Five winters of pneumococcal serotype replacement in UK carriage following PCV introduction

Rebecca A. Gladstone [a], Johanna M. Jefferies [a,b], Anna S. Tocheva [a], Kate R. Beard [a], David Garley [a], Wei Wei Chong [a], Stephen D. Bentley [c], Saul N. Faust [a,b,d], Stuart C. Clarke [a,b,*]

[a] Faculty of Medicine and Institute for Life Sciences, University of Southampton, UK
[b] Southampton NIHR Respiratory Biomedical Research Unit, University Hospital Southampton Foundation NHS Trust, Southampton, UK
[c] Pathogen Genomics, Wellcome Trust Sanger Institute, UK
[d] NIHR Wellcome Trust Clinical Research Facility, University Hospital Southampton Foundation NHS Trust, Southampton, UK

ARTICLE INFO

ABSTRACT

The seven-valent pneumococcal conjugate vaccine (PCV7) was added to the UK national immunisation programme in September 2006. PCV13 replaced PCV7 in April 2010. As carriage precedes disease cases this study collected carried pneumococci from children each winter from 2006/7 to 2010/11 over PCV introduction. Conventional microbiology and whole genome sequencing were utilised to characterise pneumococcal strains.

Overall prevalence of pneumococcal carriage remained stable. Vaccine serotypes (VT) decreased ($p < 0.0001$) with concomitant increases in non-vaccine serotypes (NVT). In winter 2010/11 only one isolate of PCV7 VT was observed (6B). PCV13 unique VTs decreased between winters immediately preceding and following PCV13 introduction ($p = 0.04$). Significant decreases for VTs 6B, 19F, 23F (PCV7) and 6A (PCV13) and increases for NVT 21, 23B, 33F and 35F were detected. The serotype replacement was accompanied by parallel changes in genotype prevalence for associated sequence types with clonal expansion contributing to replacement. By winter 2010/11, serotype coverage of PCV7 and PCV13 was 1% and 11% respectively.

VT replacement was observed for PCV7 and PCV13 serotypes. Conjugate vaccine design and use requires continuous monitoring and revision.

1. Introduction

14.5 million serious disease cases in children under five years old were caused annually by *Streptococcus pneumoniae* globally in the pre-vaccine era [1]. In England and Wales approximately 8800 invasive pneumococcal disease (IPD) cases in all age groups were estimated to occur every year before PCV7 introduction [2]. The pneumococcus is frequently carried asymptomatically in the nasopharynx. Carriage occurs in approximately 30% of children under five years old in the UK and is a precursor to disease [3]. This is largely comparable with other high-income countries although carriage rates of up to 70% have been observed within Europe [4–6]. High carriage rates are reported for resource-poor countries, correlating with higher disease rates [1].

The primary pneumococcal virulence factor is the polysaccharide capsule, which is classified into serotypes based on immunogenic properties. Serotypes vary in prevalence, geographical distribution and invasive potential [7,8]. Pneumococcal conjugate vaccines (PCVs) target a limited number of the >90 recognised pneumococcal serotypes. The 7-valent pneumococcal conjugate vaccine (PCV7, Prevenar7™ – Pfizer) includes seven pneumococcal polysaccharides particularly associated with invasive disease. PCV7 was added to the UK national immunisation programme in 2006 [9]. The 13-valent PCV (PCV13, Prevenar13™ – Pfizer) contains six additional serotypes and superseded PCV7 in the UK in April 2010 [10].

Since PCV7 introduction, serotype replacement was reported for both carriage and disease in the UK, motivating the introduction of PCV13 [2,3].

* Corresponding author at: Mailpoint 814, Level C, South Academic Block, Southampton General Hospital, Southampton SO14 6YD, UK.
Tel.: +44 02381 208895.
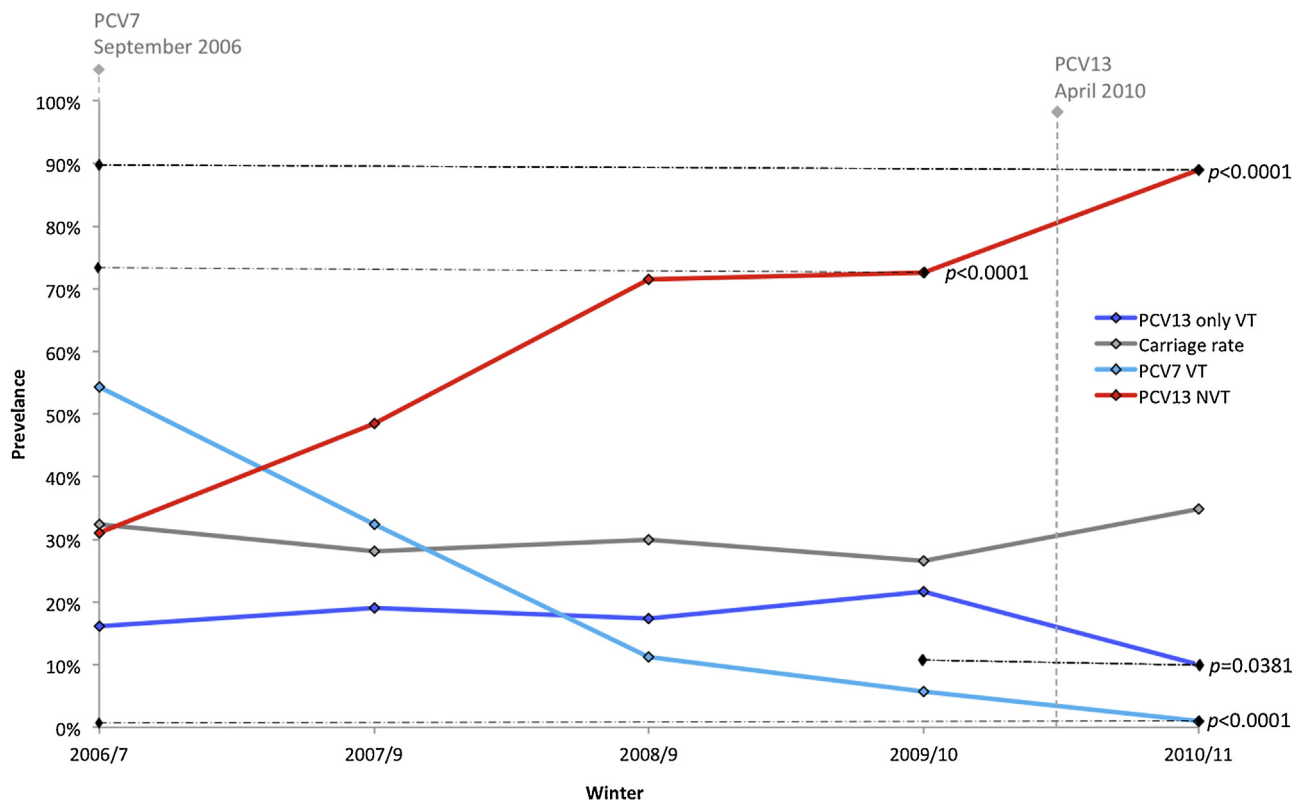E-mail address: s.c.clarke@soton.ac.uk (S.C. Clarke).

Fig. 1. Carriage rate and temporal prevalence of PCV VT and NVT. Dashed black horizontal lines denote the time period between which the significant change was observed. Dashed grey vertical lines denote vaccine introductions with respect to sampling time points

The diversity of the pneumococcus allows rapid response to selective pressure exerted by wide spread immunisation. It is therefore important to describe and monitor serotype and genotype of both carriage and disease to understand pneumococcal population dynamics. Whole genomes offer a resource capable of screening sequence types (ST) and infer serotype as well as high-resolution genome wide analysis and comparisons. We sought to integrate next generation sequencing technologies into routine microbiological surveillance. Here we report a continuation of observations made for the first three study winters during which only PCV7 had been introduced, through two further winters surrounding PCV13 introduction.

2. Methods

Nasopharyngeal swabs were collected from children aged 4 years and under during five consecutive winters, October–March 2006/7–2010/11. The initial three years of the study have been previously been described [3,11]. Briefly, guardians were approached within Southampton General Hospital outpatient department. Written informed consent was obtained with ethical approval from Southampton and South West Hampshire Research Ethics Committee 'B' (REC 06/Q1704/105). The child swabbed was not necessarily the child attending clinic but a sibling. One child per family was swabbed. Age was the primary exclusion criteria. Rayon tipped Transwabs (Medical wire, Corsham, UK) in charcoal Amies media were plated onto Columbia Colistin Naladixic Acid agar at Public Health England (PHE) regional laboratory, Southampton within 9 h. Presumptive *S. pneumoniae* were plated on Columbia blood agar with an optochin disc and confirmed with a ≥14 mm diameter inhibition zone. Up to 10 well-isolated pneumococcal colonies for each initial plate were stored at −70 °C. Sampling continued until ≥100 *S. pneumoniae* were collected each winter. In winter

2010/11 a questionnaire was completed by the guardian to determine if the participant had likely received age appropriate doses of PCVs.

Illumina sequencing was performed at the Wellcome Trust Sanger Institute with 75 bp paired-ends. Data is deposited in the European nucleotide database with accession numbers: ERR044859: ERR044953, ERR044955: ERR044987, ERR044989: ERR048210, ERR048212: ERR045394, ERR045396: ERR045400, ERR045402: ERR045413, ERR045415: ERR045431, ERR045433, ERR045434, ERR045436: ERR047110, ERR047112: ERR047149, ERR047151: ERR047163. De novo assembly was performed using Velvet assembler [12] and Velvet optimiser [13]. Serotype was inferred from whole genome data using an in silico adaptation of the Centres for Disease Control and Prevention PCR for deducing serotype [14], the presence of known serotype specific sequences in assemblies and mapping against known capsular loci as previously described [15,16]. Phenotypic serotyping was used if the genetic basis for serotype was unknown. MLST was determined from whole genomes as previously described [16].

To assess changes in prevalence, confidence intervals (CI) were calculated at 95% confidence. Significant changes in proportions between two groups were detected using Fisher's exact test with two-tailed *p*-values.

3. Results

1712 swabs were collected between winters 2006/7 and 2010/11. 519 pneumococci were isolated from which 516 genomes were available to deduce epidemiological data. The overall carriage rate was 30% (CI, 28–32%). The carriage rate remained stable with no significant differences between 2006/7 and any other year, with *p*-values of >0.1 (Fig. 1). The appropriate number of doses of PCV7 or PCV13 had been received for 91% (CI, 85–97%) of non-carriers
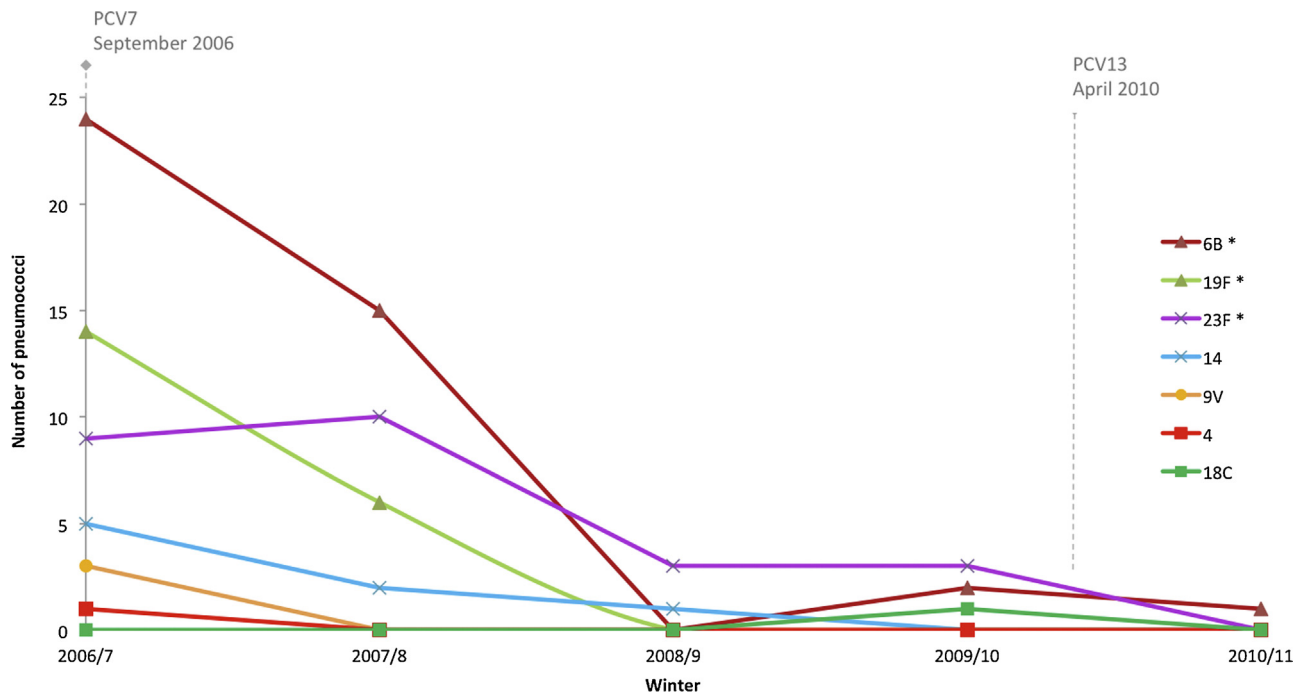
Fig. 2. Temporal prevalence of PCV7 serotypes. *Significant decrease observed between base-line winter 2006/7 and final study winter 2010/11

and 88.79% (CI, 85–97%) of carriers in winter 2010/11 according to guardian responses.

3.1. Deriving serotype from genotype

Serotyping from whole genome data allowed greater discriminatory power than PCR methods and detection of newly described serotypes for which other methods were not yet available, with the additional potential for discovery of novel variation defining putative novel serotypes. We documented sequence variation within serotypes, by looking at a number of serotype-defining genes for prevalent serotypes. Whilst within-serotype capsular polysaccharide gene variation has been previously thought to be limited [15,17] we observed considerable genetic variation in capsular specific genes within serogroup 6, serotypes 15B/C, 19A, 19F and 21 (Supplementary Table 7).

Serogroup 6 was the most prevalent serogroup, 6A ($n = 38$), 6B ($n = 45$) and 6C ($n = 44$). Genome derived serotype inference resulted in the detection of single isolates of serotype 6D [18].

Notably one non-synonymous single nucleotide polymorphism (SNP) observed in this study resulted in a change from Ala150Thr in WciNα in a 6A background. This SNP was subsequently reported by Oliver et al. to result in a hybrid serological profile putatively designated NVT 6F [19]. In total there were three WciNα 6A alleles. The dominant 6B WciNα allele was identical to the dominant 6A WciNα allele. Two additional 6B protein alleles also existed, each observed in one isolate. The single representative of the 6D WciNβ allele differed from the dominant 6C protein sequence by only one amino acid (AA). Four WciP alleles were observed within each serotype 6A, 6B and 6C with three variable positions in 6A and 6C and five in 6B.

WcrL contains the discriminatory SNP between 11A/D. PCR could not differentiate these serotypes. All WcrL alleles within this study were identical except in one isolate. No isolates contained the 11D non-synonymous SNP and all could be therefore designated 11A [20,21].

15B isolates have been reported to contain 8 repeats and 15C isolates contain 7 or 9 TA repeats, which allowed their differentiation [22]. However two additional isolates were observed in this study to have only 6 TA repeats that would also result in a non-functional O-acetyltransferase and consequently the phenotype of 15C.

Serotypes 19A and 19F structurally differ due to varying sequence and function of the Wzy repeat unit polymerase [15,23]. Three distinct Wzy alleles were observed within the 31 19A isolates. The two most prevalent Wzy alleles differing by two AAs were found in multiple ST backgrounds, both observed within ST199, the dominant 19A clone in this study. Two distinct Wzy alleles were observed for 19F isolates, both circulating in the population with equal prevalence. While for serotype 21 two equally circulating alleles of Wzx were observed belonging to different clonal complexes.

3.2. Serotype replacement

Thirty-nine serotypes were observed. The only PCV7 VT present in all winters was 6B, with a single isolate detected in winter 2010/11. Three VTs unique to PCV13 (6A, 19A and 3) and 9 NVTs were observed in all winters. PCV7 VTs significantly decreased between winters 2006/7 and 2010/11 ($p < 0.0001$) with a concomitant increase in NVTs (Figs. 1 and 2; Supplementary Table 1). The prevalence of PCV13 VTs significantly decreased between winters 2006/7 and 2010/11 ($p < 0.0001$) with a concomitant NVT increase. PCV13 unique serotypes significantly decreased (Fig. 1, $p = 0.04$) between winters 2009/10 and 2010/11 (Figs. 1 and 3; Supplementary Tables 2 and 3).

By winter 2010/11 the prevalence of PCV7 VT in carriage was only 1%, falling from 53% in 2006/7. PCV13 VT prevalence in 2006/7 was 68% (CI, 59–77%) dropping to 26% (CI, 18–35%) before PCV13 introduction, due to shared PCV7 serotypes. By winter 2010/11 only 11% of carried pneumococci were PCV13 VT (CI, 5–17%).

Significant decreases between winters 2006/7 and 2010/11 were observed for PCV7 serotypes 6B ($p < 0.01$), 19F ($p < 0.01$), 23F ($p < 0.01$) and PCV7 vaccine related type (VRT)/PCV13 VT 6A ($p < 0.05$) (Figs. 2 and 3, Supplementary Table 4). A non-significant increase ($p = 0.07$) was observed for PCV13 VT 19A between winter
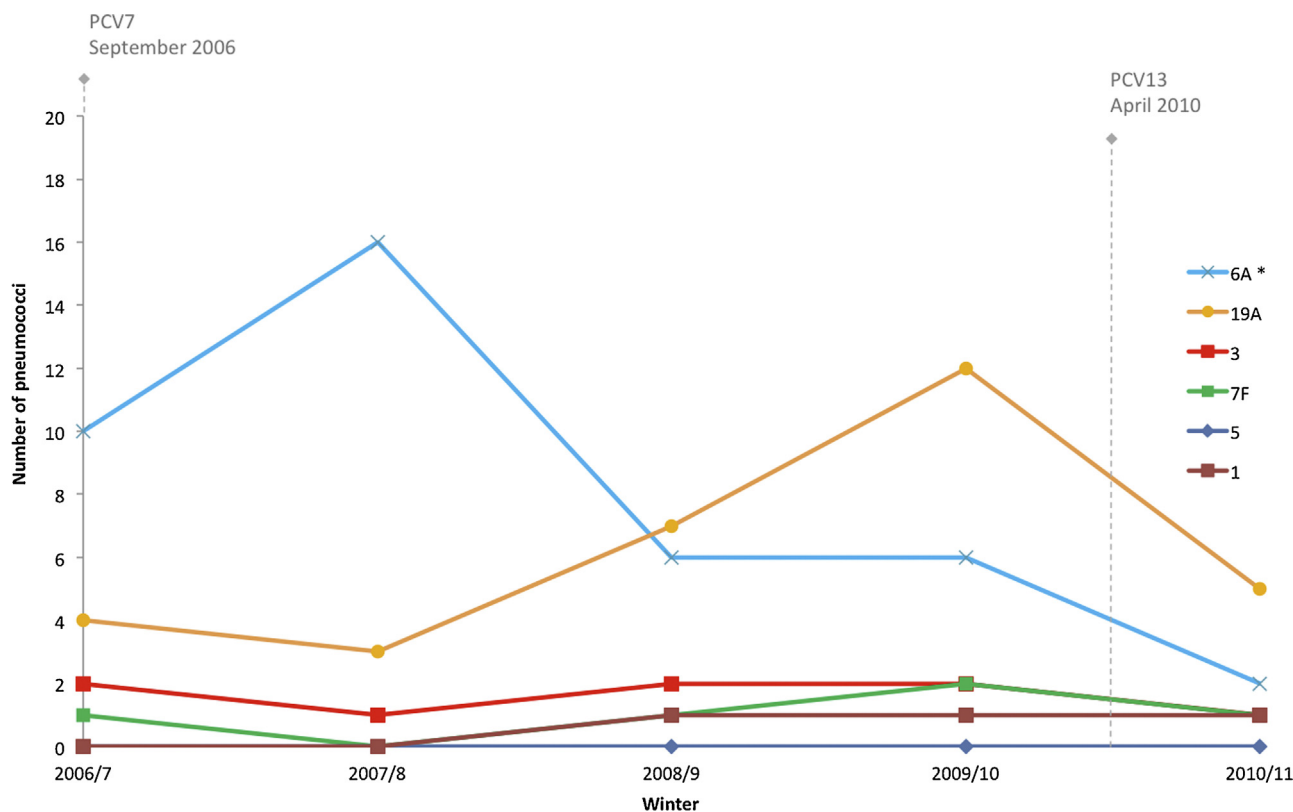
Fig. 3. Temporal prevalence of serotypes unique to PCV13. *Significant decrease observed between base-line winter 2006/7 and final study winter 2010/11

2006/7 and winter 2009/10 before PCV13 introduction. A decrease in PCV7 serotype 14 between winters 2006/7 and 2010/11 did not reach significance ($p = 0.06$), however numbers were low (2006/7 $n = 5$, 2010/11 $n = 0$). PCV13 VT 5 was not observed in the data set. Fluctuations observed for PCV serotypes 1, 3, 4, 7F, 9V and 18C were not significant ($p > 0.40$).

Twenty-seven PCV13 NVTs were observed. Of the NVTs, significant increases between winters 2006/7 and 2010/11 were observed for serotypes 21 ($p < 0.01$), 23B ($p < 0.02$), 33F ($p < 0.01$) and 35F ($p < 0.01$).

Non-significant increases were observed for serotype 15A ($p = 0.06$) and 6C ($p = 0.08$, Fig. 4). Serotypes 6C and 11A had significant fluctuations within the study period not detectable when comparing winters 2006/7 and 2010/11 (Supplementary Table 5). The increase of 6C between winters 2006/7 and 2008/9 was previously reported as a consequence of clonal expansion [24,25]. A significant decrease ($p = 0.01$) was then observed for 6C between winters 2008/9 and 2009/10 (Fig. 4).

3.3. Sequence types

127 STs were observed, derived from whole genomes. 61% of the STs were seen in only one winter, as a single or pair of isolates. Only 8 STs were observed in all winters. Significant changes between 2006/7 and 2010/11 were observed for seven ST's: 138, 162 and 176 (decreases) and 100, 432, 439 and 1635 (increases) (Supplementary Table 6). Fluctuations observed between winters 2006/7 and 2010/11 for the remaining 119 sequence types were not significant ($p > 0.10$). 15 STs were shared between the two years. Seventeen clonal complexes (CCs, share 6 of 7 alleles) were observed encompassing 54 different STs. Changes for STs remained significant when CCs were taken into account except ST439 as part of CC42 ($p > 0.10$).

3.4. Serotype and ST associations

For serotypes with significant changes between winters 2006/7 and 2010/11 there were also changes observed in the dominant ST (Fig. 5, Supplementary Table 6). A significant decrease ($p < 0.01$) could be only observed for ST65 the dominant ST for 6A between winter 2007/8 and 2010/11 after an initial increase of 6A following PCV7 introduction.

Sixteen combinations of ST and serotype and 17 combinations of CC and serotype were shared between the primary winter 2006/7 and the final winter 2010/11. These serotype lineages accounted for 39% (STs) and 59% (CC) of the isolates in 2010/11.

Isolates that shared an ST with VTs but express a NVT due to capsule switching were detected. ST 414 was observed expressing NVT 16F ($n = 1$) and VT serotype 1 ($n = 1$). ST193 was observed expressing NVT serotype 21 ($n = 11$) and VT 19A ($n = 1$). ST199 was observed expressing NVTs 15B ($n = 14$) and 15C ($n = 7$) and VT 19A ($n = 23$).

4. Discussion

Significant changes in prevalence of PCV7 VTs and NVTs previously reported continued to be observed through winter 2009/10 to the final winter (2010/11) [3]. This suggests that changes to serotype distribution are stable, in line with PCV use. Serotypes were known to fluctuate temporally before PCV introduction as part of natural cycles [26]. This phenomenon could explain within-study fluctuation of NVTs 11A and 6C. PCV7 could also have impacted prevalence of vaccine related serotype 6A during the study period through cross-reactive antibodies [27,28], post PCV7 6A increases were observed in the Netherlands. This may have contributed to the observation that PCV13 VT decreased after PCV13 introduction [6]. Conversely there is evidence that cross-protection does not occur for 19A [29]. Other observed changes in prevalence may reach
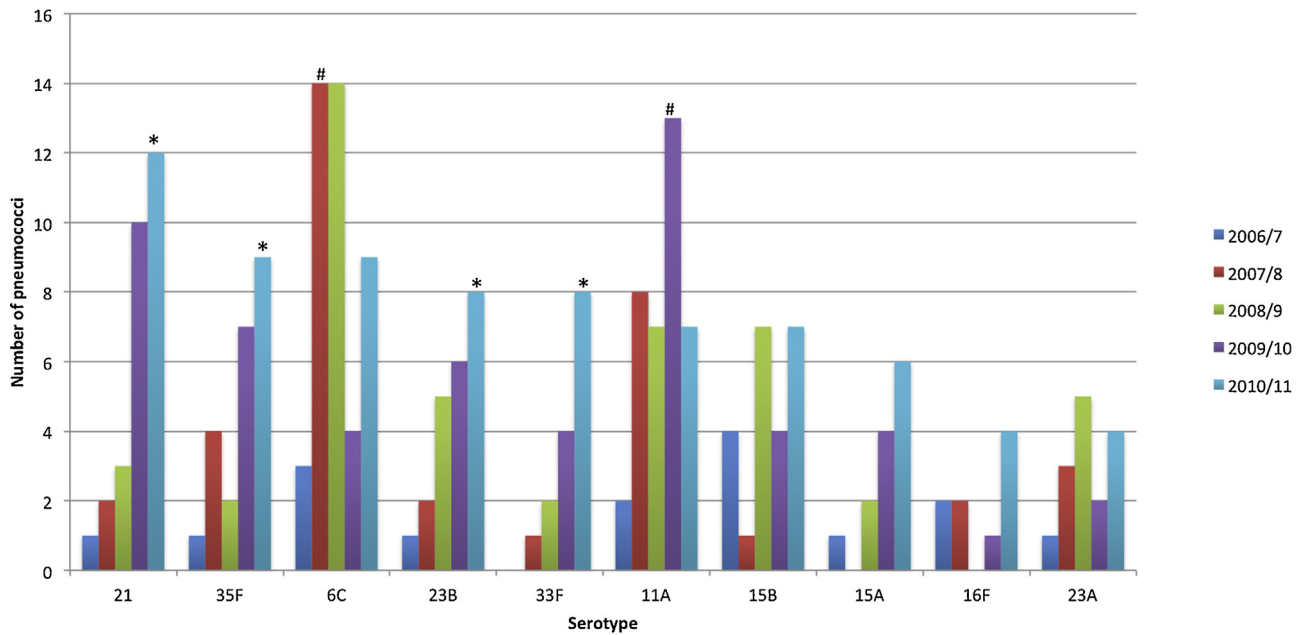
Fig. 4. Temporal prevalence for the top 10 NVT in final winter 2010/11. * Significant increase between base-line winter 2006/7 and final study winter 2010/11. # Significant increase within study not observable between base-line winter 2006/7 and final study winter 2010/11

significance in winters following 2010/11, but the ability to detect changes could be limited by the sample size. Further surveillance is needed to observe whether changes are maintained. Understanding of multiple serotype colonisation, particularly the detection of minority serotypes may further inform post PCV pneumococcal carriage dynamics. Whilst additional samples were stored, preliminary analysis of multiple colonies in the primary study winters

2006/7–2008/9 did not allow detection of multiple serotype colonisation. This traditional culture based approach is known to be less sensitive to alternative methods, such as microarray, for detecting multiple serotype colonisation [30].

The association between serotype and ST is well established and evident in this study. Clonal expansions of STs contributed to both fluctuations in serotype prevalence within the study such as 6C
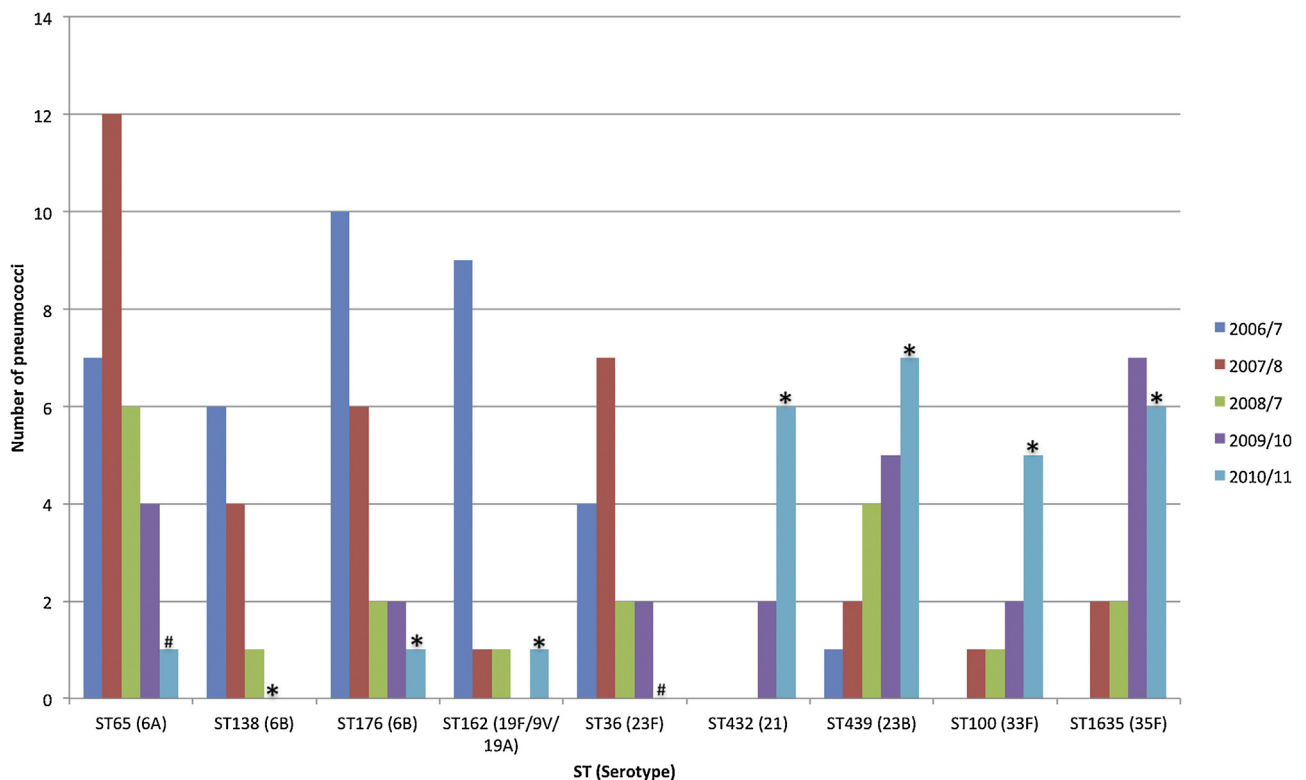

Fig. 5. Temporal prevalence of STs associated with changes in serotype prevalence between winter 2006/7 and winter 2010/11. *Significant change between base-line winter 2006/7 and final study winter 2010/11. # Significant change within study observable between winter 2007/8 and final study winter 2010/11

ST1692, and sustained changes in prevalence for example serotype 21 (ST432) and 23B (ST439) [25].

ST and serotype combinations detected that were indicative of capsule switch had historically been documented in the pneumococcal MLST database in the pre-PCV era [31]. Additionally the majority of isolates (59%) in winter of 2010/11 were of serotype and ST combinations also present in winter of 2006/7 and are therefore increased prevalence of established clones. Both findings suggest that the majority of replacement involves lineages already present in the population, complimenting previous findings that vaccine-escape isolates were a consequence of expansion of existing lineages [32].

4.1. Geographic specificity of carriage

Independent carriage studies from the UK and US also reported no detectable change in the carriage rate [32–35]. Our carriage rate better reflected observations from the US being lower than reported elsewhere in the UK but consistent with other Hampshire studies [36]. A substantial increase in 19A in both carriage and disease was reported in the US where 19A was ranked first in carriage for isolates collected in 2008/09, whereas 19A increases before PCV13 introduction were not significant in this study [34,37].

The rank order of prevalence of the top 10 PCV13 NVT in the winter 2010/11 shared eight of the top 10 NVT reported for another UK study and interestingly also in the US study despite being geographically disparate [33,35]. Furthermore when looking at the rank order of prevalence for the top 10 PCV13 NVT in post PCV carriage in children from a selection of countries (UK, US, Netherlands and Kenya), serotypes 11A, the 15B/C, 15A and US were reported in all four countries [38,39]. This was despite differing use of PCVs with Kenya introducing the 10-valent PCV where PCV13 VT 19A and 6A were still highly prevalent. The exact order of prevalence differed in each location but these comparisons illustrate similarities between distant populations that could help predict changes in other locations but also demonstrate geography-specific prevalence for which on-going local surveillance is needed.

4.2. From carriage to disease in the UK

A 98% reduction in PCV7 VT IPD in children less than two years old was reported between 2000–06 and 2009–10, mirroring the near complete removal of PCV7 VT carriage in this study [2].

NVT replacement also occurred in disease, with PHE reporting a 68% increase in NVT IPD in children less than 2 years of age [2]. In carriage, replacement was near complete and the overall carriage rate remained unchanged whilst for disease the reduction in PCV7 VT was not completely offset by increases in NVT disease. In 2009–2010 a 34% reduction in the annual incidence of all IPD cases from that of 2000–2006 was reported, a smaller reduction than of VT IPD as a result of serotype replacement [2]. The key serotypes involved in PCV7 disease replacement in 2010 were reported to be 7F, 19A and 22F [2]. Changes in these serotypes in carriage were not significant, which may be a factor of sample size. Serotypes 7F and 19A have a high invasive disease potential (odds ratios for invasive disease of 4.7 and 1.1 respectively) [8,40], whilst an odds ratio for 22F of zero was previously calculated, but isolate numbers were low [40].

Prevention of invasive disease cases through preventing carriage (herd protection) in addition to direct prevention of invasion is a key feature of PCVs that contributes to their cost effectiveness. The significant reduction in PCV13 unique serotypes in carriage observed in this study should increase the impact of PCV13 beyond its direct immunogenic effects, indirectly via herd protection. A 50% reduction in PCV13 specific IPD in children under two years of age was reported by PHE in the year following PCV13 introduction [41].

Whole genome data has proved a valuable resource for the derivation of routine pneumococcal typing to inform epidemiological analysis with the opportunity for further in-depth analysis. Additionally the availability of large collections of pneumococcal sequence data results in the opportunity to detect variation such as genetic variation affecting protein sequence within a serotype. A single point mutation within a capsular gene encoding, for example a glycosyltransferase, can result in different substrate specificity that in turn alters capsular biosynthesis. SNPs with such consequences are well documented including the 6F variant demonstrated to be immunologically distinct to 6A [19]. Therefore further novel variation also have the potential to affect capsule expression, immunological properties and ultimately current or future vaccine efficacy and effectiveness [19,42].

The data though further highlights that our understanding of serotype, particularly the genotypic contribution, is not yet complete. Documented variation could give direction to continued discrimination of serotypes and increase specificity of future genotypic inference, so that surveillance, serotype specific vaccine efficacy, modelling of future replacement, and cost effectiveness estimates are as accurate as possible.

This study shows significant changes in prevalence for individual PCV7 serotypes detected by the fifth study winter. This study similarly shows a significant decrease in PCV13 VT between winters preceding and following PCV13 introduction although this cannot be directly attributed to PCV13 as multiple factors could have contributed to this, including PCV7 serogroup cross-reactivity, natural fluctuations and movement back towards equilibrium [26,27,37]. Guardian reporting for winter 2010/11 suggests age appropriate PCV immunisation was high for our study population, further information of the vaccine status of the individuals would help tease apart the indirect and direct effects for PCV13. On-going substantial serotype replacement further indicates that continuing surveillance of carriage in addition to IPD is required to inform the future of pneumococcal vaccination.

Conflict of interest

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.vaccine.2015.03.012.

References

[1] O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, Mccall N, et al. Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates. Lancet 2009;374(9693):893–902.

[2] 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(10):760–8.

[3] Tocheva AS, Jefferies JM, Rubery H, Bennett J, Afimeke G, Garland J, et al. Declining serotype coverage of new pneumococcal conjugate vaccines relating to the carriage of *Streptococcus pneumoniae* in young children. Vaccine 2011;29(26):4400–4.

[4] Roche A, Heath PT, Sharland M, Strachan D, Breathnach A, Haigh J, et al. Prevalence of nasopharyngeal carriage of pneumococcus in preschool children attending day care in London. Arch Dis Child 2007;92(12):1073–6.

[5] Hanage WP, Bishop CJ, Huang SS, Stevenson AE, Pelton SI, Lipsitch M, et al. Carried pneumococci in Massachusetts children: the contribution of clonal expansion and serotype switching. Pediatr Infect Dis J 2011;30(4):302–8.

[6] Spijkerman J, Prevaes SM, Van Gils EJ, Veenhoven RH, Bruin JP, Bogaert D, et al. Long-term effects of pneumococcal conjugate vaccine on nasopharyngeal carriage of *S. pneumoniae*, *S. aureus*, *H. influenzae* and *M. catarrhalis*. PLoS ONE 2012;7(6):e39730.

[7] Hausdorff WP, Feikin DR, Klugman KP. Epidemiological differences among pneumococcal serotypes. Lancet Infect Dis 2005;5(2):83–93.

[8] Brueggemann AB, Peto TE, Crook DW, Butler JC, Kristinsson KG. Temporal and geographic stability of the serogroup-specific invasive disease potential of *Streptococcus pneumoniae* in Children. J Infect Dis 2004;190(7):1203–11.

[9] Spratt BG, Department of Health. Important changes to the childhood immunisation programme, PL/CMO/2006/1; 2006.

[10] Department of Health. Introduction of Prevenar 13® into the childhood immunisation programme. Department of Health; 2010 [Gateway reference: 13581].

[11] Tocheva AS, Jefferies JM, Christodoulides M, Faust SN, Clarke SC. Distribution of carried pneumococcal clones in UK children following the introduction of the 7-valent pneumococcal conjugate vaccine: a 3-year cross-sectional population based analysis. Vaccine 2013;31(31):3187–90.

[12] Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 2008;18(5):821–9.

[13] Victorian Bioinformatics Consortium. Velvet optimiser; 2008. https://github.com/Victorian-Bioinformatics-Consortium/VelvetOptimiser [accessed 2011].

[14] Centers for Disease Control and Prevention. PCR deduction of pneumococcal serotypes; 2012. http://www.cdc.gov/ncidod/biotech/strep/pcr.htm [accessed 01.12.12].

[15] Mavroidi A, Aanensen DM, Godoy D, Skovsted IC, Kaltoft MS, Reeves PR, et al. Genetic relatedness of the *Streptococcus pneumoniae* capsular biosynthetic loci. J Bacteriol 2007;189(21):7841–55.

[16] Croucher NJ, Harris SR, Fraser C, Quail MA, Burton J, Van Der Linden M, et al. Rapid pneumococcal evolution in response to clinical interventions. Science 2011;331(6016):430–4.

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

[18] Bratcher PE, Kim KH, Kang JH, Hong JY, Nahm MH. Identification of natural pneumococcal isolates expressing serotype 6D by genetic, biochemical and serological characterization. Microbiology 2010;156(Pt 2):555–60.

[19] Oliver MB, Van Der Linden MP, Kuntzel SA, Saad JS, Nahm MH. Discovery of *Streptococcus pneumoniae* serotype 6 variants with glycosyltransferases synthesizing two differing repeating units. J Biol Chem 2013.

[20] Calix JJ, Nahm MH, Zartler ER. Elucidation of structural and antigenic properties of pneumococcal serotype 11A: 11B, 11C, and 11F polysaccharide capsules. J Bacteriol 2011;193(19):5271–8.

[21] Oliver MB, Jones C, Larson TR, Calix JJ, Zartler ER, Yother J, et al. *Streptococcus pneumoniae* serotype 11D has a bi-specific glycosyltransferase and expresses two different capsular polysaccharide repeating units. J Biol Chem 2013.

[22] van Selm S, Van Cann LM, Kolkman MA, Van Der Zeijst BA, Van Putten JP. Genetic basis for the structural difference between *Streptococcus pneumoniae* serotype 15B and 15C capsular polysaccharides. Infect Immun 2003;71(11):6192–8.

[23] Morona JK, Morona R, Paton JC. Comparative genetics of capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* types belonging to serogroup 19. J Bacteriol 1999;181(17):5355–64.

[24] Tocheva A, Jefferies JM, Christodoulides M, Faust SN, Clarke SC. Increase in serotype 6C pneumococcal carriage during the implementation of the heptavalent pneumococcal conjugate vaccine. Emerg Infect Dis 2010;16(1).

[25] Loman NJ, Gladstone RA, Constantinidou C, Tocheva AS, Jefferies JM, Faust SN, et al. Clonal expansion within pneumococcal serotype 6C after use of seven-valent vaccine. PLoS ONE 2013;8(5):e64731.

[26] Jefferies JM, Smith AJ, Edwards GFS, Mcmenamin J, Mitchell TJ, Clarke SC. Temporal analysis of invasive pneumococcal clones from Scotland illustrates fluctuations in diversity of serotype and genotype in the absence of pneumococcal conjugate vaccine. J Clin Microbiol 2010;48(1):87–96.

[27] Vakevainen M, Eklund C, Eskola J, Kayhty H. Cross-reactivity of antibodies to Type 6B and 6A polysaccharides of *Streptococcus pneumoniae*: evoked by pneumococcal conjugate vaccines, in infants. J Infect Dis 2001;184(6):789–93.

[28] Lee H, Nahm MH, Burton R, Kim K. Immune response in infants to the heptavalent pneumococcal conjugate vaccine against vaccine-related serotypes 6A and 19A. Clin Vaccine Immunol 2009;16(3):376–81.

[29] van Gils EJ, Veenhoven RH, Hak E, Rodenburg GD, Keijzers WC, Bogaert D, et al. Pneumococcal conjugate vaccination and nasopharyngeal acquisition of pneumococcal serotype 19A strains. JAMA 2010;304(10):1099–106.

[30] Turner P, Hinds J, Turner C, Jankhot A, Gould K, Bentley SD, et al. Improved detection of nasopharyngeal cocolonization by multiple pneumococcal serotypes by use of latex agglutination or molecular serotyping by microarray. J Clin Microbiol 2011;49(5):1784–9.

[31] http://pubmlst.org/spneumoniae, Multi-locus sequence typing. http://pubmlst.org/spneumoniae/. 2014. Accessed 01/06/14 as MLST.net since migrated to http://pubmlst.org/spneumoniae/

[32] Croucher NJ, Finkelstein JA, Pelton SI, Mitchell PK, Lee GM, Parkhill J, et al. Population genomics of post-vaccine changes in pneumococcal epidemiology. Nat Genet 2013;45(6):656–63.

[33] Lee GM, Kleinman K, Pelton SI, Hanage W, Huang SS, Lakoma M, et al. Impact of 13-valent pneumococcal conjugate vaccination on carriage in young children in Massachusetts. J Pediatric Infect Dis Soc 2014;3(1):23–32.

[34] Flasche S, Van Hoek AJ, Sheasby E, Waight P, Andrews N, Sheppard C, et al. Effect of pneumococcal conjugate vaccination on serotype-specific carriage and invasive disease in England: a cross-sectional study. PLoS Med 2011;8(4):pe1001017.

[35] van Hoek AJ, Sheppard CL, Andrews NJ, Waight PA, Slack MP, Harrison TG, et al. Pneumococcal carriage in children and adults two years after introduction of the thirteen valent pneumococcal conjugate vaccine in England. Vaccine 2014;32(34):4349–55.

[36] Coughtrie AL, Whittaker RN, Begum N, Anderson R, Tuck A, Faust SN, et al. Evaluation of swabbing methods for estimating the prevalence of bacterial carriage in the upper respiratory tract: a cross sectional study. BMJ Open 2014;4(10):pe005341.

[37] Hanage WP, Finkelstein JA, Huang SS, Pelton SI, Stevenson AE, Kleinman K, et al. Evidence that pneumococcal serotype replacement in Massachusetts following conjugate vaccination is now complete. Epidemics 2010;2(2):80–4.

[38] Hammitt LL, Akech DO, Morpeth SC, Karani A, Kihuha N, Nyongesa S, et al. Population effect of 10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of Streptococcus pneumoniae and non-typeable Haemophilus influenzae in Kilifi, Kenya: findings from cross-sectional carriage studies. Lancet Glob Health 2014;2(7):e397–405.

[39] Spijkerman J, Van Gils EJ, Veenhoven RH, Hak E, Yzerman EP, Van Der Ende A, et al. Carriage of *Streptococcus pneumoniae* 3 years after start of vaccination program: the Netherlands. Emerg Infect Dis 2011;17(4):584–91.

[40] Brueggemann AB, Griffiths DT, Meats E, Peto TE, Crook DW, Spratt BG. Clonal relationships between invasive and carriage *Streptococcus pneumoniae* and serotype- and clone-specific differences in invasive disease potential. J Infect Dis 2003;187(9):1424–32.

[41] Miller E, Andrews NJ, Waight PA, Slack MP, George RC. Effectiveness of the new serotypes in the 13-valent pneumococcal conjugate vaccine. Vaccine 2011;29(49):9127–31.

[42] Bratcher PE, Park IH, Oliver MB, Hortal M, Camilli R, Hollingshead SK, et al. Evolution of the capsular gene locus of *Streptococcus pneumoniae* serogroup 6. Microbiology 2011;157(Pt 1):189–98.