

# SCIENTIFIC REPORTS



OPEN

## Comparison of traditional culture and molecular qPCR for detection of simultaneous carriage of multiple pneumococcal serotypes in African children

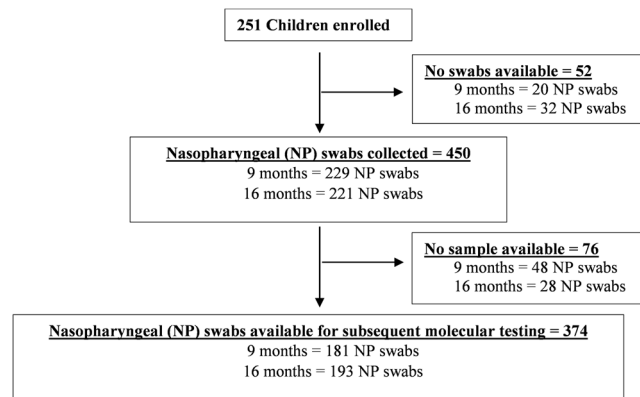
Courtney P. Olwage<sup>1,2</sup>, Peter V. Adrian<sup>1,2</sup> & Shabir A. Madhi<sup>1,2,3</sup>

*S. pneumoniae* is a common colonizer of the human nasopharynx in high income and low-middle income countries. Due to limitations of standard culture methods, the prevalence of concurrent colonization with multiple serotypes is unclear. We evaluated the use of multiplex quantitative PCR (qPCR) to detect multiple pneumococcal serotypes/group colonization in archived nasopharyngeal swabs of pneumococcal conjugate vaccine naive children who had previously been investigated by traditional culture methods. Overall the detection of pneumococcal colonization was higher by qPCR (82%) compared to standard culture (71%;  $p < 0.001$ ), with a high concordance ( $kappa = 0.73$ ) of serotypes/groups identified by culture also being identified by qPCR. Also, qPCR was more sensitive in detecting multiple serotype/groups among colonized cases (28.7%) compared to culture (4.5%;  $p < 0.001$ ). Of the additional serotypes detected only by qPCR, the majority were of lower density ( $< 10^4$  CFU/ml) than the dominant colonizing serotype, with serotype/group 6A/B, 19B/F and 23F being the highest density colonizers, followed by serotype 5 and serogroup 9A/L/N/V being the most common second and third colonizers respectively. The ability of qPCR to detect multiple pneumococcal serotypes at a low carriage density might provide better insight into underlying mechanism for changes in serotype colonization in PCV vaccinated children.

*Streptococcus pneumoniae* is a common colonizer of the human nasopharynx in both high-income and low-middle-income countries, with carriage being a pre-requisite to pathogenesis of pneumococcal disease as well as transmission between hosts<sup>1-5</sup>. To date almost 100 serotypes have been identified through their capsular polysaccharide and there is a paucity of data on concurrent carriage of multiple pneumococcal serotypes due to limitations of traditional recommended culture methods<sup>6</sup>.

Pneumococcal colonization and serotype characterization is normally investigated by culture of the dominant organism followed by serotyping with specific antisera<sup>7</sup>. Despite the ubiquitous use of culture based methods for nasopharyngeal colonization studies, detection of concurrent colonization by other serotypes of a lower density are often missed<sup>8</sup>. Furthermore, traditional culture methods are not quantitative for the colonizing serotypes. Molecular detection of pneumococci in the nasopharynx have several potential advantages over culture based methods, including detection of multiple serotypes from a single sample with high sensitivity, as well as quantitative PCR methods being able to measure the density of colonization and relative proportion of colonizing serotypes<sup>9</sup>. Interrogating beyond identification of only the dominant serotype colonizing children, could also help inform the extent to which the increase in nasopharyngeal non-vaccine serotype colonization following

<sup>1</sup>Department of Science and Technology/National Research Foundation: Vaccine Preventable Diseases, University of the Witwatersrand, Faculty of Health Sciences, Johannesburg, South Africa. <sup>2</sup>Medical Research Council: Respiratory and Meningeal Pathogens Research Unit, University of the Witwatersrand, Faculty of Health Sciences, Johannesburg, South Africa. <sup>3</sup>National Institute for Communicable Diseases: a division of National Health Laboratory Service, Johannesburg, South Africa. Correspondence and requests for materials should be addressed to S.A.M. (email: [ShabirM@nicd.ac.za](mailto:ShabirM@nicd.ac.za))



**Figure 1.** Flow Diagram of study population and sample availability. Diagram indicating the number of children initially enrolled in a PCV-unvaccinated cohort of HIV-uninfected children, as well as the number of nasopharyngeal (NP) swabs available for subsequent molecular qPCR analysis. No swabs were collected due to a missed study visit and the number of NP swabs available for molecular testing were defined by whether there was an adequate volume of sample remaining.

	qPCR (–)	qPCR (+)	Total	<i>p</i> -value	Concordance ( <i>kappa</i> )	PPV	NPV	Sensitivity (%)	Specificity (%)
Culture (+)	8(2)	258(69)	266(71)	<0.001	0.58	0.97	0.55	84	88
Culture (–)	59(16)	49(13)	108(29)						
Total	67(18)	307(82)	374(100)						

**Table 1.** Yield of culture and molecular qPCR for the detection *Streptococcus pneumoniae* in nasopharyngeal swabs ( $n = 374$ ). Numbers are value (%), *p*-values of  $<0.05$  were considered significant. PPV, positive predictive value. NPV, negative predictive value.

pneumococcal conjugate vaccine (PCV) immunization is driven by unmasking of previously minor colonizing serotypes or replacement acquisition by new serotypes.

In this study we aimed to evaluate the use of multiplex quantitative PCR (qPCR) for the detection and density characterisation of multiple pneumococcal serotypes on nasopharyngeal swabs, from PCV naive black-African children previously investigated by traditional culture methods.

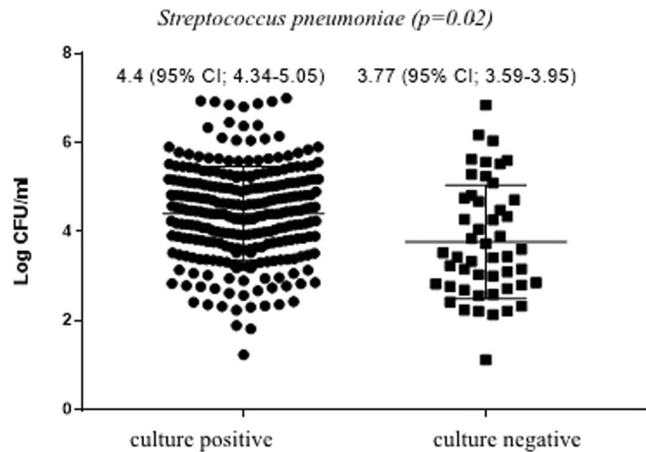
## Results

Molecular qPCR analysis involved 374 (83%) of the initial 450 nasopharyngeal swabs collected at either 9 or 15–16 months of age (Fig. 1). All children, in whom samples were available, were black-African, and 49.7% were male. Due to high similarities in findings at both 9 and 15–16 months of age, data was combined for the main analysis.

**Performance of the real-time qPCR assay.** All assays were effective in amplifying their respective targets with a high sensitivity, specificity and linearity with the exception of assay 10A/B in which only 10A was detected by the assay. However, the efficiency of all assays ranged from 90% to 105% (Supplementary Table S1). Within the linear dynamic range, the correlation coefficients ( $r^2$ ) of all the assays were 0.99. Furthermore, all assays showed high analytical sensitivities with their respective primer and probe pairs with the limit of detection equivalent 10 copies per PCR for all respective assays, with exception to primers/probes that detected serotypes/serogroups 5, 23F, 14, 6A/B/C/D, 13 and 10A in which the LLD was 10 fold less sensitive (100 copies per PCR). All primer pairs and probes were tested with genomic DNA from all pneumococcal and bacterial controls, and no cross reactions occurred (Supplementary Table S2). Both the inter-assay variation (repeatability) and inter-assay variation (reproducibility) for all respective assays was  $<0.1$ , while the accuracy for all respective assays was within  $\pm 0.1$  (Supplementary Table S1).

**Detection of pneumococcal carriage and serotype detection by culture and qPCR.** There was modest concordance between qPCR and culture for the detection of overall pneumococcus from infant NP swabs ( $kappa = 0.58$ ), with the *LytA* qPCR assay being more sensitive than culture (82% vs. 71%;  $p < 0.001$ ); Table 1. The qPCR demonstrated high sensitivity (84%) and specificity (88%) compared to culture used as a referent standard for detection of pneumococcus, with the overall density of pneumococcal carriage being higher in culture-positive (4.4, 95% CI: 4.34–5.05 CFU/ml) than culture-negative samples (3.77; 95% CI: 3.59–3.95 CFU/ml); Fig. 2.

Discordant results for identification of overall pneumococcus between qPCR and culture were strongly associated with density of pneumococcal carriage (Supplementary Table S3). The majority (47/49; 96%) of qPCR positive but culture-negative samples had estimated copy numbers of  $<10^4$  colony forming units (CFU)/ml per



**Figure 2.** Density of pneumococcal carriage as determined by quantitative PCR (qPCR) grouped by culture positive and culture negative samples.

swab. Similarly, the majority (7/8; 87.5%) of qPCR negative but culture-positive samples were reported as “scant” ( $<5$  colonies/plate) on culture. As pneumococcal density measured by qPCR increased, so did the detection by culture increase, with a perfect concordance (21/21;  $kappa = 1.0$ ) between culture and qPCR at a carriage density of  $>10^6$  CFU/ml.

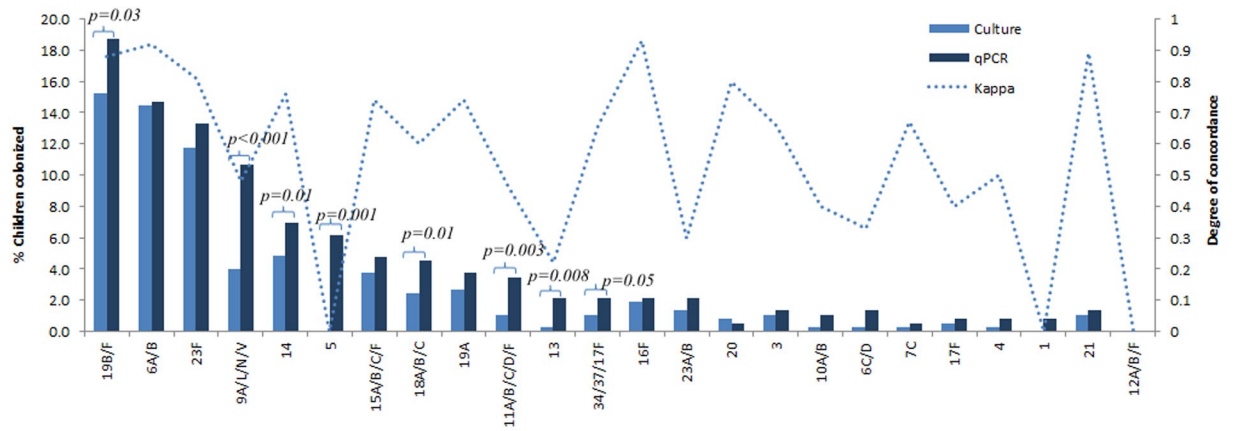
There was a high concordance ( $kappa = 0.73$ ) between serotypes/groups identified by culture and those identified by qPCR, including for those specific serotypes which were individually identifiable on qPCR ( $kappa = 0.68$ ); Supplementary Table S4. The qPCR method, however, identified an additional 150 (57.5%) serotypes/serogroups above those identified by the culture method, including higher detection prevalence by qPCR of serotypes/serogroups 5 (6.61% vs. 0%;  $p < 0.001$ ), 9A/L/N/V (10.7% vs. 4%;  $p < 0.001$ ), 11A/B/C/D/F (3.5% vs. 1.1%,  $p = 0.002$ ), 13 (2.1% vs. 0.3%,  $p = 0.008$ ), 14 (7% vs. 4.8%,  $p = 0.002$ ), 18A/B/C (4.5% vs. 2.4%,  $p = 0.011$ ), 19B/F (18.7% vs. 15.5%,  $p = 0.033$ ) and 34/37/17A (2.1% vs. 1.1%,  $p = 0.05$ ) (Fig. 3), with the sensitivity and specificity for the serotyping qPCR assays shown in Supplementary Table S5. Of these additional serotypes detected by qPCR, 98.8% had densities  $<10^4$  CFU/ml (Supplementary Table S6). Furthermore, 78% (117/150) of the additional serotypes detected by qPCR were co-carried with other pneumococcal serotypes, with only 12.8% (15/117) of these being the highest density colonizing serotype (Supplementary Table S7).

**Pneumococcal co-colonization.** Molecular qPCR was more sensitive than culture in detecting pneumococcal serotypes that were co-carried with two and three or more serotypes detected in 19.2% vs. 4.5% ( $p < 0.001$ ) and 9.4% vs. 0% ( $p < 0.001$ ) of swabs respectively (Table 2). A further 14% (43/307) of *LytA* positive NP swabs were negative for all the serogroups/serotypes we tested for, implying they were either non-typable, belonged to one of the serotypes we did not test for, or were *LytA* positive non-pneumococcal streptococcal species. However, due to the strong concordance between serotypes identified by qPCR also being identified by culture, we impute that at least 11/43 (25.6%) of NT serotypes identified by qPCR belong to one of the serotypes we did not test for.

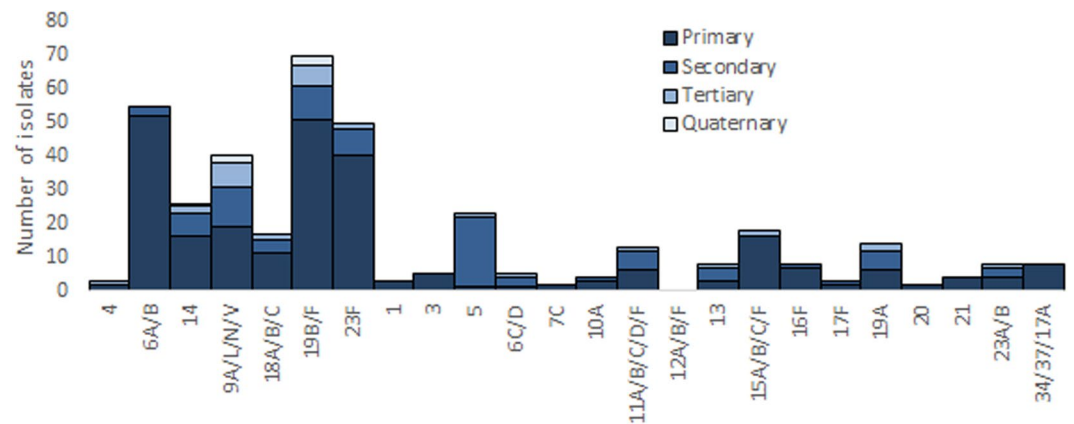
Serotypes/serogroups 6A/B, 19B/F, 23F, 5, and 9A/L/N/V were the most common serotypes to be simultaneously co-carried as detected by molecular qPCR (Fig. 4), with serogroups 6A/B and 19B/F being the most common co-colonizers (prevalence 11/88, 12.5%), followed by serogroup 19B/F and serotype 5 (prevalence 8/88, 9.1%). Serotype/groups 6A/B (prevalence 52/264, 19.7%), 19B/F (prevalence 51/264, 19.3%) and 23F (prevalence 40/264, 15.2%) were the highest density colonizers (dominant first colonizing serotype), with serotype 5 (prevalence 21/88, 23.9%) and serogroup 9A/L/N/V (prevalence 7/29, 24%) being the most common second and third colonizers respectively, as determined by density of carriage. Furthermore, serogroup 6A/B was most frequently found as a primary isolate (42/44; 95.5%) with other co-colonizing serotypes, whilst serotypes/group 1, 3, 20, and 34/37/17A were identified as primary isolates only, and serotype 17F and 10A were respectively identified as second and third colonizers only. The serotype specific propensity of whether a given serotype/serogroup is more likely to be found as a primary or non-primary isolate is shown in Fig. 5, with serotypes/serogroups 6A/B ( $p < 0.001$ ), 19B/F ( $p < 0.001$ ), 23F ( $p < 0.001$ ), 15A/B/C/F ( $p < 0.001$ ) and 34/37/17A ( $p = 0.008$ ) being more likely to be identified as a primary than a non-primary isolate and serotype 5 being more likely to be identified as a non-primary isolate ( $p < 0.001$ ).

## Discussion

The findings of our study showed that qPCR was more sensitive in detecting concurrent carriage with multiple pneumococcal serotypes, with at least 28.7% of all *LytA* positive nasopharyngeal samples having multiple serogroups/types identified compared to 4.5% of culture positive cases. Furthermore, we were able to quantify the relative proportion of these serotypes and found the majority of the additional serotypes detected by qPCR to have a lower bacterial load ( $<10^4$  CFU/ml) than the dominant colonizing serotype, indicating why these less dominant serotypes are likely missed by traditional culture methods that may be suboptimal for detecting low density carriage.



**Figure 3.** Overall prevalence of nasopharyngeal (NP) pneumococcal colonization as detected by traditional culture and quantitative PCR (qPCR) in a PCV-unvaccinated, HIV-uninfected cohort. Concordance between traditional culture and quantitative PCR (qPCR) were calculated using McNemar’s test with serogroups detected by qPCR included as serotypes. *p*-values of <0.05 were considered significant. Serogroup 6A/B (54): 35.2% Serotype 6A and 64.8% Serotype 6B on culture method. Serogroup 6C/D (1): 100% Serotype 6C on culture method. Serogroup 11A/B/C/D/F (4): 100% were serotype 11A on culture method. Serogroup 15A/B/C/F (14): 21%, 64.3%, and 14.2% were serotypes 15A, 15B and 15C on culture method respectively. Serogroup 18A/B/C (9): 100% were serotype 18C on culture method. Serogroup 19B/F (58): 8.6% were serotype 19B and 91.4% were serotype 19F on culture method. Serogroup 23A/B (5): 60% were serotype 23A and 40% were serotype 23B on culture method. Serogroup 34/37/17A (4): 100% were serotype 34 on culture method.

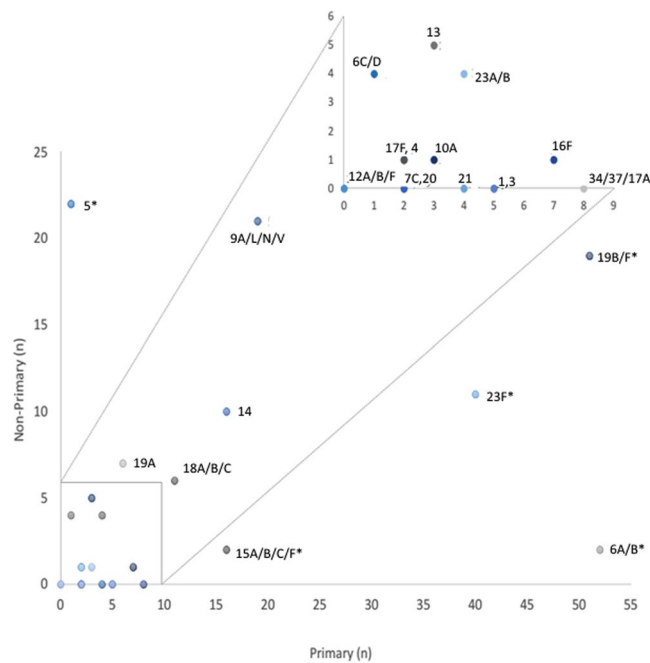


**Figure 4.** Serotype/group specific ranking of multiple pneumococcal carriage. Each isolate of *S. pneumoniae* was ranked according to its carriage density as determined by quantitative PCR (qPCR) to other isolates present in the same sample. Single colonizers were included in the analysis as primary colonizing serotypes.

	Culture	qPCR	<i>p</i> -value <sup>a</sup>
Not colonized	108(29)	67(18)	<0.001
<b>Colonized</b>	266(71)	307(82)	
1 serotype	239(99.8)	176(57.3)	<0.001
2 serotypes	12(4.5)	59(19.2)	<0.001
3+ serotypes	0	29(9.4)	<0.001

**Table 2.** Pneumococcal co-colonization as measured by traditional culture and quantitative molecular PCR (n = 374). Values are number (%). For multiple carriage percentage values are calculated based on total number of samples positive for *LytA*. PCV7, 7-valent pneumococcal conjugate vaccine. <sup>a</sup>As defined by McNemer’s test. *p*-values of < 0.05 were considered significant.

Serotypes/serogroups 6A/B, 19B/F, 23F, 5 and 9A/L/N/V were found as common co-colonizers of the nasopharynx. While serotypes/serogroups 6A/B, 19B/F, 23F and 9A/L/N/V are included among PCV7 targeted serotypes, serotype 5 is a non-PCV7 serotype, hence immunization with PCV7 could provide an opportunity for



**Figure 5.** Serotype specific propensity of whether a given serotype/serogroup is more likely to be found as a primary or non-primary isolate. Pneumococcal serotypes identified from nasopharyngeal swabs of black South-African children were classified according to whether they occurred as a primary or non-primary isolate as determined by carriage density. Single colonizers were included in the analysis as primary isolates. The box at the top of the graph is an inset of the blocked region at the bottom. \*Results were considered significant when the  $p$ -value was  $< 0.05$ .

this serotype to occupy the vacant niche and contribute to replacement disease in our setting, as it has also been shown to have high invasive disease potential in South Africa<sup>10,11</sup>. Further, a study done in Japan found serotype 3 to increase from 0% (2008–2010) to 1.3% (2011–2013) in children  $< 5$  years after PCV7 immunization, despite an overall decrease of IPD by 57%<sup>12</sup>. While in the present study we found serotype 3 to be identified as a primary isolate only. Despite the age and geographical difference between these two populations in the two above mentioned studies, the findings support the hypothesis that serotype 3 increased isolation rates after PCV immunization may be due to replacement acquisition of new serotypes.

Several models have demonstrated competition between pneumococcal serotypes for both acquisition and persistence of colonization<sup>13–15</sup>, with more prevalent and higher colonizing density serotypes being less metabolically demanding and thus better able to grow in a nutrient deficient environment. Our findings were in line with those from Dhouhadel *et al.*, in that concurrent colonizing serotypes had different bacterial loads, with one serotype dominating over the other(s), which may explain the competition among concurrent colonizing serotypes for growth due to limited nutrients and space<sup>16</sup>. Furthermore, few studies have been powered to determine if certain pairs of serotypes occur more frequently together, with most finding no significant combinations above the 5% threshold error<sup>14, 17, 18</sup>. Our study, however, found serogroups 6A/B and 19B/F to be the most common concurrent colonizers, occurring in 12.5% of co-colonised children, followed by serogroup 19B/F and serotype 5, which were found in 9.1% of co-colonised children.

A limited number of studies have investigated for concurrent colonization with pneumococcal serotypes in children, with rates varying from 1.5% to 40% according to geographical distribution and method used<sup>6, 19–21</sup>. Findings from most of these studies, however, were limited in that the majority used non quantitative traditional culture based methods that were insensitive in detecting multiple carriage, and therefore likely underestimated the true prevalence of this phenomenon<sup>6</sup>. Concurrent carriage of multiple pneumococcal serotypes in the present study was found in 28.7% of all colonised children. This finding was comparable to a recent study by Wyllie *et al.*, which also used qPCR, where concurrent colonization of multiple pneumococcal serotypes was found in 22% of colonised Nepalese children<sup>22</sup>. Other methods have also been used to investigate concurrent colonization such as conventional PCR<sup>23,24</sup>, colony blot<sup>25</sup>, RFLP<sup>26</sup> and microarray<sup>27</sup>, with the latter recently been shown to have the best performance overall for pneumococcal serotyping<sup>28</sup>. In studies done by Valente *et al.* and Kandasamy *et al.*, microarray found concurrent colonization in 20% and 22% of colonised Portuguese and Nepalese children respectively<sup>18,29</sup>. While these results are comparable to our findings, Kamng'ona and colleagues found concurrent colonization in 40% of Malawian children by microarray, a higher rate compared to our findings<sup>21</sup>. Our qPCR method, however, had the added benefit in that we were able to quantify the relative proportion of concurrent colonizing serotypes, it did not rely on an initial culture step which might change the relative proportion of the various strains, and it is less labour intensive and less expensive.

Our study was limited in that a) the qPCR method did not detect all pneumococcal serotypes and thus co-carriage in samples colonized with these non-tested serotypes could have been missed and the rate of



concurrent colonization could have been underestimated. b) The qPCR method was not able to discriminate between all serotypes within their respective serogroups. While recent studies have shown it is possible to discriminate some serotypes from their respective serogroups further (i.e. 19B and 19F from serogroups 19B/F and serogroups 9A/V and 9L/N from serogroup 9A/L/N/V)<sup>30</sup>, the high genotypic similarities between the capsular loci of certain serotypes still make it difficult to develop some serotype-specific assays and therefore certain serogroups will remain unresolved to specific serotypes. For example, serotype 6A and 6B differ by a sugar linkage due to a point mutation<sup>31</sup>, while serotypes 9A and 9V have the same synteny and differ only in their acetylation patterns<sup>32</sup>. Nonetheless, it is epidemiologically important to be able to discriminate serotypes from their respective serogroups, especially in the case where one serotype is included in the pneumococcal conjugate vaccine and the other is not, and the inability of the qPCR method to differentiate these serotypes will be an epidemiological challenge for vaccine effectiveness studies. Nevertheless, due to high concordance between culture and qPCR for the detection of serotypes in our study, the relative proportion of these serotypes could be estimated as similar to the proportion detected by culture. Further, the qPCR method was still able to identify serotypes/groups in culture-negative samples that would have otherwise have been missed; as well as being able to detect multiple serotypes from a single serotype and measure the density of the colonizing serotypes, therefore still making qPCR a useful alternative to traditional based methods. c) We were also not able to compare non-typables detected by culture and qPCR, as the qPCR method was not optimized to test for all pneumococcal serotypes and thus analysis could be misleading as a proportion of the NT detected by qPCR would rather indicate that we did not have the assay to test for all serotypes, rather than it being true pneumococcal NT. d) The loss of sensitivity by qPCR when comparing results to culture was most probably a result of degradation of DNA as a result of long-term storage and/or repeated freezing and thawing of samples. e) A recent study by the PneuCarriage project group found qPCR to have a large number of false positive results when testing pneumococcal spiked samples<sup>28</sup>. While this is concerning, the majority of false positives were for assays not included in our method. Serotype 5, however, was the only serotype to be more commonly isolated as a non-primary than a primary isolate which raises a question of whether this serotype is unable to be cultured or if the qPCR method was giving off false positives for this assay, and thus further investigation is needed. A Cq cut off <35Cqs was implemented, and both positive and negative controls were used for all assays, which should minimize the detection of false positives by the qPCR method.

In conclusion, molecular detection of pneumococcal serotypes was more sensitive than culture and allowed us to detect multiple serotypes. The ability of qPCR to detect pneumococcus at a low carriage density could allow us to gain a better understanding of the impact of PCV immunization on the changes in nasopharyngeal pneumococcal colonization and potential impact thereof on disease prevalence. Future studies could, however, aim to further discriminate additional serotypes from their respective serogroup by including additional testing such as RFLP and sequencing to differentiate between some serotypes that have high genotypic similarities.

## Material and Methods

**Study population.** We retrospectively analysed archived nasopharyngeal swab samples collected from a cohort of PCV-unvaccinated, HIV-uninfected infants from Soweto, South Africa, who had previously been investigated by standard culture methods. Detailed information of the study cohort has been described<sup>33,34</sup>. Briefly, the cohort included 251 PCV naive infants and their mothers that were enrolled between January 2007 through to October 2007, including 125 infants born to HIV-infected women but who were HIV-uninfected (HEU); and 126 infants born to HIV-uninfected mothers.

Nasopharyngeal (NP) swabs were taken at several time intervals, including at 9 and 15–16 months of age. Swabs were stored in skim milk-tryptone-glucose-glycerol (STGG) transport media at  $-70^{\circ}\text{C}$  at the Respiratory and Meningeal Pathogen Research Unit (RMPRU) in South Africa, as recommended by WHO<sup>35</sup>. The samples were previously cultured for *S. pneumoniae* using standard culture methods and pneumococcal serotyping undertaken using the Quellung method as described<sup>36</sup>.

**Multiplex qPCR methods.** Stored nasopharyngeal swab samples were thawed and processed. Briefly, NP swabs were vortexed for 30 seconds to release bacteria into the transport media and total nucleic acids were automatically extracted from STGG with the NucliSens<sup>®</sup> easyMAG<sup>®</sup> extraction system (BioMérieux, Marcy l'Etoile, France) according to the manufactures instructions. Similarly, total nucleic acids were also extracted from pneumococcal reference strains (positive control strains) grown in Todd-Hewitt broth supplemented with 5% yeast, with the NucliSens<sup>®</sup> easyMAG<sup>®</sup> extraction system according to manufactures instructions. Eluted samples and pneumococcal reference strains were stored at  $-20^{\circ}\text{C}$ .

Control strains for the pneumococcal serotypes 1, 2, 3, 4, 5, 6A, 6B, 6C, 6D, 7C, 9A, 9L, 9N, 9V, 10A, 10B, 11A, 11B, 11C, 11D, 11F, 12A, 12B, 12F, 13, 14, 15A, 15B, 15C, 15F, 16F, 17A, 17F, 18A, 18B, 18C, 19A, 19B, 19F, 20, 21, 23A, 23B, 23F, 34 and 37 were obtained from the National Institute for Communicable Diseases (NICD). DNA from these strains were used to optimise PCR assays and as positive controls.

The sequences for oligonucleotide primers and dye-labelled MGB probes were taken from previously published sequences, or designed with the ABI primer express software package, version 3.0 (Applied Biosystems, ABI, Foster City, USA). All primers and probes were designed based on previously published CpsA (Capsular polysaccharide) genes, in which all sequences were obtained from the Gene Bank database (<http://www.ncbi.nlm.nih.gov/genbank/>). Primer and Probe sequences are described in Supplementary Table S8. All PCV7 serotypes were included as targets in the assays while non-vaccine serotype targets were chosen based on the most frequently isolated non-vaccine serotypes in circulation in South Africa (based on detection by culture methods) at the time of study design. Due to the high genotypic similarities between the capsule loci of certain serotypes, the qPCR method was unable to discriminate between some serotypes within a particular serogroup (i.e. 6A/B, 6C/D, 9A/L/N/V, 10A/B, 11A/B/C/D/F, 12A/B/F, 15A/B/C/F, 18A/B/C, 19B/E, 23A/B and 34/37/17A), but was able to identify serotypes 1, 3, 4, 5, 7C, 13, 14, 16F, 17F, 19A, 20, 21 and 23F individually.

**Primer optimization, standard curves and quantification of the real-time PCR.** The qPCRs were optimized in 25  $\mu$ L reaction volumes containing 12.5  $\mu$ L 2x TaqMan gene expression master mix (ABI) and 0.25 mM MGB dye labelled probe (ABI) in a primer concentration matrix. Reactions were amplified with the ABI 7500 Real Time PCR system (Applied Biosystems, ABI, Foster City, USA) as per standard manufactures instructions with standard cycling times. Standard concentration curves and limits of detection (LOD) for each target serotype was determined by amplifying three replicates of a 10-fold serial dilution of purified genomic DNA extracted from reference strains with the linear dynamic range being  $10^1$ – $10^7$  copies per standard curve. Target control DNA was harvested from exponential phase cultures ( $OD_{600} = 0.1$ ), and the number of CFU/swab was estimated from the quantification cycle (Cq) values relative to the standard curve.

The correlation coefficients ( $r^2$ ) of each standard curve was determined and the amplification efficiency ( $e$ ) of a given primer pair was calculated by  $e = 10^{\frac{-1}{m}} - 1$ , where  $m$  is the slope of the standard curve. Limits of detection (LOD) or analytical sensitivity was determined by the minimum number of copies of DNA that could be measured repeatedly by the primers in the 10-fold serial dilutions with 95% probability. All primer pairs were tested with genomic DNA from all bacterial controls to determine the analytical specificity. Further, intra-assay variation (repeatability) was determined by running the same samples four times in the same assay run and presented as standard deviation (SD) for the copy number of the samples. A ratio of the difference between the observed copy number and the expected copy number was calculated to determine the accuracy of the assays, while the inter-assay variation (reproducibility) was determined by the SD of the copy number of the same sample tested in different runs.

Serotype targets were duplexed (Supplementary Table S9) and paired reactions were tested to ensure that the Cq values did not shift by  $>1$  Cq value, and that sensitivity and specificity remained the same as for singleplex reactions.

**Real-time PCR multiplex assay.** Target DNA was pre-screened for pneumococcus by qPCR for the *LytA* gene as described<sup>37</sup>. All *LytA* positive samples (Cq  $< 35$ ) were regarded as positive for pneumococci and tested for serotypes. Amplification data was analysed with the Applied Biosystems 7500 software, version 2.3 (Foster City, USA) with a manually defined threshold. Negative samples were defined as those with Cq values  $< 35$ . Furthermore, all *LytA* negative samples were tested for human GAPDH target to confirm the efficiency of the DNA extraction with all qPCR negative samples being positive for GAPDH.

**Statistical analysis.** Statistical analysis was performed with STATA Version 11.0 (Statacorp, Texas, USA). All *LytA* positive samples in which no serotype was identified were omitted in the multiple carriage analysis. The sensitivities of the qPCR assay and culture were tested with McNemar's test and when analysed for concordance, all serogroups were included as serotypes detected by qPCR vs. serotypes identified by culture. In the case of non-typable (NT) pneumococcus, we did not compare findings by culture and qPCR, as the qPCR method was not optimized to test for all pneumococcal serotypes and thus analysis could be misleading as a proportion of the NT detected by qPCR could belong to one of the untested serotypes. For density of carriage, geometric mean densities (GMD) and 95% confidence intervals (95% CI) of pneumococcal mean concentrations were calculated following  $\log_{10}$  transformation on the data. Furthermore, the serotype specific propensity of whether a given serotype/serogroup was more likely to be found as a primary or non-primary isolate, was analysed using a two-tail Binomial test, in which the observed proportion of primary to non-primary isolates were compared. Results were considered significant when the  $p$ -value was  $< 0.05$ .

**Ethical consent.** Ethical consent for the initial enrolment was obtained from the Medical Human Research Ethics Committee (HREC) of the University of Witwatersrand (HREC: 050705). Approval for further testing of samples was obtained from the HREC (M120972) and all methods were performed in accordance with the relevant guidelines and regulations. Written, informed consent had been obtained from the parents/guardians of all the study participants at the time of initial enrolment.

**Data availability.** The data that support the findings of this study are available from the corresponding author upon reasonable request.

## References

- Gray, B. M., Converse, G. M. 3rd & Dillon, H. C. Jr. Epidemiologic studies of *Streptococcus pneumoniae* in infants: acquisition, carriage, and infection during the first 24 months of life. *The Journal of infectious diseases* **142**, 923–933 (1980).
- Zenni, M. K. *et al.* *Streptococcus pneumoniae* colonization in the young child: association with otitis media and resistance to penicillin. *The Journal of pediatrics* **127**, 533–537 (1995).
- Faden, H. *et al.* Relationship between nasopharyngeal colonization and the development of otitis media in children. Tonawanda/Williamsville Pediatrics. *The Journal of infectious diseases* **175**, 1440–1445 (1997).
- Mastro, T. D. *et al.* Antimicrobial resistance of pneumococci in children with acute lower respiratory tract infection in Pakistan. *Lancet* **337**, 156–159 (1991).
- Lloyd-Evans, N. *et al.* Nasopharyngeal carriage of pneumococci in Gambian children and in their families. *The Pediatric infectious disease journal* **15**, 866–871 (1996).
- O'Brien, K. L. & Nohynek, H. Report from a WHO working group: standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*. *The Pediatric infectious disease journal* **22**, 133–140, doi:10.1097/01.inf.0000048676.93549.d1 (2003).
- Sorensen, U. B. Typing of pneumococci by using 12 pooled antisera. *Journal of clinical microbiology* **31**, 2097–2100 (1993).
- Robinson, J. Colonization and infection of the respiratory tract: What do we know? *Paediatrics & child health* **9**, 21–24 (2004).
- Satzke, C. *et al.* Standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*: updated recommendations from the World Health Organization Pneumococcal Carriage Working Group. *Vaccine* **32**, 165–179, doi:10.1016/j.vaccine.2013.08.062 (2013).
- Hausdorff, W. P., Siber, G. & Paradiso, P. R. Geographical differences in invasive pneumococcal disease rates and serotype frequency in young children. *Lancet* **357**, 950–952, doi:10.1016/S0140-6736(00)04222-7 (2001).

11. von Gottberg, A. *et al.* Epidemiology of invasive pneumococcal disease in the pre-conjugate vaccine era: South Africa, 2003–2008. *Vaccine* **31**, 4200–4208, doi:[10.1016/j.vaccine.2013.04.077](https://doi.org/10.1016/j.vaccine.2013.04.077) (2013).
12. Suga, S. *et al.* Nationwide population-based surveillance of invasive pneumococcal disease in Japanese children: effects of the seven-valent pneumococcal conjugate vaccine. *Vaccine* **33**, 6054–6060 (2015).
13. Lipsitch, M. *et al.* Competition among streptococcus pneumoniae for intranasal colonization in amouse model. *Vaccine* **19**, 598 (2000).
14. Lipsitch, M. *et al.* Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal serotypes in Kenya with a Markov transition model. *Epidemiology* **23**, 510–519, doi:[10.1097/EDE.0b013e31824f2f32](https://doi.org/10.1097/EDE.0b013e31824f2f32) (2012).
15. Auranen, K., Mehtala, J., Tanskanen, A. & M., S. K. Between-strain competition in acquisition and clearance of pneumococcal carriage–epidemiologic evidence from a longitudinal study of day-care children. *American journal of epidemiology* **171**, 169–176, doi:[10.1093/aje/kwp351](https://doi.org/10.1093/aje/kwp351) (2010).
16. Dhoubhadal, B. G. *et al.* Bacterial load of pneumococcal serotypes correlates with their prevalence and multiple serotypes is associated with acute respiratory infections among children less than 5 years of age. *PLoS one* **9**, e110777, doi:[10.1371/journal.pone.0111077](https://doi.org/10.1371/journal.pone.0111077) (2014).
17. Konradsen, H. B. Validation of serotyping of Streptococcus pneumoniae in Europe. *Vaccine* **23**, 1368–1373, doi:[10.1016/j.vaccine.2004.09.011](https://doi.org/10.1016/j.vaccine.2004.09.011) (2005).
18. Valente, C. *et al.* Decrease in pneumococcal co-colonization following vaccination with the seven-valent pneumococcal conjugate vaccine. *PLoS one* **7**, e30235, doi:[10.1371/journal.pone.0030235](https://doi.org/10.1371/journal.pone.0030235) (2012).
19. Huebner, R. E., Dagan, R., Porath, N., Wasas, A. D. & Klugman, K. P. Lack of utility of serotyping multiple colonies for detection of simultaneous nasopharyngeal carriage of different pneumococcal serotypes. *The Pediatric infectious disease journal* **19**, 1017–1020 (2000).
20. Hausdorff, W. P., Feikin, D. R. & Klugman, K. P. Epidemiological differences among pneumococcal serotypes. *The Lancet infectious diseases* **5**, 83–93, doi:[10.1016/S1473-3099\(05\)01280-6](https://doi.org/10.1016/S1473-3099(05)01280-6) (2005).
21. Kamng'ona, A. W. *et al.* High multiple carriage and emergence of Streptococcus pneumoniae vaccine serotype variants in Malawian children. *BMC infectious diseases* **15**, 234, doi:[10.1186/s12879-015-0980-2](https://doi.org/10.1186/s12879-015-0980-2) (2015).
22. Wylie, A. L. *et al.* Molecular surveillance of nasopharyngeal carriage of Streptococcus pneumoniae in children vaccinated with conjugated polysaccharide pneumococcal vaccines. *Scientific reports* **6**, 23809, doi:[10.1038/srep23809](https://doi.org/10.1038/srep23809) (2016).
23. Saha, S. *et al.* Detection of co-colonization with Streptococcus pneumoniae by algorithmic use of conventional and molecular methods. *Vaccine* **33**, 713–718, doi:[10.1016/j.vaccine.2014.11.040](https://doi.org/10.1016/j.vaccine.2014.11.040) (2015).
24. Rivera-Olivero, I. A., Blommaert, M., Bogaert, D., Hermans, P. W. & de Waard, J. H. Multiplex PCR reveals a high rate of nasopharyngeal pneumococcal 7-valent conjugate vaccine serotypes co-colonizing indigenous Warao children in Venezuela. *Journal of medical microbiology* **58**, 584–587, doi:[10.1099/jmm.0.006726-0](https://doi.org/10.1099/jmm.0.006726-0) (2009).
25. Bogaert, D. *et al.* Colony blot assay: a useful method to detect multiple pneumococcal serotypes within clinical specimens. *FEMS immunology and medical microbiology* **41**, 259–264, doi:[10.1016/j.femsim.2004.03.013](https://doi.org/10.1016/j.femsim.2004.03.013) (2004).
26. Brugger, S. D., 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* **5**, e11638, doi:[10.1371/journal.pone.0011638](https://doi.org/10.1371/journal.pone.0011638) (2010).
27. Turner, P. *et al.* Improved detection of nasopharyngeal cocolonization by multiple pneumococcal serotypes by use of latex agglutination or molecular serotyping by microarray. *Journal of clinical microbiology* **49**, 1784–1789, doi:[10.1128/JCM.00157-11](https://doi.org/10.1128/JCM.00157-11) (2011).
28. Satzke, C., Dunne, E. M., Porter, B. D., Klugman, K. P. & Mulholland, E. K. The PneuCarriage Project: A Multi-Centre Comparative Study to Identify the Best Serotyping Methods for Examining Pneumococcal Carriage in Vaccine Evaluation Studies. *PLoS medicine* **12**, e1001903; discussion e1001903, doi:[10.1371/journal.pmed.1001903](https://doi.org/10.1371/journal.pmed.1001903) (2015).
29. Kandasamy, R. *et al.* Multi-serotype pneumococcal nasopharyngeal carriage prevalence in vaccine naive Nepalese children, assessed using molecular serotyping. *PLoS one* **10**, e0114286, doi:[10.1371/journal.pone.0114286](https://doi.org/10.1371/journal.pone.0114286) (2015).
30. Pholwat, S., Sakai, F., Turner, P., Vidal, J. E. & Houpt, E. Development of a TaqMan array card for pneumococcal serotyping on isolates and nasopharyngeal samples. *Journal of clinical microbiology*, JCM. 00613–00616 (2016).
31. Park, I. H. *et al.* Discovery of a new capsular serotype (6C) within serogroup 6 of Streptococcus pneumoniae. *Journal of clinical microbiology* **45**, 1225–1233 (2007).
32. Mavroidi, A. *et al.* Genetic relatedness of the Streptococcus pneumoniae capsular biosynthetic loci. *Journal of bacteriology* **189**, 7841–7855 (2007).
33. Nunes, M. C. *et al.* Acquisition of Streptococcus pneumoniae in pneumococcal conjugate vaccine-naive South African children and their mothers. *The Pediatric infectious disease journal* **32**, e192–205, doi:[10.1097/INF.0b013e31828683a3](https://doi.org/10.1097/INF.0b013e31828683a3) (2013).
34. Madhi, S. A. *et al.* Effect of HIV infection status and anti-retroviral treatment on quantitative and qualitative antibody responses to pneumococcal conjugate vaccine in infants. *J Infect Dis* **202**, 355–361, doi:[10.1086/653704](https://doi.org/10.1086/653704) (2010).
35. O'Brien, K. L. *et al.* Evaluation of a medium (STGG) for transport and optimal recovery of Streptococcus pneumoniae from nasopharyngeal secretions collected during field studies. *Journal of clinical microbiology* **39**, 1021–1024, doi:[10.1128/JCM.39.3.1021-1024.2001](https://doi.org/10.1128/JCM.39.3.1021-1024.2001) (2001).
36. O'Brien, K. L. & Nohynek, H. Report from a WHO Working Group: standard method for detecting upper respiratory carriage of Streptococcus pneumoniae. *The Pediatric infectious disease journal* **22**, e1–11, doi:[10.1097/01.inf.0000049347.42983.77](https://doi.org/10.1097/01.inf.0000049347.42983.77) (2003).
37. Carvalho Mda, G. *et al.* Evaluation and improvement of real-time PCR assays targeting *lytA*, *ply*, and *psaA* genes for detection of pneumococcal DNA. *Journal of clinical microbiology* **45**, 2460–2466, doi:[10.1128/JCM.02498-06](https://doi.org/10.1128/JCM.02498-06) (2007).

## Acknowledgements

This study was supported by the National Research Foundation through the South African Research Chair Foundation.

## Author Contributions

C.P.O., P.V.A., and S.A.M., conceived and designed the study. C.P.O. was responsible for data curation and analysis. Funding acquisition was undertaken by S.A.M. C.P.O. wrote the first draft of the manuscript while all authors were involved in reviewing and editing the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at doi:[10.1038/s41598-017-04915-y](https://doi.org/10.1038/s41598-017-04915-y)

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.





**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017