

# A *Streptococcus pneumoniae* lineage usually associated with pneumococcal conjugate vaccine (PCV) serotypes is the most common cause of serotype 35B invasive disease in South Africa, following routine use of PCV

Kedibone M. Ndlangisa<sup>1,\*</sup>, Mignon du Plessis<sup>1,2</sup>, Stephanie Lo<sup>3</sup>, Linda de Gouveia<sup>1</sup>, Chrispin Chaguza<sup>3</sup>, Martin Antonio<sup>4</sup>, Brenda Kwambana-Adams<sup>5,6</sup>, Jennifer Cornick<sup>7</sup>, Dean B. Everett<sup>7,8</sup>, Ron Dagan<sup>9</sup>, Paulina A. Hawkins<sup>10,11</sup>, Bernard Beall<sup>11</sup>, Alejandra Corso<sup>12</sup>, Samanta Cristine Grassi Almeida<sup>13</sup>, Theresa J. Ochoa<sup>14</sup>, Stephen Obaro<sup>15</sup>, Sadia Shakoob<sup>16</sup>, Eric S. Donkor<sup>17</sup>, Rebecca A. Gladstone<sup>3</sup>, Pak Leung Ho<sup>18</sup>, Metka Paragi<sup>19</sup>, Sanjay Doiphode<sup>20</sup>, Somporn Srifuengfung<sup>21</sup>, Rebecca Ford<sup>22</sup>, Jennifer Moisi<sup>23</sup>, Samir K. Saha<sup>24</sup>, Godfrey Bigogo<sup>25</sup>, Betuel Sigauque<sup>26</sup>, Özgen Köseoglu Eser<sup>27</sup>, Naima Elmdaghri<sup>28</sup>, Leonid Titov<sup>29</sup>, Paul Turner<sup>30</sup>, K. L. Ravi Kumar<sup>31</sup>, Rama Kandasamy<sup>32</sup>, Ekaterina Egorova<sup>33</sup>, Margaret IP<sup>34</sup>, Robert F. Breiman<sup>35</sup>, Keith P. Klugman<sup>2,36</sup>, Lesley McGee<sup>11</sup>, Stephen D. Bentley<sup>3</sup>, Anne von Gottberg<sup>1,2</sup> and The Global Pneumococcal Sequencing Consortium<sup>37</sup>

## Abstract

Pneumococcal serotype 35B is an important non-conjugate vaccine (non-PCV) serotype. Its continued emergence, post-PCV7 in the USA, was associated with expansion of a pre-existing 35B clone (clonal complex [CC] 558) along with post-PCV13 emergence of a non-35B clone previously associated with PCV serotypes (CC156). This study describes lineages circulating among 35B isolates in South Africa before and after PCV introduction. We also compared 35B isolates belonging to a predominant 35B lineage in South Africa (GPSC5), with isolates belonging to the same lineage in other parts of the world. Serotype 35B isolates that caused invasive pneumococcal disease in South Africa in 2005–2014 were characterized by whole-genome sequencing (WGS). Multi-locus sequence types and global pneumococcal sequence clusters (GPSCs) were derived from WGS data of 63 35B isolates obtained in 2005–2014. A total of 262 isolates that belong to GPSC5 (115 isolates from South Africa and 147 from other countries) that were sequenced as part of the global pneumococcal sequencing (GPS) project were included for comparison. Serotype 35B isolates from South Africa were differentiated into seven GPSCs and GPSC5 was most common (49%, 31/63). While 35B was the most common serotype among GPSC5/CC172 isolates in South Africa during the PCV13 period (66%, 29/44), 23F was the most common serotype during both the pre-PCV (80%, 37/46) and PCV7 period (32%, 8/25). Serotype 35B represented 15% (40/262) of GPSC5 isolates within the global GPS database and 75% (31/40) were from South Africa. The predominance of the GPSC5 lineage within non-vaccine serotype 35B, is possibly unique to South Africa and warrants further molecular surveillance of pneumococci.

## DATA SUMMARY

Raw fastq data, assemblies and annotations for 262 samples that were sequenced as part of the Global Pneumococcal Sequencing project (GPS) were previously released [1] to the European Nucleotide Archive. Genome assemblies of the 22 serotype 35B isolates from South Africa that were not part of the GPS project has been deposited at GenBank under the BioProject PRJNA339372. Individual accessions for the samples are listed in Table S1 (available in the online version of this article) and are available at Figshare: <https://doi.org/10.6084/m9.figshare.16621984> [2].

## INTRODUCTION

Pneumococcal conjugate vaccines (PCV) have significantly reduced vaccine-serotype invasive pneumococcal disease (IPD) among vaccinated children in countries where PCV has been introduced [3–5]. Reduction of disease in unvaccinated children

**Impact Statement**

Increases in non-PCV serotypes due to expansion of pre-existing clones and emergence of capsule-switch strains have been reported in some countries following the introduction of PCV. In this study, we describe the epidemiology of non-vaccine serotype 35B in South Africa before and after the introduction of PCV. This serotype has become one of the emerging non-PCV serotypes in South Africa following PCV introduction and its dominance appears to be driven by a pre-existing lineage (GPSC5/CC172), which, in South Africa, was previously dominated by PCV serotype 23F prior to the introduction of PCV. These findings contribute to the general understanding of pneumococcal molecular epidemiology in the PCV era as well as the international effort to characterize replacement serotypes.

and adults because of herd effect has also been reported [3]. Despite the decrease, increases in IPD rates due to some non-PCV serotypes (serotype replacement) were observed in some countries following routine use of PCV [3–7]. Increases in non-PCV serotypes have been attributed to expansion of pre-existing clones; however, capsule-switch strains have also emerged that originated through replacement of capsular locus genes [8–11].

Serotype 35B is an important emerging non-vaccine serotype. Olarte *et al.* and Chochua *et al.* [12, 13] reported an increase in this serotype in the USA due to expansion of a pre-existing antibiotic-resistant clone CC558 (GPSC59), in addition to emergence of antibiotic-resistant CC156 (GPSC6), previously associated with PCV serotypes in the USA. Two different capsular switch events were described in the USA, occurring between 35B/CC558 donor strains and 9V/ST156 recipients, resulting in 35B/CC156 progeny [13, 14].

In South Africa, PCV7 was introduced in 2009 with a three-dose schedule at 6, 14 and 36 weeks of age and was replaced with PCV13 in 2011 using the same schedule. By 2012 (compared to 2005 through 2008), among children less than 2 years of age, the incidence of IPD due to PCV7 serotypes decreased by 89% and additional PCV13 serotypes, not in PCV7, by 57% [15]. In 2016, IPD rates of non-PCV (non-PCV13) serotypes among children <5 years of age increased 29% (4.3 to 5.6 per 100000 population) compared to 2005–2008 with non-PCV serotypes 8 and 35B as the most causes of IPD [16]. Serotype 35B IPD in children <5 years of age increased 384% (0.08 to 0.4 per 100000 population) in 2016 compared to 2005–2008.

In our previous study describing baseline genetic structure within pneumococcal serotypes in 2007, prior to routine PCV use in South Africa, we found that predominant sequence types circulating among some serotypes differed from predominant sequence

Received 15 July 2021; Accepted 22 November 2021; Published 06 April 2022

**Author affiliations:** <sup>1</sup>National Institute for Communicable Diseases (NICD), a division of the National Health Laboratory Service, Johannesburg, South Africa; <sup>2</sup>School of Pathology, University of the Witwatersrand, Johannesburg, South Africa; <sup>3</sup>Parasites and Microbes, Wellcome Sanger Institute, Hinxton, UK; <sup>4</sup>WHO Collaborating Centre for New Vaccines Surveillance, Medical Research Council Unit, The Gambia at London School of Hygiene & Tropical Medicine, Fajara, The Gambia; <sup>5</sup>NIHR Global Health Research Unit on Mucosal Pathogens, Division of Infection and Immunity, University College London, London, UK; <sup>6</sup>West Africa Partnerships and Strategies, Medical Research Council Unit The Gambia at The London School of Hygiene and Tropical Medicine, Fajara, Gambia; <sup>7</sup>Malawi-Liverpool-Wellcome-Trust, Blantyre, Malawi; <sup>8</sup>Centre for Inflammation Research, Queens Research Institute, University of Edinburgh, Edinburgh, UK; <sup>9</sup>The Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel; <sup>10</sup>Rollins School Public Health, Emory University, Atlanta, USA; <sup>11</sup>Centers for Disease Control and Prevention, Atlanta, USA; <sup>12</sup>Administración Nacional de Laboratorios e Institutos de Salud, Buenos Aires, Argentina; <sup>13</sup>Center of Bacteriology, Adolfo Lutz Institute, São Paulo, Brazil; <sup>14</sup>Instituto de Medicina Tropical, Universidad Peruana Cayetano Heredia, Lima, Peru; <sup>15</sup>University of Nebraska Medical Center, Omaha, USA; <sup>16</sup>The Aga Khan University, Karachi, Pakistan; <sup>17</sup>Department of Medical Microbiology, University of Ghana Medical School, Accra, Ghana; <sup>18</sup>Department of Microbiology and Carol Yu Centre for Infection, The University of Hong Kong, Queen Mary Hospital, Hong Kong, PR China; <sup>19</sup>National Laboratory of Health, Environment and Food, Ljubljana, Slovenia; <sup>20</sup>Hamad Medical Corporation, Doha, Qatar; <sup>21</sup>Faculty of Pharmacy, Siam University, Bangkok, Thailand; <sup>22</sup>Papua New Guinea Institute of Medical Research, Goroka, Papua New Guinea; <sup>23</sup>Agence de Médecine Préventive, Paris, France; <sup>24</sup>Child Health Research Foundation, Dhaka, Bangladesh; <sup>25</sup>Kenya Medical Research Institute, Kisumu, Kenya; <sup>26</sup>Centro de Investigação em Saúde da Manhica, Maputo, Moçambique; <sup>27</sup>Hacettepe University Faculty of Medicine, Department of Medical Microbiology, Ankara, Turkey; <sup>28</sup>Faculty of Medicine and Pharmacy & Ibn Rochd University Hospital Center, Casablanca, Morocco; <sup>29</sup>The Republican Research and Practical Center for Epidemiology and Microbiology, Minsk, Belarus; <sup>30</sup>Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, UK; <sup>31</sup>Kempegowda Institute of Medical Sciences Hospital & Research Center, Bangalore, India; <sup>32</sup>University of Oxford, and the NIHR Oxford Biomedical Research Centre, Oxford, UK; <sup>33</sup>G. N. Gabrichevsky Research Institute for Epidemiology and Microbiology, Moscow, Russia; <sup>34</sup>Department of Microbiology, Chinese University of Hong Kong, Hong Kong, PR China; <sup>35</sup>The Emory Global Health Institute, Atlanta, USA; <sup>36</sup>Hubert Department of Global Health, Rollins School of Public Health, and Division of Infectious Diseases, School of Medicine, Emory University, Atlanta, GA, USA; <sup>37</sup>The Global Pneumococcal Sequencing Project (<http://www.pneumogen.net/gps/>), UK.

\*Correspondence: Kedibone M. Ndlangisa, kedibonen@nicd.ac.za

**Keywords:** South Africa; *Streptococcus pneumoniae*; serotype 35B; global pneumococcal sequence cluster.

**Abbreviations:** CC, clonal complex; CLSI, Clinical and Laboratory Standards Institute; CSF, cerebrospinal fluid; GPS, Global Pneumococcal Sequencing project; GPSCs, global pneumococcal sequence clusters; IPD, invasive pneumococcal disease; MIC, minimum inhibitory concentration; NICD, National Institute for Communicable Diseases; BPP, Penicillin-binding protein; PCV, pneumococcal conjugate vaccine; ST, sequence type; WGS, whole genome sequencing.

**Data statement:** All supporting data, code and protocols have been provided within the article or through supplementary data files. Supplementary material is available with the online version of this article.

types identified globally [17]. Molecular surveillance of isolates from South Africa, particularly for emerging non-vaccine serotypes such as 35B, is therefore important as their emergence could be due to different lineages than those described elsewhere, and might be a sentinel for further geographic spread of this serotype. The aim of this study was to describe and compare lineages circulating among 35B isolates in South Africa before and after PCV introduction, and to compare isolates belonging to the predominant 35B lineage circulating in South Africa to those in other parts of the world.

## METHODS

### National IPD surveillance

Isolates were collected as part of GERMS-SA, a national, laboratory-based surveillance programme for IPD in all nine provinces in South Africa, initiated in 1999 [18]. In 2003, systematic collection of patient clinical data including outcome and HIV serological status from approximately 30 sentinel sites in all nine provinces commenced. Isolates and patient data for this study were from IPD cases reported from 2005 through 2014, and were submitted to the reference laboratory at the National Institute for Communicable Diseases (NICD). 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). In addition, for invasive disease rate calculations, cases included patients with normally sterile-site specimens testing positive by PCR [19], or bacterial latex antigen supported by Gram-stain microscopy.

### Serotyping and antimicrobial susceptibility testing

Serotypes of pneumococcal isolates were determined by the Quellung method using serotype-specific antisera (Statens Serum Institute, Copenhagen, Denmark) [20]. Isolates were phenotypically classified as serotype 35B if they bound to pool G, group 35 antiserum and factor sera 35 a, 35 c and 29b. Antimicrobial MIC 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. From 2009 to 2014, commercially prepared Sensititre-SASP2 panels (Trek Diagnostics, Cleveland, OH) were used for broth microdilution. Results were interpreted according to the 2014 Clinical and Laboratory Standards Institute (CLSI) guidelines and breakpoints [21]. Isolates with penicillin MICs  $\geq 0.12 \text{ mg l}^{-1}$  were considered non-susceptible to penicillin. For other antimicrobials, isolates were defined as non-susceptible if they were intermediately or fully resistant to the agent tested, according to CLSI guidelines. Multidrug resistance was defined as non-susceptibility to beta-lactams and at least two other classes of antimicrobials.

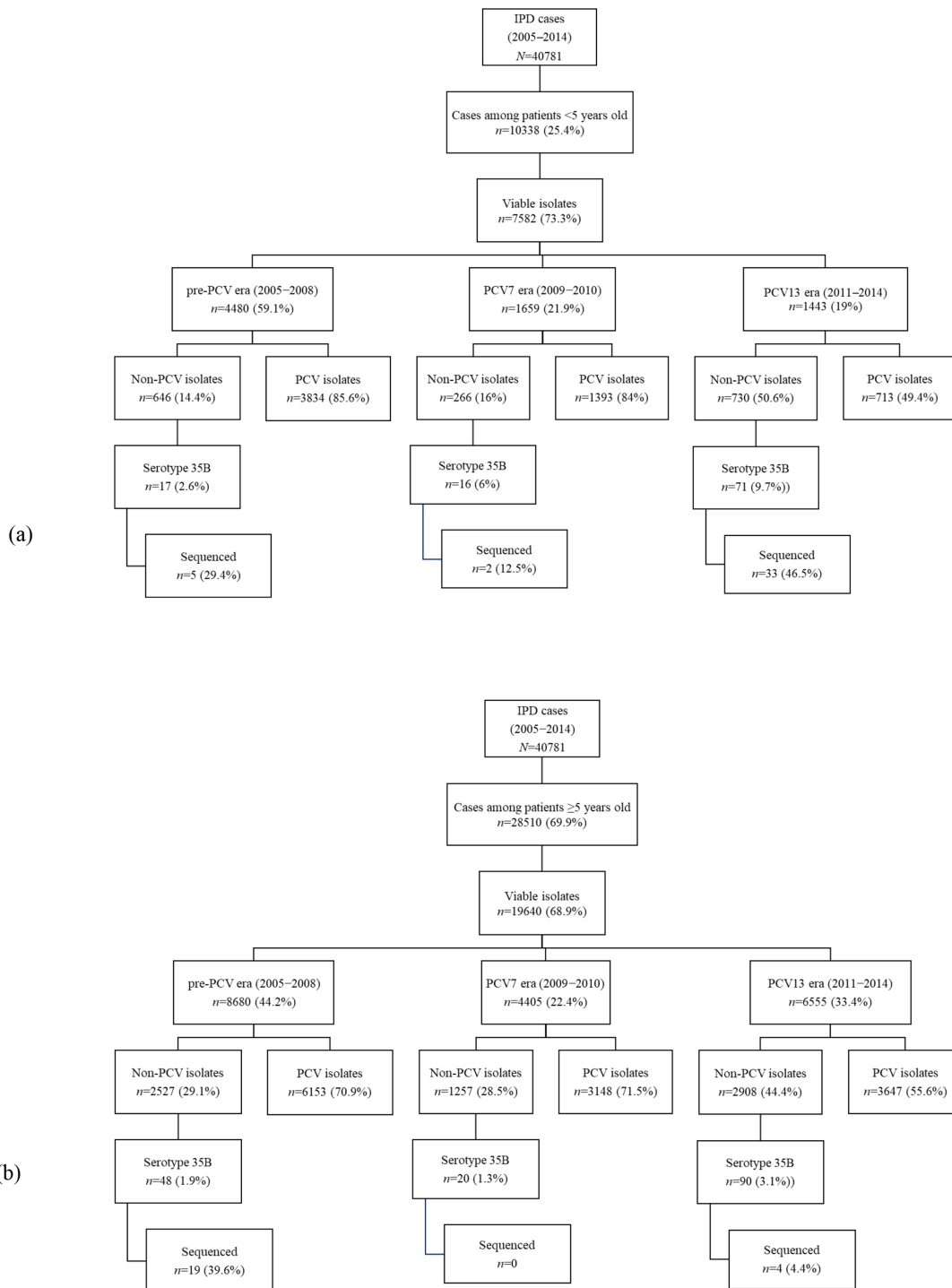
### Genetic characterization

A subset of isolates collected through our national IPD surveillance from South Africa in 2005 to 2014 were whole-genome sequenced as part of the Global Pneumococcal Sequence project (GPS) (<http://www.pneumogen.net/gps/>). The sampling strategy for the GPS project was as follows: 300 isolates per year, representing all serotypes and from patients of all ages in South Africa, were selected. Selection was random with respect to serotype and was stratified by age for each year: 150 isolates from children aged 0 to 2 years, 75 isolates from children aged 3 to 5 years and 75 isolates from individuals >5 years of age. As part of the GPS project, only 45 of the 262 isolates that were phenotypically determined to be serotype 35B were selected. We therefore randomly selected an additional 22 35B IPD isolates submitted during 2005 to 2008 (pre-PCV period) to increase the number of isolates for this study.

DNA was extracted from the 67 35B isolates using the QIAamp DNA mini kit (Qiagen, Venlo, Netherlands) and DNA extracts were quantified using the Qubit instrument and dsDNA BR Assay kit (Life Technologies, Carlsbad, CA, USA). 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.

Multi-locus sequence types (clonal complexes and sequence types), Global Pneumococcal Sequence Clusters (GPSCs), *in silico* serotypes, and resistant genotypes to penicillin, tetracycline, erythromycin, chloramphenicol and co-trimoxazole were derived from WGS data, as previously described [1, 14]. *In silico* serotype was determined from genome data using PneumoCaT and SeroBA [22]. Penicillin-binding protein (PBP) profiles were assigned based on transpeptidase domain amino acid sequences of 277–359 residues from PBPs 1a, 2b and 2x with three-number combination PBP genes referred to with standard nomenclature (pbp1a-pbp2b-pbp2x) [14]. A clonal complex was defined as a group of related STs sharing six of seven identical alleles with another ST in the group.

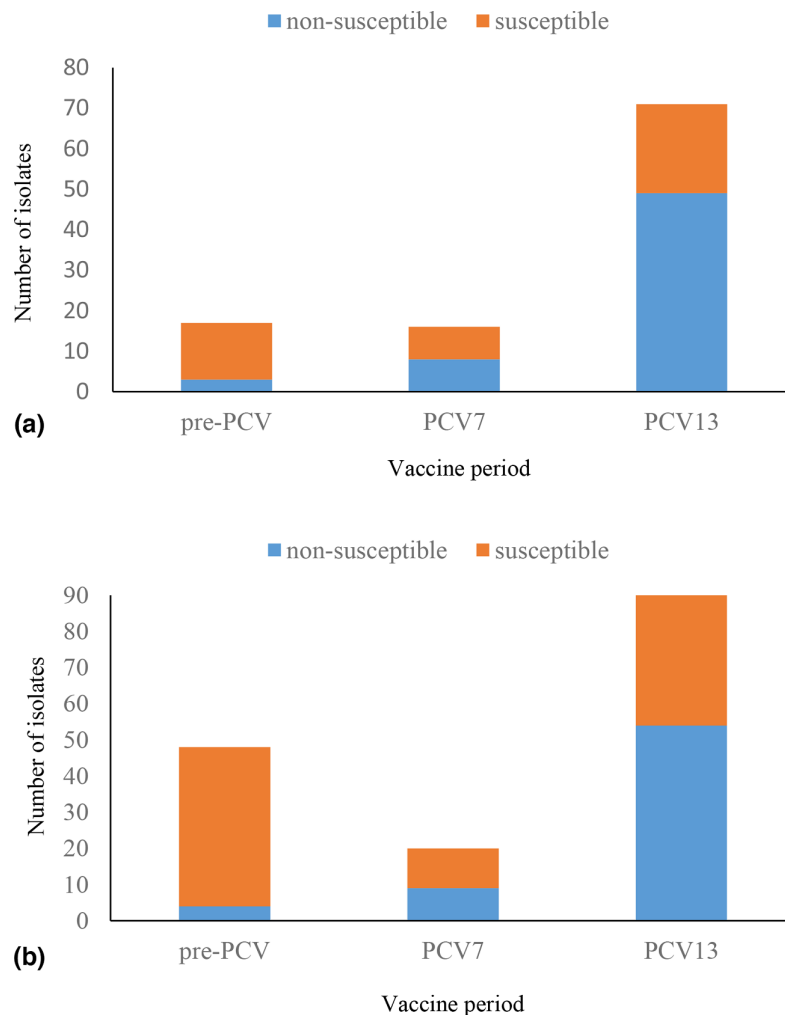
In addition to the 35B isolates from South Africa the study included 262 *S. pneumoniae* genomes representing GPSC5:115 from South Africa and 147 from other parts of the world that were sequenced as part of the larger GPS project [1]. GPSC5 is a predominant lineage within 35B in South Africa.



**Fig. 1.** Invasive pneumococcal disease cases and isolates obtained from 2005 through 2014 among South African (a) children <5 years and (b) individuals ≥5 years old.

### Statistical analysis

Average incidence for 2005–2008 (pre-PCV period), 2009–2010 (PCV7 period) and 2011–2014 (PCV13 period) was calculated by dividing the average number of cases by mid-year population estimates data from Statistics South Africa [23] and multiplying the quotient by 100000. Differences during the pre-PCV era, PCV7 period and PCV13 period, in prevalence and antimicrobial



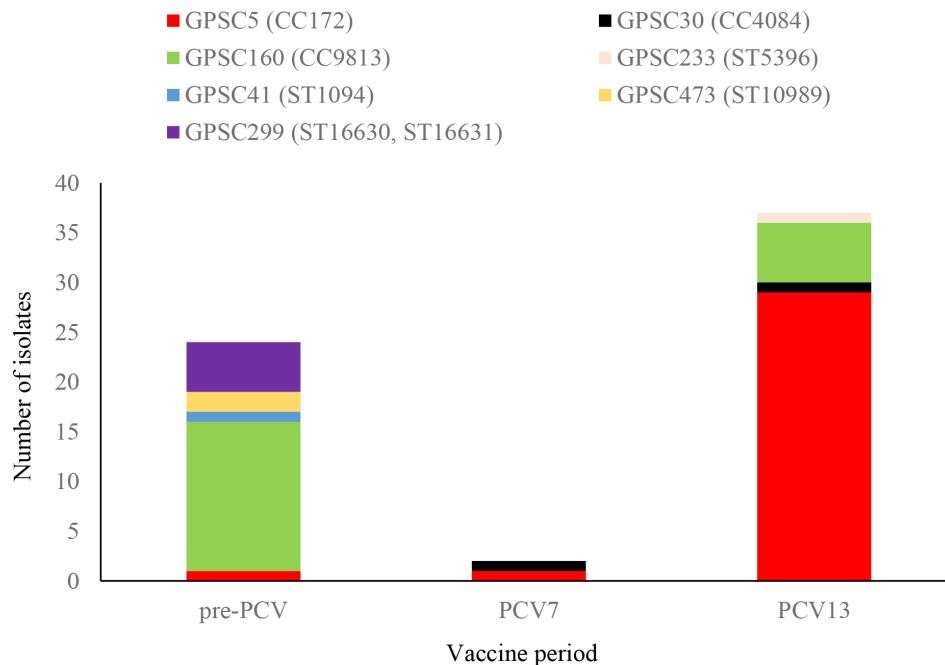
**Fig. 2.** Distribution of penicillin non-susceptible and susceptible serotype 35B pneumococcal isolates causing invasive disease in South Africa among (a) children <5 years old ( $N=104$ ) during the pre-PCV (2005–2008) ( $n=17$ ), PCV7 (2009–2010) ( $n=16$ ) and PCV13 (2011–2014) period ( $n=71$ ) and (b) patients  $\geq 5$  years old ( $N=158$ ) during the pre-PCV (2005–2008) ( $n=48$ ), PCV7 (2009–2010) ( $n=20$ ) and PCV13 (2011–2014) period ( $n=90$ ).

susceptibility of isolates belonging to clonal complexes/GPSCs were assessed using univariate analysis using Stata version 14 (StataCorp, College Station, USA).

## RESULTS

### Invasive pneumococcal disease surveillance, South Africa

During the study period, 40781 IPD cases were reported of which 28229 (69.2%) had viable isolates available for further characterization. Age was known for 27224 (96.4%) cases with viable isolates, of which 7582 (27.9%) were from children <5 years old (Fig. 1). Within this age group, 0.4% (17/4480), 1% (16/1659) and 4.9% (71/1443) of isolates obtained during the pre-PCV, PCV7 and PCV13 periods, respectively, were phenotypically identified as serotype 35B. Among individuals aged  $\geq 5$  years, serotype 35B represented 0.6% (48/8680), 0.5% (20/4405) and 1.3% (90/6555) of isolates during the pre-PCV, PCV7 and PCV13 periods, respectively. In total, 262 serotype 35B isolates were phenotypically identified among IPD isolates during the study period, from individuals of all ages in South Africa. Among children <5 years old the average incidence of serotype 35B increased by 290% (95% confidence interval [CI], 286% to 1485%) from 0.08 cases per 100000 person-years in the pre-PCV era to 0.3 cases per 100000 person-years during the PCV13 period. Among patients aged  $\geq 5$  years, the rate of serotype 35B IPD did not change (0.03 vs. 0.05 per 100000 person-years pre-PCV vs. PCV13 era). The proportion of penicillin non-susceptible serotype 35B isolates increased from 17.6% (3/17) during the pre-PCV era to 69% (49/71) during the PCV13 era among children <5 years old ( $P<0.001$ ), and from 8.3% (4/48) to 60% (54/90) among patients  $\geq 5$  years old ( $P<0.001$ ) (Fig. 2).



**Fig. 3.** Distribution of global pneumococcal sequence clusters (GPSCs), clonal complexes (CC) and sequence types (STs) among 63 serotype 35B isolates causing invasive disease in South Africa during the pre-PCV (2005–2008) ( $n=24$ ), PCV7 (2009–2010) ( $n=2$ ) and PCV13 (2011–2014) period ( $n=37$ ).

### Serotype 35B genotypes

Sixty-seven (26%) of the 262 isolates that were phenotypically identified as 35B were sequenced. *In silico* serotype assignment confirmed 63 of the 67 sequenced isolates to be serotype 35B (four isolates were identified as serotype 35D and were excluded from this analysis). The majority of 35B isolates were from children <5 years old (40/63, 63%). The 35B isolates were differentiated into seven GPSCs and belonged to three clonal complexes, namely CC172, CC9813 and CC4084, and five unrelated sequence types (Fig. 3). GPSC5 was the most common lineage and represented 49% (31/63) of the isolates, all of which were CC172. GPSC160 represented 33% (21/63) of the isolates, all of which were CC9813. Two isolates belonged to GPSC30 and were both CC4084. The remaining nine isolates belonged to four different GPSCs and five sequence types.

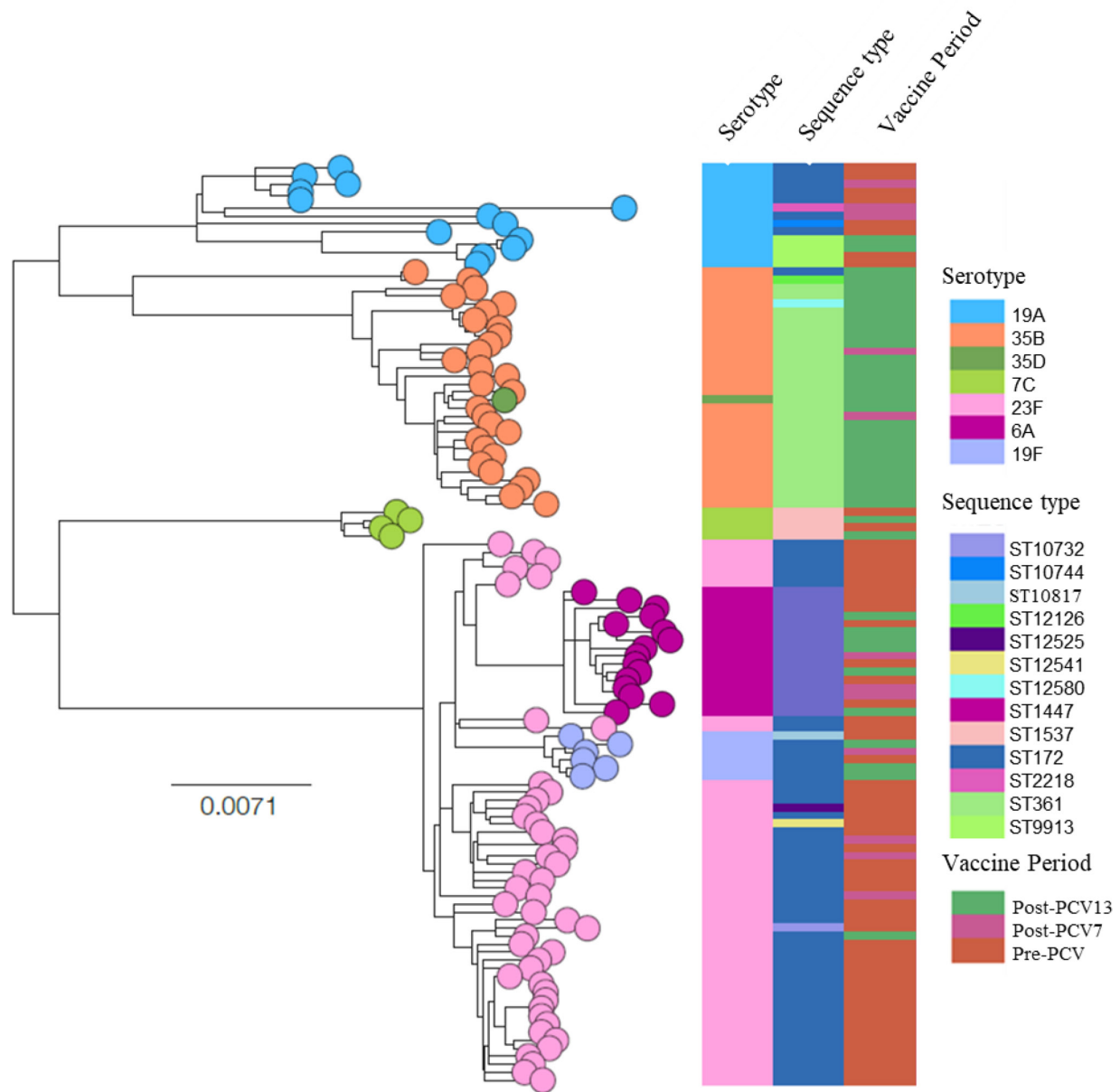
The majority of sequenced 35B isolates were from the PCV13 period (61%, 37/63). During this period, GPSC5 was represented by 78% (29/37) of the isolates, GPSC160 by 16% (6/37); GPSC30 and GPSC233 by one isolate each. Two isolates from the PCV7 era were sequenced, one belonged to GPSC5 and the other to GPSC30. Among 24 isolates from the pre-PCV period that were sequenced, GPSC160 was the most common cluster (63%, 15/24). One isolate belonged to GPSC5 (CC172) and the remaining eight isolates belonged to three different GPSCs.

### GPSC5 representation among all serotypes in South Africa

Among ~5000 IPD isolates from South Africa, from 2005 through 2014, that were sequenced, 115 isolates belonged to GPSC5 (Fig. 4). The isolates expressed one of seven serotypes, namely, 6A, 7C, 19A, 19F, 23F, 35B or 35D. While serotype 23F was the most common serotype among GPSC5 isolates during both the pre-PCV (80%, 37/46) and PCV7 periods (32%, 8/25), only 2% (1/44) of GPSC5 isolates were serotype 23F during the PCV13 period. During the PCV13 period, serotype 35B was the most common serotype among GPSC5 isolates (66%, 29/44).

### GPSC5 antibiotic resistance and penicillin-binding protein (*pbp*) profiles

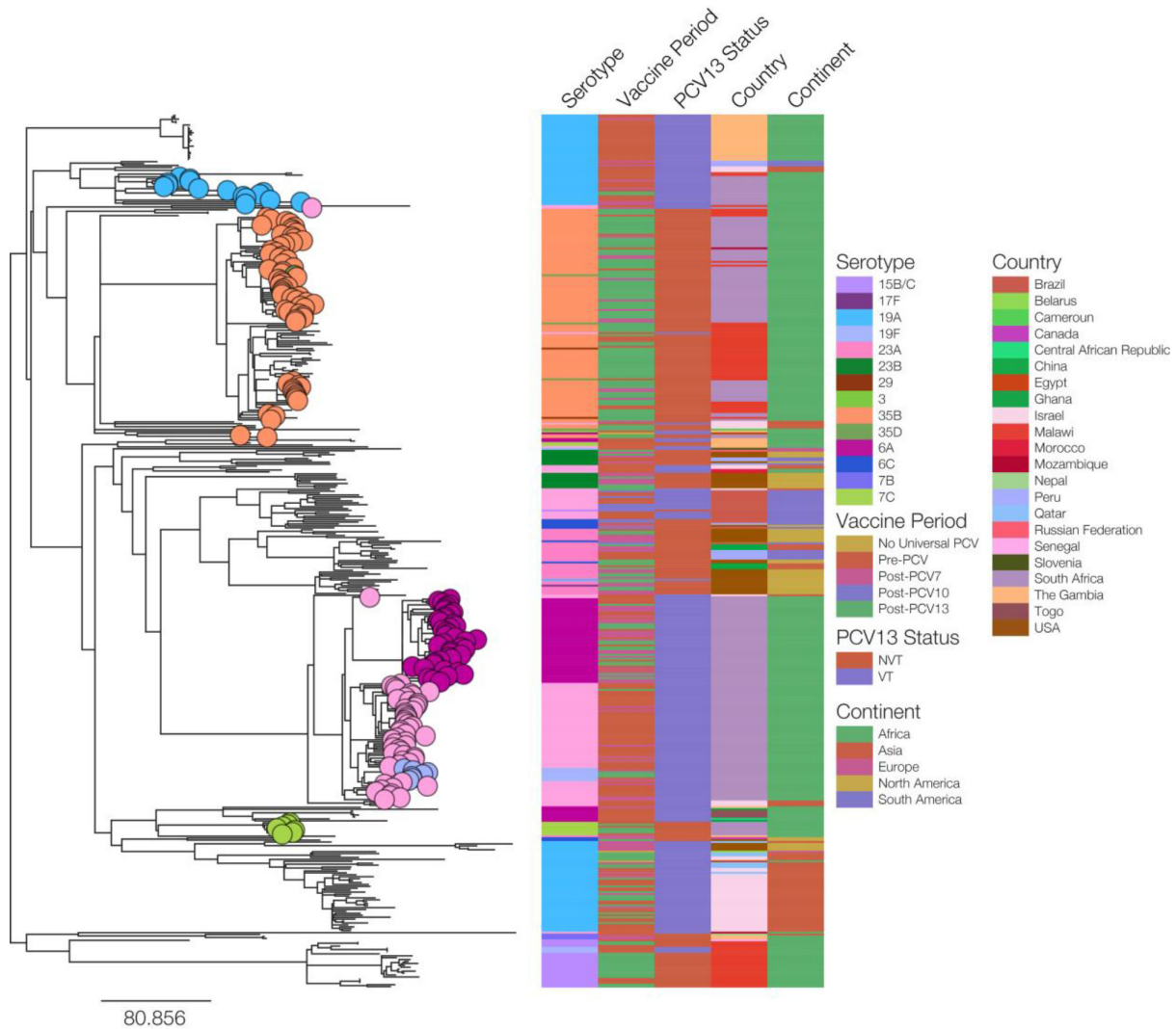
Almost all (97%, 111/115) GPSC5 isolates from South Africa were non-susceptible to penicillin ( $MIC_{50}$ : 0.25  $\mu\text{g ml}^{-1}$  and  $MIC_{90}$ : 1  $\mu\text{g ml}^{-1}$ ) and 18% (21/115) were multi-drug resistant. All serotype 35B isolates ( $n=31$ ) belonging to GPSC5 were non-susceptible to penicillin; none were multi-drug resistant. The serotype 35B GPSC5 isolates were classified into two *pbp* profiles (7-1-455 and 7-1-242). The majority of serotype 35B isolates (90%, 29/31) were 7-1-455. Neither profile was identified among non-35B GPSC5 isolates.



**Fig. 4.** Phylogeny of GPSC5 (CC172) isolates causing invasive disease in South Africa ( $N=115$ ) during the pre-PCV (2005–2008) ( $n=46$ ), PCV7 (2009–2010) ( $n=25$ ), and PCV13 (2011–2014) ( $n=44$ ) period. Leaf nodes on the tree are colour coded according to serotype. Microreact: <https://microreact.org/project/M-plmPZP5>.

### GPSC5 representation among the GPS global dataset

A total of 262 isolates belonging to GPSC5 were identified among IPD isolates in the GPS global dataset, from 2005 through 2014 (Fig. 5). The majority of GPSC5 isolates were from South Africa (44%, 115/262), followed by Israel (18%, 48/262), the USA (15%, 38/262), Brazil (7%, 19/262) and Malawi (5%, 12/262). Serotype 35B represented 15% (40/262) of GPSC5 isolates. The majority of serotype 35B isolates were from South Africa (78%, 31/40). The remaining nine serotype 35B isolates were from Malawi ( $n=4$ ), Israel ( $n=3$ ), Mozambique ( $n=1$ ) and Bangladesh ( $n=1$ ). The nine serotype 35B isolates not from South Africa were penicillin non-susceptible but none were multi-drug resistant. They were differentiated into three *pbp* profiles [7-1-455 ( $n=5$ ), 7-1-77 ( $n=3$ ) and 7-1-462]. Other than one serotype 23F isolate from Israel that was classified as *pbp* 7-1-77, the three profiles were not identified among other GPSC5 isolates.



**Fig. 5.** Phylogeny of GPSC5 isolates (N=262) causing pneumococcal disease globally, 2005–2014. Leaf nodes on the tree are colour coded according to serotype, only nodes for isolates from South Africa are shown.

## DISCUSSION

In this study we report an increase in the average incidence of serotype 35B IPD among children <5 years old in South Africa during the post-PCV13 period compared to pre-PCV. We identified GPSC5, a lineage usually associated with PCV serotypes, as the most common lineage among the non-PCV serotype 35B isolates in South Africa during the PCV13 period. While PCV serotype 23F was the most dominant serotype among isolates belonging to GPSC5 during the pre-PCV period, serotype 35B became the leading GPSC5/CC172 serotype during the PCV13 period.

Serotype 35B has been shown to be one of the non-PCV replacement serotypes following the implementation of PCV [12, 13, 24]. Demonstration of a high proportion (up to 90% in some locations) of serotype 35B isolates which are penicillin non-susceptible [24] makes replacement with serotype 35B a particular concern. In South Africa, where the incidence of serotype 35B invasive disease increased among children <5 years old during the post-PCV13 era compared to the pre-PCV era, the proportion of 35B isolates that are penicillin non-susceptible is 80%.

CC558 is reportedly the main serotype 35B genotype in Japan and in the USA and is associated with multi-drug resistance [13, 24–26]. Following PCV implementation in the USA CC558 has expanded [12, 13, 27]. In South Africa, CC558 (GPSC59) is rare among serotype 35B IPD isolates, instead, CC172 (GPSC5) is the most common lineage during the PCV13 period. Prior to PCV implementation in South Africa GPSC5 was more common among serotype 23F isolates however there was a shift towards serotype 35B as the main GPSC5 serotype during the PCV13 era. This lineage is usually associated with the PCV serotypes 6A, 6B, 23F and 19A [28–30]. The



detection of this lineage within isolates of diverse serotypes is in keeping with the higher than average propensity for recombination as measured using *r/m* previously (GPSC5 *r/m* 10.16, average 7.70) [1]. A recent analysis of 3233 IPD genomes in the GPS study, from six countries, in children <3 years of age, revealed the expansion of serotype 35B/D within GPSC5 in South Africa [31]. We corroborated this finding in our dataset, which included additional genomes from individuals of all ages. Antimicrobial resistance within this lineage could also be a key contributor to its expansion. Similar to other serotype 35B isolates, penicillin non-susceptibility was high among our 35B isolates belonging to GPSC5, which almost all had the same PBP type. GPSC5, also classified as PMEN26 (or Colombia<sup>23F</sup>-26), is an internationally disseminated penicillin-resistant clone that was first described in serotype 23F isolates in the 1990s (<https://www.pneumogen.net/pmen/index.html>). The propensity of this lineage to express different capsular types was demonstrated in the USA with the emergence of non-vaccine serotype 23A isolates associated with the 23F global clone, post-PCV7 introduction [32].

Within the global GPS dataset, isolates belonging to GPSC5 were almost exclusively detected (89%, 232/262) on the African continent and all 35B/35D isolates within this lineage were from Africa [31]. This finding is in agreement with our previous findings [17, 33] that, for some serotypes in South Africa, predominant genotypes are not the same as those circulating in other parts of the world.

Limitations to this study include the fact that not all serotype 35B isolates were sequenced and therefore the serotype 35B genetic population structure described here may not be an accurate reflection of all genotypes circulating in South Africa. Nonetheless, among a collection of over 1000 35B isolates on PubMLST, only seven belong to CC172 [34], supporting our suggestion that the predominance of the GPSC5 lineage within serotype 35B is unique to South Africa and the African continent.

The predominance of GPSC5/CC172, previously associated with serotype 23F in South Africa, among 35B isolates, following PCV introduction in South Africa highlights the importance of molecular surveillance of pneumococci.

#### Funding information

IPD Surveillance was funded by the United States Agency for International Development's Antimicrobial Resistance Initiative (cooperative agreement numbers U60/CCU022088, U62/CCU022901 and 5U2GPS001328) and whole genome sequencing by the Bill and Melinda Gates Foundation (grant code OPP1034556).

#### Acknowledgements

We would like to thank all members of the Global Pneumococcal Sequencing Consortium for their contributions to create this rich global dataset. We are also grateful for the technical support from Wellcome Sanger Institute sequencing facility and Pathogen Informatics team. For the purpose of Open Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. Global Pneumococcal Sequencing Consortium members: Diego Faccione, Paula Galletti, Abdullah W Brooks, Md Hasanuzzaman, Roly Malaker, Alexander Davydov, Maria Cristina de Cunto Brandileone, Pierra Law, Chunjiang Zhao, Hui Wang, Jeremy Keenan, Balaji Veeraraghavan, Geetha Nagaraj, Noga Givon-Lavi, Nurit Porat, Rachel Benisty, Jennifer Verani, Anmol Kiran, Maaiké Alaerts, Shamala Devi Sekaran, Stuart C Clarke, Houria Belabbès, Idrissa Diawara, Khalid Zerouali, Benild Moiane, Helio Mucavele, Andrew J Pollard, Philip E Carter, Imran Nisar, Deborah Lehmann, Anna Skoczynska, Ewa Sadowy, Waleria Hryniewicz, Elena Voropaeva, Yulia Urban, Tamara Kastrin, Mushal Ali, Nicole Wolter, Shabir A. Madhi, Susan A Nzenze, Ebenezer Foster-Nyarko, Ebrima Bojang, Peggy-Estelle Tientcheu, Michele Nurse-Lucas, Patrick E Akpaka, Alison Maguire, David Aanensen, Leon Bentley, Jyothish N Nair Thulasee Bhai, Nicholas Croucher, Rafal Mostowj, John A Lees, Rebecca Henderson, David Cleary.

#### Conflicts of interest

Professor Anne von Gottberg and Dr Mignon du Plessis received funding from Pfizer Vaccines Research and Sanofi Pasteur. Dr Gladstone reports PhD studentship from Pfizer outside the submitted work (2009–2012). The other authors declare that they have no conflicts of interest directly relating to this manuscript.

#### Ethical statement

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

#### References

- Gladstone RA, Lo SW, Lees JA, Croucher NJ, van Tonder AJ, et al. International genomic definition of pneumococcal lineages, to contextualise disease, antibiotic resistance and vaccine impact. *EBioMedicine* 2019;43:338–346.
- Ndlangisa KM, du Plessis M, Lo S, de Gouveia L, Chaguza C, et al. A *Streptococcus pneumoniae* lineage usually associated with pneumococcal conjugate vaccine (PCV) serotypes is the most common cause of serotype 35B invasive disease in South Africa, following routine use of PCV. *Figshare* 2022. DOI: 10.6084/m9.figshare.16621984.
- 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;11:760–768.
- Tyrrell GJ, Lovgren M, Chui N, Minion J, Garg S, 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;27:3553–3560.
- Pilishvili T, Lexau C, Farley MM, Hadler J, Harrison LH, et al. Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. *J Infect Dis* 2010;201:32–41.
- Muñoz-Almagro C, Jordan I, Gene A, Latorre C, Garcia-Garcia JJ, et al. Emergence of invasive pneumococcal disease caused by nonvaccine serotypes in the era of 7-valent conjugate vaccine. *Clin Infect Dis* 2008;46:174–182.
- Pai R, Moore MR, Pilishvili T, Gertz RE, Whitney CG, et al. Postvaccine genetic structure of *Streptococcus pneumoniae* serotype 19A from children in the United States. *J Infect Dis* 2005;192:1988–1995.
- Beall BW, Gertz RE, Hulkower RL, Whitney CG, Moore MR, et al. Shifting genetic structure of invasive serotype 19A pneumococci in the United States. *J Infect Dis* 2011;203:1360–1368.
- Beall B, McEllistrem MC, Gertz RE, Wedel S, Boxrud DJ, 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;44:999–1017.

10. 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;3:ofw020.
11. Scott JR, Hanage WP, Lipsitch M, Millar EV, Moulton LH, et al. Pneumococcal sequence type replacement among American Indian children: A comparison of pre- and routine-PCV7 eras. *Vaccine* 2012;30:2376–2381.
12. Olarte L, Kaplan SL, Barson WJ, Romero JR, Lin PL, et al. Emergence of Multidrug-Resistant Pneumococcal Serotype 35B among Children in the United States. *J Clin Microbiol* 2017;55:724–734.
13. Chochua S, Metcalf BJ, Li Z, Walker H, Tran T, et al. Invasive serotype 35B pneumococci including an expanding serotype switch lineage, United States, 2015–2016. *Emerg Infect Dis* 2017;23:922–930.
14. Metcalf BJ, Gertz RE, Gladstone RA, Walker H, Sherwood LK, et al. Strain features and distributions in pneumococci from children with invasive disease before and after 13-valent conjugate vaccine implementation in the USA. *Clin Microbiol Infect* 2016;22:60.
15. von Gottberg A, de Gouveia L, Tempia S, Quan V, Meiring S, et al. Effects of vaccination on invasive pneumococcal disease in South Africa. *N Engl J Med* 2014;371:1889–1899.
16. von Gottberg A, Kleynhans J, de Gouveia L, Tempia S, Meiring S, et al. Trends in invasive pneumococcal disease among children aged 5 years, South Africa, 2005–2016; Oral presentation; 11th International Symposium on Pneumococci & Pneumococcal Diseases. Melbourne, Australia, 2018.
17. Ndlangisa KM, du Plessis M, Wolter N, de Gouveia L, Klugman KP, et al. Population snapshot of *Streptococcus pneumoniae* causing invasive disease in South Africa prior to introduction of pneumococcal conjugate vaccines. *PLoS One* 2014;9:e107666.
18. Huebner RE, Klugman KP, Matai U, Eggers R, Hussey G, et al. Laboratory surveillance for *Haemophilus influenzae* type B meningococcal, and pneumococcal disease. Haemophilus Surveillance Working Group. *S Afr Med J* 1999;89:924–925.
19. Carvalho M da G, Pimenta FC, Gertz RE Jr, Joshi HH, Trujillo AA, et al. PCR-based quantitation and clonal diversity of the current prevalent invasive serogroup 6 pneumococcal serotype, 6C, in the United States in 1999 and 2006 to 2007. *J Clin Microbiol* 2009;47:554–559.
20. Austrian R. The quellung reaction, a neglected microbiologic technique. *Mt Sinai J Med* 1976;43:699–709.
21. Clinical and Laboratory Standards Institute (formerly NCCLS). Performance Standards for Antimicrobial Susceptibility Testing. Nineteenth informational supplement. 2014.
22. Lo SW, Gladstone RA, van Tonder AJ, Hawkins PA, Kwambana-Adams B, et al. Global Distribution of Invasive Serotype 35D *Streptococcus pneumoniae* isolates following Introduction of 13-Valent Pneumococcal Conjugate Vaccine. *J Clin Microbiol* 2018;56:e00228-18.
23. Statistics South Africa. Statistical release. P0302. Mid-year population estimates; (n.d.). <http://www.statssa.gov.za/publications/P0302/P03022014.pdf>
24. Miyazaki H, Shibuya R, Midorikawa N, Chang B, Ohnishi M, et al. Serotype distribution and antimicrobial susceptibility of *Streptococcus pneumoniae* strains isolated in Japan after introduction of the routine immunization program. *J Infect Chemother* 2017;23:234–240.
25. Beall B, McEllistrem MC, Gertz RE, Boxrud DJ, Besser JM, et al. Emergence of a novel penicillin-nonsusceptible, invasive serotype 35B clone of *Streptococcus pneumoniae* within the United States. *J Infect Dis* 2002;186:118–122.
26. Beall B, Chochua S, Gertz RE Jr, Li Y, Li Z, et al. A population-based descriptive atlas of invasive pneumococcal strains recovered Within the U.S. During 2015–2016. *Front Microbiol* 2018;9:2670.
27. Richter SS, Diekema DJ, Heilmann KP, Dohrn CL, Riahi F, et al. Changes in pneumococcal serotypes and antimicrobial resistance after introduction of the 13-valent conjugate vaccine in the United States. *Antimicrob Agents Chemother* 2014;58:6484–6489.
28. Donkor ES, Bishop CJ, Gould KA, Hinds J, Antonio M, et al. High levels of recombination among *Streptococcus pneumoniae* isolates from the Gambia. *mBio* 2011;2:e00040-11.
29. Gagetti P, Faccione D, Reijtman V, Fossati S, Rodriguez M, et al. Characterization of *Streptococcus pneumoniae* invasive serotype 19A isolates from Argentina (1993–2014). *Vaccine* 2017;35:4548–4553.
30. Rudolph K, Bruce MG, Bulkow L, Zulz T, Reasonover A, et al. Molecular epidemiology of serotype 19A *Streptococcus pneumoniae* among invasive isolates from Alaska, 1986–2010. *Int J Circumpolar Health* 2013;72.
31. Lo SW, Gladstone RA, van Tonder AJ, Lees JA, du Plessis M, et al. Pneumococcal lineages associated with serotype replacement and antibiotic resistance in childhood invasive pneumococcal disease in the post-PCV13 era: an international whole-genome sequencing study. *Lancet Infect Dis* 2019;19:759–769.
32. Pai R, Gertz RE, Whitney CG, Beall B. Clonal association between *Streptococcus pneumoniae* serotype 23A, circulating within the United States, and an internationally dispersed clone of serotype 23F. *J Clin Microbiol* 2005;43:5440–5444.
33. Mthibeli KM, du Plessis M, von Gottberg A, de Gouveia L, Adrian P, et al. An unusual pneumococcal sequence type is the predominant cause of serotype 3 invasive disease in South Africa. *J Clin Microbiol* 2010;48:184–191.
34. Multilocus sequence typing. *Streptococcus pneumoniae* MLST website available at; (n.d.). <https://pubmlst.org/spneumoniae/> [accessed 2019].

### Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as 'excellent' or 'very good'.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at [microbiologyresearch.org](http://microbiologyresearch.org).