



OPEN

## Changes in serotype prevalence of *Streptococcus pneumoniae* in Southampton, UK between 2006 and 2018

David W. Cleary<sup>1,2</sup>, Jessica Jones<sup>1</sup>, Rebecca A. Gladstone<sup>3,4</sup>, Karen L. Osman<sup>1</sup>, Vanessa T. Devine<sup>1</sup>, Johanna M. Jefferies<sup>1</sup>, Stephen D. Bentley<sup>3</sup>, Saul N. Faust<sup>1,2,5</sup> & Stuart C. Clarke<sup>1,2,6</sup>✉

*Streptococcus pneumoniae* continues to cause significant disease burden. Whilst pneumococcal conjugate vaccines (PCV) have substantially reduced this burden, serotype replacement partially negates this success due to increased disease associated with non-vaccine serotypes (NVTs). Continued surveillance is therefore essential to provide crucial epidemiological data. Annual cross-sectional surveillance of paediatric pneumococcal carriage was started in Southampton, UK following PCV7 roll-out in 2006. Nasopharyngeal swabs were collected from children < 5 years old each winter (October to March) from 2006/07 and for each consecutive year until 2017/18. Pneumococcal serotype was inferred from whole genome sequencing data. A total of 1429 (32.5%) pneumococci were isolated from 4093 children. Carriage ranged from 27.8% (95%CI 23.7–32.7) in 2008/09 to 37.9% (95%CI 32.8–43.2) in 2014/15. Analyses showed that carriage increased in children aged 24–35 months ( $p < 0.001$ ) and 47–60 months ( $p < 0.05$ ). Carriage of PCV serotypes decreased markedly following PCV7 and/or PCV13 introduction, apart from serotype 3 where the relative frequency was slightly lower post-PCV13 (pre-PCV13  $n = 7$ , 1.67%; post-PCV13  $n = 13$ , 1.27%). Prevalence of NVTs implicated in increased disease was low with 24F ( $n = 19$ , 1.4%) being the most common followed by 9N ( $n = 11$ , 0.8%), 8 ( $n = 7$ , 0.5%) and 12F ( $n = 3$ , 0.2%).

As a frequent coloniser of the upper respiratory tract, *Streptococcus pneumoniae*, the pneumococcus, continues to cause significant global burden of both invasive and non-invasive diseases. These include diseases such as meningitis, sepsis, pneumonia and otitis media (OM)<sup>1</sup>. The most recent estimates suggest that ~ 300,000 deaths of children under the age of five occur each year globally, a mortality rate of 45/100,000, with the majority a consequence of pneumonia<sup>1</sup>.

The introduction of pneumococcal conjugate vaccines (PCVs) into national immunisation programmes in the early 2000s substantially reduced both mortality and morbidity from this bacterium. PCVs use one of the primary virulence determinants of the pneumococcus, its polysaccharide capsule, of which there are now 100 defined types<sup>2</sup>. Initial PCV formulation targeted those serotypes causing the greatest disease burden. Continued surveillance has since seen the formulation of PCVs expand to include serotypes which subsequently increased in disease prevalence. In the USA, modelling showed that over 282,000 invasive pneumococcal disease (IPD) cases and ~ 3000 deaths have been averted, with 97 million visits to health-care providers for OM also avoided<sup>3</sup>. A similar benefit was also observed in England and Wales following the introduction of the seven-valent PCV (PCV7) in 2006. PCV7 targeted serotypes 4, 14, 18C, 19F, 23F, 6B and 9V, and was replaced in 2010 by a thirteen-valent PCV (PCV13) containing an additional six serotypes: 1, 3, 5, 19A, 6A and 7F. Accounting for both

<sup>1</sup>Faculty of Medicine & Institute for Life Sciences, University of Southampton, University Hospital Southampton Foundation NHS Trust, Southampton, UK. <sup>2</sup>NIHR Southampton Biomedical Research Centre, University Hospital Southampton Foundation NHS Trust, Southampton, UK. <sup>3</sup>Parasites and Microbes, Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, UK. <sup>4</sup>Department of Biostatistics, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway. <sup>5</sup>NIHR Southampton Clinical Research Facility, University Hospital Southampton Foundation NHS Trust, Southampton, UK. <sup>6</sup>Global Health Research Institute, University of Southampton, Southampton, UK. ✉email: S.C.Clarke@soton.ac.uk

direct protection and herd immunity in non-vaccine recipients of all age groups<sup>4,5</sup> almost 40,000 cases of IPD have been prevented in England and Wales in the eleven years after PCV7 introduction<sup>6</sup>.

Overall carriage prevalence of pneumococci has remained largely unaffected despite PCV introductions and instead the phenomenon of serotype replacement has occurred<sup>7,8</sup>. Here, non-vaccine serotypes (NVTs) inhabit the vacant ecological niche that remains following the removal of circulating VT pneumococci. Although NVTs typically have less invasive potential and overall IPD incidence is lower post-PCV, serotype replacement is nevertheless problematic as it partially negates the success of PCVs due to increases in disease associated with NVTs. For example, in England and Wales worrying increases in Serotypes 8, 12F and 9N have been observed in IPD surveillance data<sup>6</sup>. Globally 15A and 23B have been also been flagged as serotypes of concern<sup>9</sup>. Contrasts such as this highlight observed differences in serotype replacement when looking at data from individual countries. The possible reasons for this have been expertly reviewed recently<sup>10</sup>. Leaving aside study design and the way in which disease is reported, likely drivers for these differences include the pneumococcal population ecology pre-vaccination and the prospect that cross-sectional carriage studies are unable to accurately identify rarer NVT serotypes that go on to be a significant proportion of IPD. In this situation those same serotypes are hypothesised to have a high case to carrier ratio. Regardless of these issues, continued surveillance is therefore essential to provide crucial epidemiological data relating to the ever-evolving landscape of circulating pneumococci.

During the introduction of PCV7 in the UK, we started an annual surveillance study of carriage in children aged < 5 years in the Southampton area. This unique study has since yielded important data regarding asymptomatic carriage of this significant pathogen in children < 5 years of age. This has aided our understanding of both vaccine efficacy for current PCVs whilst also improving our understanding of the relationship between serotype replacement in the broader national context of invasive disease<sup>7,8,11</sup>. Here we report the carriage epidemiology of pneumococcal serotypes over a period of twelve winter seasons, from 2006/07 to 2017/18. We found potentially important increases in carriage prevalence in older children particularly. Additionally, whilst carriage of VTs was substantially diminished, there was a continued low-level circulation of serotypes 3 and 19A.

## Methods

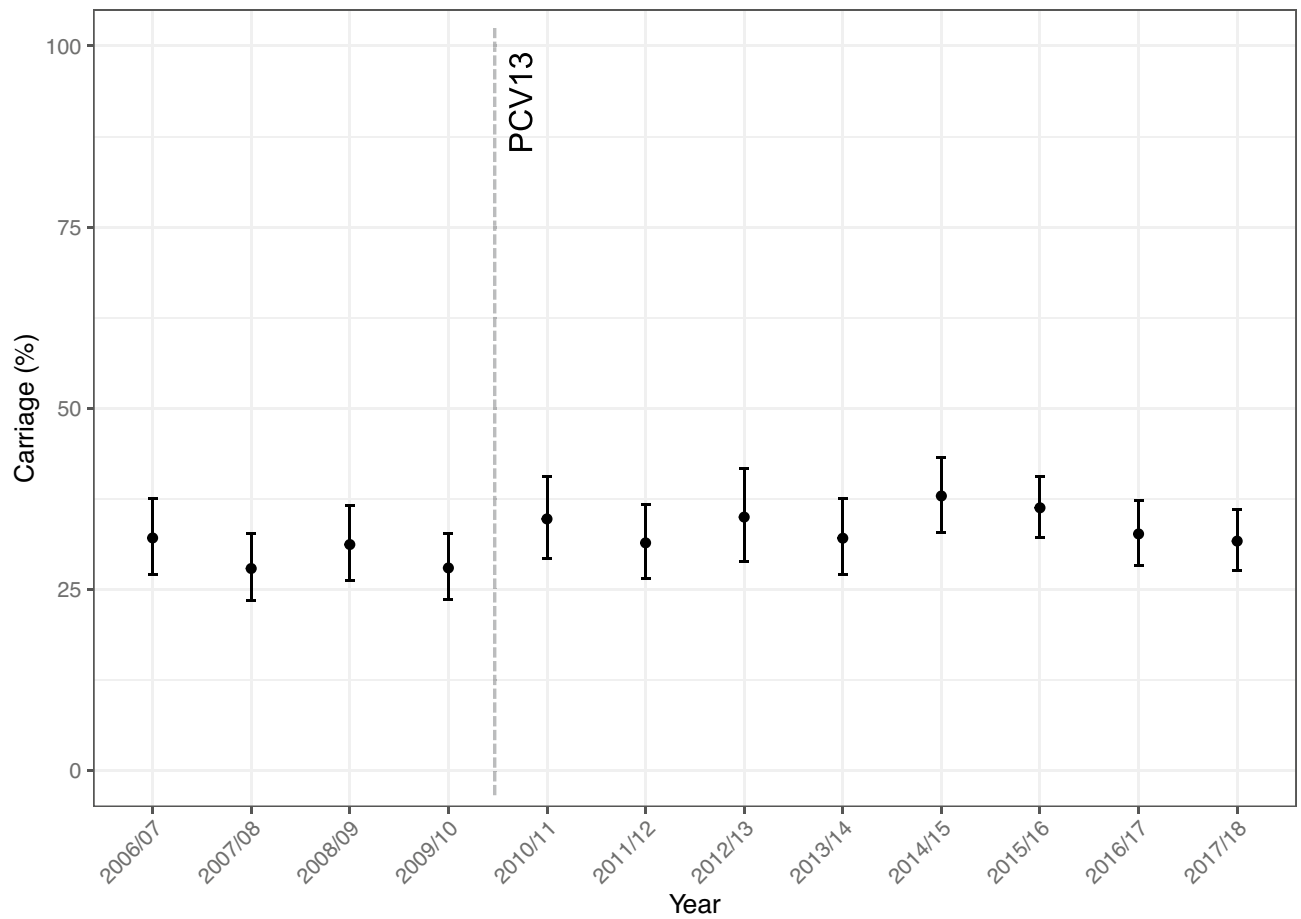
**Ethical approval.** The study was approved by the UK National Health Service (NHS) Research Ethics Service (06/Q1704/105 and 14/NS/1064). All methods and research practises outlined below were performed in accordance with relevant regulations which included the taking of informed consent from the legal guardians of all participants.

**Paediatric population.** The study site, Southampton General Hospital, is administered by University Hospital Southampton (UHS) NHS Foundation Trust, which serves a population of approximately 1.9 million in Southampton and South Hampshire. The resident population of Southampton is ~ 250,000 with ~ 16,000 being children aged < 5 years. At 6% this is in keeping with the national average<sup>12</sup> although we note that children would inevitably have also been recruited from outside the City of Southampton. Southampton is an ethnically diverse city with 78% of residents being White British or Irish in the 2011 census; a proportion that has likely decreased given that in 2017/18 nearly 40% of live births were of non-White British or Irish ethnicity<sup>12</sup>.

**Nasopharyngeal swab samples and laboratory processing.** Nasopharyngeal swabs were collected from children aged < 5 years each year commencing in the winter (October to March) of 2006/07 and for each consecutive year until 2017/18. Parents/guardians were approached for informed consent either prior to or following their child's appointment in an outpatient department of Southampton General Hospital. Aside from age, the only other exclusion criterium was that only one child per family was swabbed and that child was swabbed only once. Each year the target for isolation was  $n = 100$  pneumococci. Assuming a low carriage prevalence of 10%, this would allow the detection ~ 50% relative reduction with 80% power at a 5% significance level. Nasopharyngeal Rayon tipped Transwabs (Medical Wire, Corsham, UK) in charcoal Amies media were used for swabbing and then plated onto Columbia Colistin Naladixic Acid agar (CNA; Oxoid, Basingstoke, UK) within 9 h of swabbing. Confirmation of presumptive *S. pneumoniae* was done on 5% blood Columbia Blood Agar (CBA; Oxoid, Basingstoke, UK) using optochin sensitivity indicated by a  $\geq 14$  mm diameter inhibition zone around the disc (Thermo Scientific™, Loughborough, UK). Only one colony of *S. pneumoniae* per participant swab was selected for further analysis. Between 2006/07 and 2011/12 this was done in the Health Protection Agency Southampton Laboratory (now part of Public Health England) and from 2012/13 by technical staff in our research group.

**Questionnaire.** In the winter of 2010/11, a questionnaire was introduced for the parent/guardian of each participant to complete. This questionnaire captured such information as vaccine status (Prevenar, Bexsero, Fluenz and whether the child was on schedule for routine paediatric vaccinations), recent respiratory illness (cold, ear infection, flu-like illness, sore throat or chest infection within the preceding 30 days), and antibiotic use within the preceding 30 days, with type if known.

**Serotyping.** Isolates from skim milk, tryptone, glucose, and glycerin (STGG) stocks were cultured on CNA plates and incubated overnight at 37 °C in 5% CO<sub>2</sub> prior to DNA extraction. Extraction was carried out using QIAamp® DNA mini kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The DNA extracts were sent to the Wellcome Sanger Institute (WSI) for whole genome sequencing (WGS) using Illumina HiSeq or 10X platforms generating initially 2 × 75 bp and later 2 × 100 bp paired-end reads from libraries prepared using TruSeq chemistry. Pneumococcal serotype was inferred using PneumoCaT version 1.0<sup>13</sup>.



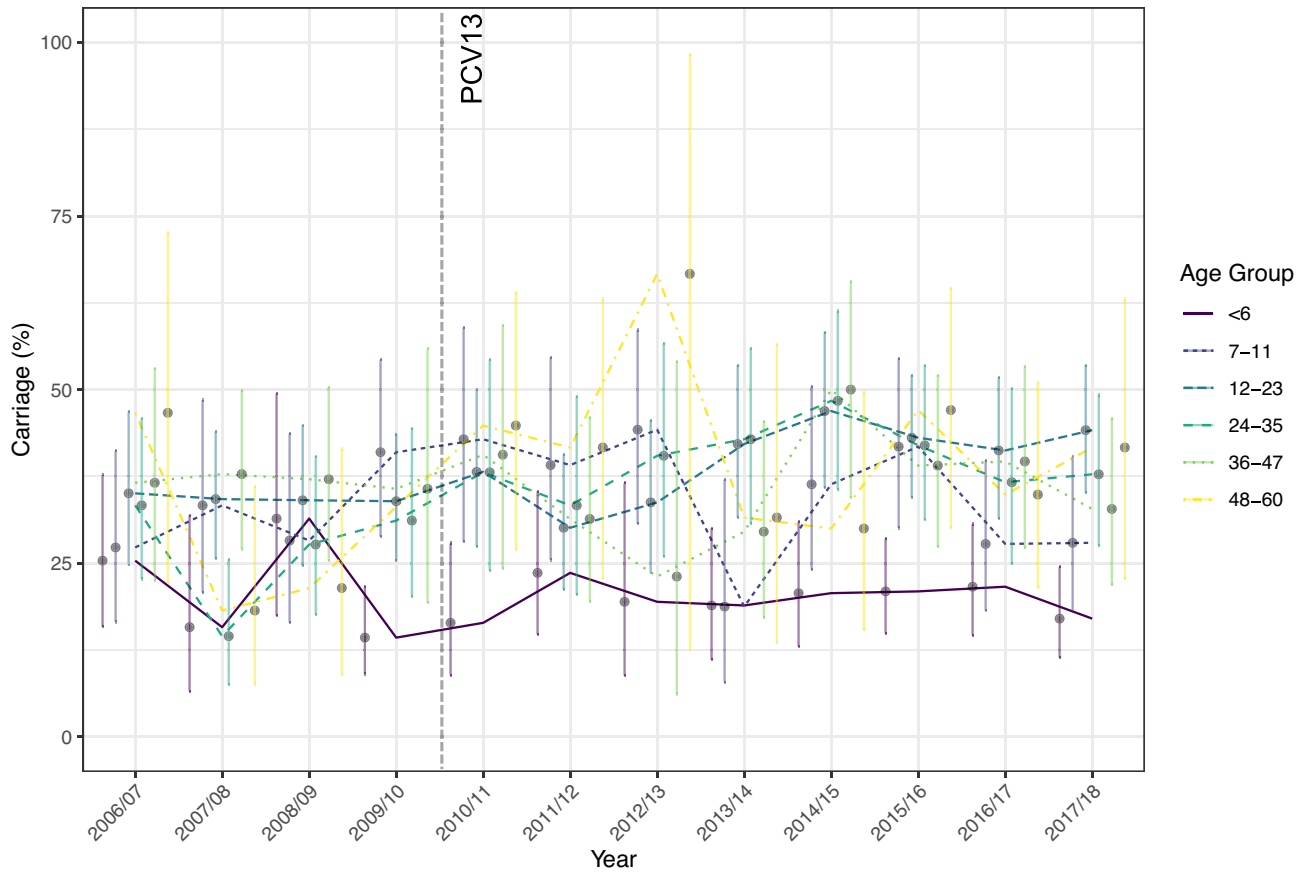
**Figure 1.** Carriage prevalence (%) of *S. pneumoniae* (all serotypes) in each year of the study. Overall carriage has remained consistent at an average of 32.5%, ranging from 27.8% in 2008/09 to 37.9% in 2014/15. Introduction of PCV13 is shown by the vertical line. Error bars represent 95% CI.

**Statistical analysis.** All statistical analysis was done in R version 3.6.0 (2019-04-26) using RStudio version 1.2.1335<sup>14,15</sup> with graphics built using the grammar of graphics package, ggplot<sup>15</sup>. Participants characterised by recruitment in the PCV7 or PCV13 era were defined as having swabbing dates of 01/01/2006–30/06/2010 inclusive and 01/07/2010 to the end of the study period respectively. To evaluate the impact of PCV dose on pneumococcal carriage a likely dose number was assigned using the child's age i.e., those > 1.0 month and ≤ 3.9 months were labelled as having one dose, those ≥ 4.0 months and ≤ 11.9 months were labelled as two doses, and those aged ≥ 12.0 months but ≤ 24 months were assumed to have received three doses. An upper age limit was used to minimise confounding effects of increasing pneumococcal carriage with age. When evaluating doses, particularly for comparisons of PCV7 versus PCV13, we excluded year one (2006/07) and year five (2010/11) to avoid confounding effects of catch-up and exclude those whose vaccinations might have spanned the period of PCV13 roll-out. Chi-squared test for trend in proportions was done using the `prop_trend_test()` in the R package `rstatix`. Simpsons index of diversity was computed using the `diversity()` function from the R package `vegan`<sup>16</sup>. Odds ratios based on multivariable logistic regression analysis and forest plots were generated using the R package `finalfit()` and `glmulti()` where the dependent variable was carriage of *S. pneumoniae*, the explanatory variables a character list derived from questionnaire data, and the random effect the year of the study.

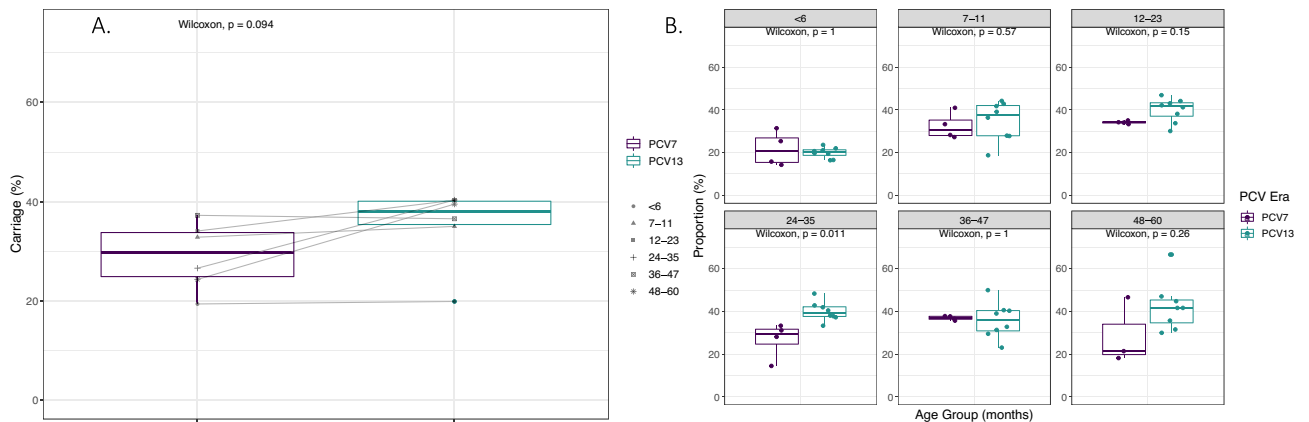
## Results

Between the winters of 2006/07 and 2017/18, from a total of 4409 study participants, 4393 children for whom age was accurately recorded as < 5 years provided NP swabs (n = 4393) for microbiological testing. Demographics of the study population are given in Supplementary Table 1. A breakdown of colonisation prevalence by age group per year with colonisation status is shown in Supplementary Fig. 1. A total of 1429 (32.5%) swabs yielded pneumococci. Pneumococcal carriage for each study year is shown in Fig. 1. Carriage ranged from a low of 27.8% (95%CI 23.7–32.7) in Year 3 (2008/09) to a high in Year 9 (2014/15) of 37.9% (95%CI 32.8–43.2).

Pneumococcal carriage in age groups is shown in Fig. 2. Participants were divided into age strata as follows: < 6 months (n = 1009, 23.0%), 7–11 months (n = 644, 14.6%), 12–23 months (n = 1148, 26.1%), 24–35 months (n = 741, 16.9%), 36–47 months (n = 567, 12.9%), and 48–60 months (n = 284, 6.5%). Trend analyses showed that carriage increased over time in children aged 24–35 months ( $p < 0.001$ ) and 47–60 months ( $p < 0.05$ ). We next

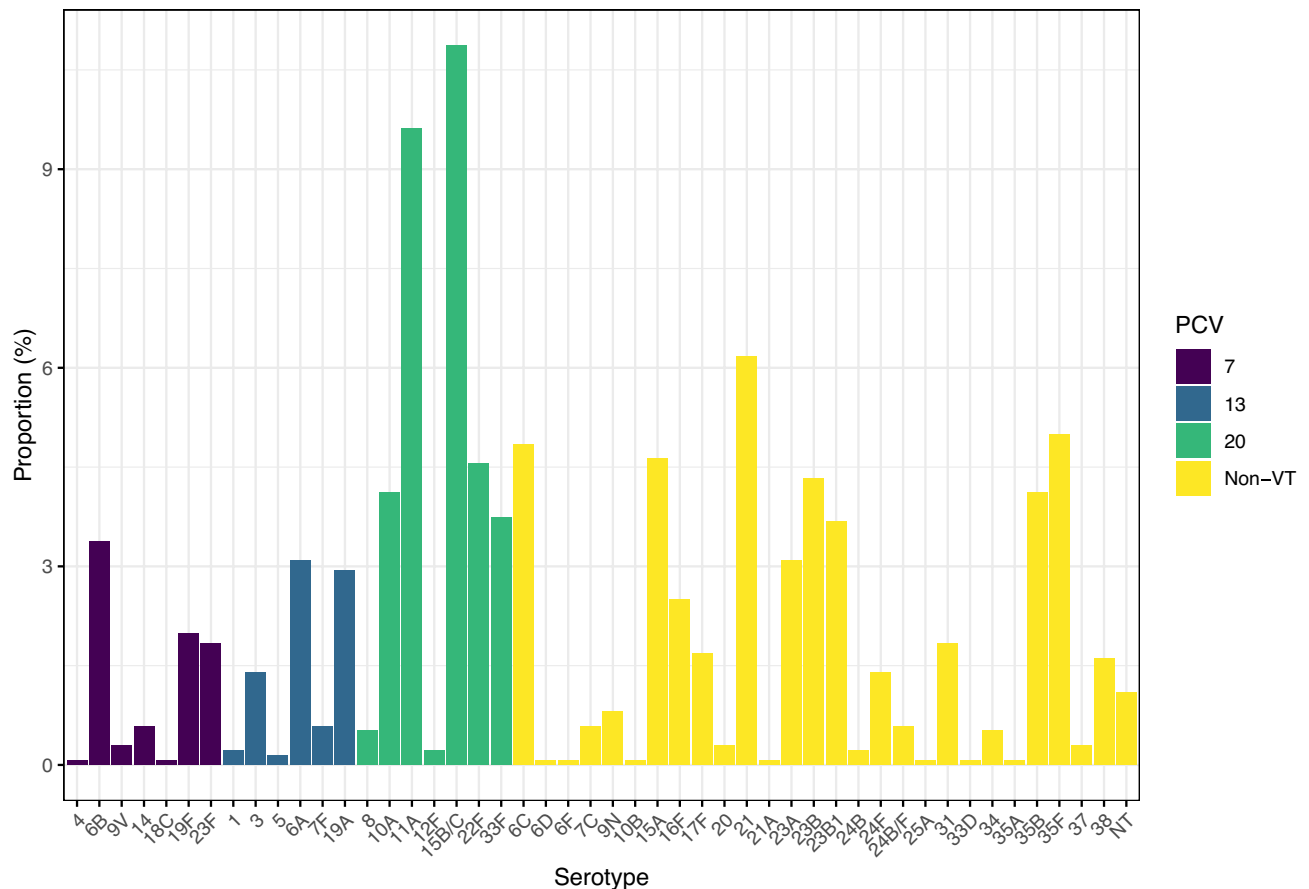


**Figure 2.** Carriage prevalence (%) of *S. pneumoniae* (all serotypes) by age group and year. Introduction of PCV13 is shown by the vertical line. Error bars represent 95% CI. Statistically significant increases in carriage in those aged 24–35 ( $p < 0.001$ ) and 48–60 ( $p < 0.05$ ) months old were shown by Chi-squared test for changes in trends of proportion.



**Figure 3.** Comparison of overall pneumococcal carriage prevalence (%) by PCV era by all ages (A), and within age groups (B). Overall, no statistically significant difference was found between the PCV7 and PCV13 era, apart from in those aged 24–35 months, in whom a significant increase in carriage in the PCV13 era was observed ( $p < 0.05$ ).

questioned whether these changes in carriage were attributable to the replacement of PCV7 with PCV13 between years four and five (2009/10 and 2010/11) (Fig. 3). As shown in Fig. 3A, there was no significant increase in carriage when all participants were grouped into either a PCV7 or PCV13 era ( $p = 0.094$ ). However, when examining the age groups individually (Fig. 3B) here again a significant increase in carriage in those aged 24–35 months old in the PCV13 compared to PCV7 era was noted ( $p = 0.004$ ).



**Figure 4.** Serotype proportion for all *S. pneumoniae* isolated between 2006/07 and 2017/18. Serotypes are colored according to the PCV in which they are currently included (PCV13), previously included (PCV7) or included in newly licensed formulations (PCV20). Serotypes not in a PCV are shown in yellow.

The distribution of serotypes carried, according to their inclusion in PCV7 or PCV13, is shown in Fig. 4.

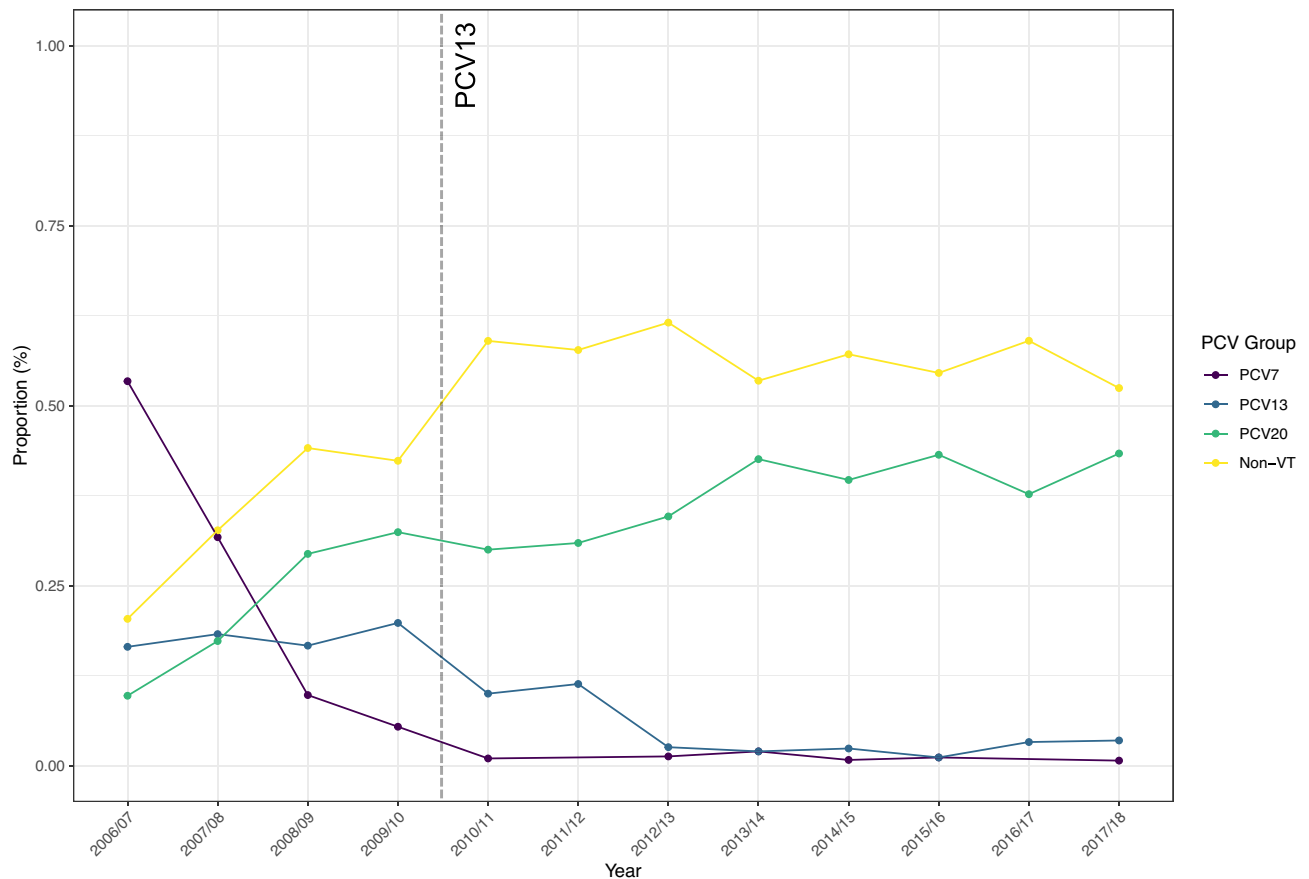
Those serotypes that will be targeted as part of PCV20, which is anticipated to be part of a childhood schedule in the future is therefore also shown. The most prevalent PCV7 serotypes were 6B ( $n = 46$ , 3.4%), 19F ( $n = 27$ , 2.0%) and 23F ( $n = 25$ , 1.8%). From PCV13, 6A ( $n = 42$ , 3.1%) and 19A ( $n = 40$ , 2.9%) were the most isolated. The NVT serotypes 15B/C and 11A were the most frequently isolated overall at 10.9% ( $n = 148$ ) and 9.6% ( $n = 131$ ) respectively.

No significant change in serotype diversity, as measured using Simpsons 1-D, was observed. Values ranged from 0.90 in 2006/07 to 0.94 in 2008/09 and 2009/10, with a mean of 0.93. When grouped in relation to PCV (Fig. 5), the reduction in PCV7 and PCV13 serotypes following PCV7 and PCV13 introduction is clear, with the concomitant increase in NVT serotypes, which in 2017/18 equated to > 95% of carriage.

This change was apparent when examining the post-PCV7/pre-PCV13 and post-PCV13 proportion of each serotype individually as shown in Fig. 6. Carriage of all PCV serotypes decreased substantially following PCV7 and/or PCV13 introduction. The exception was serotype 3 where the relative frequency was similar between eras (pre-PCV13  $n = 7$ , 1.67%; post-PCV13  $n = 13$ , 1.27%). NVTs showed generally the reverse with increases in frequency post-PCV introduction. Notable exceptions were 22F and 6C which decreased in the post-PCV period in more recent years. Since 2010/11  $n = 47$  isolates of a PCV7 or PCV13 serotype have been isolated (Fig. 7) and have included serotypes 6B and 19F from PCV7 and 3, 5, 6A, 7F and 19A from PCV13. No carriage of 5, 6B or 6A has been seen since 2012/13. Single occurrences of 7F carriage occurred in 2014/15 and 2016/17, however 19A and 3 were isolated in most years ( $n = 7$  and  $n = 5$  respectively since 2014/15 inclusive).

The impact of PCV dose was examined in relation to carriage (Fig. 8). Again, by examining the post-PCV7/pre-PCV13 (Fig. 8A) and post-PCV13 (Fig. 8B) eras separately, a statistically significant increase of carriage in children expected to have received multiple doses was observed: one vs. three doses  $p = 0.0021$ , one vs. two doses  $p = 0.0021$ , and two vs. three doses  $p = 0.026$ . Although a similar trend for PCV7 was observed, smaller counts meant this was not significant.

The odds of *S. pneumoniae* carriage was examined in relation to age, gender, vaccinations, illness in the 30-days prior to swabbing and antibiotic use in the same period (Fig. 9). The impact of increasing age was seen with both those aged 7–11 and 12–23 months having increased odds for carriage (1.90, 95%CI 1.18–3.06,  $p = 0.008$  & 2.72, 95%CI 1.44–5.16,  $p = 0.002$  respectively) relative to those < 6 months. Those that reported a cold-like illness in the previous 30 days were significantly more likely to be culture positive for pneumococci



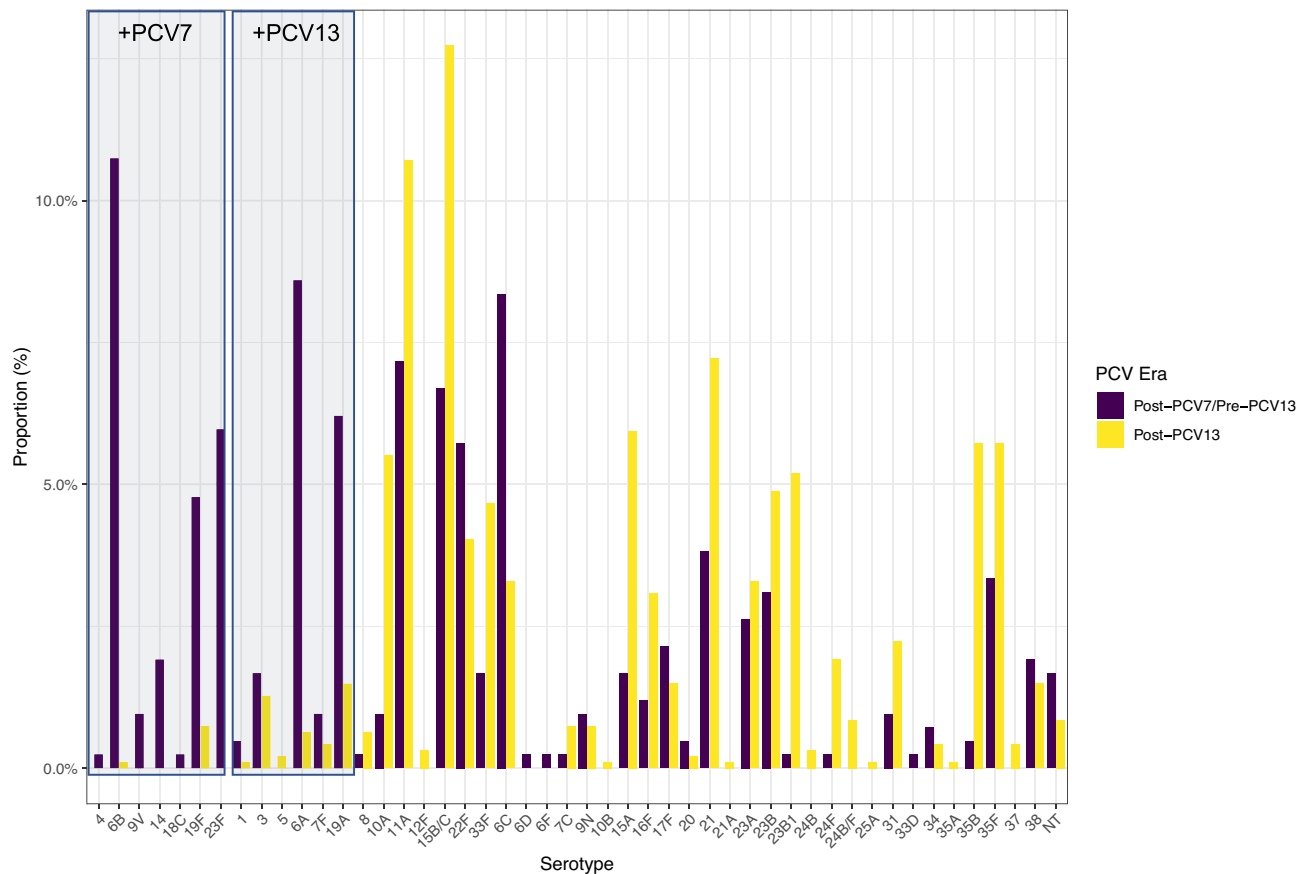
**Figure 5.** Prevalence of serotypes grouped by inclusion in PCVs, noting the reduction in PCV7 serotypes post-PCV7 from >50% (purple) and of PCV13 serotypes (blue) from a high of ~20% post-PCV13. Serotype replacement with NVTs, here split to illustrate those that will be targeted by PCV20, equates to 95% of carriage in 2017/18.

(OR 1.74, 95%CI 1.37–2.21,  $p < 0.001$ ). In contrast, those that had received antibiotics were less likely to carry *S. pneumoniae* (0.68, 95%CI 0.46–1.00) although at  $p = 0.051$  this did not pass the significance threshold. No other significant relationships were observed.

## Discussion

Monitoring the epidemiology of *S. pneumoniae* is a continuing requirement to safeguard against shifts that might negate vaccine efficacy, or to highlight expansions of disease associated NVTs. These important data may then be used to inform strategies for the design and implementation of increased valency vaccines. Here we present twelve years of data from the unique Southampton pneumococcal carriage study; the only cross-sectional, annual paediatric surveillance study that has been running since the introduction of PCV7 in the UK. We show how the pneumococcal serotype epidemiology of asymptomatic carriage in children < 5 years has altered in this defined geographic region.

Although VT serotypes have significantly decreased, our study shows they have not been eliminated from paediatric carriage completely. Analysis of VT carriage shows that PCV7 serotype 19F and PCV13 serotypes 3, 7F and 19A have all been observed post-PCV13 introduction. This is in keeping with carriage studies in the UK<sup>17</sup> as well as in the USA, where 19A accounted for 5% of pneumococci 5 years after the introduction of PCV13<sup>18</sup>. Here as well, the frequency of serotype 3, although low, remained unchanged<sup>18</sup>. Further, persistence of 19A in carriage has been noted in the Gambia<sup>19</sup>, 19A and 3 in South African mother–child pairs<sup>20</sup>, and 19A, in particular, in Sweden<sup>21</sup>. In the UK serotype 3 still causes 9.4% of all age IPD, with 19A causing 5.6% and 7F causing 1.6%. The persistence of serotype 3 as a major cause of disease has been noted in many other countries as well<sup>22,23</sup>. Previous work has determined that both limited direct (immunological) protection from childhood immunisations with knock-on effects for herd immunity are the main causes<sup>24,25</sup>. However, there is a bias towards carriage studies in young children and consequently much less is known about serotype distributions in older children and healthy adults. For example, recently Adler et al. (2019) found serotype 3 to be the most common serotype found in healthy adults during pre-screening for experimental human pneumococcal colonisation (EHPC)<sup>26</sup>. Serotype 3 was also shown to be carried by older children in England<sup>27</sup> and both serotype 3 and 19A carriage has been seen in older adults<sup>28,29</sup>. Clearly this may have important implications for our understanding of invasive potential. Nevertheless, our finding that PCV13 has had little impact on paediatric carriage of



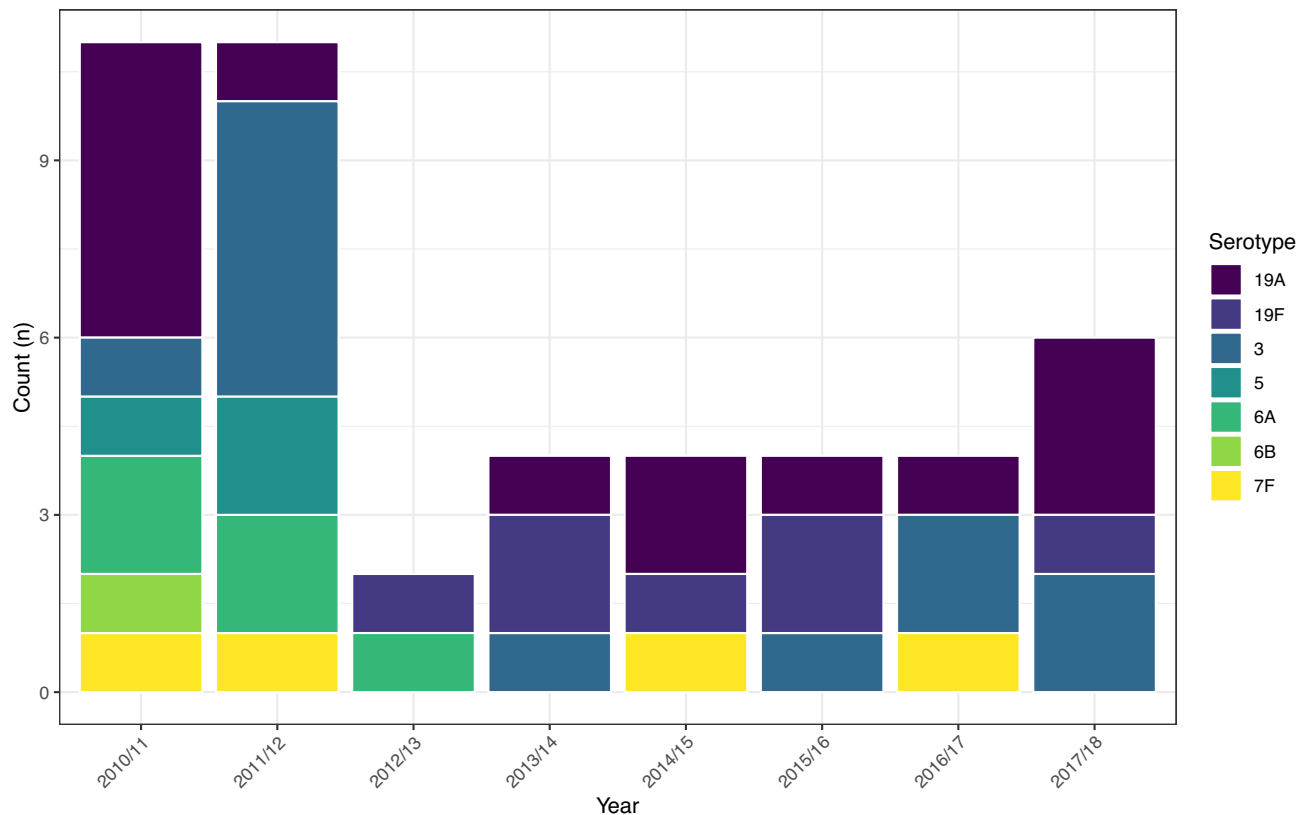
**Figure 6.** Differences in serotype proportion between the post-PCV7/pre-PCV13 era (purple) and post-PCV13 era (yellow). All PCV serotypes show a marked decrease due to PCV implementation, aside from serotype 3 which has also decreased but to a lesser extent. Notable exceptions to the increase in NVTs are 22F and 6C which have also decreased in the post-PCV13 era, although this followed earlier expansions.

serotype 3 is an important observation and is in keeping with other data<sup>21,30</sup>. Given the recent examination of a new clade of serotype 3, one that is more antibiotic resistant<sup>31</sup>, this warrants further investigation to confirm the phylogenomics of these isolates<sup>32</sup>.

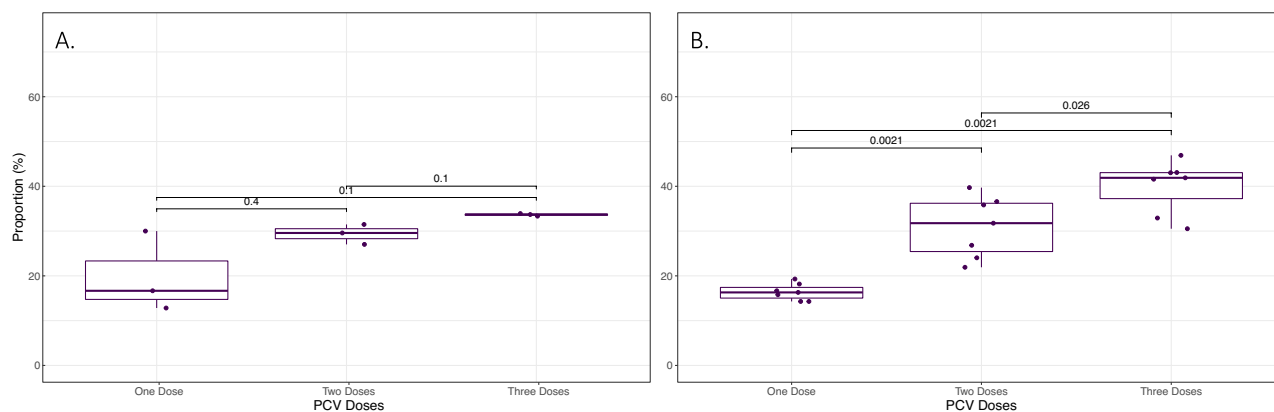
The issues of serotype replacement, plagued by unpredictability when comparing national trends<sup>10</sup>, supports the necessity for continued surveillance. Recent data on the increase in IPD caused by NVTs in adults in England and Wales attributed this burden to serotypes 8, 12F and 9N<sup>6</sup>. Elsewhere, serotype 24F has also been highlighted<sup>30</sup>. Whilst all four of these were identified in carriage, the prevalence was low with 24F ( $n = 19$ , 1.4%) being the most common followed by 9N ( $n = 11$ , 0.8%), 8 ( $n = 7$ , 0.5%) and 12F ( $n = 3$ , 0.2%). These ranked 18th, 20th, 23rd and 27th respectively in terms of the frequency of NVTs. That these rarely carried serotypes accounted for >40% of IPD in England and Wales in 2016/17 suggests high invasive potential and the need for continued surveillance. In contrast, serotypes 15A and 33F, which ranked 7th and 8th in terms of IPD, were the 6th and 11th most isolated serotypes and show a marked increase between post-PCV7/pre-PCV13 and post-PCV13 eras and suggests burden here is due to increased prevalence as opposed to invasive potential.

Our findings that recent respiratory tract infection was associated with increased risk of pneumococcal carriage has been highlighted previously<sup>33</sup>. Whilst not significant there was an indication that recent antibiotic use was associated with lower odds of carriage, again supported by previous studies<sup>34</sup>. Caution should be exercised when interpreting the findings of vaccination and number of PCV doses with increased carriage as it is difficult to disentangle the increases in carriage associated with age in the first three years of life<sup>35</sup>. The analysis presented here is clouded further by having to infer dose number from age as opposed to having immunisation dates for each participant.

The strength of this work is the considerable time over which pneumococcal carriage has been monitored and the annual nature of this surveillance. Consequently, we have been able to examine the impacts of the introduction of both PCV7 and PCV13, making important observations on, for example, the expansion of pneumococcal clones of serotype 6C and 22F<sup>36,37</sup>. It is also prudent to consider that these data provide important baselines against which to assess changes in vaccine schedule and new formulations. Nevertheless, there are several important limitations to this study. Firstly, any inference to national carriage epidemiology is just that, an inference. Community-level serotype prevalence from a proxy-population will naturally be a poorer substitute for national surveillance and will be hampered by potential geographic and demographic biases. In addition, the analysis of only one isolate from each culture positive individual ignores the impact of colonisation



**Figure 7.** Isolation of PCV-type serotypes following PCV13 introduction in 2010. Serotypes 5, 6B, 6A have not been isolated since 2012/13. In recent years, 19A and 3 have been the most frequently isolated PCV serotypes ( $n = 7$  and  $n = 5$  respectively).

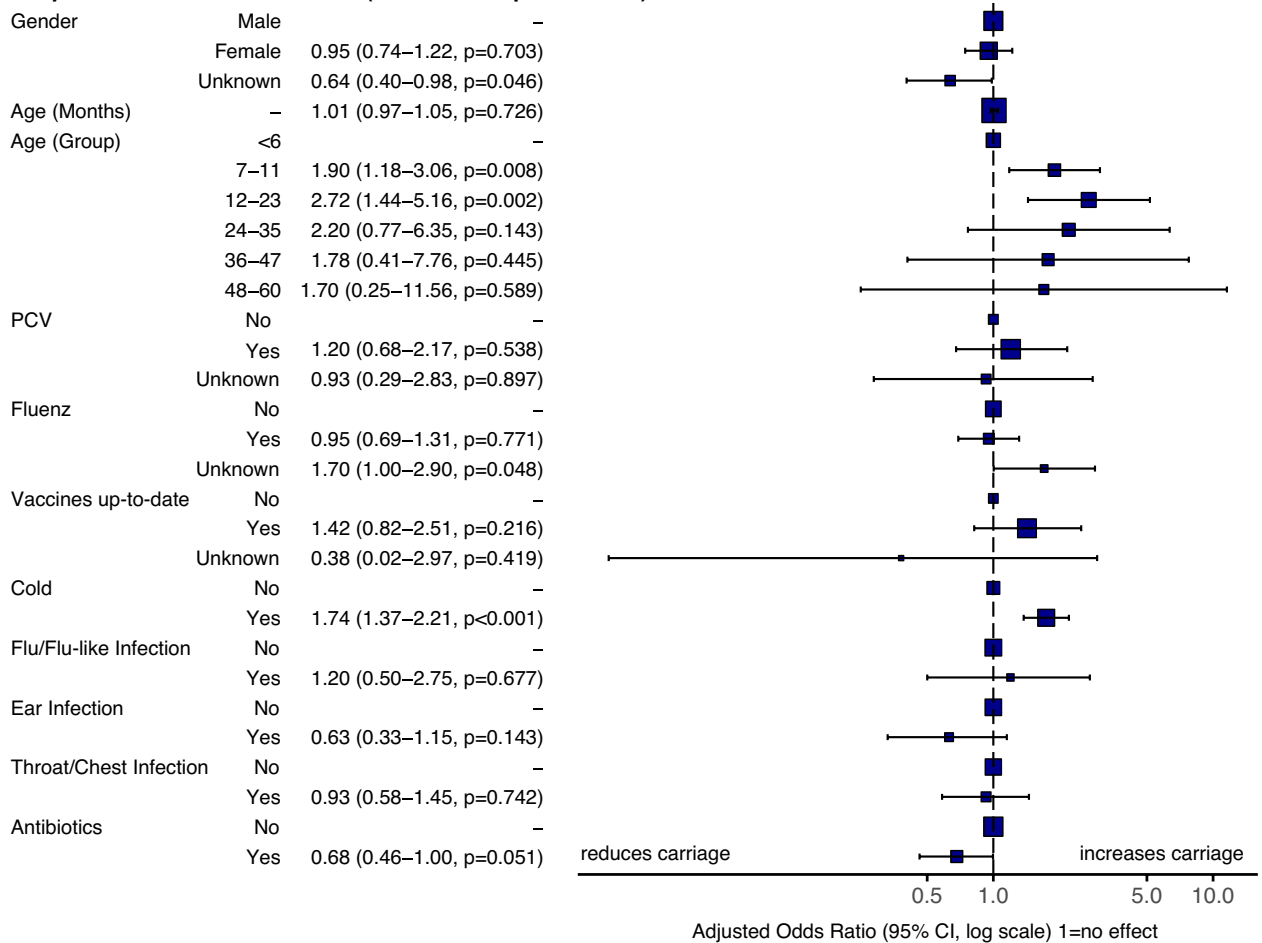


**Figure 8.** Impact of the estimated number of PCV doses on pneumococcal carriage for PCV7 (A.) and PCV13 (B.). Carriage as a proportion of number of children within each dose category was significantly higher as dose number increased for PCV13 but not PCV7. Error bars represent 95% CI. Significance values shown are from Kruskal–Wallis comparison of means.

density and multi-serotype carriage as shown elsewhere<sup>38, 39</sup>. Clearly, there are further data such as immunisation dates, co-carriage of other pathogens and/or viruses that could be collected. Indeed some has been (co-carriage for example), however, the data presented here are the most complete for this twelve-year period i.e., since the initial study design in 2006/07. Finally, the cross-sectional nature means that changes in an individual's carriage of serotypes over time are missed.

In conclusion we have shown that there is a continued circulation of serotypes 19A and 3 eight years after the introduction of PCV13. Moreover, we have demonstrated that carriage of NVT serotypes, which are now causing significant concern in IPD, are infrequently carried in our paediatric population.



***S. pneumoniae*: aOR (95% CI, p-value)**

**Figure 9.** Forest plot showing variables associated with pneumococcal carriage. Odds were calculated using a multivariable logistic regression model. Participants reporting a cold in the previous 30 days were significantly more likely to carry *S. pneumoniae* (OR 1.73:1.36–2.21,  $p < 0.001$ ). Age groups 7–11 and 12–23 months were also more likely to be pneumococcal carriers relative to those <6 months of age.

### Data availability

All sequencing data (fastqs) has been deposited in the European Nucleotide Archive under study accession PRJEB2417 (Whole genome sequencing of carried *Streptococcus pneumoniae* during the implementation of pneumococcal conjugate vaccines in the UK).

Received: 17 November 2021; Accepted: 27 July 2022

Published online: 03 August 2022

### References

1. Wahl, B. *et al.* Burden of *Streptococcus pneumoniae* and *Haemophilus influenzae* type b disease in children in the era of conjugate vaccines: Global, regional, and national estimates for 2000–15. *Lancet Glob. Health* **6**, e744–e757. [https://doi.org/10.1016/S2214-109X\(18\)30247-X](https://doi.org/10.1016/S2214-109X(18)30247-X) (2018).
2. Ganaie, F. *et al.* A new pneumococcal capsule type, 10D, is the 100th serotype and has a large cps fragment from an oral streptococcus. *MBio* **11**, e00937–e920. <https://doi.org/10.1128/mBio.00937-20> (2020).
3. Wasserman, M. *et al.* Twenty-year public health impact of 7- and 13-valent pneumococcal conjugate vaccines in US children. *Emerg. Infect. Dis.* **27**, 1627–1636 (2021).
4. Miller, E., Andrews, N. J., Waight, P. A., Slack, M. P. & George, R. C. Herd immunity and serotype replacement 4 years after seven-valent pneumococcal conjugate vaccination in England and Wales: An observational cohort study. *Lancet Infect. Dis.* **11**, 760–768. [https://doi.org/10.1016/s1473-3099\(11\)70090-1](https://doi.org/10.1016/s1473-3099(11)70090-1) (2011).
5. Tsaban, G. & Ben-Shimol, S. Indirect (herd) protection, following pneumococcal conjugated vaccines introduction: A systematic review of the literature. *Vaccine* **35**, 2882–2891. <https://doi.org/10.1016/j.vaccine.2017.04.032> (2017).
6. Ladhani, S. N. *et al.* Rapid increase in non-vaccine serotypes causing invasive pneumococcal disease in England and Wales, 2000–17: A prospective national observational cohort study. *Lancet Infect. Dis.* **18**, 441–451. [https://doi.org/10.1016/S1473-3099\(18\)30052-5](https://doi.org/10.1016/S1473-3099(18)30052-5) (2018).
7. Devine, V. T. *et al.* The rise and fall of pneumococcal serotypes carried in the PCV era. *Vaccine* **35**, 1293–1298. <https://doi.org/10.1016/j.vaccine.2017.01.035> (2017).

8. Gladstone, R. A. *et al.* Five winters of pneumococcal serotype replacement in UK carriage following PCV introduction. *Vaccine* **33**, 2015–2021. <https://doi.org/10.1016/j.vaccine.2015.03.012> (2015).
9. Løchen, A., Croucher, N. J. & Anderson, R. M. Divergent serotype replacement trends and increasing diversity in pneumococcal disease in high income settings reduce the benefit of expanding vaccine valency. *Sci. Rep.* **10**, 18977. <https://doi.org/10.1038/s41598-020-75691-5> (2020).
10. Lewnard, J. A. & Hanage, W. P. Making sense of differences in pneumococcal serotype replacement. *Lancet Infect. Dis.* **19**, e213–e220. [https://doi.org/10.1016/S1473-3099\(18\)30660-1](https://doi.org/10.1016/S1473-3099(18)30660-1) (2019).
11. Tocheva, A. S. *et al.* Declining serotype coverage of new pneumococcal conjugate vaccines relating to the carriage of *Streptococcus pneumoniae* in young children. *Vaccine* **29**, 4400–4404. <https://doi.org/10.1016/j.vaccine.2011.04.004> (2011).
12. Southampton Data Observatory, <<https://data.southampton.gov.uk/population/pop-structure/>>
13. Kapatai, G. *et al.* Whole genome sequencing of *Streptococcus pneumoniae*: Development, evaluation and verification of targets for serogroup and serotype prediction using an automated pipeline. *PeerJ* **4**, e2477. <https://doi.org/10.7717/peerj.2477> (2016).
14. R: A language and environment for statistical computing. *Computing, Vienna, Austria* (2017).
15. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, 2016).
16. Oksanen, J., Blanchet, F., Kindt, R., Legendre, P. & O'Hara, R. Vegan: Community ecology package. R Packag. 2.3-3 (2016).
17. Flasche, S. *et al.* Effect of pneumococcal conjugate vaccination on serotype-specific carriage and invasive disease in England: A cross-sectional study. *PLoS Med.* **8**, e1001017. <https://doi.org/10.1371/journal.pmed.1001017> (2011).
18. Kaur, R., Casey, J. R. & Pichichero, M. E. 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.* **35**, 901 (2016).
19. Usuf, E. *et al.* Maternal pneumococcal nasopharyngeal carriage and risk factors for neonatal carriage after the introduction of pneumococcal conjugate vaccines in The Gambia. *Clin. Microbiol. Infect.* **24**, 389–395. <https://doi.org/10.1016/j.cmi.2017.07.018> (2018).
20. Nzenze, S. A. *et al.* Temporal changes in pneumococcal colonization in HIV-infected and HIV-uninfected mother-child pairs following transitioning from 7-valent to 13-valent pneumococcal conjugate vaccine, Soweto, South Africa. *J. Infect. Dis.* **212**, 1082–1092. <https://doi.org/10.1093/infdis/jiv167> (2015).
21. Galanis, I. *et al.* Effects of PCV7 and PCV13 on invasive pneumococcal disease and carriage in Stockholm, Sweden. *Eur. Respir. J.* **47**(4), 1208–1218 (2016).
22. Lansbury, L., Lim, B., McKeever, T. M., Lawrence, H. & Lim, W. S. Non-invasive pneumococcal pneumonia due to vaccine serotypes: A systematic review and meta-analysis. *eClinicalMedicine* **44**, 101271. <https://doi.org/10.1016/j.eclinm.2022.101271> (2022).
23. Sings, H. L. *et al.* Effectiveness of 13-valent pneumococcal conjugate vaccine against invasive disease caused by serotype 3 in children: A systematic review and meta-analysis of observational studies. *Clin. Infect. Dis.* **68**, 2135–2143. <https://doi.org/10.1093/cid/ciy920> (2019).
24. Dagan, R. *et al.* Comparative immunogenicity and efficacy of 13-valent and 7-valent pneumococcal conjugate vaccines in reducing nasopharyngeal colonization: A randomized double-blind trial. *Clin. Infect. Dis.* **57**, 952–962. <https://doi.org/10.1093/cid/cit428> (2013).
25. Moore, M. R. *et al.* Effect of use of 13-valent pneumococcal conjugate vaccine in children on invasive pneumococcal disease in children and adults in the USA: Analysis of multisite, population-based surveillance. *Lancet Infect. Dis.* **15**, 301–309. [https://doi.org/10.1016/S1473-3099\(14\)71081-3](https://doi.org/10.1016/S1473-3099(14)71081-3) (2015).
26. Adler, H. *et al.* Pneumococcal colonization in healthy adult research participants in the conjugate vaccine era, United Kingdom, 2010–2017. *J. Infect. Dis.* **219**, 1989–1993. <https://doi.org/10.1093/infdis/jiz034> (2019).
27. Southern, J. *et al.* Pneumococcal carriage in children and their household contacts six years after introduction of the 13-valent pneumococcal conjugate vaccine in England. *PLoS ONE* **13**, e0195799. <https://doi.org/10.1371/journal.pone.0195799> (2018).
28. Zanella, R. C. *et al.* Nasopharyngeal carriage of *Streptococcus pneumoniae*, *Haemophilus influenzae*, and *Staphylococcus aureus* in a Brazilian elderly cohort. *PLoS ONE* **14**, e0221525–e0221525. <https://doi.org/10.1371/journal.pone.0221525> (2019).
29. Almeida, S. T. *et al.* Low prevalence of pneumococcal carriage and high serotype and genotype diversity among adults over 60 years of age living in Portugal. *PLoS ONE* **9**, e90974. <https://doi.org/10.1371/journal.pone.0090974> (2014).
30. Kandasamy, R. *et al.* Persistent circulation of vaccine serotypes and serotype replacement after 5 years of infant immunization with 13-valent pneumococcal conjugate vaccine in the United Kingdom. *J. Infect. Dis.* **221**, 1361–1370. <https://doi.org/10.1093/infdis/jiz178> (2020).
31. Azarian, T. *et al.* Global emergence and population dynamics of divergent serotype 3 CC180 pneumococci. *PLoS Pathog.* **14**, e1007438. <https://doi.org/10.1371/journal.ppat.1007438> (2018).
32. Groves, N. *et al.* Evolution of *Streptococcus pneumoniae* serotype 3 in England and Wales: A major vaccine evader. *Genes* **10**, 845. <https://doi.org/10.3390/genes10110845> (2019).
33. Ricketson, L. J. *et al.* Trends in asymptomatic nasopharyngeal colonization with *Streptococcus pneumoniae* after introduction of the 13-valent pneumococcal conjugate vaccine in Calgary, Canada. *Pediatr. Infect. Dis. J.* **33**, 724–730. <https://doi.org/10.1097/inf.000000000000267> (2014).
34. Özdemir, H. *et al.* Risk factors for nasopharyngeal carriage of *Streptococcus pneumoniae* in healthy Turkish children after the addition of heptavalent pneumococcal conjugate vaccine (PCV7) to the national vaccine schedule. *Turk. J. Pediatr.* **55**, 575–583 (2013).
35. Bogaert, D. *et al.* Colonisation by *Streptococcus pneumoniae* and *Staphylococcus aureus* in healthy children. *The Lancet* **363**, 1871–1872. [https://doi.org/10.1016/S0140-6736\(04\)16357-5](https://doi.org/10.1016/S0140-6736(04)16357-5) (2004).
36. Loman, N. J. *et al.* Clonal expansion within pneumococcal serotype 6C after use of seven-valent vaccine. *PLoS ONE* **8**, e64731. <https://doi.org/10.1371/journal.pone.0064731> (2013).
37. Cleary, D. W. *et al.* Comparative genomics of carriage and disease isolates of *Streptococcus pneumoniae* serotype 22F reveals lineage-specific divergence and niche adaptation. *Genome Biol. Evol.* **8**, 1243–1251. <https://doi.org/10.1093/gbe/evw066> (2016).
38. Chaguza, C. *et al.* Carriage dynamics of pneumococcal serotypes in naturally colonized infants in a rural African setting during the first year of life. *Front. Pediatr.* <https://doi.org/10.3389/fped.2020.587730> (2021).
39. Wyllie, A. L. *et al.* Molecular surveillance of nasopharyngeal carriage of *Streptococcus pneumoniae* in children vaccinated with conjugated polysaccharide pneumococcal vaccines. *Sci. Rep.* **6**, 23809. <https://doi.org/10.1038/srep23809> (2016).

## Acknowledgements

We would like to acknowledge the support of our technical colleagues, namely Rebecca Anderson, Denise E. Morris, Stephen Gomer and Emily J. Dineen in the collection and processing of samples for the study. We are grateful to staff in both the Pathogen Genomics team and those associated with the Global Pneumococcal Sequencing Project at the WSI for their support in genome sequencing of collected isolates. We are indebted to staff at the NIHR Southampton Wellcome Trust Clinical Research Facility for their assistance in the collection of samples and to staff at Public Health England for swab processing between 2006/07 and 2011/12. The authors also acknowledge the use of the IRIDIS High Performance Computing Facility, and associated support services

at the University of Southampton, in the completion of this work. Finally, we acknowledge both the guardians and participants without whom this study would not have been possible.

### Author contributions

S.C.C. and S.N.F. conceived and secured funding for the study with assistance from J.M.J. S.C.C. and S.N.F. were the site primary investigators. V.T.D., J.J. and R.A.G. undertook participant recruitment and microbiological isolation. Whole genome sequencing was facilitated by S.D.B., D.W.C., J.J., R.A.G., V.T.D. and K.L.O. analysed the whole genome sequencing data. D.W.C. wrote the manuscript. All authors reviewed the manuscript.

### Funding

This work was supported by Pfizer as an investigator-led research grant. Funding for whole genome sequencing was provided by the Wellcome Sanger Institute.

### Competing interests

DWC was a post-doctoral researcher on GSK funded projects in 2014/15, and currently receives grant support from Pfizer and the National Institute for Health via the NIHR Southampton Biomedical Research Centre. SNF receives support from the National Institute for Health Research funding via the NIHR Southampton Wellcome Trust Clinical Research Facility and the NIHR Southampton Biomedical Research Centre. SNF and SCC act as principal investigators for clinical trials and other studies conducted on behalf of University Hospital Southampton NHS Foundation Trust/University of Southampton that are sponsored by vaccine manufacturers. No personal payments are received from them. SNF, JMJ and SCC have participated in advisory boards for vaccine manufacturers but receive no personal payments for this work. SNF, SCC and JMJ have received financial assistance from vaccine manufacturers to attend conferences. All grants and honoraria are paid into accounts within the respective NHS Trusts or Universities, or to independent charities. RAG, VTD and JJ received PhD studentships via the University of Southampton from Pfizer. RAG received post-doctoral support in 2012 on a GSK funded research project via the University of Southampton. KLO received PhD studentship support from GSK, again via the University of Southampton. All other authors have no conflicts of interest.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-17600-6>.

**Correspondence** and requests for materials should be addressed to S.C.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022