). CC218 represented 72% (21/29) of serotype 7F isolates. The major clonal complex (CC4881) among serotype 9V isolates represented 57% (40/70) of the isolates. Of the clonal complexes identified among serotype 14 isolates, CC230 was the largest and represented 43% (49/115) of the isolates. Serotype 18C isolates were differentiated into 4 main clonal complexes. The largest clonal complex among serotype 18C isolates (CC1016) represented 30% (17/57) of the isolates. The majority of serotype 19A isolates were CC2062 (109/131, 83%). CC320 was the most prevalent clonal complex among serotype 19F isolates (23/46, 50%). For serotype 23F, CC242 (25/59, 42%) and CC156 (12/59, 20%) were the major clonal complexes. For non-PCV serotypes, ST7052 which belongs to CC346, was predominant (7/12, 58%) among serotype 15B/C isolates, while ST53 was prevalent (9/14, 64%) among serotype 8 isolates (Table 3.2).

3.3.4 Antimicrobial susceptibility

For those PCV isolates with complete MLST profiles, 640 belonged to clonal complexes/singletons representing at least one penicillin non-susceptible isolate (Table 3.3). Reduced susceptibility to penicillin was detected in 481 of those 640 isolates, of which 192 were

multidrug resistant. Serotype 14, 19A and 19F had the largest proportion of isolates that were non-susceptible to penicillin (95% [109/115], 91% [119/131] and 89% [41/46], respectively) (Table 3.1 and 3.3). Non-susceptible isolates were common among serotype 14 and 19F isolates belonging to multiple clonal complexes. Except for one serotype 6B isolate, all other isolates belonging to CC230 (n=49 for serotype 14 and 1 each for 19F and 23F) were multidrug resistant. All isolates belonging to CC320 were also multidrug resistant (n=23 and 1 isolate for serotype 19F and 23F, respectively).

Of the 134 non-PCV serotype isolates recovered from children <2 years, 16% (21/134) were non-susceptible to penicillin and, of these, 19% (4/21) were multidrug resistant. The majority (14/21, 67%) of penicillin non-susceptible non-PCV isolates belonged to serotype 15B/C and were also non-susceptible to trimethoprim/sulphamethoxazole (13/14, 93%) but none were multidrug resistant. Of the 74 isolates with MLST results, 9 were non-susceptible to penicillin, 8 of which belonged to serotype 15B/C (Table 3.2). A non-serotypeable ST344 isolate, extremely rare in IPD, was multidrug resistant.

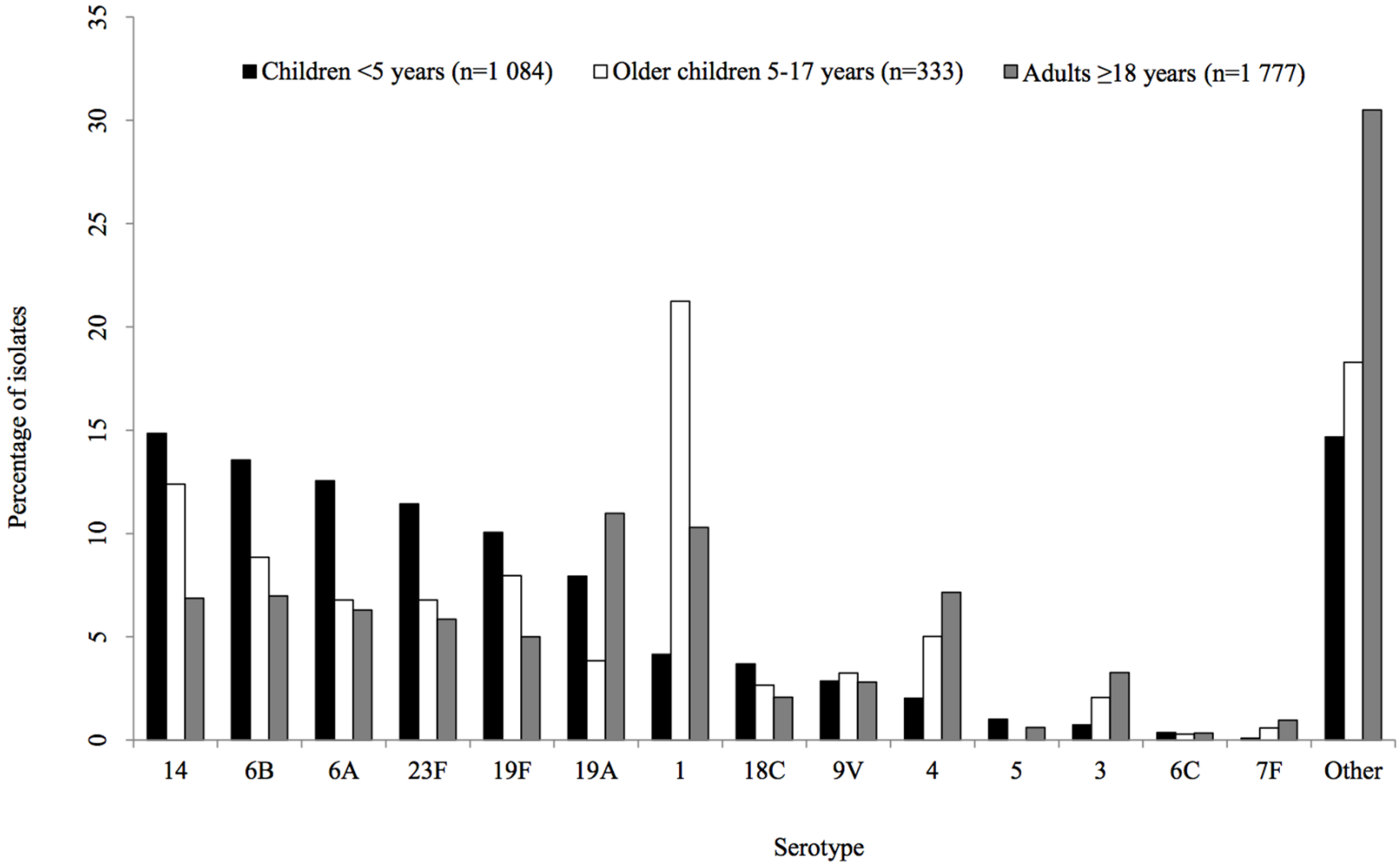


Figure 3.1 Distribution of pneumococcal serotypes causing invasive disease in South Africa, by age group, 2007 (n = 3,194).

'Other' indicates serotypes not included in pneumococcal conjugate vaccines (PCV13 and 6C).

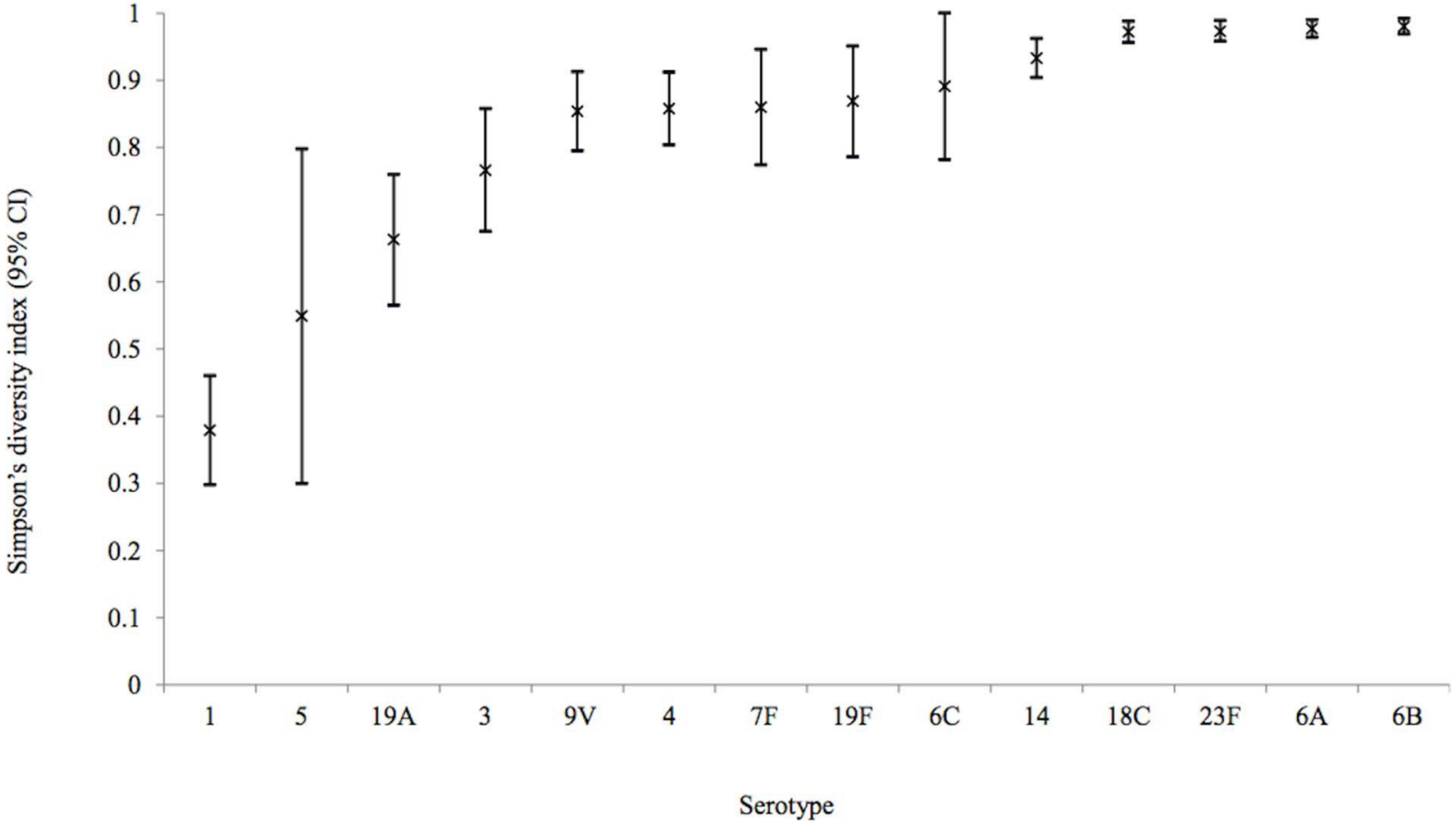


Figure 3.2 Diversity of sequence types within serotypes.

Simpson's index of diversity was used to assess diversity of isolates that caused invasive pneumococcal disease in 2007 among individuals of all ages, South Africa.

Table 3.1 Clonal complex and sequence type distribution of pneumococcal conjugate vaccine serotype isolates causing invasive pneumococcal disease in South Africa, 2007 (n=1,064)

Serotype (No. of isolates)	Clonal complex (No. of isolates)	Sequence type (No. of isolates)
1 (146)	217 (143)	217 (111), 612 (31), 3577 (1)
	other (3)	611 (3)
3 (93)	458 (56)	458 (44), 6366 (1), 7348 (1), 7349 (1), 7350 (6), 7353 (3)
	180 (14)	180 (6), 3842 (1), 4118 (2), 7356 (1), 7357 (4)
	378 (11)	378 (1), 7362 (1), 7363 (1), 7364 (1), 7365 (1), 7367 (1), 7368 (1), 7369 (3), 7366 (1)
	Other (12)	700 (1), 1765 (1), 2837 (5), 7347 (1), 7351 (1), 7352 (1), 7378 (1), 7997 (1)
4 (81)	32 (31)	5410 (16), 5655 (4), 6367 (1), 6368 (1), 7370 (4), 7372 (1), 7373 (1), 7374 (1), 7375 (1), 7376 (1)
	Other (50)	63 (1), 205 (3), 271 (1), 1221 (25), 1222 (6), 2213 (1), 2290 (5), 6369 (1), 7354 (1), 7355 (2), 7358 (1), 7359 (1), 7371 (1), 7727 (1)
5 (18)	289 (18)	289 (3), 5659 (12), 6370 (1), 7377 (1), 7745 (1)

6A (95)	473 (37)	1096 (1), 2285 (10), 2295 (2), 4087 (3), 5073 (4), 6372 (2), 6373 (1), 6374 (1), 6380 (1), (1), 6381 (1), 6382 (1), 6385 (1), 6388 (1), 6390 (1), 7298 (1), 7300 (1), 7301 (1), 7310 (1), 7735 7736 (1), 7744 (1)
	156 (23)	138 (2), 185 (5), 1447 (1), 2987 (1), 3207 (1), 6375 (1), 6376 (1), 6379 (2), 6389 (3), 7304 (1), 7307 (1), 7309 (1), 7313 (1), 7740 (1), 7741 (1)
	1094 (19)	1094 (5), 2289 (7), 4941 (2), 6371 (1), 6383 (1), 6384 (2), 6387 (1)
	Other (16)	2283 (1), 2909 (1), 5251 (1), 5412 (1), 6377 (1), 6378 (1), 6386 (1), 6391 (1), 7084 (1), 7299 (1), 7306 (1), 7308 (1), 7312 (1), 7721 (1), 7732 (1), 7743 (1)
6B (113)	2421 (26)	2421 (12), 2909 (2), 4929 (2), 6290 (1), 6291 (1), 6299 (1), 6307 (1), 6309 (2), 7305 (1),7724 (1), 7731 (1), 8231 (1)
	156 (26)	138 (1), 185 (7), 1447 (1), 3473 (1), 6235 (1), 6259 (1), 6293 (3), 6296 (1), 6297 (5), 6304 (1), 6526 (1), 6529 (1), 7302 (1), 7738 (1)
	473 (12)	2285 (2), 4087 (4), 5073 (1), 6308 (2), 6289 (1), 7303 (1), 7311 (1)
	1094 (8)	1094 (4), 2289 (1), 6292 (1), 4941 (1), 6298 (1)
	242 (6)	242 (2), 6279 (1), 6288 (1), 6303 (1), 6305 (1)

	1381 (4)	6239 (2), 6243 (2)
	230 (4)	230 (2), 6231 (1), 6234 (1)
	Other (27)	193 (1), 439 (1), 2185 (1), 3594 (1), 6230 (1), 6232 (1), 6233 (1), 6236 (2), 6245 (1), 6251 (2), 6276 (1), 6278 (1), 6294 (2), 6295 (1), 6300 (1), 6301 (1), 6302 (1), 6306 (2), 6310 (1), 6393 (1), 6525 (1), 7084 (1), 7733 (1)
6C (11)	2185 (9)	2185 (3), 2283 (1), 6310 (3), 6311 (1), 7345 (1)
	Other (2)	6297 (1), 7344 (1)
7F (29)	218 (21)	218 (8), 3544 (9), 6838 (2), 7884 (1), 7886 (1)
	Other (8)	2416 (1), 7048 (1), 7360 (1), 7361 (1), 7379 (1), 7883 (1), 7885 (1), 7887 (1)
9V (70)	4881 (40)	3454 (14), 4881 (22), 6314 (1), 6316 (1), 6530 (1), 7889 (1)
	156 (19)	156 (6), 162 (4), 557 (1), 609 (5), 644 (1), 2280 (1), 7891 (1)
	Other (11)	53 (1), 473 (1), 706 (1), 3983 (1), 5408 (1),6312 (1), 6313 (1), 6315 (1), 6531 (1), 7888 (1), 7890 (1)
14 (115)	230 (49)	230 (25), 1701 (1), 2707 (4), 3980 (1), 6227 (1), 6228 (1), 6229 (1), 6231 (2), 6234 (7), 6532 (1), 7267 (2), 7283 (1), 7291 (2)

	15 (30)	1492 (11), 2652 (7), 7282 (2), 7271 (1), 7276 (1), 7278 (1), 7280 (1), 7284 (1), 7287 (1) 7288 (1), 7726 (1), 7728 (1), 7729 (1)
	63 (25)	63 (7), 861 (2), 2414 (6), 5004 (1), 5187 (1), 7270 (1), 7277 (1), 7279 (1), 7281 (1), 7293 (1), 7722 (1), 4902 (2)
	Other (11)	124 (1), 2421 (1), 6226 (1), 6393 (1), 6533 (1), 7266 (1), 7269 (1), 7273 (1), 7292 (1), 7295 (1), 8233 (1)
18C (57)	1016 (17)	102 (4), 1016 (3), 6245 (2), 6247 (1), 7262 (2), 7265 (2), 7290 (1), 7297 (1), 7730 (1)
	1381 (13)	6239 (5), 7272 (2), 7274 (1), 7275 (3), 7289 (1), 6240 (1)
	193 (12)	4893 (1), 6241 (1), 6242 (4), 6246 (1), 6248 (1), 7268 (1), 7723 (1), 6237 (1), 7263 (1)
	6236 (10)	6236 (6), 7264 (1), 7286 (1), 7296 (1), 7742 (1)
	Other (5)	6238 (1), 6244 (1), 6250 (1), 7294 (1), 7346 (1)
19A (131)	2062 (109)	2062 (76), 4872 (1), 6121 (5), 6253 (1), 6254 (1), 6255 (3), 6256 (2), 6258 (1), 6260 (1), 6261 (1), 6262 (2), 6263 (1), 6264 (1), 6266(1), 6268 (1), 7716 (1), 7718 (1), 7720 (1), 7894 (1), 7895 (1), 7897 (2), 7898 (1), 7899 (1), 7903 (1), 8230 (1)
	156 (9)	75 (1), , 6265 (1), 7734 (1), 7737 (1), 7861 (1)

	Other (17)	172 (4)1381 (1), 2416 (1), 2421 (1), 6257 (1), 6267 (1), 7715 (1), 7717 (1), 7739 (1), 7896 (1), 7901 (1), 7902 (1), 7904 (1)
19F (46)	320 (23)	271 (16), 1467 (1), 6270 (4), 6275 (1), 7725 (1)
	347 (6)	347 (6)
	420 (5)	420 (2), 3217 (1), 7857 (1), 7858 (1)
	Other (12)	81 (1), 230 (1), 2067 (2), 6271 (1), 6272 (1), 6273 (1), 6278 (1), 6281 (1), 7131 (1), 7719 (1), 7860 (1)
23F (59)	242 (25)	2059 (5), 5653 (5), 6279 (5), 6285 (2), 242 (1), 362 (1), 6280 (1), 6284 (1), 7862 (1), 7868 (1), 8232 (1), 8233 (1)
	156 (12)	172 (5), 361 (1), 1535 (1), 6281 (1), 6287 (2), 7861 (1), 7882 (1)
	Other (22)	81 (4), 193 (1), 230 (1), 320 (1), 439 (2), 802 (2), 4087 (1), 63 (1), 6286 (1), 6389 (1), 7863 (1), 7865 (1), 7866 (1), 8032 (1), 8033 (1), 8034 (1), 8035 (1)

Table 3.2 Sequence type distribution among pneumococcal non-conjugate vaccine serotype pneumococci causing invasive disease in South African children <2 years (n=74), 2007

Serotype	Sequence type (no. of isolates)	Sequence type (other serotypes listed in the MLST database)[1]
15B/C (12)	411 (1), 4693 (1), **5808** (1), **7052 (7),** 8687 (1), 8820 (1)	4693 (19A), 8687 (6A)
8 (14)	53 (9), 1480 (1), 3847 (2), 7888 (1), 8831 (1)	7888 (9V)
9N (6)	66 (2), 3983 (2), 5624 (2)	66 (4, 14. 19A, 19F, 23F)
16 (4)	4088 (2), 8827 (1), 8834 (1)	none
10A (4)	2068 (1), 8821 (1), 8825(1), 8839 (1)	none
29 (4)	5483 (1), 8826 (3)	none
13(3)	5647 (3)	none
12F (3)	989 (1), 2416 (1), 8828 (1)	2416 (7F, 19A)
34 (2)	8818 (1), 8822 (1)	none
17F (2)	8829 (1), 8835 (1)	none
7C (2)	8838 (2)	none
23A (2)	2319 (1), 5080 (1)	5080 (23F)
33D (1)	8714 (1)	8714 (6A)
33F (1)	8837 (1)	none
35B (1)	8830 (1)	none
9L (1)	8833 (1)	none
NT (1)	**344 (1)**	344 (6A, 19F and 35B)

10F (1)	8841 (1)	none
20 (1)	8824 (1)	none
28 (1)	8823 (1)	none
38 (1)	393 (1)	393 (25A)
11A (1)	8836 (1)	none
11C (1)	8836 (1)	none
18B (1)	6239 (1)	ST6239 (18C)
18F (1)	8840 (1)	none
23B (1)	2911 (1)	2911 (23F)
33B (1)	4084 (1)	4084 (3, 10A, NT)
35F (1)	8832 (1)	none

[1]Other serotypes associated with this sequence type (as listed in the MLST global database (59)) but not identified in this study, are within parantheses.

Sequence types representing isolates with reduced susceptibility to penicillin are in bold typeface.

Table 3.3 Serotype and clonal complex distribution of penicillin non-susceptible (PNS) and multidrug-resistant (MDR) isolates causing invasive pneumococcal disease in South Africa, 2007

Serotype	Clonal complex[1]	No. of isolates	No. of PNS isolates (%)[2]	No. of MDR isolates (%)[3]
4	32	32	32 (100)	2 (6)
	63	1	1 (100)	1 (100)
6A	473	37	2 (5)	0
	1094	19	14 (74)	2 (11)
	156	23	8 (35)	1 (4)
	7084	2	1	0
	Singletons	8	3	0
6B	2421	26	19 (73)	4 (15)
	156	27	20 (74)	4 (15)
	473	12	9 (75)	2 (17)
	1094	8	6 (75)	0
	230	4	3 (75)	0
	242	7	6 (86)	4 (57)
	6236	2	2 (100)	0
	1381	5	3 (60)	0
	6393	2	2 (100)	0
	2185	2	1 (50)	0
	193	1	1 (100)	0
	439	1	1 (100)	0

	1016	1	1 (100)	0
	3594	1	1 (100)	0
	Singletons	15	10	0
9V	4881	40	14 (35)	1 (3)
	156	19	13 (68)	0
	473	1	1 (100)	1 (100)
	706	1	1 (100)	0
14	230	49	49 (100)	49 (100)
	63	25	22 (88)	20 (80)
	15	30	29 (97)	29 (97)
	156	2	1 (100)	1 (100)
	242	1	1 (100)	1 (100)
	2421	1	1 (100)	1
	6393	1	1 (100)	1
	Singletons	6	5	4
19A	2062	109	100 (92)	7 (6)
	156	9	9 (100)	1 (11)
	1381	1	1 (100)	0
	2416	1	1 (100)	0
	2421	1	1 (100)	0
	Singletons	10	7	1
19F	320	23	23 (100)	23 (100)
	347	6	6 (100)	1 (17)
	420	5	4 (80)	0

	2067	2	1 (50)	0
	81	1	1 (100)	0
	230	1	1 (100)	1 (100)
	Singletons	6	5	3
23F	242	25	16 (64)	16 (64)
	156	12	8 (67)	1 (8)
	81	4	4 (100)	4 (100)
	230	1	1 (100)	1 (100)
	63	1	1	0
	320	1	1 (100)	1 (100)
	473	1	1 (100)	1 (100)
	Singletons	8	5	3
Total		640	481	192

[1]Only clonal complexes/singletons with at least one penicillin non-susceptible (PNS) isolate are reflected in the table; [2]Percentage of isolates that are penicillin non-susceptible within the specified clonal complex, [3]Percentage of isolates that are multi-drug resistant (MDR) within the specified clonal complex.

3.4 Discussion

Serotypes with low carriage detection rates such as serotype 1 are normally less diverse while the opposite is true for serotypes that are associated with carriage. Consistent with the literature, our serotype 1 and 5 isolates were the least diverse while the remaining PCV serotypes were more heterogeneous. Several globally disseminated clones described by the pneumococcal molecular epidemiology network (PMEN) (60), were identified among our isolates. Sweden^{1}-27 (CC217) represented nearly all (98%) of our serotype 1 isolates. We have previously shown the sustained circulation of this clone among our serotype 1 isolates (61). South Africa^{6B}-8 (ST185), which was identified among our serotype 6B isolates recovered in the 1990's (62), continues to circulate. Denmark^{14}-32 (CC230) was the most prevalent clone among serotype 14 isolates. ST271, which is a single-locus variant of the multidrug resistant Taiwan^{19F}-14 (ST236), was prevalent among our serotype 19F isolates.

Some serotypes were represented mainly by clonal complexes that appear to be uncommon in other parts of the world. CC458 was the most common genotype among serotype 3 isolates, while Netherlands^{3}-31 (CC180) only represented 15% of our isolates. Our earlier study on a random selection of South African serotype 3 isolates, collected from 2000 through 2005 (63) also showed the predominance of CC458 which accounted for 54% (84/155) of serotype 3 isolates tested, while CC180 represented only 12% (18/155) of the isolates. CC458 has also been documented among serotype 3 isolates from Ghana, Egypt, Israel, Costa Rica and Lithuania (64-66). ST5410 (CC32), which represented the majority of our serotype 4 isolates, has only been reported in Mozambique (67). Interestingly, the majority (83%) of our serotype 19A isolates were CC2062, which, to our knowledge, has only been reported in Ireland, as an emerging serotype 19A clone post PCV introduction (68). In our earlier study, CC2062 was a

novel genotype identified among 20% (5/25) of multidrug resistant serotype 19A isolates that caused IPD from 1998 through 2004 (unpublished data). Other than CC156, which represented only 7% (9/131) of our serotype 19A isolates, no other common serotype 19A genotype including CC199 (predominant in the US), CC230 (common in Spain and Kenya) and CC320 (common in Spain, Asia and France and the US), was identified among our serotype 19A isolates (32;69-71).

We identified clonal complexes among our PCV serotypes that were represented by more than one capsular type. Although some of these clonal complexes have previously been reported in more than one serotype, we identified new serotype-ST combinations. CC156 has been identified among isolates belonging to multiple serotypes including 6A, 6B, 6C, 19A, 19F and 23F (67). Similarly, CC156 was represented among our isolates belonging to multiple serotypes but was more common among serotype 6A and 23F isolates. The majority of our serotype 7F isolates were CC218, which is mainly associated with serotype 12F (67). This clonal complex has been rarely identified among 7F isolates from other parts of the world (30;67). Although ST63 is a globally disseminated serotype 15A clone, it has also been identified among other serotypes including serotype 14 (67;72-74). It is therefore not surprising that more than 20% of our serotype 14 isolates were CC63. We also identified clonal complexes that are normally associated with PCV serotypes among some of our non-PCV isolates. A limitation of this study is that data only includes isolates collected over one year; hence we were unable to determine the sustainability of unusual serotype-clonal complex combinations have been circulating in South Africa. This is addressed in chapter 4, where isolates recovered over a longer period were analyzed.

PMEN was initially established for surveillance of globally disseminated pneumococcal clones that are antibiotic resistant. The majority of South African non-susceptible isolates were related to PMEN clones. All serotype 14 and 23F isolates related to Denmark^{14}-32 were multidrug resistant. All serotype 19F and 23F isolates belonging to CC320, which also represent Taiwan^{19F}-14 (ST236), were also multidrug resistant. The majority (80%) of serotype 14 isolates related to Sweden^{15A}-25 (CC63) were multidrug resistant. The majority (73%) of serotype 6B isolates related to South Africa^{6B}-8 [ST185 (CC156)] were non-susceptible to penicillin.

In this study, serotype 15B/C was the most common serotype among non-PCV isolates from children <2 years and had the highest proportion of isolates that were non-susceptible to penicillin. Serotype 15B was among the top 20 serotypes that caused IPD among South African children in 2003 through 2008, prior to the introduction of PCV (48), and may therefore be poised to fill the niche vacated by PCV serotypes.

4 Chapter Four

Population snapshot of *Streptococcus pneumoniae* isolates causing invasive disease before and after introduction of pneumococcal conjugate vaccines in South Africa

4.1 Introduction

Increases in IPD rates due to non-PCV serotypes have been observed in some countries following routine use of the PCV (23-27). These increases have been attributed to expansion of pre-existing clones; however, capsule switching has also been reported (30-33). Predominant sequence types circulating in South Africa before PCV introduction among some serotypes (3, 4 and 19A) differed from predominant sequence types identified in other parts of the world (75). In this study genotypes were compared among isolates causing IPD prior to and following routine PCV7 and PCV13 use in South Africa.

4.2 Materials and methods

4.2.1 Incidence rates

Average incidence for the periods 2005-2008 (pre-PCV period), 2009-2010 (PCV7 period) and 2011-2013 (PCV13 period) was calculated by dividing the number of cases by mid-year population estimates data from Statistics South Africa (76) and multiplying the quotient by 100,000.

4.2.2 Selection of isolates for genetic characterization

Isolates from cases reported from 2005 through 2013 were used in this study. Isolates were selected for two separate projects with different sample selection criteria. In the first project all viable PCV13 serotypes (1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F and 23F) and serotype 6C isolates from infants and young children (<5 years of age) were selected for characterization. Serotype 6C was included due to cross-protection from 6A and 6B. Due to the high numbers of isolates for older individuals (≥5 years of age), 50% of all viable PCV13 serotypes were selected for characterization. PCV13 isolates were stratified by collection date and every second isolate was selected.

For the second project, 300 isolates per year representing all serotypes collected from 2005 through 2013 were sampled. Isolate selection was stratified as follows: 150 isolates from children aged 0 to 2 years, 75 isolates from children aged 3 to 5 years and 75 isolates from individuals who were >5 years of age.

4.2.3 Genetic characterization

Isolates selected for the first project from isolates collected in 2007 and 2011 were characterized by conventional MLST as described in chapter 2. For the remaining isolates from the first project and for isolates selected for the second project, MLST genotypes were obtained through whole-genome sequencing as described in chapter 2. Relationships between sequence types and diversity by serotype were assessed as described in chapter 2.

4.3 Results

4.3.1 IPD cases (all ages)

From 2005 through 2013, 36,267 IPD cases were reported of which 70% (25,508/36,267) had viable isolates available for further characterization. Infants and young children aged <5 years represented 29% (7,279/25,508) of IPD cases with viable isolates.

4.3.2 IPD in infants and young children (<5 years)

During the pre-PCV period (2005-2008), PCV7 period (2009-2010) and PCV13 period 5,928, 2,246 and 1,703 cases of IPD were reported of which 76% (4,480/5,928), 74% (1,659/2,246) and 69% (1,140/1,703) had viable isolates. Serotypes 14, 6B, 6A, 23F and 19F were the most prevalent serotypes during the pre-PCV period accounting for 10-16% of the 4,480 isolates each, while 19A, 1, 6A, 8 and 23F were the most common serotypes during the PCV13 period accounting for 7-9% of the 1,140 isolates. Overall PCV13 serotypes and 6C caused 85% (3,816/4,480), 83% (1,385/1,659) and 55% (625/1,140) of IPD during the pre-PCV, PCV7 and PCV13 period, respectively. The average incidence of PCV (PCV13 and 6C) serotype-disease declined 79% (95% CI, -82% to -75%) from 18.6/100,000 population in the pre-PCV period to 4.0/100,000 population in the PCV13 period (Figure 4.1a). The percentage of non-PCV serotypes increased from 15% (664/4,480) during the pre-PCV period to 45% (515/1,140) during the PCV13 period however, the average incidence for non-PCV serotypes remained the same during the pre-PCV and PCV13 period (3.3/100,000 population vs. 3.3/100,000 population) (Figure 4.1a). Incidence rates increased among some non-PCV serotypes (8, 12F, 16F, 15A, 17F and 35B) however 35B was the only serotype that increased significantly. Serotype 35B increased 76% (95% CI, 26% to 94%) from 0.078/100,000 population in the pre-

PCV period to 0.32/100,000 population in the PCV13 period (Figure 4.2a). Serotype 8 increased 25% (95% CI, -41% to 61%) from 0.37/100,000 during the pre-PCV period to 0.49/100,000 population during the PCV13 period. Serotype 12F incidence increased 25% (95% CI, -60% to 67%) from 0.25/100,000 population in the pre-PCV period to 0.34/100,000 population in the PCV13 period. Serotype 16F incidence increased 37% (95% CI, -64% to 77%) from 0.16/100,000 population in the pre-PCV period to 0.25/100,000 population in the PCV13 period.

4.3.3 IPD in older children and adults (≥5 years)

Among older children and adults, 12,410, 6,412 and 7,576 IPD cases were reported during the pre-PCV, PCV7 and PCV13 period, respectively. Viable isolates were available for 70% (8,679/12,410), 69% (4,403/6,412) and 68% (5,147/7,576) of cases from the three periods, respectively. Serotypes 1, 19A, 4, 6A and 14 were the most prevalent serotypes during the pre-PCV period each accounting for 6-14% of the 8,679 isolates, while 1, 19A, 12F, 4, 3, 6A, and 8 were the most common serotypes during the PCV13 period each accounting for 5-14% of the 5,147 isolates. PCV serotypes were responsible for 70% (6,108/8,679), 71% (3,132/4,403) and 58% (2,990/5,147) of IPD during the three periods, respectively. The average incidence of PCV serotype-disease declined 70% (95% CI, -72% to -68%) from 7.2/100,000 population in the pre-PCV period to 2.1/100,000 population in the PCV13 period (Figure 4.1b). Non-PCV serotypes were responsible for 30% (2,571/8,679), 29% (1,271/4,403) and 42% (2,175/5,147) of IPD during the pre-PCV, PCV7 and PCV13 period, respectively. Average incidence for non-PCV serotypes declined 49% (95% CI, -54% to -54%) from 3.0/100,000 population during the pre-PCV period to 1.5/100,000 population PCV13 period (Figure 4.2b).

4.3.4 Genotypes and population structure

Sequence types were available for 17% (2,203/13,159), 8% (511/6,062) and 37% (2,324/6,287) of isolates obtained during the pre-PCV, PCV7 and PCV13 periods, respectively (appendix A). Sequence types and clonal complexes are listed in appendix B and C and eBURST population snapshots are depicted in figures 4.3 and 4.4.

4.3.4.1 PCV serotypes

Serotype 1: Pneumococcal Molecular Epidemiology Network (PMEN) clone Sweden^{1}-1 (ST217) and SLVs, represented >90% of serotype 1 isolates during the pre-PCV and PCV13 periods for infants and young children and for older children and adults (appendix B and C, figure 4.3).

Serotype 3: CC458 was the main genotype among serotype 3 isolates and represented 53% to 79% of isolates from infants and young children and 47% to 55% from older children and adults dung each of the three periods (appendix B and C, figure 4.3). CC180 (Netherlands^{3}-33 (ST180) and SLVs) accounted for 7% to 21% of isolates from infants and young children and 16% to 22% from older children and adults during each period. CC230 (all of which belong to ST700) was common (18%) among isolates from older children and adults during the PCV13 period but was not detected in the pre-PCV period. A non-significant increase in sequence type diversity was observed for isolates from infants and young children during the pre-PCV period compared to the PCV13 period (appendix F).

Serotype 4: CC5410 and CC1221 were the main genotypes among serotype 4 isolates during the three periods. The two respectively accounted for 33% to 68% and 21% to 50% among isolates from infants and young children and 9% to 38% and 14% to 64% among isolates from older children and adults (appendix B and C, figure 4.3). Sweden^{4}-38 (ST205) was detected during the pre-PCV period (n=4) and PCV13 period (n=9). A significant decline occurred in ST diversity for isolates from older children and adults during the PCV13 compared to the pre-PCV period (appendix G).

Serotype 5: During each of the three periods, over 90% of serotype 5 isolates from both age groups belonged to CC5659 (Columbia^{5}-19 (ST289) and SLV's) (appendix B and C, figure 4.3).

Serotype 6A: For both age groups, CC2285 was the most common genotype among 6A isolates, accounting for 39% to 47% of isolates during each of the three periods (appendix B and C, figure 4.3). S. Africa^{6B}-8 (ST185) was detected during the pre-PCV (n=5) period.

Serotype 6B: CC2421 was common among serotype 6B isolates, accounting for 19% to 33% for isolates from infants and young children and 22% to 38% for older children and adults, during each of the three periods. South Africa^{6B}-8 (ST185) and SLVs were also common during all three periods and represented 10% to 44% of isolates from infants and young children and 17% to 26% from older children and adults.

Serotype 7F: CC3544/218 (Netherlands^{12F}-24 (ST218) and SLVs) was prevalent among serotype 7F isolates from both age groups, representing 67% to 100% of isolates during each period.

Serotype 9V: CC4881 represented the majority of 9V isolates, accounting for 58% to 72% for isolates from infants and young children and 54% to 67% for older children and adults during each period. CC156 (Spain^{9V}-3 (ST156) and SLVs) represented 10% to 17% of isolates from infants and young children and 23% to 33% of isolates from older children and adults during each period.

Serotype 14: CC230 (Denmark^{14}-32 (ST230) and SLVs) was predominant among serotype 14 isolates, accounting for 29% to 52% of isolates for both age groups during each period. For both age groups, Sweden^{15A}-25 (ST63) and SLVs were also common during each of the three periods.

Serotype 18C: The majority of 18C isolates belonged to three genotypes (CC6239, CC6245 and CC4893). During each period and for both age groups, together the three represented 12% to 41% of isolates.

Serotype 19A:

CC2062 was the most common 19A genotype, accounting for ≥80% of isolates during each period for both age groups.

Serotype 19F: CC347, CC271 and CC420 were the most predominant genotypes during each period and each represented 10% to 35% of isolates for both age groups.

Serotype 23F: CC6279/242 (Taiwan^{23F}-15 and SLVs) was common among 23F isolates, accounting for 54% to 85% of isolates from infants and young children and 36% to 93% from older children and adults during each period. For both age groups, CC172 was more common during the pre-PCV7 period (representing 27% of isolates from infants and young children and 28% from older children and adults) than during the PCV13 period (1% and 5% of isolates for the two age groups, respectively). Spain^{23F}-1 (ST81) was detected among isolates from both age groups [pre-PCV (n=5) and PCV13 (n=3)].

4.3.4.2 Non-PCV serotypes.

Serotype 8: Netherland^{8}-33 (ST53) and SLVs were common among serotype 8 isolates during each of the three periods and accounted for 93% to 100% of isolates from infants and young children and 79% to 96% from older children and adults (appendix B and C, figure 4.4).

Serotype 12F: Two genotypes (CC5026/989 and CC2416) were predominant among serotype 12F isolates. Together the two represented 39% to 56% and 37% to 63% of 12F isolates for the two age groups respectively during each period. Isolates belonging to Netherland^{8}-33 (ST53) were detected during the pre-PCV period (Table 4.1).

Serotype 15A: For both age groups, the majority of 15A isolates belonged to unrelated STs. CC10588 accounted for 33% and 44% and Sweden^{15A}-25 (ST63) for 18% and 11% of isolates

from infants and young children and from older children and adults, respectively (appendix B and C, figure 4.4).

Serotype 15B/C: During the pre-PCV period the majority (60%) of 15B/C isolates from infants and young children belonged to ST7052, while ST8687 (50%) and ST7052 (27%) were common during the PCV13 period (appendix B). For the older age group, during the pre-PCV period, isolates belonged to CC199 (50%, 3/6), ST7052 (n=2) and ST8687 (n=1). ST7052 (n=2) and ST8687 (n=2) were the only genotypes identified among isolates from older children and adults during the PCV13 period (appendix C). A significant decline in ST diversity occurred between the pre-PCV and PCV13 period among 15B/C isolates from older children and adults (appendix G).

Serotype 35B: For both age groups, majority (>70%) of 35B isolates belonged to singletons during the pre-PCV period, while CC172 was common during the PCV13 period [56% for infants and young children and 67% for older children and adults] (appendix B and C). During the pre-PCV period CC172 represented 13% of isolates from infants and young children and 8% from older children and adults.

4.3.4.3 Genotypes represented by two or more serotypes

The most common genotype among serotype 23F (CC6279) was also detected among 6B isolates throughout the study period [5% for each of the two age groups] (Table 4.2). CC172, a common genotype among serotype 23F isolates, was also common among 6A, 19A and 35B isolates and was identified among a few 6B isolates (Table 4.2). For these four serotypes, CC172 was detected throughout all periods. CC445 was detected among serotype 33F and 22F isolates

from both age groups during each period. The main 19A genotype (CC2062) was identified among serotype 19F isolates during the pre-PCV and PCV13 period. The most common 7F genotype (CC3544/218) was detected among 12F isolates during the pre-PCV period (6% for infants and young children and 15% for older children and adults).

4.3.4.4 Antimicrobial susceptibility

Reduced susceptibility to penicillin was detected in 53% (2,390/4,480), 62% (1,035/1659) and 39% (448/1140) of isolates from infants and young children during the pre-PCV, PCV7 and PCV13 period, respectively. For the older age group, reduced susceptibility to penicillin was detected in 27% (2,373/8,679), 36% (1574/4403) and 29% (1502/5147) of isolates during the three periods, respectively. For infants and young children, penicillin non-susceptible isolates were predominantly serotype 14, 6B and 19F during both the pre-PCV (89% [625/700], 72% [438/607], and 83% [365/440], respectively) and the PCV13 period (100% [43/43], 80% [45/56] and 97% [57/59], respectively). For the older age group serotype 14 and 19F had high proportions of penicillin non-susceptible isolates during both the pre-PCV [86% (450/524) and 80% (307/382), respectively] and the PCV13 period [96% (110/106), and 94% (146/154), respectively]. For both age groups penicillin non-susceptible isolates were common among serotype 14 belonging to CC63, CC230 and CC2652 and 19F isolates belonging to CC271, CC347 and CC420 (appendix D and E). The proportion of penicillin non-susceptible 19A isolates increased between the pre-PCV and PCV13 period [68% (238/351) to 99% (102/103) ($P<0.0001$) for infants and young children and 63% (475/750) to 96% (522/542) ($P<0.0001$) for the older age group. Among non-PCV serotypes, the proportion of penicillin non-susceptible 35B isolates increased between the pre-PCV and PCV13 period [(18% (3/17) to 56% (28/50)

(p=0.02)] for infants and young children and [8% (2/44) to 50% (32/64) (p<0.0001)] for older children and adults. Penicillin non-susceptible 35B isolates were predominantly CC172 (appendix D and E). Increases in proportions of serotype 15A penicillin non-susceptible isolates occurred between the pre-PCV and PCV13 period [13% (2/16) to 42% (13/31) (p=0.05) for infants and young children and 9% (6/70) to 40% (48/120) (p<0.0001) for older children and adults].

Multidrug resistance was detected in 21% (941/4,480), 24% (412/1,659) and 13% (143/1140) of isolates from infants and young children and in 9% (799/8679), 13% (553/4403) and 9% (468/5147) of isolates from the older age group during the pre-PCV, PCV7 and PCV13 period, respectively. Multidrug resistance was common among serotype 14 isolates during both the pre-PCV and the PCV13 period [83% (579/7000) and 85% (39/46) in infants and young children and 80% (420/524) and 81% (101/114) in the older age group]. Multidrug resistance was common among serotype 14 isolates belonging to CC63, CC230 and CC2652 (appendix D and E). The proportion of multidrug resistant serotype 15A isolates increased between the pre-PCV and PCV13 period [6% (1/16) to 23% (7/31) (p=0.2) and 3% (2/70) to 18% (22/120) (p=0.03) for the two age groups respectively]. Multidrug resistant 15A isolates belonging to CC63 and ST11766 were detected (appendix D and E). Overall serotype 3 isolates were susceptible to tested antimicrobial agents however nearly all serotype 3 isolates belonging to CC230 were multidrug resistant (appendix D and E).

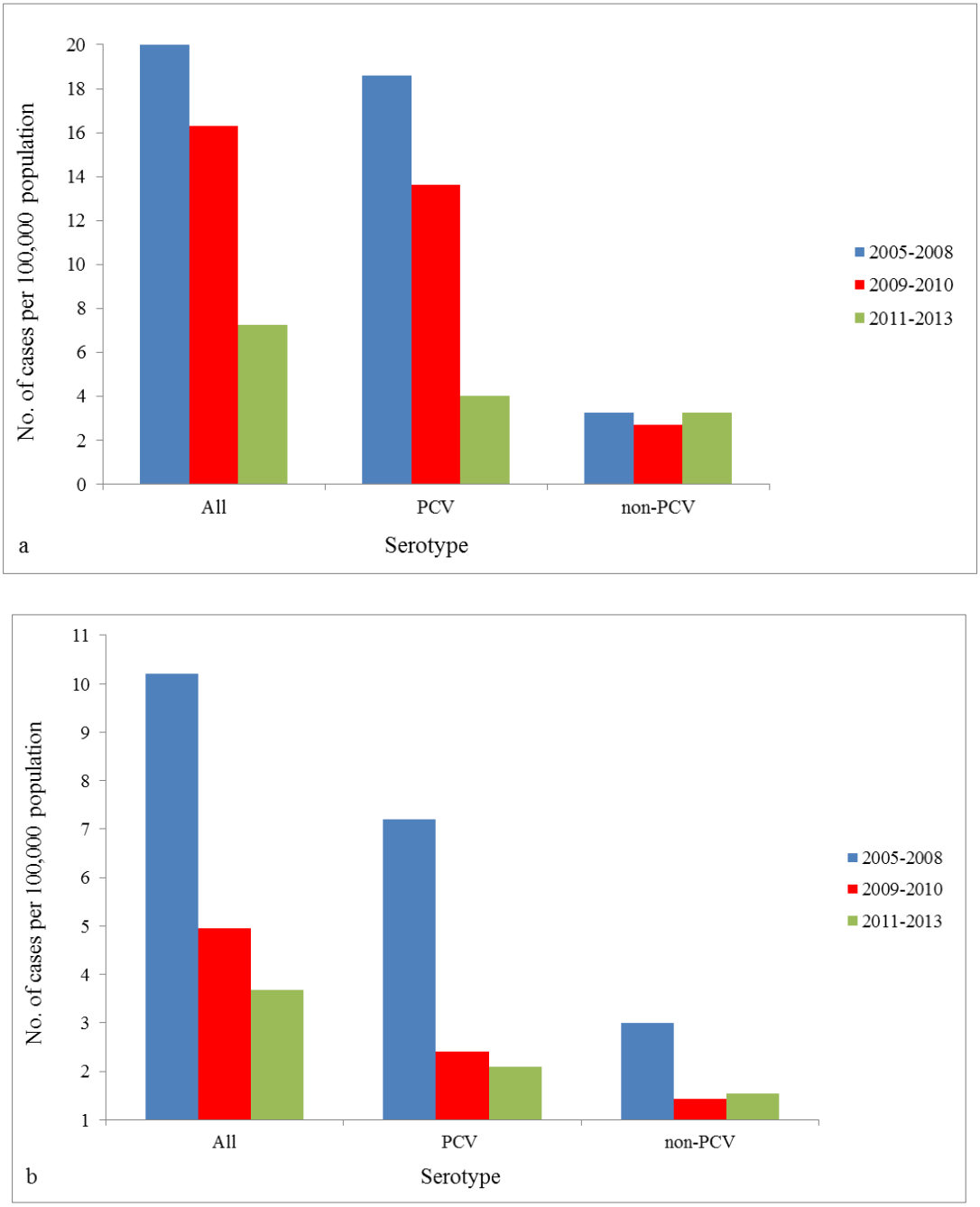


Figure 4.1 Average incidence rates of invasive pneumococcal disease in South Africa.

Rates were calculated by period [2005-2008 (pre-PCV), 2009-2010 (PCV7) and 2011-2013 (PCV13)] and serotype; (a) children <5 years old, (b) individuals ≥5 years old.

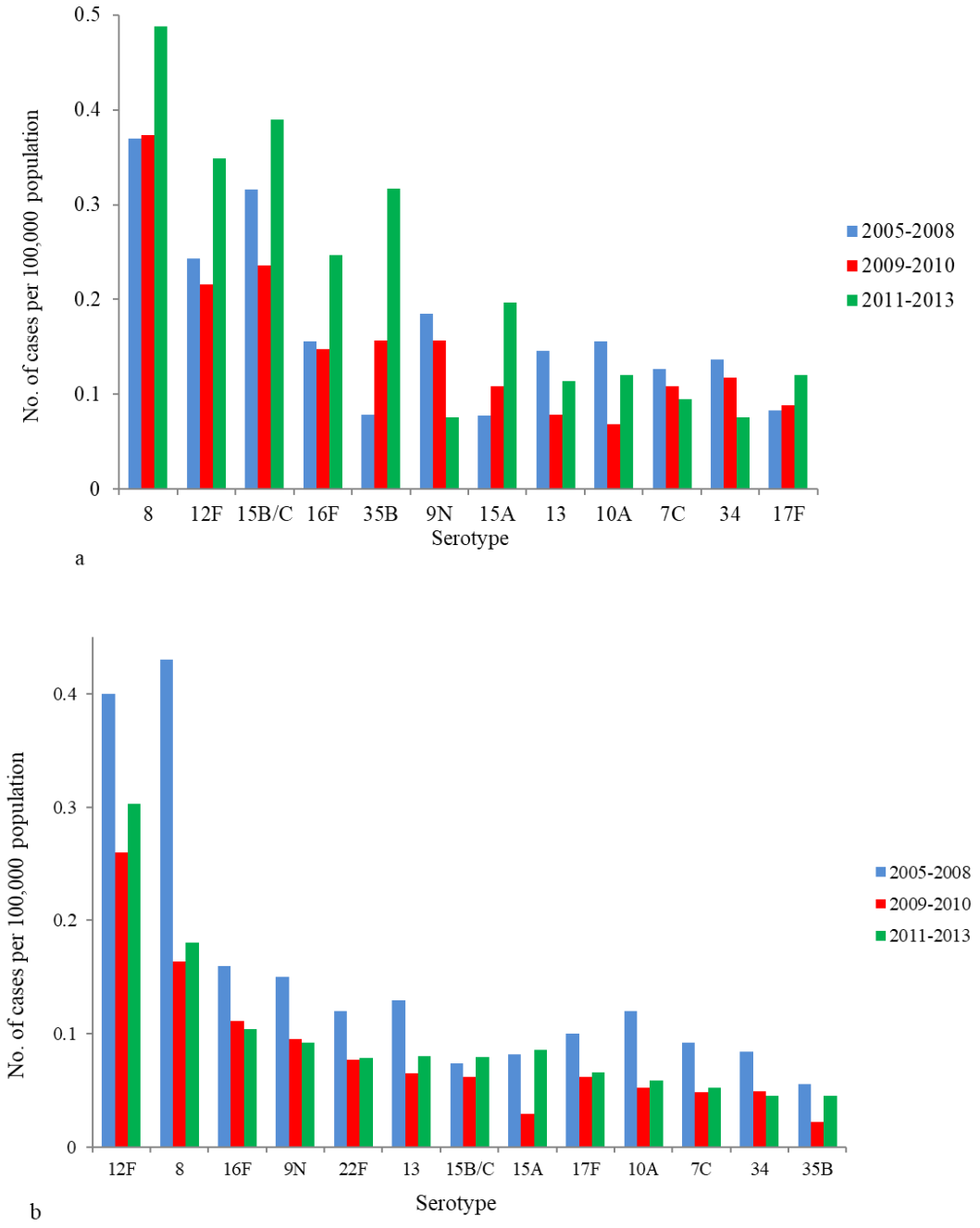


Figure 4.2 Average incidence rates of invasive pneumococcal disease caused by common non-PCV serotypes in South Africa.

Rates were calculated by period [2005-2008 (pre-PCV), 2009-2010 (PCV7) and 2011-2013 (PCV13)] and serotype; (a) children <5 years old, (b) individuals ≥5 years old.

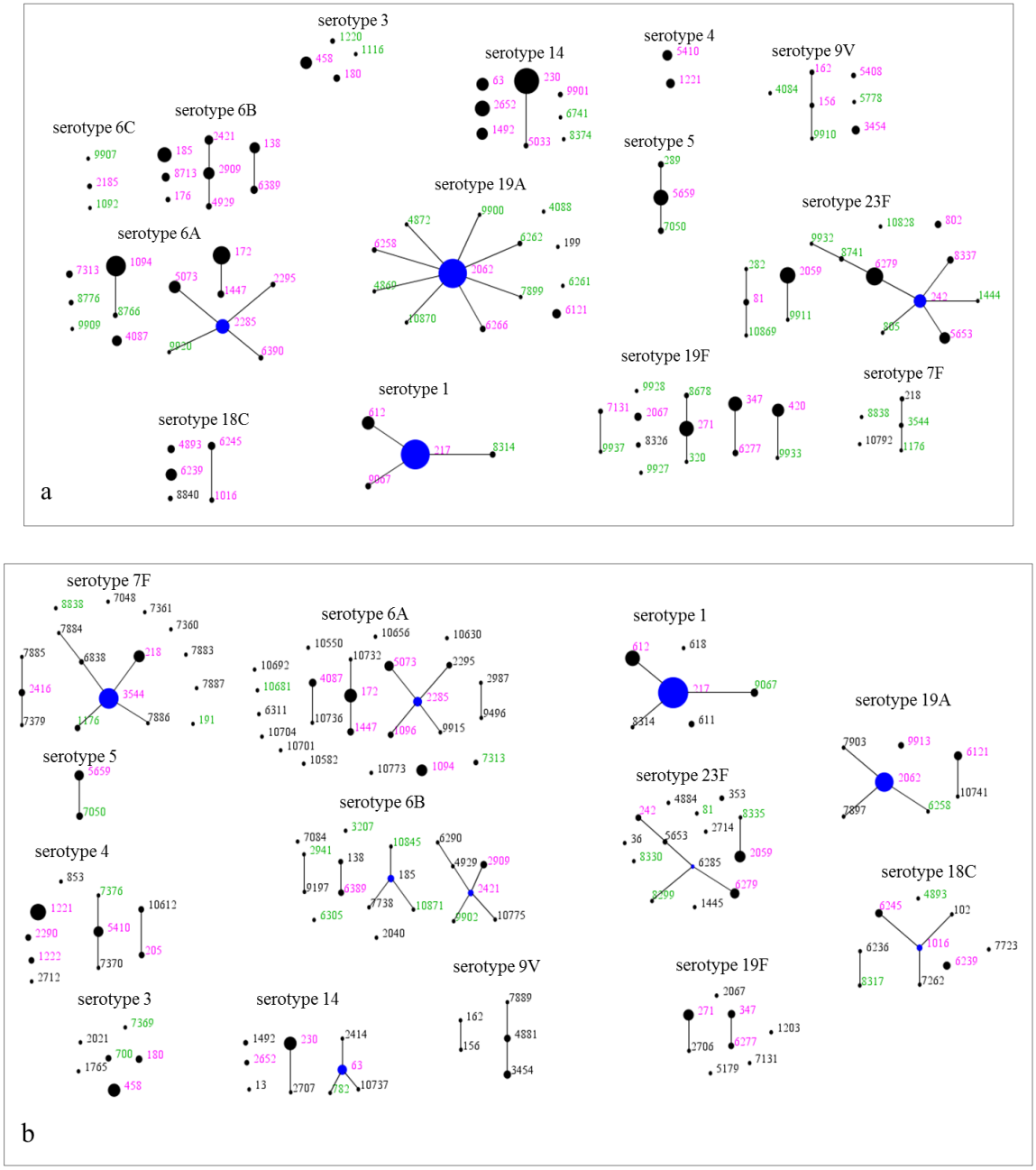


Figure 4.3 Population snapshot depicting relationship between PCV serotypes causing invasive pneumococcal disease in South Africa.

(a) children <5 years old [pre-PCV (2005-2008) (n=953) and PCV13 period (2011-2013) (n=512)], (b) older individuals [pre-PCV (n=905) and PCV13 period (n=1584)]. Related STs differing by one allele are connected by a line. STs found only during the pre-PCV period are depicted in black, STs found during the PCV13 period are in green and STs found during both periods are in pink.

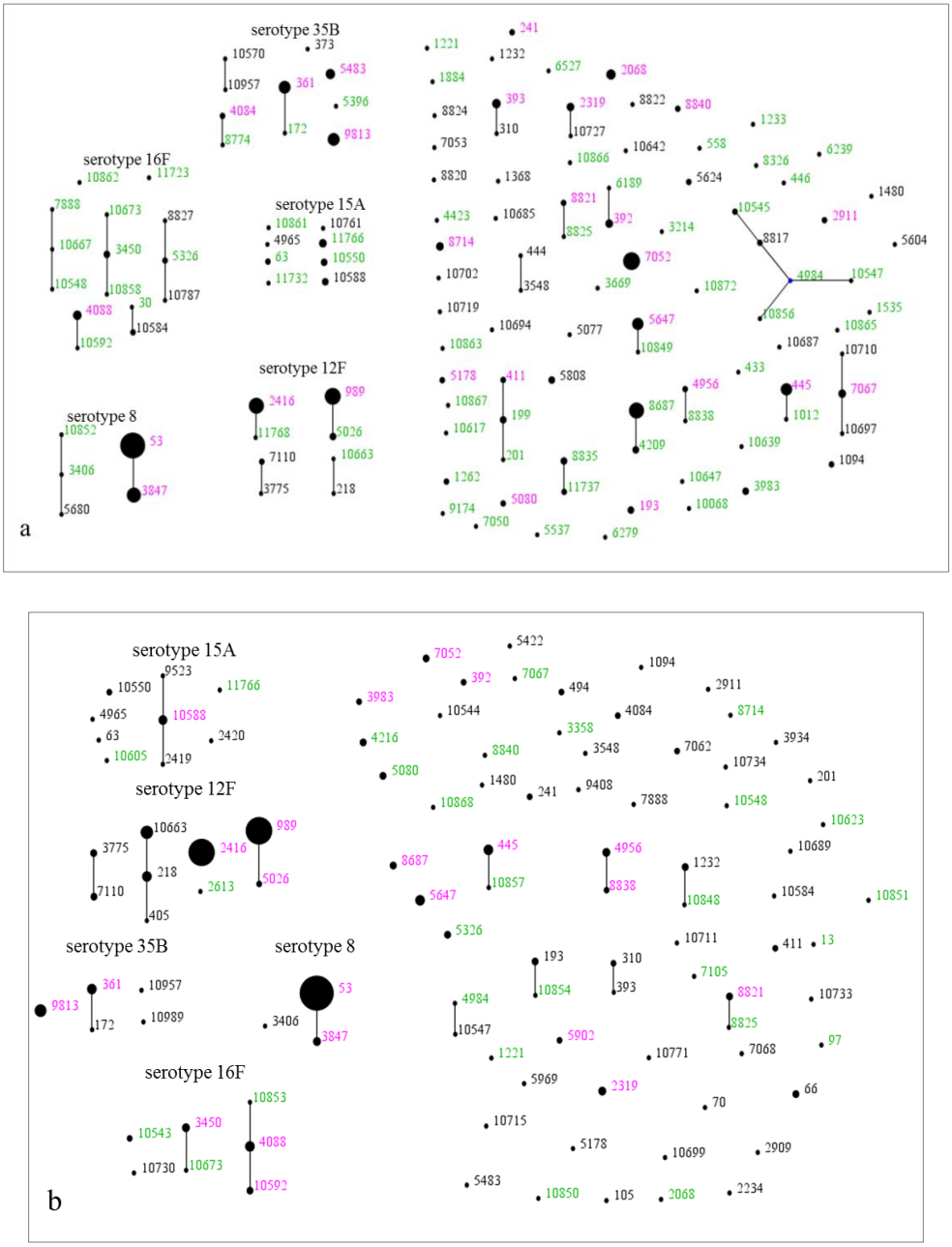


Figure 4.4 Population snapshot depicting relationship between non-PCV serotype isolates causing invasive pneumococcal disease in South Africa.

(a) children <5 years old [pre-PCV (2005-2008) (n=120) and the PCV13 period (2011-2013) (n=167)], (b) older individuals [pre-PCV (n=274) and the PCV13 period (n=107)]. Each circle represents a sequence type (ST) and the size is proportional to the number isolates for the ST. Related STs differing by one allele are connected by a line. STs found during the pre-PCV era are depicted in black, STs found during the PCV13 era are in green and STs found during both periods are in pink. Common serotypes are indicated.

Table 4.1 Pneumococcal molecular epidemiology network (PMEN) clones identified among isolates causing invasive pneumococcal disease in South Africa during the pre-PCV (2005-2008), PCV7 (2009-2010) and PCV13 period (2011-2013)

	Serotype (no. of isolates)		
PMEN clone	pre-PCV period	PCV7 period	PCV13 period
Spain^{23F}-1 (ST81)	23F (5)	-	23F (2), 6A (3)
Spain^{6B}-2 (ST90)	-	6B (1)	6B (1)
Spain^{9V}-3 (ST156)	9V (8)	9V (2)	9V (8)
S. Africa^{19A}-7 (ST75)	19A (1)	-	-
S. Africa^{6B}-8 (ST185)	6A (4), 6B (22)	6B (3)	6B (27)
Taiwan^{23F}-15 (ST242)	23F (10), 6B (4)	23F (3), 6B (1)	23F (11), 6B (1)
Columbia^{5}-19 (ST289)	5 (2)	-	5 (3)
Netherlands^{12F}-24 (ST218)	7F (9), 12F (6)	7F (2)	7F (12)
Sweden^{15A}-25 (ST63)	14 (2)	-	14 (3)
Sweden^{1}-27 (ST217)	1 (204)	1 (54)	1 (185)
Sweden^{1}-28 (ST306)	1 (1)	-	1 (5)
USA^{1}-29 (ST615)	1 (1)	-	1 (1)
Creece^{21}-30 (ST193)	21 (4)	-	21 (2)

Netherlands^{3}-31 (ST180)	3 (9)	3 (2)	3 (38)
Demark^{14}-32 (ST230)	14 (78),18C (1), 23F (1)	14 (34)	14 (61), 19F (2)
Netherland^{8}-33 (ST53)	8 (80)	8 (7)	8 (40)
Netherlands^{14}-35 (ST124)	14 (2)	-	-
Netherlands^{15B}-37 (ST199)	19A (1)	-	15B/C (1)
Netherlands^{7F}-39 (ST191)	-	-	7F (3)
Sweden^{4}-38 (ST205)	4 (4)	-	4 (9)

Table 4.2 Clonal complexes detected in two or more pneumococcal serotypes of isolates causing invasive disease in South Africa, 2005-2013.

		No. of isolates (%): pre-PCV/PCV7/PCV13 period	
Clonal complex	Serotype	children <5 years of age	individuals ≥5 years of age
172	6A	4 (4)/7 (12)/4 (74)	1 (1)/0/24 (16)
	19A	3 (3)/3 (8)/1 (1)	7 (8)/1 (7)/5 (3)
	23F	27 (26)/8 (21)/1 (1)	17 (25)/0/8 (5)
	35B	0/1 (50)/9 (53)	1 (7)/0/4 (67)
230	3	1 (7)/0/0	0/1 (25)/82 (47)
	14	77 (50)/22 (44)/19 (48)	30 (41)/2 (29)/45 (49)
445	22F	3 (100)/0/5 (71)	2 (100)/1 (100)/2 (67)
	33F	0/0/1 (50)	1 (100)/1 (100)/2 (100)
6279	6B	8 (6)/0/1 (3)	2 (3)/1 (25)/5 (5)
	23F	54 (52)/26 (67)/57 (84)	22 (32)/14 (88)/134 (83)
2062	19A	97 (86)/32 (84)/77 (91)	68 (81)/12 (80)/168 (90)
	19F	3 (4)/0/1 (2)	3 (5)/0/3 (3)
3544/218	7F	2 (67)/3 (75)/2 (67)	20 (63)/1 (100)/61 (75)
	12F	1 (6)/0/0	15 (15%)/0/0

199	15B/C	0/0/4 (13)	3 (43)/0/0
	16F	0/0/1 (5)	0/0/0
	26	1 (100)0/0	0/0/0
63	14	25 (16)/8 (16)/5 (13)	20 (27)/0/11 (12)
	15A	0/0/2 (18)	1 (8)/0/0
	23F	1 (17)/0/0	0/0/0

4.4 Discussion

In this study, sequence types of isolates from South African patients with IPD from 2005 through 2013 are described. During this period, PCV7 and PCV13 were introduced in 2009 and 2011, respectively. For both infants and young children aged <5 years and older individuals, incidence of PCV serotype disease declined during the PCV13 period compared to the pre-PCV period. For infants and young children, overall incidence of non-PCV serotype disease did not change between the two periods; however, increases occurred among some serotypes (8, 12F, 16F, 15A, 17F and 35B). For the older age group, incidence of non-PCV serotype disease declined. For most serotypes sequence type diversity and predominant sequence types did not change between the pre-PCV and PCV13 period and in most instances, isolates from infants and young children and those from older individuals consisted of the same genotypes. Similar to the baseline data presented in chapter 3, serotypes 1, 5, 6B, 7F, 8, 9V, 12F, 19F, 14 and 23F were predominantly globally disseminated clones, while serotypes 3, 4, 18C, 19A were mainly represented by genotypes rarely identified elsewhere.

The observed decrease in incidence of disease due to PCV serotypes during the PCV13 period is in agreement with reduction in incidence of disease due to PCV7 serotypes previously reported for 2012 compared to the pre-vaccine period, among South African children less than 2 years and adults 25-44 years age, (11). An increase in the proportions and number of non-PCV IPD cases, including serotype 12F, 15A and 35B, following routine vaccination with PCV has been reported by others (33;72;77-79). In this study, serotype 35B increased significantly among infants and young children while non-significant increases occurred in serotype 8, 12F, 15A, 16F and 17F. For older individuals a non-significant increase occurred in serotype 15A.

Since sequence type diversity for each of the serotypes did not change, increases appear to be due to expansion of genotypes that existed in the pre-PCV period. Olarte and co-authors (44) reported an increase in serotype 35B due to expansion of a pre-existing clone (CC558), in addition to emergence of CC156, previously associated with PCV serotypes. Later, Chochua et al. (80) described capsular switching between a PCV13-9V/CC156 donor strain and a 35B/CC558 recipient strain, resulting in 35B/CC156 progeny. In this study, the serotype 35B increase appears to be driven by expansion of CC172 which accounted for 13% and 56% of 35B isolates from infants and young children during the pre-PCV and PCV13 period, respectively. To our knowledge, the serotype 35B/CC172 combination is rare with Malawi being the only other country to have described this combination in two isolates (44;67). Like CC156, CC172 is associated with a number of PCV serotypes (1, 6A, 6B, 23F and 19A) (67). Among isolates in this study, CC172 was also common among several PCV serotypes (6A, 19A, 19F and 23F). For serotype 15A, two sequence types were detected in the PCV13 period that were not detected prior to PCV use. We cannot confirm the contribution of these genotypes to the 15A increase due to the limited number of serotype 15A isolates that were analyzed in this study. The decline in incidence of non-PCV serotype-disease among older individuals may be due to improved care for HIV-infected individuals in South Africa (81).

Similar to the findings of our baseline genetic study (75), PMEN clones were predominant among some serotypes. Sweden^{1}-217 (ST217), a major serotype 1 clone in Africa and Israel (82-84), represented the majority of our serotype 1 isolates. Among serotype 5 isolates, the widely distributed Columbia^{5}-19 clone (67;72;74;85) and associated single-locus variants was prevalent. The South African 6B PMEN clone [S. Africa^{6B}-8 (ST185)], rarely detected

elsewhere (30;67), was common among serotype 6B isolates. A few of serotype 6A isolates (n=5) isolates during the pre-PCV period were ST185. To our knowledge ST185 has not been described among 6A isolates (67). Two PMEN clones [Demark^{14}-32 (ST230) and Sweden^{15A}-25 (ST63)] were common among serotype 14 isolates. Demark^{14}-32 is commonly associated with serotype 14 (67) and Sweden^{15A}-25 with serotype 15A but has been detected among serotype 14 isolates (67;72-74). Other PMEN clones detected include Netherlands^{3}-31 (ST180) which represents the main serotype 3 clone in most parts of the world but is not the main clone in Africa (30;34;67;72;73). Netherlands^{8}-33 (ST53), a common serotype 8 clone in most places (73;85;86), was common among serotype 8 isolates. Spain^{9V}-3 (ST156) was the second most common genotype among serotype 9V isolates. Among serotype 23F isolates, Spain^{23F}-1 (ST81) and Taiwan^{23F}-15 (ST242) were common. ST81 and ST242 were also detected in low frequencies among our 6A and 6B isolates, respectively. Both these PMEN clones are commonly detected among serotype 23F isolates but have also been detected among serotype 6A and 6B (34;67;73;87). Denmark^{12F}-34 (ST218) was prevalent among our serotype 7F isolates. This serotype/ST combination has been rarely reported elsewhere with one or two isolates reported from the USA, Germany, Iceland and Saudi Arabia (30;67). Among our serotype 12F isolates, Denmark^{12F}-34 was detected in low frequency only in the pre-PCV period. Like most genotypes detected in more than one serotype in this study, CC218 circulated among serotype 7F isolates prior to routine PCV use and in the PCV13 period, it is therefore unlikely that the 7F/CC218 combination in our study is a consequence of vaccine pressure.

Similar to the baseline data presented in chapter 3, some serotypes were predominately represented by sequence types that are rare in other parts of the world. The majority of the

serotype 3 isolates during each of the three period belonged to CC458 which is rare in other parts of the world (34;67;73). The third most common serotype 3 clonal complex was CC230. All serotype 3 isolates belonging to this clonal complex were ST700, a double locus variant of Denmark ^{14}32 (ST230). ST700 has been reported exclusively in serotype 3 but only from Africa and Asia (30;34;67;73;84). In agreement with data presented in chapter 3, during each period the majority of 19A isolates were CC2062 (at least 80% during each period).

Capsular switching between serotype 19A and 19F has previously been documented (32). We identified CC2062, a common serotype 19A clone in South Africa, among our 19F isolates from the pre-PCV and the PCV13 period. Since CC2062 has been circulating within our serotype 19A isolates for nearly two decades (unpublished data); we believe serotype 19A may be the likely ancestral serotype. However, we cannot confirm if this is true since we did not genotype serotype 19F isolates obtained prior to 2005.

5 Chapter Five

Invasive disease caused simultaneously by dual serotypes of *Streptococcus pneumoniae*, South Africa

5.1 Introduction

Pneumococcal disease caused by more than one serotype is rarely reported (16-18). The Quellung reaction remains the gold standard method for pneumococcal serotyping but is reliant on bacterial isolation and requires serotyping of multiple colonies to detect multiple serotypes in a specimen. Molecular techniques, such as real-time PCR and microarray determine serotype directly from a clinical specimen and therefore have increased sensitivity to detect multiple serotypes present in a specimen (88-90).

Apart from Hiller *et al.* (42) who analysed six isolates belonging to two distinct pneumococcal strains obtained over seven months from a patient with chronic pneumococcal infection, we are aware of only two other studies that characterized pneumococcal isolates obtained simultaneously from the same disease episode (16;17). Using pulsed-field gel electrophoresis, both studies showed that the isolates had different serotypes and genotypes.

Through laboratory-based surveillance for IPD in South Africa from 2005 – 2014, we identified individuals co-infected with more than one pneumococcal serotype. In this study, we describe characteristics of these patients and identify factors associated with dual-serotype IPD. We

sequenced genomes of isolate pairs collected from these individuals to determine their molecular characteristics.

5.2 Material and methods

5.2.1 Identification of patients co-infected with more than one serotype (dual serotype IPD)

Cases reported for IPD surveillance from 2005 through 2014 were reviewed. A mixed culture was defined as the simultaneous identification of at least two serotypes (including non-serotypeable isolates) either from the same normally sterile-site specimen of an IPD patient or from two or more normally sterile-site specimens obtained from an IPD patient within 21 days of each other. Serotypes of serotypeable isolates and mixed cultures where isolates of different serotypes could not be separated were confirmed using real-time PCR (91).

5.2.2 Epidemiological data analysis

Univariate and multivariable analysis were performed with Stata version 14 (StataCorp Limited, College Station, United States). We assessed factors associated with dual serotype infection among patients with IPD from enhanced sites, where additional information such as HIV status was available, using unconditional logistic regression. Factors significant at $p<0.1$ on univariate analysis were evaluated for model inclusion and non-significant factors at $p>0.05$ were dropped from the multivariable model using stepwise forward selection. All two-way interactions in the final multivariable additive model were evaluated. Two-sided p-values <0.05 were considered significant throughout.

5.2.3 Genome sequencing and analysis

Paired-end reads quality checks, trimming, assembly, annotations of contiguous sequences and MLST were performed as described in chapter 2. Using phylogenetic similarity analysis,

genomes of paired isolates were compared to at least two selected genomes of invasive isolates from South Africa with corresponding serotype and sequence type, collected during the same period (2005-2014). Phylogenetic similarity analysis, based on the core genomes, was performed using the Rapid large-scale prokaryote pan genome analysis (Roary) (92) and the maximum likelihood trees were generated using RaxML version 8 (93). Furthermore, for each co-infecting pair of isolates, Roary was performed using these IPD isolate sequences (of the same serotype and sequence type) (94) to identify core and accessory genes.

5.3 Results

5.3.1 National surveillance

A total of 40,205 cases of IPD were reported from 2005 through 2014, of which 70% (28,213/40,205) had viable isolates. Serotype 1 (n=3,340), 19A (n=2,541), 6A (n=2,030), 14 (n=1,898), 23F (n=1,891), 6B (n=1,681), 4 (n=1,494), 19F (n=1,470), 12F (n=1,387) and 8 (n=1,109) were the most common serotypes identified. IPD co-infection with more than one serotype was identified in 35 (0.1%) patients with available viable isolates. In two of the patients, infections were caused by isogenic strains described in chapter 6. In this chapter we describe the remaining 33 cases and corresponding isolates.

5.3.2 Characteristics of patients and isolates causing dual serotype IPD

Two serotypes were identified in each of the 33 patients (Table 5.1) and in 31 patients both isolates (of different serotypes) were obtained from the same specimen. For the remaining two patients different serotypes were identified from two different specimen types (Table 5.1). Isolates were predominantly recovered from blood (20/33, 61%) followed by CSF (13/33, 39%). Serotype 6B (n=10), 14 (n=8), 6A (n=7), 19A (n=5), 19F (n=5) and 1 (n=4) were the most common serotypes identified. For 30 (91%) of the 33 patients at least one of the detected serotype in each pair was a PCV13 serotype (serotype 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F or 23F). For 21 (62%) patients both isolates were PCV13 serotypes, 9 (32%) patients had infection with one PCV13 and one non-PCV13 serotype, and for 3 (9%) patients both isolates were non-PCV13 serotypes. Serotype 6B/14 (n=3), 6A/14 (n=2) and 6A/19F (n=2) combinations were detected more than once, whereas all other cases (n=27) had unique serotype

pairs (Table 5.1). For the majority of patients (24/33, 73%), the co-infecting isolates had different antibiotic susceptibility profiles (Table 5.1). For the remaining nine patients the pairs had the same antibiograms, including a patient where the pair was susceptible to all tested antimicrobial agents, one patient where both isolates were non-susceptible to cotrimoxazole only and seven patients where the pairs of isolates were non-susceptible to penicillin and cotrimoxazole.

5.3.3 Factors associated with dual serotype IPD infection

IPD patients from enhanced sites, where additional clinical data were collected, represented 41% (11,701/28,213) of patients with viable isolates. Dual serotypes were identified in 0.2% (21/11,701) of patients from enhanced sites (Table 5.2). On multivariable analysis, dual serotype was associated with patients less than 5 years of age (adjusted odds ratio (aOR), 4.7; 95% CI, 1.8-11.7, underlying illness other than HIV (aOR, 2.8; 95% CI, 1.1-6.6) and death (aOR, 2.5; 95% CI, 1.08-6.09) (Table 5.2).

5.3.4 Genotypic characteristics of dual serotype IPD isolates

Paired isolates were available for whole genome sequencing in 20 of the 33 cases (n=40 isolates). For 13 patients one (n=10) or both isolates (n=3) isolates were no longer viable on subculture from long term storage. The genome coverage for the isolates ranged from 31X to 607X and assemblies contained contiguous sequences with total lengths ranging from 1.9 to 2.5 million base pairs.

Among each pair of isolates from the same patient none were identical by MLST (Table 5.1). No predominant genotype was identified among serotype 14 isolates (n=7 unrelated STs), however CC1094, CC185, CC2062 and CC347 were common among 6A (5/5), 6B (2/6), 19A (3/5) and 19F (4/6) isolates, respectively. Core genomes of all 40 isolates clustered with corresponding isolates of the same serotype and sequence type (Figure 5.1).

Among the 40 dual serotype isolates, 1,794 accessory genes were identified. A total of 701 (39%) of the 1,794 genes coded for known proteins and the remaining 1,093 coded for hypothetical proteins. *QueT* and *infC* (coding for queuosine precursor transporter QueT and translation initiation factor IF-3 respectively) were identified among 39 and 40 co-infecting isolates respectively, and were not found in other invasive isolates. Beside *queT* and *infC*, only two other genes (*ald 2* and *serB*) were identified among pairs of isolates from the same patient. *Ald 2* shared by the 6A and 12F pair (patient 25 in Table 5.1) was also identified among other 6A isolates (n=4) from IPD patients with dual serotypes. *SerB* was shared by a serotype 22F/33F pair from the same patient (patient 28 in Table 5.1). CHY zinc finger, streptococcal histidine triad protein and DNA alkylation repair enzyme (each encoded by more than one accessory gene) were common proteins, distributed among at least ten serotypes, respectively.

Table 5.1 Characteristics of patients and isolates causing invasive pneumococcal dual serotypes disease, South Africa, 2005-2014

Patient	Age (years)	HIV status	In hospital outcome	Source specimen (s)	Year	Serotype1	Antibiogram1	ST(CC) 1	Serotype2	Antibiogram2	ST(CC) 2
1	1	unknown	unknown	blood	2005	6A	P,COT	1094 (1094)	14	P,COT	7270
2	1	negative	recovered	blood	2005	6A	P,COT	1094 (1094)	19F	P,COT	347 (347)
3	1	positive	recovered	blood and CSF	2005	14	P,E,CD,T	NV	18C	S	NV
4	29	positive	recovered	CSF	2005	14	COT	10237	19A	P	2062 (2062)
5	1	positive	died	blood and CSF	2006	6B	P,COT	7738	19F	P,COT	10536 (347)
6	1	unknown	unknown	blood	2006	15B	P,COT	7052	19F	P,COT	6277
7	4	unknown	unknown	blood	2006	22	COT	5577	23F	COT	36
8	35	positive	recovered	blood	2006	9N	COT	NV	18A	NV	NV

9	85	unknown	died	CSF	2007	3	S	NV	35B	P,T,COT	NV
10	5	unknown	recovered	blood	2007	6A	P,E,CD,T,COT	1094 (1094)	14	P,C,T,COT	9901
11	1	positive	recovered	blood	2007	6A	P,COT	1094 (1094)	19F	P,COT	347 (347)
12	3	unknown	died	blood and CSF	2007	6A	S	NV	23F	P,E,CD,C,COT	NV
13	50	negative	recovered	blood	2007	8	S	NV	23F	P,E,T,COT	NV
14	34	unknown	unknown	CSF	2008	3	S	458 (458)	19A	P,E,COT	2062
15	28	unknown	unknown	blood	2008	6B	P,T,COT	10537	14	S	10239
16	0	positive	died	CSF	2008	6B	COT	138	14	P,T,COT	230
17	0	unknown	unknown	CSF	2008	9V	P,COT	3454	15A	S	10065
18	0	positive	recovered	blood and CSF	2008	19A	P,COT	2062 (2062)	22	S	445 (445)
19	0	negative	died	CSF	2009	1	S	NV	22A	COT	NV

20	1	unknown	died	blood	2009	6A	P,E,COT	NV	16F	S	NV
21	32	positive	died	blood	2009	6B	P,E,T,COT	6285	14	COT	2914
22	12	positive	recovered	CSF	2009	6B	P,COT	10066	19F	P,COT	8326
23	0	negative	recovered	CSF	2009	9V	S	1492	14	P,E,CD,T,COT	8782
24	30	unknown	unknown	blood	2010	1	S	NV	18C	P,COT	NV
25	0	positive	recovered	CSF	2010	6A	P,COT	1094 (1094)	12F	T,C,COT	989
26	0	unknown	unknown	CSF	2010	6B	P,COT	NV	18C	S	NV
27	37	unknown	unknown	blood	2010	6B	P,COT	185 (185)	19A	P,COT	6261
28	unknown	unknown	unknown	BPS	2010	22F	S	ST445 (445)	33F	S	10068
29	9	unknown	unknown	blood and CSF	2011	1	S	NV	6B	P,E,T,COT	NV
30	40	negative	died	blood	2011	1	S	9067 (217)	19A	P,COT	2062 (2062)

31	20	unknown	died	blood	2011	3	S	NV	6B	P,E,CD,T	NV
32	49	unknown	unknown	blood and pleural fluid	2011	6B	P,COT	NV	23F	P,E,T,COT	NV
33	1	negative	died	blood	2012	16F	COT	NV	34	S	NV

[1]Enhanced surveillance sentinel sites are hospitals located in all nine provinces where additional information, including HIV status and outcome are collected. ST, Sequence type, CC, clonal complex, CSF, cerebrospinal fluid; BPS, brain pus swab; NT, non-serotypeable; C, chloramphenicol; CD, clindamycin; COT, cotrimoxazole; E, Erythromycin; P, penicillin; T, tetracycline; S, susceptible to all tested antibiotics; NV, non-viable (for 13 cases isolates were non-viable on subculture for both isolates or one of the co-detected isolates).

Table 5.2 Univariate and multivariable analysis of factors associated with dual serotype infection among patients with invasive pneumococcal disease at enhanced sites, South Africa, 2005 – 2014

	No.(%) of cases			Univariate analysis		Multivariable analysis	
Characteristic	Total	Cases with single serotype	Cases with dual serotypes	Odds ratio (95% CI)	[a]p-value	Adjusted odds ratio (95% CI)	[a]p-value
Year	N=11701	N=11680	N=21				
2005	1583	1580 (14)	3 (14)	Reference	-		
2006	1521	1519 (13)	2 (10)	0.6 (0.1-4.1)	0.6		
2007	1416	1411 (12)	5 (24)	1.86 (0.4-7.8)	0.3		
2008	1418	1416 (12)	2 (10)	0.7 (0.7-0.1)	0.7		
2009	1458	1453 (12)	5 (24)	1.8 (0.4-7.5)	0.4		
2010	1173	1172 (10)	1 (5)	0.4 (0.04-4.3)	0.4		
2011	997 (9)	995 (9)	2 (10)	1.05 (0.1-6.3)	0.9		
2012	804 (7)	803 (7)	1 (5)	0.6 (0.06-6.3)	0.7		
2013	732 (6)	732 (6)	0				
2014	599 (5)	599 (5)	0				
Gender	N=11690	N=11669	N=21				
Male	5760	5746 (49)	14 (67)	Reference			
Female	5930	5923 (51)	7 (33)	0.5 (0.1-1.2)	0.1		
Age group (years)	N=11691	N=11670	N=21				
≥5	7976	7969(68)	7 (33)	Reference	-	Reference	-
<5	3715	3701 (32)	14 (67)	4.3 (1.7-10.6)	**0.002**	4.7 (1.8-11.7)	**0.001**

HIV status	N=9006	N=8990	N=16				
Negative	2472	2466 (27)	6 (38)	Reference			
Positive	6534	6524 (73)	10 (63)	0.6 (0.2-1.7)	0.3		
[b]Underlying illness other than HIV	N=11701	N=11680	N=21				
No	7942	7966 (68)	9 (43)	Reference	-	Reference	-
Yes	3759	3747 (32)	12 (57)	2.8 (1.1-6.7)	**0.01**	2.8 (1.1-6.6)	**0.01**
Outcome	N=11470	N=11449	N=21				
Recovered	8115	8104 (70)	11 (52)	Reference	-	Reference	-
Died	3355	3345 (29)	10 (48)	2.2 (0.9-5.1)	0.07	2.5 (1.08-6.09)	**0.03**
Specimen type	N=11701	N=11680	N=21				
Cerebrospinal fluid	3384	3376 (29)	8 (38)	Reference	-		
Blood	7645	7632 (65)	13 (62)	0.7 (0.2-1.7)	0.46		
Other	672 (5)	672 (6)	0	-	-	-	-
[c]Serotype	N=11701	N=11680	N=21				
non-PCV13	3585	3583 (31)	2 (10)	Reference	-		
PCV13	8116	8097 (69)	19 (90)	4.2 (0.9-18.05)	0.05		
[d]Penicillin susceptibility	N=11701	N=11680	N=21				
Susceptible	7333	7326 (63)	7 (33)	Reference	-		
Non-susceptible	4368	4354 (37)	14 (67)	3.3 (1.3-8.3)	**0.009**		
[d]Multidrug resistance	N=11701	N=11680	N=21				
No	9501	9487 (81)	14 (67)	Reference	-		
Yes	2200	2193 (19)	7 (33)	2.1 (0.8-5.3)	0.09		

[a]Significant p values ($p<0.05$) are in bold type

[b]Underlying illness includes: asthma, chronic lung disease, cirrhosis/liver failure, chronic renal failure, heart failure, heart failure, valvular heart disease, coronary heart disease, immunosuppressive therapy, splenectomy, diabetes, burns, kwashiorkor/marasmus, nephrotic syndrome, spinal cord injury, seizure disorder and emphysema or cancer.

[c]For individuals with dual serotypes the pair was classified as PCV13 if at least one of the isolates was a PCV13 serotype (serotype 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F or 23F) or non-PCV13 if both isolates were serotypes other than PCV13. [d]For individuals with dual serotypes, the pair was classified as penicillin non-susceptible and/or multidrug resistant if at least one of the isolates was penicillin non-susceptible and/or multidrug resistant.

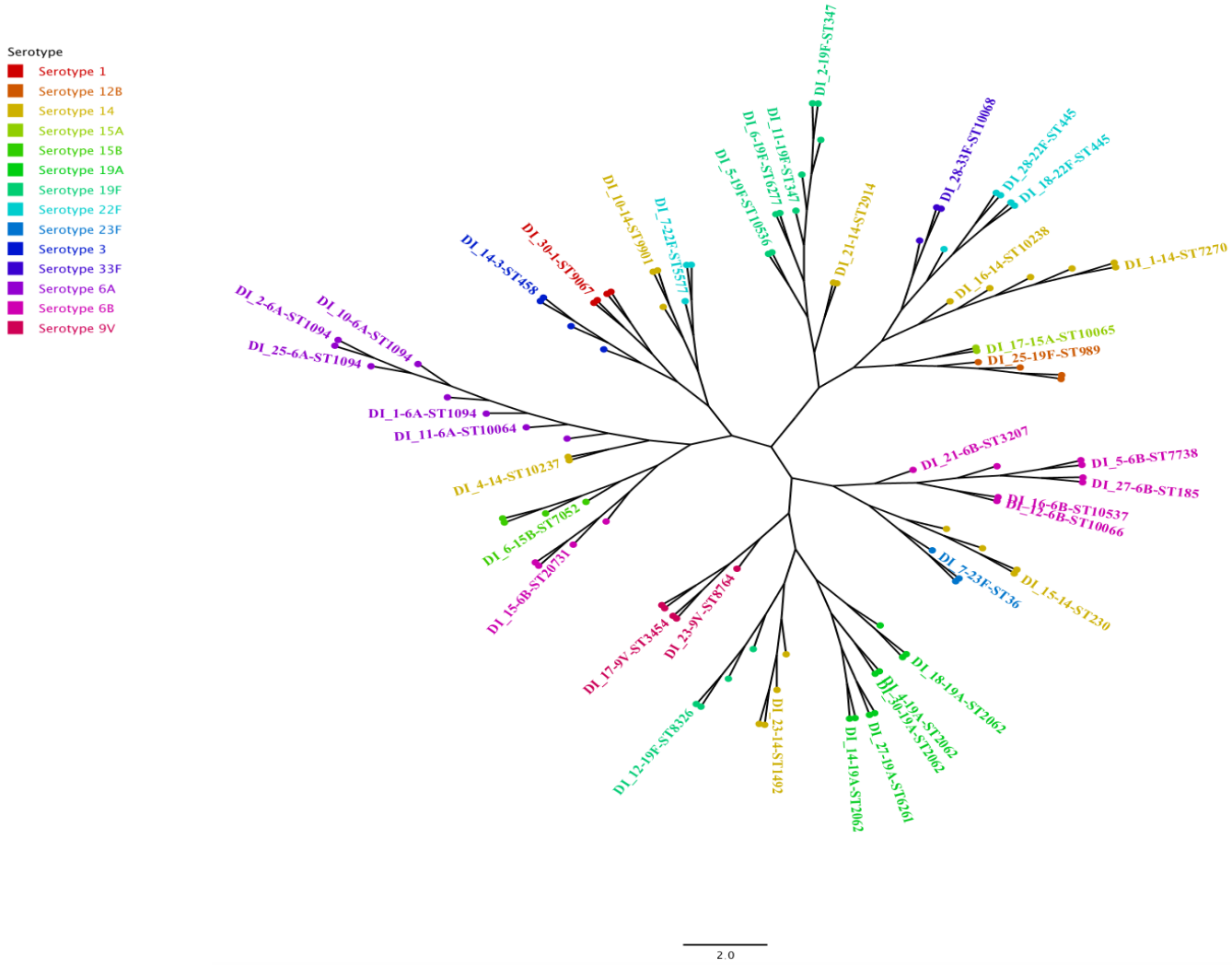


Figure 5.1 Phylogenetic comparison of core genomes of 40 isolates from South African patients with invasive pneumococcal disease caused by dual serotypes.

Genomes of isolates (n=59) expressing the same serotype and with the same sequence type (ST) as each of the 40 isolates, 2005-2014 are included for comparison. Colour coding is according to serotype. Only isolates from dual infections (DI) are labelled on the tree. The label includes DI/patient number, serotype and ST of the isolate, respectively.

5.4 Discussion

Cases of invasive pneumococcal disease caused simultaneously by two different serotypes have been reported previously by others, mainly in immuno-compromised patients or in young children (16-18). Similarly, in our study children <5 years of age and patients with underlying illness other than HIV were more likely to have dual serotype infection compared to older individuals and those without underlying illness, respectively. We identified increased odds of death among patients with dual serotype infections compared to those with single infections. The association between infection with dual serotypes and young children is probably a reflection of underdeveloped immune systems and the higher carriage rates of the pneumococcus in children, with increased chances of invasive disease by more than one serotype in children carrying multiple serotypes (95).

This study has some limitations. Dual serotypes were not verified from original clinical specimens as primary isolation is carried out at the diagnostic laboratories, following which, isolates are submitted to the NICD for surveillance purposes. However, given the rarity of dual serotypes in IPD, such cases were investigated carefully and several measures were taken to rule out laboratory error. The detection of multiple serotypes from the same clinical specimen has been shown to be improved by detecting directly from specimens rather than from culture plates (96). As such, some cases may have been missed because serotype was not identified directly from specimens. Additionally, serotypes (particularly those present in low copy numbers in original specimens), may have been missed because more sensitive methods such as molecular techniques were not used for serotyping.

The most common serotypes identified among patients with dual infection were serotypes included in PCV13 (1, 6A, 6B, 14, 19A and19F). This is in line with the overall IPD serotype prevalence prior to routine use of PCV13 in South Africa (11;12;97) during which the majority of these cases occurred, where over 80% of IPD cases were due to PCV13 serotypes (11). It is however interesting that the majority of co-infecting serotypes (6A, 6B, 19F) identified in our study are known to have low invasive disease potential (98). The significant association between dual infection and underlying illness could explain the vulnerability of these patients to invasive disease with two serotypes.

Similar to dual serotype IPD cases reported previously where co-infecting isolates from the same patient were genotypically unrelated (16;17), the majority of our isolate pairs were different strains as indicated by unrelated sequence types. The core genome of each isolate clustered with genomes of other South African IPD isolates of the same serotype and sequence type. Genotypes identified were common genotypes among those serotypes circulating in South Africa prior to routine use of PCV (75).

A study in The Gambia (99), with a high prevalence (19%) of multiple serotype carriage, reported evidence of recombination in 34% (49/144) of single serotype and co-carried pneumococcal isolates. Two accessory genes identified in all genomes (*queT* and *infC*) and two from genomes of two individual pairs (*ald2* and *serB*) were the only accessory genes common between pairs of isolates from the same IPD cases. The lack of sequence similarity between the majority of accessory genes is evidence that during each episode of invasive disease in this study, there was little or no genetic exchange between co-infecting pairs of isolates. For these

patients, the isolates could have invaded too soon after acquisition, giving limited time for co-carriage and recombination.

6 Chapter Six

Two cases of serotypeable and non-serotypeable variants of *Streptococcus pneumoniae* detected simultaneously during invasive disease

6.1 Introduction

Reduced levels of capsule expression are required for effective attachment to epithelial cells and colonization (100). Thus NT pneumococci are commonly isolated during carriage studies and very rarely cause invasive disease (16;17).

While reviewing national surveillance data to identify episodes of mixed IPD caused by two or more serotypes (described in chapter 5), we identified two cases (subsequently referred to as case A and case B) whereby one of the co-detected isolates was non-serotypeable. In each of the two cases, the serotypeable and the NT variant was identified from a culture recovered from the same specimen (cerebrospinal fluid for case A and blood for case B). Given the rarity of NT isolates in causing IPD we hypothesized that these NT isolates were variants of their co-detected serotypeable isolates. We therefore compared the genomes of these co-detected NT and serotypeable isolates to determine their relationships.

6.2 Materials and methods

6.2.1 Identification of cases with mixed cultures of a non-serotypeable and a serotypeable isolate

We reviewed cases and used isolates collected for IPD surveillance in South Africa from 2005 through 2014. A case was defined and identified as described in chapter 5. In this chapter, isolates from cases where at least one of the co-detected isolates was non-serotypeable were characterized. Such cases were detected during routine serotyping by Quellung if an isolate reacted partially with a specific antiserum pool. In such cases, the stored isolate and Dorset transport medium were sub-cultured for single colonies on 5% horse blood agar plates (DMP). Following 24 hour incubation at 37°C in 5% CO_2, the plates were examined for differences in colony morphology and the Quellung reaction was repeated on different colony types. The colonies were sub-cultured, re-identified and serotyped to determine if they were true mixtures. Real-time PCR, targeting the *lytA* gene, was used to confirm non-serotypeable isolates as *S. pneumoniae* (49). Non-serotypeable isolates were defined as pneumococcal isolates confirmed by *lytA* real-time PCR and for which a serotype could not be assigned by the Quellung method.

6.2.2 Transmission electron microscopy (TEM)

TEM was conducted to visualize the presence of capsular material. Isolates from overnight cultures were fixed *in situ* on agar plates, and processed according to the protocol developed by Hammerschmidt *et al.* with ruthenium red and L-lysine acetate fixation (101). In addition, three clinical isolates with confirmed serotyping results (one serotype 1 and two NT) were selected from the GERMS-SA collection and were included as controls for TEM analysis.

6.2.3 Genome analysis

Whole genome sequencing was performed at the NICD core sequencing facility as described in chapter 2. Paired-end reads quality checks, trimming, assembly, and annotations of contiguous sequences were performed as described in chapter 2. In addition, sequenced reads were mapped against references [GenBank: CR931632 and CR931673] for capsular regions of co-detected serotypes.

Variations within the capsular polysaccharide biosynthesis genes were detected using the fixed ploidy variant detection tool within the CLC Genomics Workbench. This model detects the variants whose representation in the reads is in accordance with the assumed ploidy, discards variants whose representation in the reads is likely due to sequencing errors or mapping artifacts and then reports on positions where there may be single nucleotide variation (SNV) including insertions, deletions and substitutions. Identified variations were further investigated using CLC Genomic workbench to identify amino acid changes.

MLST analysis of bacterial genomes was performed as described in chapter 2. Furthermore, contiguous sequences were uploaded to the *S. pneumoniae* PubMLST isolates database (94), which runs the Bacterial Isolate Genome Sequence database (BIGSdb) platform to facilitate the ribosomal MLST (rMLST) analysis (102).

Core genome maximum likelihood trees were generated as described in chapter 5. The genomes of case A and case B isolates were compared to serotype 1 (n=54) and 18C (n=58) genomes randomly selected from sequences available on *S. pneumoniae* PubMLST database (67). The selected sequences were of serotype 1 and 18C isolates that caused IPD in South Africa from 1989 to 2013 and 2005 to 2013, respectively. For the analysis, core genomes consisted of genes

that were present as single copies in at least 90% of isolates being analyzed and, in addition, had the same sequence length in every isolate. The BLAST Ring Image Generator (BRIG) was used to visualize similarities between genomes of mixed culture isolates (103).

6.3 Results

In two of the 35 cases of mixed cultures with more than one serotype detected (described in chapter 5), one of the co-detected isolates was NT. These cases were reported in 2009 (case A) and 2010 (case B), respectively and the age range for both was 10-14 year-old (Table 6.1). These two mixed culture cases were identified during routine serotyping of cultures recovered from cerebrospinal fluid (case A) and blood (case B). Case A was diagnosed with meningitis and clinical information was not available for case B to confirm diagnosis.

For case A, serotype 1 and NT *S. pneumoniae* were identified from the same culture, with the NT isolate representing approximately 98% of the culture (using the Quellung reaction). The case B isolates were identified as NT and 18C. In each of the two cases, the same results were obtained by two different laboratory staff performing the Quellung reaction on fresh overnight cultures grown from the original transport medium. After several attempts to separate the two variants, only the NT isolate could be obtained for case A, however, real-time PCR confirmed the presence of the serotype 1 *wzy* gene in the mixed culture. For case B, pure cultures were obtained for both the 18C and the NT isolate. The three isolates were susceptible to all tested antimicrobial agents.

TEM confirmed the absence of capsular material for the case A non-serotypeable isolate (Figure 6.1). For case B, 50 to 120 nm thick capsular material was observed around cells of the serotype 18C isolate, while there was no capsular material identified for the NT isolate (Figure 6.1).

The genome coverage for the case A NT isolate was 320X and assemblies contained 33 contiguous sequences with a total length of 2,023,465 bp. For the case B NT isolate, the coverage was 355X and assemblies contained 91 contiguous sequences with a total length of 2,110,682 bp. The case B serotype 18C isolate genome coverage was 310X and assemblies contained 78 contiguous sequences with total length of 2,115,055 bp.

Sequence analysis of the capsular locus of the case A NT isolate revealed the absence of all but four capsular genes [rhamnose biosynthesis genes (*rmlA, B, C* and *D*)] (Figure 6.2). The isolate was identified as ST217 and rST3462. Because we did not obtain a pure culture for the case A serotype 1 isolate, we compared the case A NT genome sequence to genomes (n=54) of ST217 serotype 1 isolates that caused invasive disease in South Africa, available on *S. pneumoniae* PubMLST database (94) to further confirm that this NT isolate was indeed a serotype 1 variant. Except for the absence of genes in the capsular region; the NT core genome was 100% similar (in 100% of the bootstrap replications) to sequences of a serotype 1 isolate with the same rMLST profile (rST3462). It also clustered with ST217 clinical isolates collected from 1989 through 2013 (Figure 6.3).

For case B, the NT and 18C isolates shared the same new ST (ST9817) and rMLST profile. The NT isolate had all serotype 18C capsule-specific genes (2), however variations were identified

in the capsular locus compared to the co-detected 18C isolate and three other genotypically related invasive 18C isolates from South Africa. A nucleotide insertion (472_473 insG) that resulted in a translational frame shift, introducing a stop codon in the *wchA* gene, was identified. A substitution (149C→A) in the *wze* gene resulted in an amino acid change (T50K). Genome sequence comparison of the case B isolates revealed high similarity between the two genomes with 100% identity between their core genomes.

Table 6.1 Characteristics of two patients (case A and B) with invasive *Streptococcus pneumoniae* serotypeable and non-serotypeable co-disease, South Africa

	Case A	Case B
Age range in years	10-14	10-14
Gender	male	male
HIV status	unknown	unknown
Specimen	cerebrospinal fluid	blood
Year identified	2009	2010
Diagnosis	meningitis	unknown
Co-detected serotypes	non-serotypeable and serotype 1	non-serotypeable and serotype 18C
Isolated serotypes	[1] non-serotypeable	non-serotypeable and 18C

[1] Pure culture for serotype 1 could not be obtained due to its low representation in the mixed culture (see results for more information).

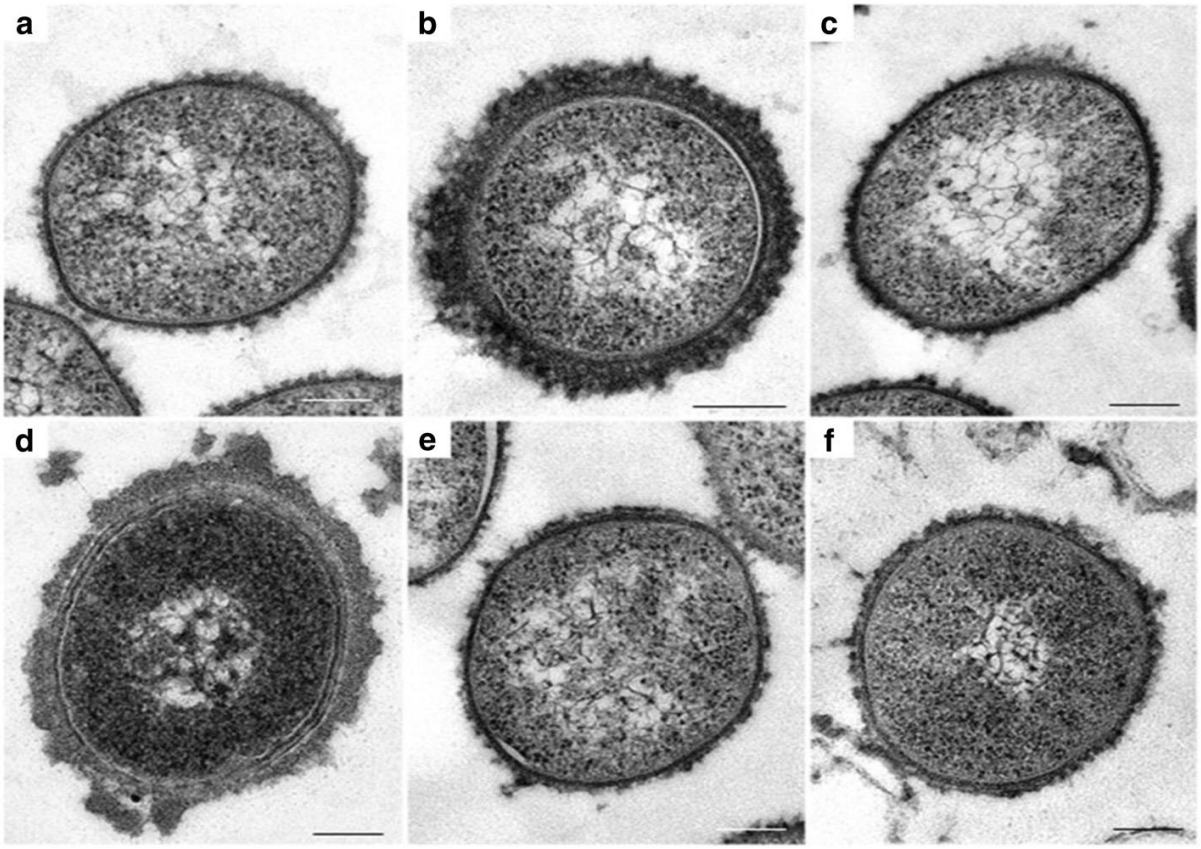


Figure 6.1 Visualization of pneumococcal isolates using transmission electron microscopy (TEM).

Capsular materials of non-serotypeable pneumococcal isolates causing mixed infections in two patients in South Africa were compared to capsular materials of serotypeable isolates. The two cases were reported in 2009 (case A) and 2010 (case B). For case A non-serotypeable and a serotype 1 isolate were identified and for case B a non-serotypeable and 18C isolates were identified. TEM of the case A non-serotypeable isolate is shown in a, TEM of serotype 1 clinical isolate used as a control in b, TEMs of two non-serotypeable clinical isolates used as a controls in c and f, TEM of case B serotype 18C isolate in d, and case B non-serotypeable isolate in e. Scale bar = 175 nm.

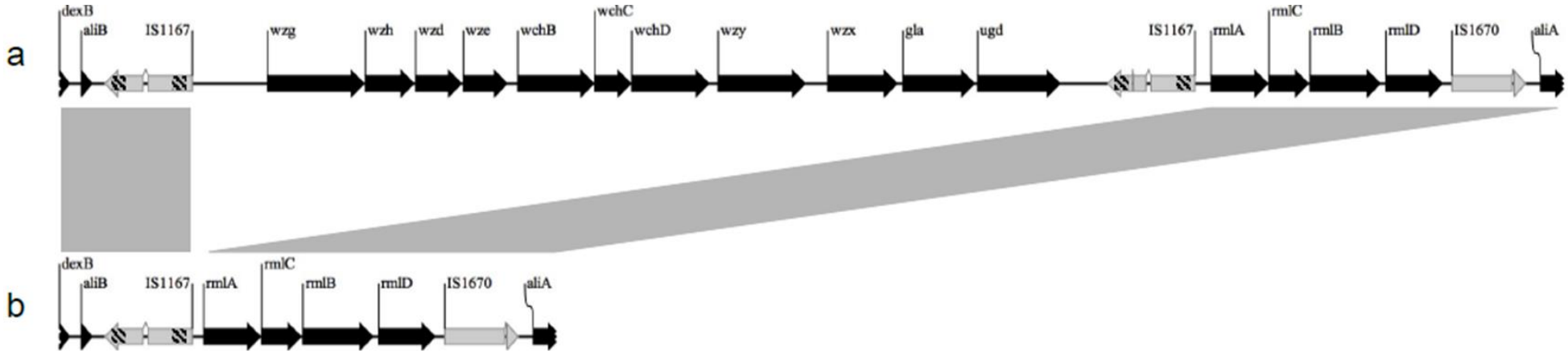


Figure 6.2 Schematic diagram representing capsular polysaccharide loci of a pneumococcal non-serotypeable and serotype 1 isolate.

The schematic diagram represents comparison between the capsular polysaccharide (cps) locus of a non-serotypeable isolate (B) co-detected with a serotype 1 isolate during an invasive disease episode in South Africa, 2009, and the cps locus of an invasive serotype 1 clinical isolate from South Africa (A). Transposases are indicated by grey arrows and flanking repeated sequences by lines within the arrow delimitation.

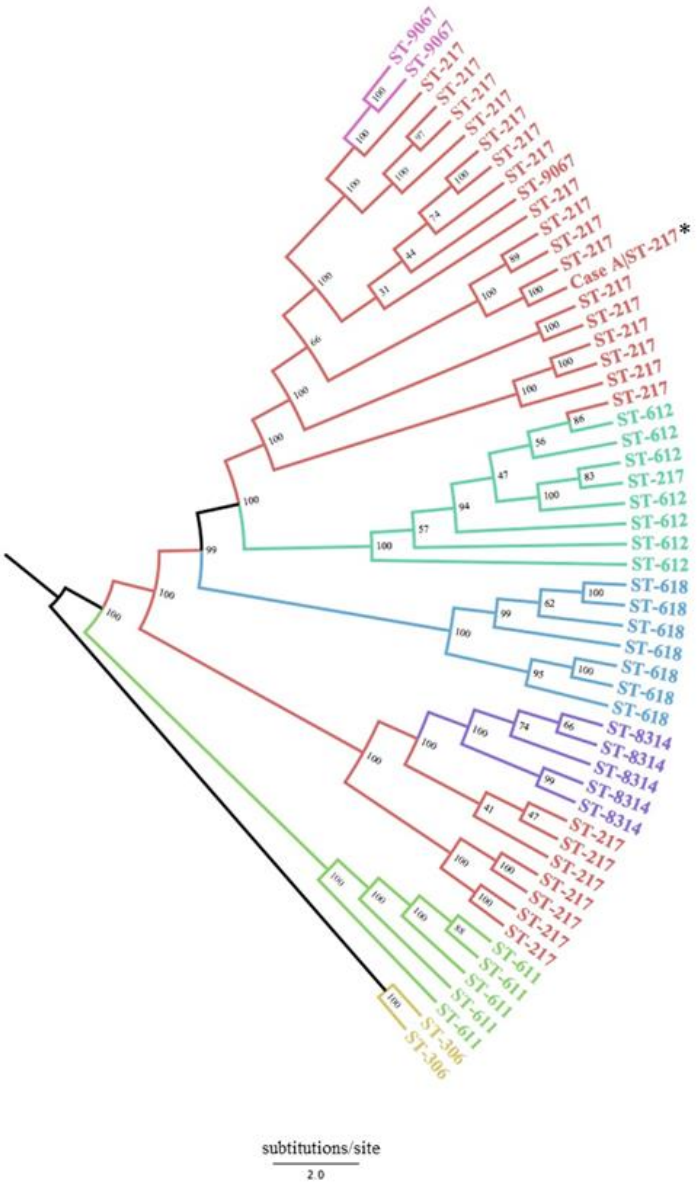


Figure 6.3 Core genome phylogenetic tree of Case A non-serotypeable isolate and serotype 1 isolates from South Africa.

The tree represents phylogenetic comparison of the core genome of a non-serotypeable co-detected with a serotype 1 isolate in 2009 in South Africa and genomes of serotype 1 isolates (n=54) that caused invasive disease in South Africa from 1989 to 2013. This maximum likelihood tree was built using core gene sequences which were concatenated and aligned, using MUSCLE (EMBL-EBI). The numbers at the nodes indicate RAxML bootstrap values. Branch lengths are proportional to the number of substitutions per site (see the scale bars). The color coding is according to sequence type (ST).

6.4 Discussion

The polysaccharide capsule is an important component of the pneumococcus that aids in evading the immune system. Hence un-encapsulated (non-serotypeable) pneumococci rarely cause invasive disease. Nevertheless, we describe two cases of invasive pneumococcal disease caused simultaneously by what appears to be serotypeable and non-serotypeable variants of the same strain. The hypothesis, that the NT isolates were variants of their encapsulated counterparts that either lost their capsules or had reduced levels of capsular expression during the infection, is supported by the lack of capsule for both NT isolates from the two cases. Both NT isolates were sequence types of major clones circulating among encapsulated isolates in South Africa (75). One NT isolate was ST217, the dominant serotype 1 clone in South Africa (75). The second isolate was a new sequence type (ST9817) that belonged to the ST1016 clonal complex, a predominant clonal complex among serotype 18C isolates in South Africa (75).

Since the NT isolates from both cases were sequence types commonly associated with serotype 1 and 18C, we aligned their *cps* sequences to that of serotype 1 and 18C, respectively. The non-serotypability of the ST217 isolate appears to be caused by the deletion of genes within the *cps* region. This mechanism is common among NT isolates with a serotype 1 genetic background. Scott *et al.* (104), analyzed 28 NT isolates from IPD patients and identified eight isolates with partial deletion of the capsular locus. Four of the eight isolates were a sequence type associated with serotype 1 (ST227) and, similar to our ST217 isolate, only had the rhamnose biosynthesis genes (*rmlA, B, C* and *D*) in their capsular region. Similarly, four IPD NT isolates analyzed by Salter *et al.* (105) that were ST227 only had *rmlA, B, C* and *D* present in their capsular locus. Five additional ST217 NT isolates that were identified among our collection of IPD surveillance

isolates and were analyzed as part of another study also had only the four rhamnose biosynthesis genes present in the capsular locus (106). The deletion of the *cps* genes in our ST217 non-serotypeable isolate could be transposon mediated as the deleted serotype 1 genes are flanked by transposase genes (IS1167). We were unable to isolate the encapsulated variant (serotype 1) from this case due to its low representation (approximately 2%) in the mixed culture. Nevertheless, detection of the serotype 1 *wzy* gene (absent in the genome sequence of the NT isolate) in the mixed culture by real-time PCR supports the presence of serotype 1 in that culture. Our genotypic results also confirm that the NT isolate is a serotype 1 variant as both the sequence type and ribosomal MLST profile are exclusively associated with serotype 1 (ST217 and rST3462). In addition, phylogenetic analysis showed that its core genome clustered with those of other invasive ST217 serotype 1 isolates circulating in South Africa during the same period.

In the second case, the NT isolate had an intact *cps* locus with all serotype 18C capsular genes with the exception of two capsular genes that were variable, compared to the serotype 18C isolate from the same patient. Variation in the *wze* gene resulted in an amino acid change within the protein. Together with *wzd, wze* encodes proteins responsible for the export of the mature polysaccharide to the cell surface (2). The *wze* gene encodes an auto-phosphorylating protein-tyrosine kinase which is inhibited by aerobic microenvironments, and is thought to be one of the factors causing decreased capsular expression in host environments with higher oxygen levels such as the nasopharynx (100;107). Previous studies have demonstrated that mutations in the *wze* gene can inhibit capsule production (100;107;108). An insertion that introduced a premature stop codon in the *wchA* gene was identified in our NT isolate. We suspect that this

insertion may be responsible for the loss of capsule expression in our isolate as *wchA* encodes the initial transferase that catalyzes the transfer of glucose-phosphate to the lipid carrier during capsular polysaccharide biosynthesis (109). Melchior *et al* (110) described two mutations in the *wchA* gene (a nucleotide deletion at position 910 and a substitution at position 1081) in two invasive non-serotypeable isolates with serotype 7F capsular genes both of which introduced premature stop codons in the *wchA* gene. Although the positions of these mutations differed from that of our isolate, both resulted in a premature stop codon. In another study, Schaffner *et al.* (7) recovered a NT isolate, with serotype 18C capsular locus, from the nasopharynx of a child with otitis media. The variations identified in our NT isolate with serotype 18C capsular genes differ from the *wchA* single base substitution (C1135G) described in the Schaffner study.

Repeated sub-culturing in the laboratory may lead to a change in phenotype (111) or the cultures may have become mixed as a result of laboratory error. It is unlikely that this occurred in this instance given the rarity of these cases in our laboratory and the correlating genotypic data between the case B variants. It is favorable for the pneumococcus to have reduced levels of capsule for effective attachment to epithelial cells in the nasopharynx however the capsule is required for survival of the pneumococcus during invasive disease as it provides protection against complement-mediated phagocytosis (100). We therefore suspect that both variants may have been present in the nasopharynx and invaded simultaneously. Unfortunately we don't have corresponding nasopharyngeal isolates from these patients to confirm this hypothesis. Another possibility is that the serotypeable variant could have invaded and mutated in the host to lose the capsule. Survival of the NT isolates in the blood and meninges without protection of the capsule is highly irregular but may be related to host factors such as suppressed immune status

due to underlying diseases such as HIV infection (112). Unfortunately we did not have all the relevant patient information to assess if this was true.

7 Chapter Seven

Conclusion

Epidemiological studies of pneumococcal disease are necessary to determine the impact of routine use of serotype-specific vaccines such as PCV7 and PCV13. Additionally, strain characterization gives an indication of how the pneumococcal population may be affected by such an intervention, and whether the intervention creates a niche for other serotypes and genotypes to emerge. Following introduction of the pneumococcal conjugate vaccine, some countries reported increases in non-PCV serotypes. Although cases of capsular switching have been reported, in most cases increases were driven by pre-existing clones that were circulating prior to PCV introduction.

Although a lot of research has been done on molecular epidemiology of the pneumococcus in Europe and the USA, little is known about pneumococcal genotypes circulating in South Africa and other African countries in the era of pneumococcal conjugate vaccines. From the limited data available, it appears as if some predominant sequence types circulating among isolates from the African countries are different to major sequence types circulating in other parts of the world (72;84). In this study, predominant genotypes that appear to be unique to South Africa were identified. This study provides additional data on these unique genotypes of the pneumococcus and adds new information on the spread of these genotypes. In addition, unique genotypes that have also been reported in other African countries (e.g. ST700) were identified. The uniqueness of our isolates emphasizes the need to characterize isolates from South Africa

and other African countries as the impact of PCV on the molecular epidemiology may be different to what has been experienced elsewhere. Some of the serotypes in this study were predominantly represented by globally disseminated clones. As expected, the majority of these PMEN clones were non-susceptible to penicillin or were multidrug resistant.

Pneumococcal disease incidence is usually higher in children compared to adults and disease is usually transmitted from children to adults. It is therefore not surprising that in this study, genotypes circulating among isolates from children and adults did not differ. For most serotypes, predominant genotypes circulating before and during PCV use did not differ. It is possible that our PCV13 period (2011-2013) was too soon after PCV13 introduction (2011) for changes to be detected. Additionally, due to high numbers, we could not characterize all isolates, particularly non-PCV isolates where vaccine escape variants are likely to occur. Nonetheless our data strongly suggest that the post PCV genotype distribution is predictable based on the pre-PCV genotypes circulating in South Africa.

In this study, we identified MLST genotypes that were shared by more than one serotype, indicating possible capsular switching between some isolates. Similar to findings from elsewhere, these capsular switches were identified before PCV was routinely used in South Africa, suggesting that PCV pressure did not initiate capsular switching events in these isolates. Nevertheless, vaccine pressure could favor capsular switches to non-PCV serotypes and hence may contribute to increases of these strains. This may have been the case with penicillin non-susceptible serotype 35B, which significantly increased during the PCV13 period compared to the pre-PCV period in South Africa. The increase in our serotype 35B isolates appears to be due

to expansion of CC172, usually associated with antimicrobial non-susceptible PCV serotypes 6A, 6B, 19A and 23F. Genome comparison of 35B and CC172 isolates expressing other capsular types may provide some information on the origin and evolution of the 35B/CC172 combination.

We identified co-infecting dual serotypes of predominantly vaccine serotype isolates that were not genetically related. The findings of this study suggest that IPD due to dual serotypes is rare at least when using conventional culture for detection. Nevertheless, it is important to monitor such infections as they are more likely to occur in children probably because of their immature immunity and in patients with underlying illness because of impaired immune response. Additionally, compared to patients with a single serotype, these patients were more likely to die; the presence of two serotypes potentially exacerbated the disease. If IPD cases with more than one serotype become more important or prevalent in the future, more sensitive methods should be used routinely in laboratories to increase the ability to detect such cases.

In this study whole genome data allowed a more detailed analysis of the relationship between co-infecting isolates. Two different vaccine-serotype isolates that appear to have lost their capsules during two separate episodes of invasive disease were identified. Non-serotypeable pneumococci are not targeted by current pneumococcal vaccines and thus present a potential mechanism whereby the organism could escape vaccine pressure. Unencapsulated isolates have higher recombination efficiency than encapsulated pneumococci and therefore provide a reservoir of antibiotic resistance genes for genetic exchange (113). Antibiotics coupled with routine PCV use may drive the selection of non-serotypeable variants and thus prevalence and

genotypes of such pneumococci should be monitored to identify capsular switching or other genetic exchange that may occur under vaccine pressure.

The biggest limitation of PCV is its inability to protect against pneumococcal disease caused by non-vaccine serotypes and hence the possibility of replacement disease. Alternative pneumococcal vaccines that provide broad protection against as many serotypes as possible are needed. Pneumococcal components that are being considered as vaccine candidates include protein antigens (including IgA binding proteins, factor H inhibitor of complement and pneumolysin), whole-cell (killed cells of unencapsulated pneumococci) and DNA vaccines (114). Nonetheless, polysaccharide conjugate vaccines are currently the only licenced pneumococcal vaccines in children and their use has resulted in significant reductions in disease not only in vaccinated children, but also in unvaccinated children and adults through herd protection. Their use may however result in the emergence of new, or more often, according to our data, the expansion of existing clones.

8 Appendices

Appendix A

Invasive pneumococcal disease isolates with genotyping, by year of specimen collection and age group, isolates were obtained during the Pre-PCV period, PCV7 period and PCV13 period in South Africa

	Pre-PCV period (2005-2008)	PCV7 period (2009-2010)	PCV13 period (2011-2013)	2005-2013
Number of cases [all ages]	18330	8659	9279	36267
Viable isolates [all ages], n/N (%)	13159/18330 (72)	6062/8659 (70)	6287/9279 (68)	25508/36267 (70)
Number of viable isolates [children <5 years]	4480	1659	1140	7279
Isolates with genotypic results [children <5 years], n/N (%)	1048/4480 (23)	339/1659 (20)	674/1140 (59)	2061/7279 (28)
Number of viable isolates [individuals ≥5 years], N	8679	4403	5147	18229
Isolates with genotypic results [individuals ≥5 years], n/N (%)	1155/8679 (13)	172/4403 (4)	1650/5147 (32)	2977/18229 (16)
Total isolates with genotypic results [all ages], n/N (%)	2203/13159 (17)	511/6062 (8)	2324/6287 (37)	5038/25508 (20)

Appendix B

Clonal complex and sequence type (ST) distribution of pneumococcal isolates causing invasive disease among infants and children <5 years of age in South Africa during pre-PCV period (2005-2008), PCV7 period (2009-2010) and PCV13 period (2011-2013)

Serotype	No. of isolates: pre-PCV/PCV7/PCV13 period	No. with genotypic results: pre-PCV/PCV7/PCV13 period	Clonal complex (ST)	pre-PCV period n/N (%)	PCV7 period n/N (%)	PCV13 period n/N (%)
1	218/105/86	90/30/59	217 (217, 612, 8314)	90/90 (100)	30/30 (100)	57/59 (97)
			Singleton (306)	0	0	2/59 (3)
3	47/21/26	14/3/19	458 (458, 8772)	11/14 (79)	2/3 (67)	10/19 (53)
			180 (180)	1/14 (7)	1/3 (33)	4/19 (21)
			230 (700), 271, 378	2/14 (14)	0	1/19 (5)
			Singletons (1116, 1220, 1765)	0	0	4/19 (21)
4	91/39/13	28/12/12	5410 (5410, 5655, 6367, 6367, 6368, 6369)	19/28 (68)	5/12 (42)	4/12 (33)
			1221	6/28 (21)	6/12 (50)	8/12 (67)

			Singletons (10757, 2290, 7727)	3/28 (11)	1/12 (8)	0
5	54/30/36	13/3/31	5659 (289, 5659, 7050, 7745)	13/13 (100)	3/3 (100)	30/31 (97)
			Singleton (8719)	0	0	1/31 (3)
6A	147/204/84	105/47/64	2285 (n=15 STs)	44/105 (42)	14/47 (30)	30/64 (47)
			2289 (1094, 2289, 6371, 6383, 6387)	31/105 (30)	16/47 (34)	12/64 (19)
			1094 (4941)	10/105 (10)	4/47 (9)	4/64 (6)
			172 (1447, 6379)	5/105 (5)	6/47 (12)	4/64 (6)
			6389 (6376, 6389, 7740, 7741)	8/105(8)	3/47 (6)	1/64 (2)
			473 (471, 473)	2/105 (2)	1/47 (2)	6/64 (9)
			81, 185	1/105 (1)	0	1/64 (2)
			Singletons (5412, 7313, 10582, 10713, 10755, 6384, 6391, 7131, 7313, 7743, 8714, 8718, 8776)	4/105 (4)	3/47 (6)	6/64 (9)
6B	607/204/55	136/29/32	2421 (2421, 2907, 4929, 6290, 6290, 7724, 7731, 8391, 8768)	45/136 (33)	6/29 (21)	6/32 (19)
			185 (185, 6235, 6293, 6296, 6297, 6526, 6529, 7730)	34/136 (25)	3/29 (10)	14/32 (44)
			6389 (6376, 6389, 7740, 7741)	8/136 (6)	2/29 (10)	1/32 (3)

			7084	4/136 (3)	4/29 (14)	0
			6279 (6279, 6288, 6303, 6305)	8/136 (6)	0	1/32 (3)
			2285 (2285, 6289)	4/136 (3)	0	1/32 (3)
			2289 (2289, 1094, 6298)	4/136 (3)	0	2/32 (9)
			6393 (6301, 6393, 8767)	1/136 (1)	2/29 (7)	2/32 (6)
			6239, 172, 230	4/136 (4)	0	0
			Singletons (386, 7733, 8713, 8776, 10720, 10726, 10742, 10743, 10751, 10760, 10770, 10772, 10791, 10815, 2156, 3207, 3473, 386, 4368, 6292, 6294, 6295, 6300, 6302, 6308, 6525, 7313, 8715, 90, 9600, 9923)	24/136 (18)	8/29 (30)	7/32 (22)
6C	12/2/3	3/1/3	2185 (6310)	2/3 (77)	1/1 (100)	1/3 (33)
			2283 (6311)	1/3 (33)	0	1/3 (33)
			Singletons (1092)	0	0	1/3 (33)
7F	10/5/5	3/4/3	3544 (218, 3544)	2/3 (67)	3/4 (75)	2/3 (67)
			4956 (8838)	0	0	1/3 (33)
			Singletons (10792, 5027)	1/3 (33)	1/4 (25)	0

9V	140/40/13	39/12/12	4881 (3454, 4881, 6314, 6316, 6530, 7889)	28/39 (72)	7/12 (58)	7/12 (58)
			156 (162, 559, 644)	4/39 (10)	2/12 (17)	2/12 (17)
			5408 (5408, 6313)	3/39 (7)	1/12 (8)	2/12 (17)
			4084, 4902	2/39 (5)	0	1/12 (8)
			Singletons (10753, 10797, 2280, 6317)	2/40 (5)	2/12 (17)	0
14	700/223/43	150/50/39	230 (1701, 230, 2707, 6229, 6234)	77/150 (51)	22/50 (44)	19/39 (49)
			2652 (1492, 2652, 6230, 7726, 7728, 7729, 8370, 8785)	39/150 (25)	17/50 (34)	8/39 (21)
			63 (2414, 5004, 5187, 63)	25/150 (17)	8/50 (16)	5/39 (13)
			156, 271, 2062, 4902, 6279, 10588	3/150 (2)	2/50 (4)	3/39 (8)
			Singletons (6741, 8374, 9901, 10714, 124, 6226, 6228, 6533, 8371, 9901)	6/150 (4)	3/150 (6)	4/39 (10)
18C	125/35/26	50/7/17	6239 (6239, 6243, 6249, 8372)	13/50 (26)	3/7 (43)	7/17 (41)
			6245 (1016, 102, 6245)	13/50 (26)	0	6/17 (35)
			4893 (4893, 6241, 6242, 6246, 7723)	12/50 (24)	2/7 (29)	4/17 (24)
			6236 (6236)	5/50 (10)	2/7 (29)	0

			Singletons (10783, 6238, 6244, 6250, 6251, 6252, 8840)	7/50 (13)	0	0
19A	350/153/103	111/35/80	2062 (n=17 STs)	97/111 (87)	32/35 (91)	77/80 (96)
			172 (6265, 7737)	3/111 (3)	3/35 (6)	1/80 (1)
			199, 230, 271	2/111 (2)	0	1/80 (1)
			Singletons (416, 10690, 10703, 10744, 10790, 6095, 6257, 6267, 7734, 7739, 9913)	9/111 (8)	1/35 (3)	1/80 (1)
19F	440/155/59	79/36/36	347 (347)	25/79 (32)	12/36 (33)	8/36 (22)
			271 (1467, 271, 6270, 8670)	18/79 (22)	6/36 (17)	11/36 (28)
			420 (420)	9/79 (11)	7/36 (19)	10/41 (24)
			2067 (2067, 10716, 10823)	5/79 (6)	2/36 (6)	1/36 (2)
			2062, 172, 6279	1/79 (1)	1/36 (3)	3/36 (8)
			Singletons (10619, 10683, 10709, 10740, 10784, 10795, 10817, 10827, 1203, 2834, 5179, 6271, 6272, 6273, 6274, 7131, 7719, 7725, 8326, 8328, 8332, 87, 9927, 9928, 9937)	21/79 (27)	8/36 (22)	3/36 (8)
23F	503/171/76	101/37/68	6279 (n=12 STs)	54/101 (54)	26/37 (70)	58/68 (85)

			172 (172, 6281, 6287)	27/101 (27)	8/37 (22)	1/68 (1)
			81 (81)	2/101 (2)	1/37 (3)	2/68 (3)
			802 (802)	1/101 (1)	0	2/68 (3)
			63, 445, 4084, 4881, 10684	5/101 (5)	0	2/68 (3)
			Singletons (10828, 10712, 10728, 10754, 2290, 2651, 2714, 353, 36, 4090, 507, 6286, 7865)	12/101 (12)	2/37 (5)	2/68 (3)
7C	26/11/15	3/3/4	2289	1/3 (33)	0	0
			4956	1/3 (33)	2/3 (67)	2/4 (50)
			7053	1/3 (33)	0	0
			Singletons (10639, 1535)	0	1/3 (33)	2/4 (50)
8	76/38/77	27/3/34	53 (53, 3847)	25/27 (93)	3/3 (100)	32/34 (94)
			3406 (10852, 3406, 2680)	1/27 (4)	0	2/34 (6)
			Singleton (1480)	1/27 (4)	0	0
9N	38/16/12	4/0/5	6239	0	0	1/5 (20)
			Singletons (10642, 3983, 5624)	4/4 (100)	0	4/5 (80)
11A	11/4/11	3/0/3	8817	2/3 (67)	0	3/3 (100)

			Singletons (10694, 5537)	1/3 (33)	0	0
12F	50/22/55	18/2/25	989 (989, 5026)	7/18 (39)	0	14/25 (56)
			2416 (1178, 2416)	7/18 (39)	2/2 (100)	10/25 (40)
			218 (10663, 218)	1/18 (6)	0	1/25 (4)
			Singletons (3775, 7110)	3/18 (17)	0	0
13	30/8/18	5/1/7	5647	4/5 (80)	1/1 (100)	6/7 (86)
			10849	0	0	1/7 (14)
			Singleton (10702)	1/5 (20)	0	0
15A	16/11/31	3/0/11	63	0	0	2/11 (18)
			10588	1/3 (33)	0	0
			Singletons (10550, 10761, 10861, 11732, 11766, 4965)	2/3 (67)	0	9/11 (82)
15B/C	69/25/64	15/5/31	199	0	0	4/31 (13)
			2421	0	0	1/31 (3)
			10588	1/15 (7)	0	0
			Singletons (10617, 1262, 361, 3669, 4209, 5808, 7052, 8687, 8820)	14/15 (93)	5/5 (100)	26/31 (84)

16	17/0/0	3/0/0	4088	2/3 (67)	0	0
			5326	1/3 (33)	0	0
16F	32/15/39	5/4/19	3450	0	0	5/19 (26)
			4088	0	1/4 (25)	4/19 (21)
			10667	0	1/4 (25)	3/21 (16)
			5326	1/5 (20)	1/4 (25)	2/19 (11)
			30	2/5 (40)	0	1/19 (5)
			199	0	0	1/19 (5)
			5659	0	0	1/19 (5)
			Singletons (10543, 10862, 11723, 310)	2/5 (40)	1/4 (25)	2/19 (11)
17F	17/9/19	1/4/12	8835	0	2/4 (50)	5/12 (42)
			392	1/1 (100)	1/4 (25)	4/12 (33)
			Singletons (10623, 10647, 10867, 10872)	0	1/4 (25)	3/12 (25)
18A	8/3/2	3/0/1	185	1/3 (33)	0	0
			Singletons (1232, 241)	2/3 (67)	0	1/1 (100)

18B	5/1/1	0/1/0	6239	0	1/1 (100)	0
20	4/4/3	1/0/2	8817	0	0	2/2 (100)
			Singleton (8824)	1/1 (100)	0	0
21	8/6/10	1/0/4	1221	0	0	1/4 (25)
			Singletons (1233, 193)	1/1 (100)	0	3/4 (75)
22A	4/2/0	1/0/1	445	1/1 (100)	0	0
			6279	0	0	1/1 (100)
22F	11/6/13	3/2/7	445	3/3 (100)	0	5/7 (71)
			6275	0	0	1/7 (14)
			Singletons (10819, 433)	0	2/2 (100)	1/7 (14)
23A	11/4/6	4/1/4	53	0	0	1/4 (25)
			2319	3/4 (75)	1/1 (100)	2/4 (50)
			Singletons (5080)	1/4 (25)	0	1/4 (25)
23B	11/2/5	3/1/2	2289	1/3 (33)	0	0
			Singletons (10719, 2911, 4423)	2/3 (67)	1/1 (100)	2/2 (100)
26	1/0/0	1/0/0	199	1/1 (100)	0	0

29	27/5/5	4/1/3	217	0	0	1/3 (33)
			Singletons (10826, 5483, 558)	4/4 (100)	1/1 (100)	2/3 (67)
31	4/0/0	4/0/0	3548	2/4 (50)	0	0
33D	11/11/8	1/1/1	4084	1/1 (100)	1/1 (100)	1/1 (100)
			10588	2/4 (50)	0	0
33F	7/3/6	0/1/2	445	0	0	1/2 (50)
			Singletons (100, 10068)	0	1/1 (100)	1/2 (50)
34	28/12/12	5/1/5	7067	4/5 (80)	0	2/5 (40)
			Singletons (10544, 10863, 10865, 1884, 8822)	1/5 (20)	1/1 (100)	3/5 (60)
35B	17/16/50	8/2/16	172 (172, 361)	1/8 (13)	1/2 (50)	9/16 (56)
			4084 (8774)	0	1/2 (50)	1/16 (6)
			Singletons (10957, 373, 5396, 5483, 9813)	7/8 (87)	0	6/16 (38)

Appendix C

Clonal complex and sequence type (ST) distribution of pneumococcal isolates causing invasive disease among older children ≥5 years of age in South Africa during pre-PCV period (2005-2008), PCV7 period (2009-2010) and PCV13 period (2011-2013)

Serotype	No. of isolate: pre-PCV/PCV7/PCV13 period	No. with genotypic results: pre-PCV/PCV7/PCV13 period	Clonal complex (ST)	pre-PCV period n/N (%)	PCV7 period n/N (%)	PCV13 period n/N (%)
1	1288/711/696	177/38/208	217 (217, 612, 9067, 2839, 8313, 8314, 8315, 8319, 8321, 8385)	170/177 (96)	38/38 (100)	202/208 (97)
			2289, 2652, 3544	3/177 (2)	0	1/208 (1)
			Singletons (611, 618, 306, 615, 8684)	4/177 (2)	0	6/208 (2)
3	408/179/275	94/4/173	458 (458, 7348, 7349, 7350, 7353, 8386)	55/94 (55)	1/4 (25)	82/173 (47)
			180 (180,181, 3842, 4118, 7356, 7357, 8388)	15/94 (16)	1/4 (25)	38/173 (22)
			230 (700)	0	1/4 (25)	32/173 (18)
			378 (7369, 233, 378, 7362, 7364, 7365,7367, 7368, 7369)	10/94 (11)	0	7/173 (4)
			445, 1221, 2289, 2416, 2421, 4084, 5026, 5410	2/94 (2)	0	7/173 (4)

			Singletons (1220, 1765, 2021, 2837, 6193, 7347, 7351, 7352 7366)	12/94 (13)	1/4 (25)	7/173 (4)
4	603/298/311	84/11/175	1221 (1221, 7370, 7354,7355)	32/84 (38)	7/11 (64)	119/175 (68)
			5410 (5410, 7370, 5655, 7372, 7373, 7374, 7375, 7376)	25/84 (30)	1/11 (9)	25/175 (14)
			271, 5026, 5659, 6279	1/84 (1)	0	3/175 (2)
			Singletons (1222, 2290, 10612, 205, 2213, 2712, 2941, 7358, 7359, 7371, 853, 8762, 9197, 9408, 9554)	26/84 (31)	3/11 (27)	28/175 (16)
5	93/69/52	10/5/24	5659 (289, 5659, 7050, 8370)	10/10 (100)	5/5 (100)	24/24 (100)
6A	552/286/264	82/7/146	2285 (n=23 STs)	32/82 (39)	2/7 (29)	61/146 (42)
			2289 (1094, 2289, 7312, 7721, 10656, 4941, 8690, 8788)	15/82 (18)	2/7 (28)	23/146 (16)
			172 (1447, 8689)	1/82 (1)	0	24/146 (16)
			6389 (138, 6389, 7304, 7309)	8/82 (10)	0	5/146 (3)
			473 (471, 473, 7298, 8695)	1/82 (1)	0	9/146 (6)
			185 (185, 8307)	5/82 (6)	0	0
			81, 180, 1221, 2283, 2421, 4084, 6279, 7084	7/82 (9)	2/7 (29)	7/146 (5)
			Singletons (2987, 10550, 10582, 10681, 10692, 10701, 10704, 10773, 2987, 3207, 3821, 4957, 5251, 5412, 7299,	13/82 (16)	1/7 (14)	17/146 (12)

			7313, 8687, 8688, 8692, 8694, 8700, 9496)			
6B	410/183/121	58/3/90	2421 (n=STs)	13/58 (22)	0	34/90 (38)
			185 (185, 7738, 10845, 10871, 6293, 7302, 7307, 7738, 8381)	10/58 (17)	0	23/90 (26)
			2285 (2285, 4087, 5073, 7303)	9/58 (16)	0	4/90 (4)
			6389 (138, 6389, 7304, 7309)	4/58 (7)	1/3 (33)	4/90 (4)
			6279 (6305, 242, 6219, 6305)	2/58 (3)	1/3 (33)	4/90 (4)
			7084	2/58 (3)	0	5/90 (6)
			2289, 2185, 172, 230, 1094, 4893, 6393	11/58 (19)	0	9/90 (3)
			Singletons (1518, 2040, 2156, 3207, 3473, 4250, 6232, 6251, 6282, 6294, 7311, 7733, 8383, 8387, 8713, 90, 9600)	8/58 (14)	1/3 (33)	12/90 (13)
6C	43/14/20	7/3/13	2185 (2185, 6310, 7345, 7346, 8382)	5/7 (71)	3/3 (100)	10/13 (69)
			2283 (2283)	1/7 (14)	0	0
			Singletons (1669, 1692, 398, 4879, 7344)	1/7 (14)	0	4/13 (31)
7F	117/58/99	28/1/81	3544/218 (218, 3544, 1176, 6838, 7884, 7886, 8679)	20/28 (71)	1/1 (100)	61/81 (75)
			2416	3/28 (11)	0	9/81 (11)
			4956	0	0	1/81 (1)

			Singletons (191, 5027, 7048, 7360, 7361, 7883, 7887, 8669, 8679)	5/28 (22)	0	10/81 (12)
9V	287/129/84	53/3/61	4881 (3454, 4881, 6314, 6530, 7889, 8760, 8763)	31/53 (58)	2/3 (67)	33/61 (54)
			156 (156, 162, 557, 644, 7891)	12/53 (23)	1/3 (33)	16/61 (26)
			5408, 4902, 53, 458, 2285, 8817, 10667	4/53 (8)	0	6/61 (10)
			Singletons (2280, 280, 3983, 609, 6317, 8761)	6/53 (11)	0	6/61 (10)
14	524/212/114	73/7/87	230 (230, 2707, 3980, 6227, 6231, 62314, 6267, 7283, 7291, 8750)	30/73 (41)	2/7 (29)	45/87 (52)
			2652 (n=17 STs)	17/73 (23)	5/7 (71)	24/87 (28)
			63 (63, 10707, 2414, 7270, 7277, 7279, 7722, 861, 7293)	20/73 (27)	0	11/87 (13)
			156, 217, 2062, 4902 and 6279	2/73 (3)	0	4/87 (5)
			Singletons (124, 1262, 143, 7266, 7273, 7292, 782)	4/73 (5)	0	3/87 (3)
18C	171/94/76	43/2/53	6245 (102, 1016, 6245, 7262, 7265, 7297, 7730, 8650)	14/43 (33)	1/2 (50)	14/53 (26)
			6239 (6239, 7272, 7274, 7275, 7289, 8316, 8820)	12/43 (28)	0	13/53 (25)
			4893 (1221, 7370, 7354, 7355)	5/43 (12)	1/2 (50)	11/53 (21)
			6236 (8317, 6236, 6264, 7286, 7296, 7742, 8317, 8318, 8381)	8/43 (19)	0	8/53 (15)

			217, 230 and 6279	1/43 (2)	0	2/53 (4)
			Singletons (1381, 5483, 7263, 7290, 7294, 8311, 8682, 8840)	3/43 (7)	0	5/53 (9)
19A	750/453/542	84/15/180	2062 (n=25 STs)	69/84 (82)	12/15 (80)	168/180 (93)
			172 (172, 2218, 7861, 8300)	7/84 (8)	1/15 (7)	4/180 (2)
			347, 2416, 4088, 6239	1/84 (1)	2/15 (13)	2/180 (1)
			Singletons (2345, 306, 3111, 75, 7896, 7900, 7901, 7902, 7904, 8309, 9913)	7/84 (8)	0	6/180 (6)
19F	382/165/154	37/9/94	271 (271, 10825, 2706, 6270, 6271, 8678)	16/37 (43)	2/9 (22)	33/94 (35)
			347 (347, 6277)	5/37 (14)	2/9 (22)	17/94 (18)
			420 (3217, 420, 7857, 7858)	4/37 (11)	2/9 (22)	9/94 (10)
			2067 (2067, 7859, 8373)	3/37 (8)	0	5/94 (5)
			2062 (2062, 6121, 7897)	2/37 (5)	0	3/94 (5)
			6279, 230, 172, 6389	2/37 (5)	1/9 (22)	7/94 (7)
			Singletons (10762, 1203, 179, 2345, 5179, 6278, 7131, 7860, 8236, 8326, 8328, 8331, 8332, 8333, 8665, 8670, 8673, 8676, 8677)	5/37 (14)	2/9 (22)	20/94 (25)
23F	518/291/202	61/15/161	6279 (n=23 STs)	22/61 (36)	14/15 (93)	134/161 (83)

			172 (172, 6287, 7861,7866, 7882)	17/61 (28)	0	8/161 (5)
			802 (802, 6120)	2/61 (3)	0	5/161 (3)
			81 (81, 83, 7867)	4/61 (7)	0	2/161 (1)
			230, 271, 1221, 2285, 3450, 4881 and 4956	4/61 (7)	0	3/161 (2)
			Singletons (8330, 10732, 1535, 193, 2651, 2714, 353, 36, 4884, 507, 66, 7863, 7865, 8330, 8667, 8668, 8675, 9518)	12/61 (24)	1/15 (7)	9/161 (6)
7C	77/4373	3/3/4	4956	3/3 (100)	2/3 (67)	3/4 (75)
			7053	0	0	1/4 (75)
			Singleton (1535)	0	1/3 (33)	0
8	364/146/253	69/5/19	53 (53, 3847)	66/69 (96)	4/5 (80)	15/19 (79)
			3406	1/69 (1)	1/5 (20)	0
			Singletons (10868, 1480, 2234, 4216)	2/69 (3)	0	4/19 (21)
9N	128/85/129	4/5/1	Singletons (3983, 66)	4/4 (100)	5/5 (100)	1/1 (100)
10A	100/47/83	1/1/7	5026	0	0	1/7 (14)
			Singletons (10824, 10864, 2068, 8821, 8825, 97)	1/1 (100)	1/1 (100)	6/7 (85)
10F	26/11/17	1/0/0	Singleton (5422)	1	0	0

11A	53/30/39	4/3/2	8817 6279 Singleton (10734, 10771)	2/4 (50) 0 2/4 (50)	2/3 (67) 1/3 (33)	2/2 (100) 0 0
11B	14/3/6	1/0/1	Singletons (10850, 5969)	1	0	1/1 (100)
12F	344/232/424	99/10/16	5026/989 (989, 5026) 2416 (2416) 3544/218 (218, 10663) Singletons (3775, 7062, 7110)	37/99 (37) 40/99 (44) 15/99 (15) 7/99 (7)	5/10 (50) 5/10 (50) 0 0	10/16 (63) 6/16 (38) 0 0
13	108/58/112	6/2/3	5647 172 and 3544 Singleton (70)	3/6 (50) 2/6 (33) 1/6 (17)	2/2 (100) 0 0	3/3 (100) 0 0
15A	70/26/120	9/0/3	10588 63 Singletons (10550, 10605, 11766, 2420, 4965, new ST)	4/9 (44) 1/9 (11) 4/9 (44)	0 0 0	1/3 (33) 0 2/3 (67)
15B/C	133/58/111	8/4/5	199 Singletons (4209, 7052, 8687, 1262)	4/8 (50) 4/8 (50)	0 4/4 (100)	0 5/5 (100)
16	179/99/146	10/4/15	4088	5/10 (50)	1/4 (25)	5/15 (33)

			3450	1/9 (11)	0	4/15 (27)
			30, 5326 and 10667	3/9 (33)	2/4 (50)	4/15 (27)
			Singletons (10543, 10730)	1/9 (11)	1/4 (25)	2/15 (13)
17F	87/55/93	2/1/2	392 Singletons (10564, 10623, 10689)	1/2 (50) 1/2 (50)	0 1/1 (100)	1/2 (100) 1/2 (100)
18A	52/17/16	7/1/1	185 Singletons (10715, 10848, 1232, 241)	1/7 (14) 6/7 (86)	0 1/1 (100)	0 1/1 (100)
18F	23/6/6	1/0/1	10588 8840	1/1 (100) 0	0 0	0 1/1 (100)
20	15/7/13	0/0/1	8817	0	0	1/1 (100)
21	11/7/9	4/0/1	Singletons (10854, 193, 7068)	4/4 (100)	0	1/1 (100)
22A	17/6/6	1/0/0	10588 Singleton (5266)	1/1 (100) 0	0 0	0 0
22F	104/69/110	2/1/3	445 Singleton (10851)	2/2 (100) 0	1/1 (100) 0	2/3 (67) 1/3 (33)
23A	38/26/41	3/1/4	2319 2652	3/3 (100) 0	0 0	0 1/4 (25)

			Singleton (5080)	0	1/1 (100)	3/4 (75)
23B	23/10/14	1/0/0	Singleton (2911)	1	0	0
25	79/0/0	1/0/0	Singleton (105)	1	0	0
28	15/1/0	2/0/0	Singleton (494)	2	0	0
28A	0/2/0	0/1/0	Singleton (494)	0	1/1 (100)	0
29	50/12/5	2/1/1	1221 Singletons (5178, 5483)	0 2/2 (100)	0 1/1 (100)	1/1 (100) 0
31	27/9/27	1/1/2	445 3548	1/1 (100) 0	1/1 (100) 0	1/2 (50) 1/2 (50)
33D	35/13/13	3/0/1	4084 Singletons (7062, 8714)	2/3 (67) 1/3 (33)	0 0	0 1/1 (100)
33F	3212/18	1/1/2	445	1/1 (100)	1/1 (100)	2/2 (100)
34	71/44/63	1/4/1	7067 Singletons (10544, 1884, 8822)	0 1/1 (100)	2/4 (50) 2/4 (50)	1/1 (100) 0
35A	0/1/5	0/0/1	217	0	0	1/1 (100)
35B	48/20/64	12/0/6	172 (361) 2289 (1094)	1/12 (8) 1/12 (8)	0 0	4/6 (67) 0

			2416 (2416)	1/12 (8)	0	0
			Singletons (10957, 10989, 9813)	9/12 (75)	0	2/6 (33)
35F		2/0/0	Singletons (10699, 9408)	2	0	0
38	47/18/1	3/0/0	Singletons (310, 393)	3/3 (100)	0	0
39	1/1/1	0/0/1	Singleton (10873)	0	0	1/1 (100)

Appendix D

Serotype and clonal complex distribution of penicillin non-susceptible (PNS) and multidrug-resistant (MDR) isolates causing invasive pneumococcal disease among infants and children <5 years of age in South Africa during the pre-PCV (2005-2008), PCV7 (2009-2010) and PCV13 period (2011-2013)

		pre-PCV			PCV7			PCV13		
Serotype	Clonal complex[1]	No. of isolates	No. (%) of PNS isolates	No. (%) of MDR isolates	No. of isolates	No. (%) of PNS isolates	No. (%) of MDR isolates	No. of isolates	No. (%) of PNS isolates	No. (%) of MDR isolates
1	217	90	0	0	30	0	0	57	2 (4)	2 (4)
3	230	2	1 (50)	1 (50)	0	0	0	1	0	0
4	5410	19	17 (89)	0	5	5 (100)	1 (20)	4	4 (100)	0
5	5659	13	1 (8)	0	3	0	0	30	3 (10)	0
6A	81	0	0	0	0	0	0	0	1	1 (100)
	172	5	5 (100)	0	6	6 (100)	0	4	4 (100)	0
	185	1	1 (100)	0	0	0	0	0	0	0
	473	2	2 (100)	0	1	0	0	6	3 (50)	0
	2285	44	1 (3)	0	14	0	0	30	1 (3)	0
	2289	41	27 (66)	3 (7)	20	20 (100)	2 (10)	16	16 (100)	0
	6389	8	1 (13)	0	3	0	0	1	0	0
	other	4	2 (50)	0	3	1(33)	0	6	5 (83)	0
6B	185	34	34 (100)	0	3	3 (100)		14	4 (100)	0
	230	1	1 (100)	0	0	0	0	0	0	0

	2285	4	1 (25)	0	0	0	0	1	0	0
	2289	4	3 (75)	0	3	3 (100)	0	0	0	0
	2421	45	32 (71)	0	6	6 (100)	0	6	6 (100)	0
	6239	2	1 (50)	0	0	0	0	0	0	0
	6279	8	2 (25)	5 (63)	0	0	0	1	1 (100)	0
	6389	8	0	0	3	1 (33)	0	1	1 (100)	0
	6393	1	1 (100)	0	2	2 (100)	0	2	2 (100)	0
	7084	4	0	3 (75)	4	1 (25	3 (75)	0		0
	other	24	12 (50)	3 (13)	8	1 (13)	5 (63)	7	3 (43)	0
9V	156	4	2 (50)	0	2	1 (50)	0	2	1 (50)	0
	4881	28	9 (32)	0	7	1 (14)	0	7	1 (14)	0
	other	2	2 (100)	0	2	1 (50)	0	0	0	0
14	63	25	17 (68)	15 (60)	8	8 (100)	7 (88)	5	5 (100)	5 (100)
	230	77	77 (100)	77 (100)	22	22 (100)	22 (100)	19	19 (100)	19 (100)
	2652	39	38 (97)	37 (95)	17	17 (100)	16 (94)	8	8 (100)	7 (88)
	other	6	5 (83)	3 (50)	3	1 (33)	0	4	4 (100)	3 (75)
15A	63	0	0	0	0	0	0	2	2 (100)	2 (100)
18C	6239	13	1 (8)	0	3	0	0	17	0	0
19A	172	3	3 (100)	1 (33)	3	2 (67)	1 (33)	1	1 (100)	0
	230	1	0	1 (100)	0	0	0	0	0	0
	271	0	0	0	0	0	0	1	0	1 (100)
	2062	97	79 (81)	1 (1)	32	32 (100)		77	77 (100)	
	other	9	6 (67)	1 (11)	1	0	0	1	0	0
19F	172	0	0	0	1	1 (100)	0	1	1 (100)	0
	271	18	18 (100)	10 (56)	6	6 (100)	6 (100)	11	11 (100)	11 (100)

	347	25	25 (100)	1 (4)	12	12 (100)	0	8	8 (100)	0
	420	9	9 (100)	8 (89)	7	7 (100)	0	10	10 (100)	0
	2062	1	0	0	0	0	0	1	1 (100)	0
	other	21	14 (67)	3 (14)	8	7 (88)	0	3	3 (100)	1 (33)
23F	63	1	1 (100)	1 (100)	0	0	0	0	0	0
	81	2	2 (100)	2 (100)	1	1 (100)	1 (100)	2	2 (100)	2 (100)
	172	27	23 (85)	1 (4)	8	8 (100)	2 (25)	1	1 (100)	0
	802	1	0	0	0	0	0	2	0	1
	6279	54	31 (57)	30 (56)	26	23 (88)	23 (88)	58	52 (90)	49 (84)
	10684	3	3 (100)	0	0	0	0	0	0	0
	other	12	6 (50)	1 (8)	2	1 (50)	0	2	1 (50)	1 (50)
35B	172	1	0	0	1	1 (100)	0	9	9 (100)	0
	other	7	2 (29)	1 (14)	0	0	0	6	1 (17)	0

[1]Other: singletons/ unrelated sequence types

Appendix E

Serotype and clonal complex distribution of penicillin non-susceptible (PNS) and multidrug-resistant (MDR) isolates causing invasive pneumococcal disease among children and adults ≥5 years of age in South Africa during the pre-PCV (2005-2008), PCV7 (2009-2010) and PCV13 period (2011-2013).

		pre-PCV			PCV7			PCV13		
Serotype	Clonal complex[1]	No. of isolates	No. (%) of PNS isolates	No. (%) of MDR isolates	No. of isolates	No. (%) of PNS isolates	No. (%) of MDR isolates	N	No. (%) of PNS isolates	No. (%) of MDR isolates
1	217	178	1 (0.5)	0	38	0	0	202	3 (1)	0
3	230	0	0	0	1	0	1 (100)	32	32 (100)	32 (100)
	458	55	0	0	1	0	0	82	3 (4)	0
	2416	1	1 (100)	0	0	0	0	1	0	0
	5410	0	0	0	0	0	0	1	1 (100)	0
	other	12	5 (42)	5 (42)	1	0	0	17	1 (17)	0
4	1221	32	0	0	7	0	0	119	2 (6)	0
	5410	25	24 (96)	1 (4)	1	1 (100)	0	25	25 (100)	0
	other	26	0	0	3	0	0	28	2 (7)	0
6A	81	0	0	0	0	0	0	3	3 (100)	3 (100)
	172	1	1 (100)	0	0	0	0	24	24 (100)	0
	185	5	3 (60)	0	0	0	0	0	0	0
	473	1	1 (100)	0	0	0	0	9	7 (78)	0
	2285	32	2 (6)	1 (3)	2	0	0	61	2 (3)	0

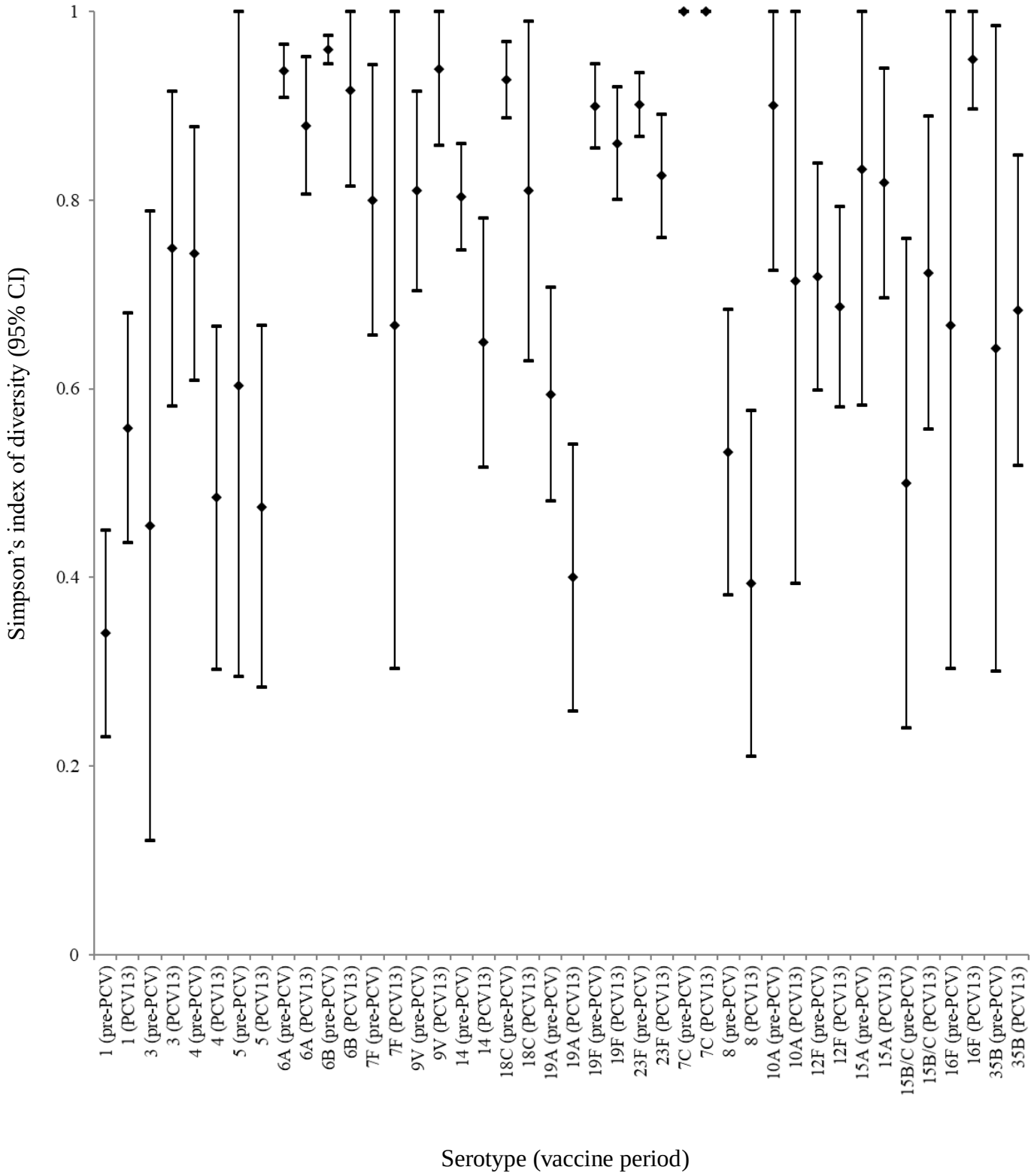
	2289	15	9 (60)	0	3	3 (100)	1 (33)	23	23 (23)	0
	2421	4	3 (75)	0	0	0	0	1	0	0
	6279	0	0	0	0	0	0	1	1 (100)	1 (100)
	other	13	3 (23)	1 (8)	1	0	0	17	4 (24)	1 (6)
6B	172	1	1 (100)	0	0	0	0	0	0	0
	185	10	8 (80)	0	0	0	0	23	22 (96)	0
	230	2	1 (50)	0	0	0	0	1	0	0
	1094	1	1 (100)	0	0	0	0	0	0	0
	2285	9	5 (56)	0	0	0	0	4	1 (25)	1 (25)
	2289	3	2 (67)	1 (33)	0	0	0	1	1 (100)	0
	2421	13	7 (54)	1 (8)	0	0	0	34	31 (91)	1 (3)
	4893	1	1 (100)	0	0	0	0	0		0
	6279	2	2 (100)	2 (100)	1	1 (100)	1 (100)	4	4 (100)	2 (50)
	6389	4	1 (25)	0	1	0	0	4	0	0
	6393	2	2 (100)	0	0	0	0	2	2 (100)	0
	7084	2	2 (100)	1 (50)	0	0	0	5	5 (100)	3 (60)
	other	8	6 (75)	0	1	1 (100)	1 (100)	12	11 (92)	6 (50)
6C	other	1	0	0	0	0	0	4	1 (25)	0
8	53	66	1 (2)	0	4	0	0	15	0	0
9V	156	12	7 (58)	0	1	1 (100)	0	16	11 (69)	3 (19)
	2285	1	1 (100)	0	0	0	0	0	0	0
	4881	31	12 (39)	0	2	1 (50)	0	33	9 (27)	0
	other	6	5 (83)	0	0	0	0	6	4 (67)	0
14	63	20	17 (85)	16 (80)	0	0	0	11	11 (100)	11 (100)
	156	1	1 (100)	1 (100)	0	0	0	2	2 (100)	2 (100)

	230	30	30 (100)	30 (100)	2	2 (100)	2 (100)	45	45 (100)	44 (98)
	2652	17	16 (94)	13 (76)	5	2 (100)	5 (100)	24	3 (13)	2 (8)
	other	4	2 (50)	1 (25)	0	0	0	3	2 (67)	1 (33)
15B/C	other	3	2 (67)	0	4	1 (25)	0	4	2 (50)	0
16F	other	1	0	0	1	1 (100)	0	2	2 (100)	0
18C	6245	14	0	0	1	0	0	14	1 (7)	0
	6279	0	0	0	0	0	0	1	1 (100)	1 (100)
19A	172	7	6 (86)	0	1	1 (100)	0	4	4 (100)	0
	347	0	0	0	0	0	0	1	1 (100)	0
	2062	69	58 (84)	1 (1)	12	12 (100)	0	168	163 (97)	1 (97)
	2416	1	1 (100)	0	0	0	0	0	0	0
	4088	0	0	0	1	1 (100)	0	0	0	0
	other	7	5 (71)	2 (29)	0	0	0	6	5 (83)	2 (33)
19F	172	0	0	0	0	0	0	2	2 (100)	0
	230	1	1 (100)	1 (100)	0	0	0	2	2 (100)	1 (50)
	271	16	16 (100)	10 (63)	2	2 (100)	1 (50)	33	33 (100)	32 (67)
	347	5	3 (60)	0	2	2 (100)	0	17	17 (100)	0
	420	4	3 (75)	0	2	2 (100)	0	9	9 (100)	0
	2062	2	1 (50)	0	0	0	0	3	3 (100)	0
	2067	3	1(33)	0	0	0	0	5	4 (80)	0
	6279	0	0	0	0	0	0	3	3 (100)	3 (100)
	6389	0	0	0	1	1 (100)	0	0	0	0
	other	5	3 (60)	1 (20)	2	1 (50)	0	20	19 (95)	2 (10)
23F	81	4	4 (100)	4 (100)	0	0	0	2	2 (100)	2 (100)
	172	17	14 (82)	2 (12)	0	0	0	8	8 (100)	0

	230	2	2 (100)	1 (50)	0	0	0	1	1 (50)	0
	271	1	1 (50)	0	0	0	0	0	0	0
	802	2	0	0	0	0	0	5	2 (40)	2 (40)
	1221	0	0	0	0	0	0	1	1 (100)	1 (100)
	6279	22	16 (73)	16 (73)	14	11 (79)	11 (79)	134	122 (91)	122 (91)
	other	12	7 (58)	3 (25)	1	1 (100)	0	9	5 (55)	4 (44)
35B	172	1	1 (100)	0	0	0	0	4	4 (100)	0
	2289	1	1 (100)	0	0	0	0	0	0	0
	2416	1	1 (100)	0	0	0	0	0	0	0

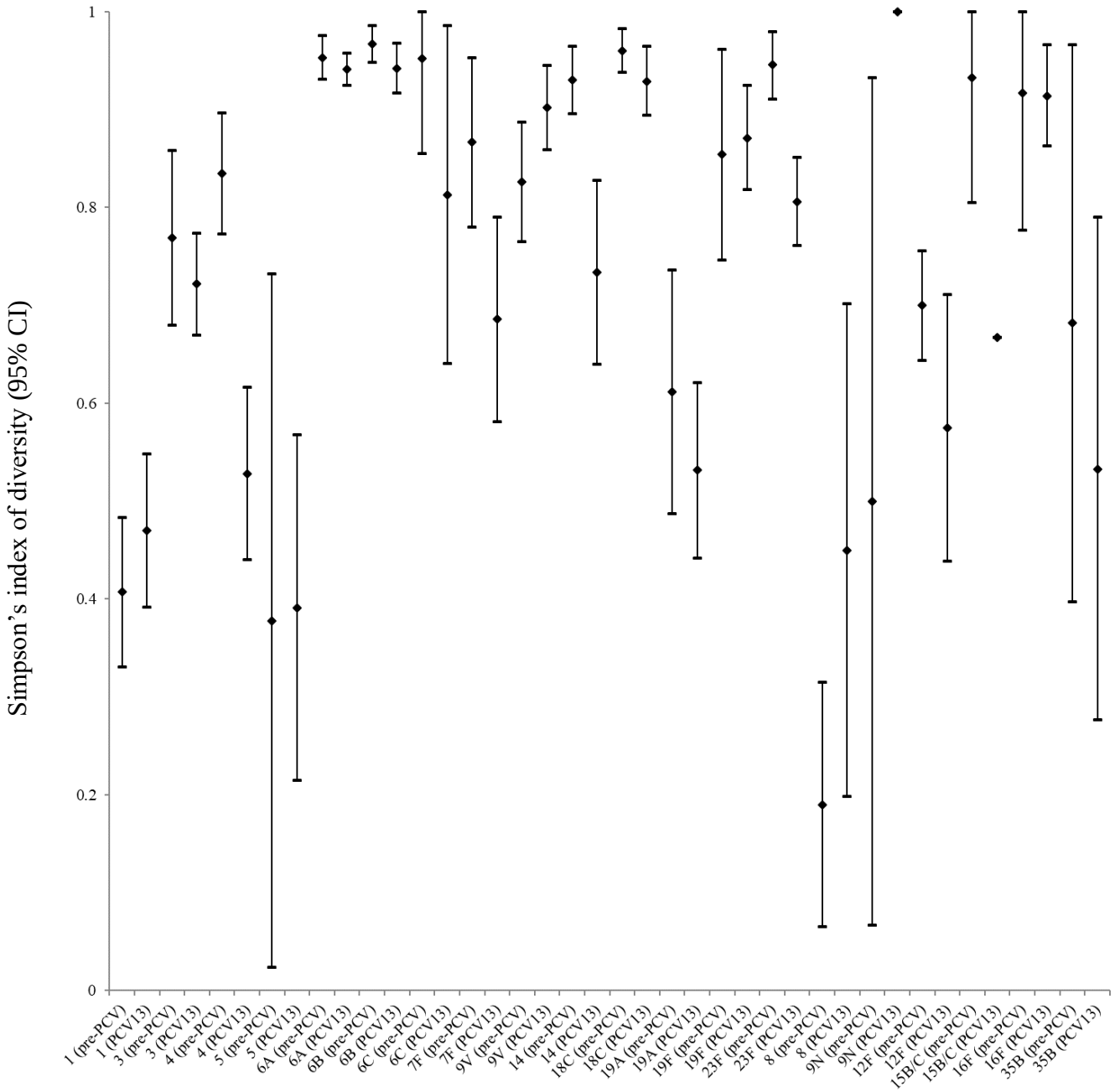
[1]Other: singletons/ unrelated sequence types

Appendix F



Diversity of sequence types within serotypes of isolates causing invasive pneumococcal disease among infants and children <5 years of age in South Africa during the pre-PCV (2005-2008) and PCV13 (2011-2013) periods.

Appendix G



Diversity of sequence types within serotypes of isolates causing invasive pneumococcal disease among older children and adults ≥5 years of age in South Africa during the pre-PCV (2005-2008) and PCV13 (2011-2013) periods.

Appendix H

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG
Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Kedibone M Mothibeli et al

CLEARANCE CERTIFICATE	M111008
PROJECT	The Impacy of Pneumococcal 7-and 13-Valent Conjugate Vaccines on the Molecular Epidemiology in Invasive Streptococcus pneumoniae from Children <2 Years of Age in South Africa
INVESTIGATORS	Kedibone M Mothibeli et al.
DEPARTMENT	Division of Virology & Communicable Diseases
DATE CONSIDERED	28/10/2011
M1110080DECISION OF THE COMMITTEE*	Approved unconditionally

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

DATE 20/02/2012 **CHAIRPERSON** .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable
cc: Supervisor : Dr Mignon du Plessis

DECLARATION OF INVESTIGATOR(S)
To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**
PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...

Appendix I

Declaration: Student's contribution to article(s) and agreement of co-author(s)

I, Kedibone Maria Ndlangisa, student number 9703712H, declare that this Thesis is my own work and that I contributed adequately towards research findings published in the article(s) stated below which are included in my Thesis.

Signature of Student ..... **Date** 17/03/2017

Name of Primary Supervisor Dr Mignon du Plessis

Signature of Primary Supervisor ..... **Date** 22/01/17

Agreement by co-authors: By signing this declaration, the co-authors listed below agree to the use of the article(s) by the student as part of his/her Thesis/Dissertation/Research Report. In cases where the student is not the 1st author of a published article, the primary supervisor must explain (under comments) why the student is entitled to use the paper for his/her degree purposes.

Article 1: Two cases of serotypeable and non-serotypeable variants of *Streptococcus pneumoniae* detected simultaneously during invasive disease, BMC Microbiology, 2016, 16(1):126

Authors	Name	Signature	Date
1st author	du Plessis M		17.03.17
2nd author	Allam M		22/03/2017
3rd author	Wolter N		22/03/2017
4th author	Mohale T		22/03/2017
5th author	de Gouveia L		22/3/2017
6th author	Birkhead M		22/3/2017
7th author	Klugman KP		31/3/2017
8th author	von Gottberg A		17/3/2017

Comments by primary supervisor:

The student is the 1st author

Article 2: Title: Population snapshot of Streptococcus pneumoniae causing invasive disease in South Africa prior to introduction of pneumococcal conjugate vaccines. PLoS One. 2014;9(9):e10766

Authors	Name	Signature	Date
1st author	du Plessis M		17.03.17
2nd author	Wolter N		22/03/2017
3rd author	de Gouveia L		22/3/2017
4th author	Klugman KP		31/3/2017
5th author	von Gottberg A		17/3/2017
6th author			

Comments by primary supervisor:

The student is the 1st author

1

Appendix I

Declaration: Student's contribution to article(s) and agreement of co-author(s)

Article 3: Invasive disease caused simultaneously by dual serotypes of *Streptococcus pneumoniae*. J Clin Microbiol. 2018, 56 (1): e01149-17

Authors	Name	Signature	Date
1st author	du Plessis M		22/01/18
2nd author	Allam M		22/01/2018
3rd author	Wolter N		22/01/2018
4th author	de Gouveia L		22/1/2018
5th author	Klugman KP		18/1/18
6th author	Cohen C		18/1/18
7th author	Gladstone RA	Gladstone	12/01/18
8th author	von Gottberg A		22/1/2018

Comments by primary supervisor:

The student is the 1st author

2

Appendix J

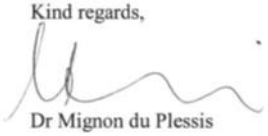
University of the Witwatersrand
Faculty of Health Sciences, Postgraduate Office
Parktown, 2193
Johannesburg, South Africa

To whom it may concern

RE: Turnitin report for Kedibone Ndlangisa (student no. 9703712H) PhD thesis

Kedibone Ndlangisa is a part-time student doing PhD at the University of the Witwatersrand. She published three papers in peer-reviewed journals from her research thesis during the course of her PhD in the Faculty of Health Sciences (see page iv). The overall high similarity index score (42%) in the Turnitin plagiarism report is partly due to the three published manuscripts included as chapters in this thesis.

Kind regards,



Dr Mignon du Plessis
Lecturer, Clinical Microbiology and Infectious Diseases
Faculty of Health Sciences, University of the Witwatersrand
and
Senior Medical Scientist, Centre for Respiratory Diseases and Meningitis, National Institute for Communicable Disease

PHDthesis250317KN.docx

ORIGINALITY REPORT

%42	%38	%39	%11
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	bmcmicrobiol.biomedcentral.com Internet Source	%11
2	journals.plos.org Internet Source	%10
3	Ndlangisa, Kedibone M., Mignon du Plessis, Nicole Wolter, Linda de Gouveia, Keith P. Klugman, Anne von Gottberg, and for GERMS-SA. "Population Snapshot of Streptococcus pneumoniae Causing Invasive Disease in South Africa Prior to Introduction of Pneumococcal Conjugate Vaccines", PLoS ONE, 2014. Publication	%4
4	Ndlangisa, Kedibone M., Mignon du Plessis, Mushal Allam, Nicole Wolter, Thabo Mohale, Linda de Gouveia, Monica Birkhead, Keith P. Klugman, and Anne von Gottberg. "Two cases of serotypeable and non-serotypeable variants of Streptococcus pneumoniae detected simultaneously during invasive disease", BMC Microbiology, 2016. Publication	%3

5	Submitted to London School of Hygiene and Tropical Medicine Student Paper	%1
6	Submitted to American Intercontinental University Online Student Paper	%1
7	wwwnc.cdc.gov Internet Source	%1
8	www.julkari.fi Internet Source	%1
9	www.ncbi.nlm.nih.gov Internet Source	<%1
10	www.biomedcentral.com Internet Source	<%1
11	Submitted to University of Witwatersrand Student Paper	<%1
12	Kamng'ona, Arox W., Jason Hinds, Naor Bar-Zeev, Katherine A. Gould, Chrispin Chaguza, Chisomo Msefula, Jennifer E. Cornick, Benard W. Kulohoma, Katherine Gray, Stephen D. Bentley, Neil French, Robert S. Heyderman, and Dean B. Everett. "High multiple carriage and emergence of Streptococcus pneumoniae vaccine serotype variants in Malawian children", BMC Infectious Diseases, 2015. Publication	<%1

13	www.inspq.qc.ca Internet Source	<%1
14	bmcinfectdis.biomedcentral.com Internet Source	<%1
15	pneumonia.org.au Internet Source	<%1
16	bmcpediatr.biomedcentral.com Internet Source	<%1
17	nwhu.on.ca Internet Source	<%1
18	du Plessis, Mignon, Mushal Allam, Stefano Tempia, Nicole Wolter, Linda de Gouveia, Claire von Mollendorf, Keith A. Jolley, Nontombi Mbelle, Jeannette Wadula, Jennifer E. Cornick, Dean B. Everett, Lesley McGee, Robert F. Breiman, Rebecca A. Gladstone, Stephen D. Bentley, Keith P. Klugman, and Anne von Gottberg. "Phylogenetic Analysis of Invasive Serotype 1 Pneumococcus in South Africa, 1989 to 2013", Journal of Clinical Microbiology, 2016. Publication	<%1
19	pdfs.semanticscholar.org Internet Source	<%1
20	www.science.gov Internet Source	<%1
	Submitted to Nottingham Trent University	

21	Student Paper	<%1
22	www.circumpolarhealthjournal.net Internet Source	<%1
23	www.cdc.gov Internet Source	<%1
24	Submitted to National University of Singapore Student Paper	<%1
25	www.amedeo.com Internet Source	<%1
26	lra.le.ac.uk Internet Source	<%1
27	Submitted to University of Cape Town Student Paper	<%1
28	www.deepdyve.com Internet Source	<%1
29	www.activ-france.com Internet Source	<%1
30	www.rmpru.com Internet Source	<%1
31	www.mrc.ac.za Internet Source	<%1
32	projectes.camfic.cat Internet Source	<%1

33	Submitted to The University of Manchester Student Paper	<%1
34	www.ice.gov Internet Source	<%1
35	www.mrc.co.za Internet Source	<%1
36	rivm.nl Internet Source	<%1
37	issuu.com Internet Source	<%1
38	von Gottberg, Anne, Linda de Gouveia, Stefano Tempia, Vanessa Quan, Susan Meiring, Claire von Mollendorf, Shabir A. Madhi, Elizabeth R. Zell, Jennifer R. Verani, Katherine L. O'Brien, Cynthia G. Whitney, Keith P. Klugman, and Cheryl Cohen. "Effects of Vaccination on Invasive Pneumococcal Disease in South Africa", New England Journal of Medicine, 2014. Publication	<%1
39	www.wits.ac.za Internet Source	<%1
40	www.sajei.co.za Internet Source	<%1
41	erj.ersjournals.com Internet Source	<%1

42	Submitted to Imperial College of Science, Technology and Medicine Student Paper	<%1
43	www.biorxiv.org Internet Source	<%1
44	Submitted to University of Birmingham Student Paper	<%1
45	Ben-Shimol, S., N. Givon-Lavi, E. Leibovitz, S. Raiz, D. Greenberg, and R. Dagan. "Near elimination of otitis media caused by the PCV13 serotypes in Southern Israel shortly after sequential introduction of PCV7/PCV13", Clinical Infectious Diseases, 2014. Publication	<%1
46	del Amo, Eva, Cristina Esteva, Susanna Hernandez-Bou, Carmen Galles, Marian Navarro, Goretti Sauca, Alvaro Diaz, Paula Gassiot, Carmina Marti, Nieves Larrosa, Pilar Ciruela, Mireia Jane, Raquel Sá-Leão, and Carmen Muñoz-Almagro. "Serotypes and Clonal Diversity of Streptococcus pneumoniae Causing Invasive Disease in the Era of PCV13 in Catalonia, Spain", PLoS ONE, 2016. Publication	<%1
47	www.scielo.sa.cr Internet Source	<%1

48	Wolter, N., S. Tempia, C. Cohen, S. A. Madhi, M. Venter, J. Moyes, S. Walaza, B. Malope-Kgokong, M. Groome, M. du Plessis, V. Magomani, M. Pretorius, O. Hellferscee, H. Dawood, K. Kahn, E. Variava, K. P. Klugman, and A. von Gottberg. "High Nasopharyngeal Pneumococcal Density, Increased by Viral Coinfection, Is Associated With Invasive Pneumococcal Pneumonia", The Journal of Infectious Diseases, 2014. Publication	<%1
49	jjmicrobiol.com Internet Source	<%1
50	f1000.com Internet Source	<%1
51	www2.hu-berlin.de Internet Source	<%1
52	www.recentmedicalfindings.com Internet Source	<%1
53	bmcgenomics.biomedcentral.com Internet Source	<%1
54	posstrictosensu.iptsp.ufg.br Internet Source	<%1
55	jid.oxfordjournals.org Internet Source	<%1
56	www.visualmaterialculture.ac.nz Internet Source	<%1

57	www.google.nl Internet Source	<%1
58	tesisenxarxa.net Internet Source	<%1
59	Park, In Ho, K. Aaron Geno, Logan K. Sherwood, Moon H. Nahm, and Bernard Beall. "Population-Based Analysis of Invasive Nontypeable Pneumococci Reveals That Most Have Defective Capsule Synthesis Genes", PLoS ONE, 2014. Publication	<%1
60	Submitted to Oregon State University Student Paper	<%1
61	tampub.uta.fi Internet Source	<%1
62	pneumonia.biomedcentral.com Internet Source	<%1
63	Mothibeli, K.M.. "Distribution of factor H binding protein beyond serogroup B: Variation among five serogroups of invasive Neisseria meningitidis in South Africa", Vaccine, 20110303 Publication	<%1
64	www2.kenes.com Internet Source	<%1
65	clinicaltrials.gov Internet Source	<%1

66	Boyles, Salynn Hattersley, Andrew T. "Street Youths Should Be Focus of HBV Vaccine Efforts.(Statistical Data Included)", Vaccine Weekly, Oct 25 1999 Issue Publication	<%1
67	Submitted to Cardiff University Student Paper	<%1
68	run.unl.pt Internet Source	<%1
69	xipe.insp.mx Internet Source	<%1
70	Wolter, N., C. Cohen, S. Tempia, S. A. Madhi, M. Venter, J. Moyes, S. Walaza, B. M. Kgokong, M. Groome, M. du Plessis, M. Pretorius, H. Dawood, K. Kahn, E. Variava, K. P. Klugman, and A. von Gottberg. "HIV and Influenza Infection Are Associated with Increased Blood Pneumococcal Load, South Africa, 2009-2011", The Journal of Infectious Diseases, 2013. Publication	<%1
71	von Gottberg, Anne, Cheryl Cohen, Linda de Gouveia, Susan Meiring, Vanessa Quan, Andrew Whitelaw, Penny Crowther-Gibson, Shabir A. Madhi, Cynthia G. Whitney, and Keith P. Klugman. "Epidemiology of invasive pneumococcal disease in the pre-conjugate vaccine era: South Africa, 2003–2008",	<%1

	Vaccine, 2013. Publication	
72	apropos.mcw.edu Internet Source	<%1
73	f1000research.com Internet Source	<%1
74	Submitted to University of Warwick Student Paper	<%1
75	www.surv.esr.cri.nz Internet Source	<%1
76	www.nicd.ac.za Internet Source	<%1
77	www.eurosurveillance.org Internet Source	<%1
78	Submitted to nyack college Student Paper	<%1
79	Submitted to East Tennessee State University Student Paper	<%1
80	www.rhdhv.co.za Internet Source	<%1
81	Cohen, Cheryl, Sibongile Walaza, Jocelyn Moyes, Michelle Groome, Stefano Tempia, Marthi Pretorius, Orienka Hellferscee, Halima Dawood, Meera Chhagan, Fathima Naby, Summaya Haffejee, Ebrahim Variava,	<%1

	Kathleen Kahn, Susan Nzenze, Akhona Tshangela, Anne von Gottberg, Nicole Wolter, Adam L. Cohen, Babatyi Kgokong, Marietjie Venter, and Shabir A. Madhi. "Epidemiology of Viral-Associated Acute Lower Respiratory Tract Infection among Children <5 Years of Age in a High HIV Prevalence Setting, South Africa, 2009-2012 :", The Pediatric Infectious Disease Journal, 2014. Publication	
82	Steven Black. "Hemagglutination Inhibition Antibody Titers as a Correlate of Protection for Inactivated Influenza Vaccines in Children :", The Pediatric Infectious Disease Journal, 12/2011 Publication	<%1
83	hivthisweek.unaids.org Internet Source	<%1
84	etd.lsu.edu Internet Source	<%1
85	amedeo.com Internet Source	<%1
86	iai.asm.org Internet Source	<%1
87	haveyoursay.hamilton.govt.nz Internet Source	<%1
	bioinformatics.net.au	

88	Internet Source	<%1
89	docs.di.fc.ul.pt Internet Source	<%1
90	Munoz-Almagro, C.. "Emergence of invasive pneumococcal disease caused by multidrug-resistant serotype 19A among children in Barcelona", Journal of Infection, 200908 Publication	<%1
91	Song, Joon Young, Moon H. Nahm, and M. Allen Moseley. "Clinical Implications of Pneumococcal Serotypes: Invasive Disease Potential, Clinical Presentations, and Antibiotic Resistance", Journal of Korean Medical Science, 2013. Publication	<%1
92	kce.fgov.be Internet Source	<%1
93	Langenberger, David (Akademische Oberrätin Dr. Kay Nieselt, Prof. Dr. Peter F. Stadler and Universität Leipzig, Fakultät für Mathematik und Informatik). "High-throughput sequencing and small non-coding RNAs", Universitätsbibliothek Leipzig, 2013. Publication	<%1
94	pub.uni-bielefeld.de Internet Source	<%1
	onlinelibrary.wiley.com	

95	Internet Source	<%1
96	www.ahcmedia.com Internet Source	<%1
97	Submitted to University of Florida Student Paper	<%1
98	B. Beall. "Pre- and Postvaccination Clonal Compositions of Invasive Pneumococcal Serotypes for Isolates Collected in the United States in 1999, 2001, and 2002", Journal of Clinical Microbiology, 03/01/2006 Publication	<%1
99	eprints.qut.edu.au Internet Source	<%1
100	journals.lww.com Internet Source	<%1
101	conservancy.umn.edu Internet Source	<%1
102	etheses.bham.ac.uk Internet Source	<%1
103	Waure, Chiara de. "Effectiveness of 7-valent pneumococcal conjugate vaccine: A meta-analysis of post-marketing studies", World Journal of Meta-Analysis, 2015. Publication	<%1
104	evaluation.nrf.ac.za Internet Source	<%1

105	dyuthi.cusat.ac.in Internet Source	<%1
106	193.113.179.203 Internet Source	<%1
107	arno.unimaas.nl Internet Source	<%1
108	F. Mahjoub-Messai. "Population Snapshot of Streptococcus pneumoniae Serotype 19A Isolates before and after Introduction of Seven-Valent Pneumococcal Vaccination for French Children", Journal of Clinical Microbiology, 01/14/2009 Publication	<%1
109	G. Brick. "Non-susceptibility trends and serotype distributions among Streptococcus pneumoniae from community-acquired respiratory tract infections and from bacteraemias in the UK and Ireland, 1999 to 2007", Journal of Antimicrobial Chemotherapy, 11/01/2008 Publication	<%1
110	www.docstoc.com Internet Source	<%1
111	boris.unibe.ch Internet Source	<%1
112	"Pneumococcal Mastoiditis in Children and the Emergence of Multidrug-Resistant	<%1

Serotype 19A Isolates", PEDIATRICS, 07/01/2008
Publication

113	mro.massey.ac.nz Internet Source	<%1
114	Reinert, R.. "Pneumococcal disease caused by serotype 19A: Review of the literature and implications for future vaccine development", Vaccine, 20100611 Publication	<%1
115	Sharma, Dolly, Wendy Baughman, Amy Holst, Stephanie Thomas, Delois Jackson, Maria da Gloria Carvalho, Bernard Beall, Sarah Satola, Robert Jerris, Shabnam Jain, Monica M. Farley, and J. Pekka Nuorti. "Pneumococcal Carriage and Invasive Disease in Children Before Introduction of the 13-valent Conjugate Vaccine : Comparison With the Era Before 7-valent Conjugate Vaccine", The Pediatric Infectious Disease Journal, 2013. Publication	<%1
116	Wiese, Andrew D. Grijalva, Carlos G. Zhu. "Changes in childhood pneumonia hospitalizations by race and sex associated with pneumococcal conjuga", Emerging Infectious Diseases, June 2016 Issue Publication	<%1
117	www.jourlib.org	

	Internet Source	<%1
118	S. H. Yeh. "Immunogenicity and Safety of 13-Valent Pneumococcal Conjugate Vaccine in Infants and Toddlers", PEDIATRICS, 09/01/2010 Publication	<%1
119	on behalf of the Group for Enteric, Respiratory, Meningeal Disease Surveillance in South Africa (GERMS-SA). "Invasive disease due to Haemophilus influenzae serotype b ten years after routine vaccination, South Africa, 2003-2009", Vaccine, 20120111 Publication	<%1
120	von Mollendorf, Claire Cohen, Cheryl Tem. "Epidemiology of serotype 1 invasive pneumococcal disease, South Africa, 2003-2013.(RESEARCH)(Report)", Emerging Infectious Diseases, Feb 2016 Issue Publication	<%1
121	www.plospathogens.org Internet Source	<%1
122	mic.sgmjournals.org Internet Source	<%1
123	wiredspace.wits.ac.za Internet Source	<%1
	helvia.uco.es	

124	Internet Source	<%1
125	jcm.highwire.org Internet Source	<%1
126	Sahni, Vanita Naus, Monika Hoang, Linda . "The epidemiology of invasive pneumococcal disease in British Columbia following implementation of an", Canadian Journal of Public Health, Jan-Feb 2012 Issue Publication	<%1
127	jb.asm.org Internet Source	<%1
128	Submitted to Universiti Putra Malaysia Student Paper	<%1
129	Andrade, A.L.. "Pneumococcal disease manifestation in children before and after vaccination: What's new?", Vaccine, 20110914 Publication	<%1
130	von Mollendorf, C., C. Cohen, L. de Gouveia, V. Quan, S. Meiring, C. Feldman, K. P. Klugman, and A. von Gottberg. "Factors associated with ceftriaxone non-susceptibility of Streptococcus pneumoniae: analysis of South African national surveillance data, 2003-2010", Antimicrobial Agents and Chemotherapy, 2014. Publication	<%1

131	Reingold, A. Hadler, J. Faley, M.M. Harr. "Invasive pneumococcal disease in children 5 years after conjugate vaccine introduction--eight states", Morbidity and Mortality Weekly Report, Feb 15 2008 Issue Publication	<%1
132	Nzenze, Susan; Klugman, Keith and Madhi, Shabir. "The potential impact of the 13-valent conjugate pneumococcal vaccine on antibiotic resistance in pneumococci", Southern African Journal of Epidemiology & Infection, 2011. Publication	<%1
133	Vestrheim, D.F.. "Indirect effect of conjugate pneumococcal vaccination in a 2+1 dose schedule", Vaccine, 20100302 Publication	<%1
134	Pittet, L. F., and K. M. Posfay-Barbe. "Pneumococcal vaccines for children: a global public health priority : Pneumococcal vaccines for children", Clinical Microbiology and Infection, 2012. Publication	<%1
135	Ener Cagri Dinleyici. "Pneumococcal conjugated vaccine: PHiD-CV", Expert Review of Anti-infective Therapy, 11/2009 Publication	<%1
136	Mohale, Thabo, Nicole Wolter, Mushal Allam, Kedibone Ndlangisa, Penny Crowther-Gibson,	<%1

Mignon du Plessis, and Anne von Gottberg. "Genomic analysis of nontypeable pneumococci causing invasive pneumococcal disease in South Africa, 2003–2013", BMC Genomics, 2016.
Publication

137	annals.org Internet Source	<%1

EXCLUDE QUOTES	ON	EXCLUDE MATCHES	OFF
EXCLUDE BIBLIOGRAPHY	ON		

9 References

Reference List

(1) Geno KA, Saad JS, Nahm MH. Discovery of Novel Pneumococcal Serotype 35D, a Natural WciG-Deficient Variant of Serotype 35B. J Clin Microbiol 2017 May;55(5):1416-25.

(2) Bentley SD, Aanensen DM, Mavroidi A, Saunders D, Rabbinowitsch E, Collins M, et al. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. PLoS Genet 2006 Mar;2(3):e31.

(3) Park IH, Kim KH, Andrade AL, Briles DE, McDaniel LS, Nahm MH. Nontypeable pneumococci can be divided into multiple cps types, including one type expressing the novel gene pspK. MBio 2012 Apr 24;3(3):e00035-12.

(4) Hathaway LJ, Stutzmann MP, Battig P, Aebi S, Muhlemann K. A homologue of aliB is found in the capsule region of nonencapsulated *Streptococcus pneumoniae*. J Bacteriol 2004 Jun;186(12):3721-9.

(5) Hilty M, Wuthrich D, Salter SJ, Engel H, Campbell S, Sa-Leao R, et al. Global phylogenomic analysis of nonencapsulated *Streptococcus pneumoniae* reveals a deep-branching classic lineage that is distinct from multiple sporadic lineages. Genome Biol Evol 2014 Dec 4;6(12):3281-94.

(6) Waite RD, Penfold DW, Struthers JK, Dowson CG. Spontaneous sequence duplications within capsule genes cap8E and tts control phase variation in *Streptococcus pneumoniae* serotypes 8 and 37. Microbiology 2003 Feb;149(Pt 2):497-504.

(7) Schaffner TO, Hinds J, Gould KA, Wuthrich D, Bruggmann R, Kuffer M, et al. A point mutation in cpsE renders *Streptococcus pneumoniae* nonencapsulated and enhances its growth, adherence and competence. BMC Microbiol 2014 Aug 28;14:210. doi: 10.1186/s12866-014-0210-x.:210-0210.

(8) Park IH, Geno KA, Sherwood LK, Nahm MH, Beall B. Population-based analysis of invasive nontypeable pneumococci reveals that most have defective capsule synthesis genes. PLoS One 2014 May 15;9(5):e97825.

(9) Van Dyke MK, PirÃ§on JY, Cohen R, Madhi SA, RosenblÃ¼t A, Macias PM, et al. Etiology of Acute Otitis Media in Children Less Than 5 Years of Age: A Pooled Analysis of 10 Similarly Designed Observational Studies. Pediatr Infect Dis J 2017 Mar;36(3):274-81.

(10) Iroh Tam PY, Thielen BK, Obaro SK, Brearley AM, Kaizer AM, Chu H, et al. Childhood pneumococcal disease in Africa - A systematic review and meta-analysis of incidence, serotype distribution, and antimicrobial susceptibility. Vaccine 2017 Mar 8;(17):10.

(11) von Gottberg A., de Gouveia L., Tempia S, Quan V, Meiring S, von Mollendorf C., et al. Effects of vaccination on invasive pneumococcal disease in South Africa. N Engl J Med 2014 Nov 13;371(20):1889-99.

(12) Meiring S, Cohen C, Quan V, de Gouveia L., Feldman C, Karstaedt A, et al. HIV Infection and the Epidemiology of Invasive Pneumococcal Disease (IPD) in South African Adults and Older Children Prior to the Introduction of a Pneumococcal Conjugate Vaccine (PCV). PLoS One 2016 Feb 10;11(2):e0149104.

(13) Geno KA, Gilbert GL, Song JY, Skovsted IC, Klugman KP, Jones C, et al. Pneumococcal Capsules and Their Types: Past, Present, and Future. Clin Microbiol Rev 2015 Jul;28(3):871-99.

(14) Hausdorff WP, Bryant J, Kloek C, Paradiso PR, Siber GR. The contribution of specific pneumococcal serogroups to different disease manifestations: implications for conjugate vaccine formulation and use, part II. Clin Infect Dis 2000 Jan;30(1):122-40.

(15) Hausdorff WP, Feikin DR, Klugman KP. Epidemiological differences among pneumococcal serotypes. Lancet Infect Dis 2005 Feb;5(2):83-93.

(16) Chaves F, Campelo C, Sanz F, Otero JR. Meningitis due to mixed infection with penicillin-resistant and penicillin-susceptible strains of *Streptococcus pneumoniae*. J Clin Microbiol 2003 Jan;41(1):512-3.

(17) de Andrade AL, Pimenta FC, Laval CA, de Andrade JG, Guerra ML, Brandileone MC. Invasive pneumococcal infection in a healthy infant caused by two different serotypes. J Clin Microbiol 2004 May;42(5):2345-6.

(18) Jordens JZ, Paul J, Bates J, Beaumont C, Kimari J, Gilks C. Characterization of *Streptococcus pneumoniae* from human immunodeficiency virus-seropositive patients with acute and recurrent pneumonia. J Infect Dis 1995 Oct;172(4):983-7.

(19) Hausdorff WP, Bryant J, Paradiso PR, Siber GR. Which pneumococcal serogroups cause the most invasive disease: implications for conjugate vaccine formulation and use, part I. Clin Infect Dis 2000 Jan;30(1):100-21.

(20) Fenoll A, Granizo JJ, Aguilar L, Gimenez MJ, Aragoneses-Fenoll L, Hanquet G, et al. Temporal trends of invasive *Streptococcus pneumoniae* serotypes and antimicrobial resistance patterns in Spain from 1979 to 2007. J Clin Microbiol 2009 Apr;47(4):1012-20.

(21) 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 Rep 2010 Dec 10;59(RR-11):1-18.

(22) Madhi SA, Cohen C, von Gottberg A. Introduction of pneumococcal conjugate vaccine into the public immunization program in South Africa: translating research into policy. Vaccine 2012 Sep 7;30 Suppl 3:C21-7. doi: 10.1016/j.vaccine.2012.05.055.:C21-C27.

(23) 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 1;201(1):32-41.

(24) Miller E, Andrews NJ, Waight PA, Slack MP, George RC. Herd immunity and serotype replacement 4 years after seven-valent pneumococcal conjugate vaccination in England and Wales: an observational cohort study. Lancet Infect Dis 2011 Oct;11(10):760-8.

(25) Tyrrell GJ, Lovgren M, Chui N, Minion J, Garg S, Kellner JD, et al. Serotypes and antimicrobial susceptibilities of invasive *Streptococcus pneumoniae* pre- and post-seven valent pneumococcal conjugate vaccine introduction in Alberta, Canada, 2000-2006. Vaccine 2009 Jun 2;27(27):3553-60.

(26) Munoz-Almagro C, Jordan I, Gene A, Latorre C, Garcia-Garcia JJ, Pallares R. Emergence of invasive pneumococcal disease caused by nonvaccine serotypes in the era of 7-valent conjugate vaccine. Clin Infect Dis 2008 Jan 15;46(2):174-82.

(27) Pai R, Moore MR, Pilishvili T, Gertz RE, Whitney CG, Beall B. Postvaccine genetic structure of *Streptococcus pneumoniae* serotype 19A from children in the United States. J Infect Dis 2005 Dec 1;192(11):1988-95.

(28) van der Linden M, Falkenhorst G, Perniciaro S, Imohl M. Effects of Infant Pneumococcal Conjugate Vaccination on Serotype Distribution in Invasive Pneumococcal Disease among Children and Adults in Germany. PLoS One 2015 Jul 1;10(7):e0131494.

(29) Aguiar SI, Brito MJ, Goncalo-Marques J, Melo-Cristino J, Ramirez M. Serotypes 1, 7F and 19A became the leading causes of pediatric invasive pneumococcal infections in Portugal after 7 years of heptavalent conjugate vaccine use. Vaccine 2010 Jul;%19;28(32):5167-73.

(30) Beall B, McEllistrem MC, Gertz RE, Jr., Wedel S, Boxrud DJ, Gonzalez AL, et al. Pre- and postvaccination clonal compositions of invasive pneumococcal serotypes for isolates collected in the United States in 1999, 2001, and 2002. J Clin Microbiol 2006 Mar;44(3):999-1017.

(31) Scott JR, Hanage WP, Lipsitch M, Millar EV, Moulton LH, Hinds J, et al. Pneumococcal sequence type replacement among American Indian children: a comparison of pre- and routine-PCV7 eras. Vaccine 2012 Mar 16;30(13):2376-81.

(32) Beall BW, Gertz RE, Hulkower RL, Whitney CG, Moore MR, Brueggemann AB. Shifting genetic structure of invasive serotype 19A pneumococci in the United States. J Infect Dis 2011 May 15;203(10):1360-8.

(33) 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 Feb 2;3(1):ofw020.

(34) Nakano S, Fujisawa T, Ito Y, Chang B, Suga S, Noguchi T, et al. Serotypes, antimicrobial susceptibility, and molecular epidemiology of invasive and non-invasive *Streptococcus pneumoniae* isolates in paediatric patients after the introduction of 13-

valent conjugate vaccine in a nationwide surveillance study conducted in Japan in 2012-2014. Vaccine 2016 Jan 2;34(1):67-76.

(35) Dagan R, Givon-Lavi N, Leibovitz E, Greenberg D, Porat N. Introduction and proliferation of multidrug-resistant *Streptococcus pneumoniae* serotype 19A clones that cause acute otitis media in an unvaccinated population. J Infect Dis 2009 Mar 15;199(6):776-85.

(36) Choi EH, Kim SH, Eun BW, Kim SJ, Kim NH, Lee J, et al. *Streptococcus pneumoniae* serotype 19A in children, South Korea. Emerg Infect Dis 2008 Feb;14(2):275-81.

(37) Harboe ZB, Benfield TL, Valentiner-Branth P, Hjuler T, Lambertsen L, Kaltoft M, et al. Temporal trends in invasive pneumococcal disease and pneumococcal serotypes over 7 decades. Clin Infect Dis 2010 Feb 1;50(3):329-37.

(38) Jefferies JM, Smith AJ, Edwards GF, McMenamin J, Mitchell TJ, Clarke SC. Temporal analysis of invasive pneumococcal clones from Scotland illustrates fluctuations in diversity of serotype and genotype in the absence of pneumococcal conjugate vaccine. J Clin Microbiol 2010 Jan;48(1):87-96.

(39) Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A 1998 Mar 17;95(6):3140-5.

(40) 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 11):3049-60.

(41) Brueggemann AB, Pai R, Crook DW, Beall B. Vaccine escape recombinants emerge after pneumococcal vaccination in the United States. PLoS Pathog 2007 Nov;3(11):e168.

(42) Hiller NL, Ahmed A, Powell E, Martin DP, Eutsey R, Earl J, et al. Generation of genic diversity among *Streptococcus pneumoniae* strains via horizontal gene transfer during a chronic polyclonal pediatric infection. PLoS Pathog 2010 Sep 16;6(9):e1001108.

(43) Jefferies JM, Smith A, Clarke SC, Dowson C, Mitchell TJ. Genetic analysis of diverse disease-causing pneumococci indicates high levels of diversity within serotypes and capsule switching. J Clin Microbiol 2004 Dec;42(12):5681-8.

(44) Olarte L, Kaplan SL, Barson WJ, Romero JR, Lin PL, Tan TQ, et al. Emergence of Multidrug-Resistant Pneumococcal Serotype 35B among Children in the United States. J Clin Microbiol 2017 Mar;55(3):724-34.

(45) Croucher NJ, Harris SR, Fraser C, Quail MA, Burton J, van der Linden M, et al. Rapid pneumococcal evolution in response to clinical interventions. Science 2011 Jan 28;331(6016):430-4.

(46) Metcalf BJ, Chochua S, Gertz RE, Jr., Li Z, Walker H, Tran T, et al. Using whole genome sequencing to identify resistance determinants and predict antimicrobial resistance phenotypes for year 2015 invasive pneumococcal disease isolates recovered in the United States. Clin Microbiol Infect 2016 Dec;22(12):1002.

(47) Huebner RE, Klugman KP, Matai U, Eggers R, Hussey G. Laboratory surveillance for *Haemophilus influenzae* type B, meningococcal, and pneumococcal disease. Haemophilus Surveillance Working Group. S Afr Med J 1999 Sep;89(9):924-5.

(48) von Gottberg A., Cohen C, de Gouveia L., Meiring S, Quan V, Whitelaw A, et al. Epidemiology of invasive pneumococcal disease in the pre-conjugate vaccine era: South Africa, 2003-2008. Vaccine 2013 Aug 28;31(38):4200-8.

(49) Carvalho MG, 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.

(50) AUSTRIAN R. The quellung reaction, a neglected microbiologic technique. Mt Sinai J Med 1976 Nov;43(6):699-709.

(51) Clinical and laboratory Standards Institute (formerly NCCLS). Performance Standards for Antimicrobial Susceptibility Testing. Clinical and Laboratory Institute,Wayne, Pennsylvania, USA; 2013. Report No.: M100-S14 [ISBN1-56238-516-X].

(52) Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. Reordering contigs of draft genomes using the Mauve aligner. Bioinformatics 2009 Aug 15;25(16):2071-3.

(53) Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics 2014 Jul 15;30(14):2068-9.

(54) 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.

(55) Jefferies JM, Smith AJ, Edwards GF, McMenamin J, Mitchell TJ, Clarke SC. Temporal analysis of invasive pneumococcal clones from Scotland illustrates fluctuations in diversity of serotype and genotype in the absence of pneumococcal conjugate vaccine. J Clin Microbiol 2010 Jan;48(1):87-96.

(56) Choi EH, Kim SH, Eun BW, Kim SJ, Kim NH, Lee J, et al. *Streptococcus pneumoniae* Serotype 19A in Children, South Korea. Emerg Infect Dis 2008 Feb;14(2):275-81.

(57) Dagan R, Givon-Lavi N, Leibovitz E, Greenberg D, Porat N. Introduction and proliferation of multidrug-resistant *Streptococcus pneumoniae* serotype 19A clones that cause acute otitis media in an unvaccinated population. J Infect Dis 2009 Mar 15;199(6):776-85.

(58) Fenoll A, Granizo JJ, Aguilar L, Gimenez MJ, Aragoneses-Fenoll L, Hanquet G, et al. Temporal Trends of Invasive *Streptococcus pneumoniae* Serotypes and Antimicrobial Resistance Patterns in Spain from 1979 to 2007. J Clin Microbiol 2009 Apr;47(4):1012-20.

(59) Multilocus sequence typing. *Streptococcus pneumoniae* advance query form available at: http://spneumoniae.mlst.net/sql/burstspadvanced.asp, accessed 6 June. 2013.

(60) 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.

(61) du Plessis M., Allam M, Tempia S, Wolter N, de Gouveia L., von Mollendorf C., et al. Phylogenetic Analysis of Invasive Serotype 1 Pneumococcus in South Africa, 1989 to 2013. J Clin Microbiol 2016 May;54(5):1326-34.

(62) Smith AM, Klugman KP. Three predominant clones identified within penicillin-resistant South African isolates of *Streptococcus pneumoniae*. Microb Drug Resist 1997;3(4):385-9.

(63) Mothibeli KM, du Plessis M., von Gottberg A., de Gouveia L., Adrian P, Madhi SA, et al. An unusual pneumococcal sequence type is the predominant cause of serotype 3 invasive disease in South Africa. J Clin Microbiol 2010 Jan;48(1):184-91.

(64) Leimkugel J, Adams FA, Gagneux S, Pfluger V, Flierl C, Awine E, et al. An outbreak of serotype 1 *Streptococcus pneumoniae* meningitis in northern Ghana with features that are characteristic of *Neisseria meningitidis* meningitis epidemics. J Infect Dis 2005 Jul 15;192(2):192-9.

(65) Porat N, Soley C, Marengolciene MM, Greenberg D, Givon-Lavi N, Trefler R, et al. An international serotype 3 clone causing pediatric noninvasive infections in Israel, Costa Rica, and Lithuania. Pediatr Infect Dis J 2008 Aug;27(8):709-12.

(66) Wasfy MO, Pimentel G, Abdel-Maksoud M, Russell KL, Barrozo CP, Klena JD, et al. Antimicrobial susceptibility and serotype distribution of *Streptococcus pneumoniae*

causing meningitis in Egypt, 1998-2003. J Antimicrob Chemother 2005 Jun;55(6):958-64.

(67) Jolley KA, Maiden MC. *Streptococcus pneumoniae* MLST Databases available at http://pumlst.org/spneumoniae accessed 11 January 2017. 2017.

(68) Corcoran M, Vickers I, Fitzgerald M, Mereckiene J, Muchan S, Cotter S, et al. The persitance of serotype 19A - despite introduction of PCV13 vaccine. 10th International Symposium on Pneumococci and Pneumococcal Diseases , 75. 2016.

(69) Picazo J, Ruiz-Contreras J, Hernandez B, Sanz F, Gutierrez A, Cercenado E, et al. Clonal and clinical profile of Streptococcus pneumoniae serotype 19A causing pediatric invasive infections: a 2-year (2007-2009) laboratory-based surveillance in Madrid. Vaccine 2011 Feb 17;29(9):1770-6.

(70) Mahjoub-Messai F, Doit C, Koeck JL, Billard T, Evrard B, Bidet P, et al. Population snapshot of *Streptococcus pneumoniae* serotype 19A isolates before and after introduction of seven-valent pneumococcal Vaccination for French children. J Clin Microbiol 2009 Mar;47(3):837-40.

(71) Shin J, Baek JY, Kim SH, Song JH, Ko KS. Predominance of ST320 among *Streptococcus pneumoniae* serotype 19A isolates from 10 Asian countries. J Antimicrob Chemother 2011 May;66(5):1001-4.

(72) Antonio M, Dada-Adegbola H, Biney E, Awine T, O'Callaghan J, Pfluger V, et al. Molecular epidemiology of pneumococci obtained from Gambian children aged 2-29

months with invasive pneumococcal disease during a trial of a 9-valent pneumococcal conjugate vaccine. BMC Infect Dis 2008 Jun 11;8:81. doi: 10.1186/1471-2334-8-81.:81-8.

(73) 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 Mar 8;11(3):e0151125.

(74) Moore CE, Giess A, Soeng S, Sar P, Kumar V, Nhoung P, et al. Characterisation of Invasive *Streptococcus pneumoniae* Isolated from Cambodian Children between 2007 - 2012. PLoS One 2016 Jul 22;11(7):e0159358.

(75) Ndlangisa KM, du Plessis M., Wolter N, de Gouveia L., Klugman KP, von Gottberg A. Population snapshot of *Streptococcus pneumoniae* causing invasive disease in South Africa prior to introduction of pneumococcal conjugate vaccines. PLoS One 2014 Sep 18;9(9):e107666.

(76) Statistics South Africa.StatsOnline. P0302 - Mid-year population estimates, (http://www .statssa .gov .za/publications/P0302/ P03022013 .pdf). 2013.

(77) Gertz RE, Jr., Li Z, Pimenta FC, Jackson D, Juni BA, Lynfield R, et al. Increased penicillin nonsusceptibility of nonvaccine-serotype invasive pneumococci other than serotypes 19A and 6A in post-7-valent conjugate vaccine era. J Infect Dis 2010 Mar;201(5):770-5.

(78) Gonzalez BE, Hulten KG, Lamberth L, Kaplan SL, Mason EO, Jr. *Streptococcus pneumoniae serogroups* 15 and 33: an increasing cause of pneumococcal infections in children in the United States after the introduction of the pneumococcal 7-valent conjugate vaccine. Pediatr Infect Dis J 2006 Apr;25(4):301-5.

(79) Kaur R, Casey JR, Pichichero ME. Emerging *Streptococcus pneumoniae* Strains Colonizing the Nasopharynx in Children After 13-valent Pneumococcal Conjugate Vaccination in Comparison to the 7-valent Era, 2006-2015. Pediatr Infect Dis J 2016 Aug;35(8):901-6.

(80) Chochua S, Metcalf BJ, Li Z, Walker H, Tran T, McGee L, et al. Invasive Serotype 35B Pneumococci Including an Expanding Serotype Switch Lineage, United States, 2015-2016. Emerg Infect Dis 2017 Jun;23(6):922-30.

(81) Nunes MC, von Gottberg A., de Gouveia L., Cohen C, Moore DP, Klugman KP, et al. The impact of antiretroviral treatment on the burden of invasive pneumococcal disease in South African children: a time series analysis. AIDS 2011 Feb;%20;25(4):453-62.

(82) Antonio M, Hakeem I, Awine T, Secka O, Sankareh K, Nsekpong D, et al. Seasonality and outbreak of a predominant *Streptococcus pneumoniae* serotype 1 clone from The Gambia: expansion of ST217 hypervirulent clonal complex in West Africa. BMC Microbiol 2008 Nov 17;8:198. doi: 10.1186/1471-2180-8-198.:198.

(83) Brueggemann AB, Spratt BG. Geographic distribution and clonal diversity of *Streptococcus pneumoniae* serotype 1 isolates. J Clin Microbiol 2003 Nov;41(11):4966-70.

(84) Brueggemann AB, Muroki BM, Kulohoma BW, Karani A, Wanjiru E, Morpeth S, et al. Population Genetic Structure of *Streptococcus pneumoniae* in Kilifi, Kenya, Prior to the Introduction of Pneumococcal Conjugate Vaccine. PLoS One 2013 Nov 25;8(11):e81539.

(85) 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 May 11;11(5):e0153602.

(86) Clarke SC, Scott KJ, McChlery SM. Serotypes and sequence types of pneumococci causing invasive disease in Scotland prior to the introduction of pneumococcal conjugate polysaccharide vaccines. J Clin Microbiol 2004 Oct;42(10):4449-52.

(87) Miernyk KM, Bulkow LR, Case SL, Zulz T, Bruce MG, Harker-Jones M, et al. Population structure of invasive *Streptococcus pneumoniae* isolates among Alaskan children in the conjugate vaccine era, 2001 to 2013. Diagn Microbiol Infect Dis 2016 Oct;86(2):224-30.

(88) Kamng'ona AW, Hinds J, Bar-Zeev N, Gould KA, Chaguza C, Msefula C, et al. High multiple carriage and emergence of *Streptococcus pneumoniae* vaccine serotype variants in Malawian children. BMC Infect Dis 2015 Jun;%20;15:234. doi: 10.1186/s12879-015-0980-2.:234-0980.

(89) Kandasamy R, Gurung M, Thapa A, Ndimah S, Adhikari N, Murdoch DR, et al. Multi-serotype pneumococcal nasopharyngeal carriage prevalence in vaccine naive Nepalese children, assessed using molecular serotyping. PLoS One 2015 Feb 2;10(2):e0114286.

(90) Satzke C, Turner P, Virolainen-Julkunen A, Adrian PV, Antonio M, Hare KM, et al. Standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*: updated recommendations from the World Health Organization Pneumococcal Carriage Working Group. Vaccine 2013 Dec 17;32(1):165-79.

(91) Pimenta FC, Roundtree A, Soysal A, Bakir M, du Plessis M, Wolter N, et al. Sequential triplex real-time PCR assay for detecting 21 pneumococcal capsular serotypes that account for a high global disease burden. J Clin Microbiol 2013 Feb;51(2):647-52.

(92) Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics 2015 Jul;20. pii: btv421.:btv421.

(93) Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 2014 May 1;30(9):1312-3.

(94) PubMLST database. http://pubmlst.org/. Accessed Jul 2016. 2016.

(95) 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 16;5(7):e11638.

(96) Kaltoft MS, Skov Sorensen UB, Slotved HC, Konradsen HB. An easy method for detection of nasopharyngeal carriage of multiple *Streptococcus pneumoniae* serotypes. J Microbiol Methods 2008 Dec;75(3):540-4.

(97) Cohen C, Naidoo N, Meiring S, de Gouveia L., von Mollendorf C., Walaza S, et al. *Streptococcus pneumoniae* Serotypes and Mortality in Adults and Adolescents in South Africa: Analysis of National Surveillance Data, 2003 - 2008. PLoS One 2015 Oct 13;10(10):e0140185.

(98) Brueggemann AB, Peto TE, Crook DW, Butler JC, Kristinsson KG, Spratt BG. Temporal and geographic stability of the serogroup-specific invasive disease potential of *Streptococcus pneumoniae* in children. J Infect Dis 2004 Oct 1;190(7):1203-11.

(99) Donkor ES, Bishop CJ, Gould K, Hinds J, Antonio M, Wren B, et al. High levels of recombination among *Streptococcus pneumoniae* isolates from the Gambia. MBio 2011 Jun 21;2(3):e00040-11.

(100) Weiser JN, Bae D, Epino H, Gordon SB, Kapoor M, Zenewicz LA, et al. Changes in availability of oxygen accentuate differences in capsular polysaccharide expression by phenotypic variants and clinical isolates of *Streptococcus pneumoniae*. Infect Immun 2001 Sep;69(9):5430-9.

(101) Hammerschmidt S, Wolff S, Hocke A, Rosseau S, Muller E, Rohde M. Illustration of pneumococcal polysaccharide capsule during adherence and invasion of epithelial cells. Infect Immun 2005 Aug;73(8):4653-67.

(102) Jolley KA, Maiden MC. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 2010 Dec 10;11:595. doi: 10.1186/1471-2105-11-595.:595-11.

(103) Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics 2011 Aug 8;12:402. doi: 10.1186/1471-2164-12-402.:402-12.

(104) Scott JR, Hinds J, Gould KA, Millar EV, Reid R, Santosham M, et al. Nontypeable pneumococcal isolates among Navajo and White Mountain Apache communities: are these really a cause of invasive disease? J Infect Dis 2012 Jul 1;206(1):73-80.

(105) Salter SJ, Hinds J, Gould KA, Lambertsen L, Hanage WP, Antonio M, et al. Variation at the capsule locus, cps, of mistyped and non-typable *Streptococcus pneumoniae* isolates. Microbiology 2012 Jun;158(Pt 6):1560-9.

(106) Mohale T, Wolter N, Allam M, Ndlangisa K, Crowther-Gibson P, du Plessis M., et al. Genomic analysis of nontypeable pneumococci causing invasive pneumococcal disease in South Africa, 2003-2013. BMC Genomics 2016 Jun 22;17:470. doi: 10.1186/s12864-016-2808-x.:470-2808.

(107) Morona JK, Paton JC, Miller DC, Morona R. Tyrosine phosphorylation of CpsD negatively regulates capsular polysaccharide biosynthesis in *Streptococcus pneumoniae*. Mol Microbiol 2000 Mar;35(6):1431-42.

(108) Morona JK, Miller DC, Morona R, Paton JC. The effect that mutations in the conserved capsular polysaccharide biosynthesis genes cpsA, cpsB, and cpsD have on virulence of *Streptococcus pneumoniae*. J Infect Dis 2004 May 15;189(10):1905-13.

(109) Kolkman MA, Wakarchuk W, Nuijten PJ, van der Zeijst BA. Capsular polysaccharide synthesis in *Streptococcus pneumoniae* serotype 14: molecular analysis of the complete cps locus and identification of genes encoding glycosyltransferases required for the biosynthesis of the tetrasaccharide subunit. Mol Microbiol 1997 Oct;26(1):197-208.

(110) Melchiorre S, Camilli R, Pietrantoni A, Moschioni M, Berti F, Del GM, et al. Point mutations in wchA are responsible for the non-typability of two invasive *Streptococcus pneumoniae* isolates. Microbiology 2012 Feb;158(Pt 2):338-44.

(111) Shainheit MG, Valentino MD, Gilmore MS, Camilli A. Mutations in Pneumococcal cpsE Generated via In Vitro Serial Passaging Reveal a Potential Mechanism of Reduced Encapsulation Utilized by a Conjunctival Isolate. J Bacteriol 2015 May 15;197(10):1781-91.

(112) Madhi SA, Petersen K, Madhi A, Wasas A, Klugman KP. Impact of human immunodeficiency virus type 1 on the disease spectrum of Streptococcus pneumoniae in South African children. Pediatr Infect Dis J 2000 Dec;19(12):1141-7.

(113) Chewapreecha C, Harris SR, Croucher NJ, Turner C, Marttinen P, Cheng L, et al. Dense genomic sampling identifies highways of pneumococcal recombination. Nat Genet 2014 Mar;46(3):305-9.

(114) Alderson MR. Status of research and development of pediatric vaccines for *Streptococcus pneumoniae*. Vaccine 2016 Jun 3;34(26):2959-61.