Validation of Fourier Transform Infrared Spectroscopy for Serotyping of *Streptococcus pneumoniae*

I. Passaris,[a] N. Mauder,[b] M. Kostrzewa,[b] I. Burckhardt,[c] S. Zimmermann,[c] N. M. van Sorge,[d,e] H.-C. Slotved,[f] S. Desmet,[g,h] P.-J. Ceyssens[a]

aBacterial Diseases Unit, Sciensano, Brussels, Belgium
bBruker Daltonic GmbH, Bremen, Germany
cDepartment of Infectious Diseases, Microbiology and Hygiene, University Hospital of Heidelberg, Heidelberg, Germany
dDepartment of Medical Microbiology and Infection Prevention, UMC Amsterdam, University of Amsterdam, Amsterdam, The Netherlands
eNetherlands Reference Laboratory for Bacterial Meningitis, UMC Amsterdam, University of Amsterdam, Amsterdam, The Netherlands
fDepartment of Bacteria, Parasites and Fungi, Statens Serum Institut, Copenhagen, Denmark
gNational Reference Centre for (invasive) S. pneumoniae, UZ Leuven, Leuven, Belgium
hDepartment of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium

ABSTRACT Fourier transform infrared (FT-IR) spectroscopy (IR Biotyper; Bruker) allows highly discriminatory fingerprinting of closely related bacterial strains. In this study, FT-IR spectroscopy-based capsular typing of *Streptococcus pneumoniae* was validated as a rapid, cost-effective, and medium-throughput alternative to the classical phenotypic techniques. A training set of 233 strains was defined, comprising 34 different serotypes and including all 24 vaccine types (VTs) and 10 non-vaccine types (NVTs). The acquired spectra were used to (i) create a dendrogram where strains clustered together according to their serotypes and (ii) train an artificial neural network (ANN) model to predict unknown pneumococcal serotypes. During validation using 153 additional strains, we reached 98.0% accuracy for determining serotypes represented in the training set. Next, the performance of the IR Biotyper was assessed using 124 strains representing 59 non-training set serotypes. In this setting, 42 of 59 serotypes (71.1%) could be accurately categorized as being non-training set serotypes. Furthermore, it was observed that comparability of spectra was affected by the source of the Columbia medium used to grow the pneumococci and that this complicated the robustness and standardization potential of FT-IR spectroscopy. A rigorous laboratory workflow in combination with specific ANN models that account for environmental noise parameters can be applied to overcome this issue in the near future. The IR Biotyper has the potential to be used as a fast, cost-effective, and accurate phenotypic serotyping tool for *S. pneumoniae*.

KEYWORDS FT-IR spectroscopy, *Streptococcus pneumoniae*, machine learning, pneumococcus, serotyping

Streptococcus pneumoniae is a Gram-positive opportunistic pathogen that can cause a wide variety of diseases, ranging from more mild infections, such as sinusitis, otitis media, and conjunctivitis, to more severe diseases, such as community-acquired pneumonia (CAP), bacteremia, and meningitis (1–5). The World Health Organization (WHO) estimated that *S. pneumoniae* infections still killed close to 300,000 children under 5 years old worldwide in 2015 (6). Moreover, in 2016, and thus before the COVID-19 pandemic, *S. pneumoniae* was the leading cause of pneumonia-induced mortality worldwide, accounting for more deaths than all other etiologies causing lower respiratory tract infections combined (7).

The cells of most *S. pneumoniae* strains are surrounded by a polysaccharide capsule that protrudes into the extracellular space (8, 9). The capsule has been shown to be an important virulence determinant and has been linked to different roles in the infection process

(10), including (i) inhibition of both the classic and alternative complement pathways, which reduces opsonophagocytic killing (11, 12), and (ii) promotion of nasopharyngeal colonization (13). Currently, around 100 different chemical capsular structures have been described and categorized into serotypes (8, 14). However, only a limited number of pneumococcal serotypes cause the large majority of disease around the world, and different vaccines have been developed to target the ones dominant in invasive diseases (15, 16). The pneumococcal polysaccharide vaccine (PPSV23) has been on the market since the 1980s and protects against 23 serotypes but with a moderate long-term immunity (16–18). More importantly, this vaccine does not protect infants below the age of two, which is the age group with the highest incidence of severe disease and the highest level of carriage. The protein conjugate vaccines (PCVs) were first implemented in the U.S. childhood vaccination programs in the early 2000s because of their projected efficacy in infants, as well as their promising potential in providing long-lasting immunity (19–22). The current PCVs protect against 10 (PCV10) or 13 (PCV13) different serotypes, but in the meantime, higher valency PCVs (PCV15 and PCV20) have been licensed for adults and are (or soon will be) available (23–27). The surveillance of circulating serotypes in the population is therefore a critical epidemiological tool for assessing the impact of vaccination on the distribution of serotypes (28, 29).

Historically, the identification and classification of pneumococcal serotypes is based on reaction with specific rabbit polyclonal antisera (8). Visual inspection using light microscopy and the observation of "swollen" capsules (when antibodies have reacted with and bound a specific capsule) defines a positive reaction. This particular serotyping method is known as the quellung reaction and has been used for over 100 years to provide information on the chemical structure of the capsule (30, 31). More recently, the quellung reaction has been complemented by other phenotypic techniques, of which latex agglutination has been the more widely adopted because of its speed of use and simplicity compared to the quellung reaction (32). Although these phenotypic methods are still used routinely in national reference centers as the gold standard for serotyping *S. pneumoniae* isolates, there are disadvantages linked to them (33). First, the interpretation can be subjective and requires well-trained and experienced technicians (34–36). Second, performing these techniques is time consuming and labor intensive (typically 5 to 30 min per strain is required, depending on the specific serotype) and they are not well suited for high-throughput serotyping. Finally, the required typing antisera are expensive. Overall, this has limited the use of these phenotypic methods as a routine method to a limited number of clinical laboratories (mostly national reference centers) (37).

Over the years, various molecular alternatives have been developed, mostly using the genetic information of a particular pneumococcal strain and translating this into a serotype (33). First of all, *cpsB* "sequetyping" has been proposed as a simple genetic approach (only one standard PCR is required) based on the genetic diversity between *cpsB* alleles of different serotypes (38). Albeit easy to use, the technique suffers from relatively low sensitivity and specificity (around 85%), is often limited to identification to the serogroup (SG) level, and only works for a subset of serotypes/groups (33, 38). Multiplex real-time (RT)-PCR overcomes some of these limitations, targeting up to 34 serotypes and 13 small serogroups in 12 reactions and achieving high sensitivity (39). The high-throughput nature of multiplex RT-PCR is an advantage compared to whole-genome sequencing (WGS), the most recent genetic approach that has gained momentum in the last decade (34, 35). Different pneumococcal serotype prediction algorithms have been developed that perform with high accuracies, such as PneumoCat (35) (>99% sensitivity), SeroBA (34) (98% sensitivity), and SeroCall (40) (100% sensitivity for major serotypes and up to 86% for minor serotypes). However, to date, WGS remains a rather slow procedure, and this approach, similar to *cpsB* sequetyping and RT-PCR, does not measure the chemical composition of the capsule as such. Overall, capsule gene sequencing is able to accurately detect genetic changes like single-nucleotide polymorphisms (SNPs) or recombination events, but the impact of these changes on the capsule structure or expression cannot be deduced.

A promising technique that does provide information on the capsule structure itself is Fourier transform infrared (FT-IR) spectroscopy (41). With FT-IR spectroscopy, biological material is exposed to infrared light and an absorption spectrum is generated (42). The region

of interest for capsular typing (fingerprinting region) is located between wavenumbers 800 and 1,300 cm^{-1}, where the carbohydrates are predominantly absorbing (C-O stretching and O-H bending). This technique has shown potential in serotyping (or serogrouping) different bacteria, such as *Salmonella* (43–45), *Klebsiella* (46, 47), *Enterobacter* (47, 48), *Acinetobacter* (47), and *Pseudomonas* (47), and has the advantage of being easy to perform, cost effective and medium throughput. Furthermore, FT-IR spectroscopy has already been described in the literature as a promising serotyping technique for *S. pneumoniae*, but it is only quite recently that Burckhardt et al. (37) have provided a first comprehensive validation of the technique.

In this study, FT-IR spectroscopy was further evaluated as a serotyping technique for *S. pneumoniae*. An accurate serotyping workflow was developed through analysis of an in-house-constructed training set of 233 international reference strains (representing 34 serotypes) in combination with a validation set of 277 *S. pneumoniae* strains (representing 93 different serotypes), using dimensionality reduction analysis, hierarchical cluster analysis (HCA), and a machine-learning algorithm. The robustness of the method was evaluated, and possible bottlenecks for the standardization and wide implementation of FT-IR spectroscopy for pneumococcal serotyping were investigated.

MATERIALS AND METHODS

Strains and growth conditions. *S. pneumoniae* reference and validation strains were either part of our in-house collection (49, 50) or obtained from the reference centers for pneumococcal diseases of Belgium, The Netherlands, and Denmark (Table S1 in the supplemental material). Isolates from Belgium and The Netherlands were serotyped using the quellung reaction, while isolates from Denmark were serotyped either by the quellung reaction alone or by the pneumotest latex kit (SSI Diagnostica, Denmark) combined with the quellung reaction. For the quellung reaction, type-specific rabbit pneumococcal antisera (SSI Diagnostica, Denmark) were used, and the method was performed as described previously (36, 51, 52). Isolates from Belgium and Denmark were stored in brain heart infusion (BHI) broth with 10% glycerol at $-80°C$, while strains from The Netherlands were received and stored in peptone broth with 8% glycerol at $-80°C$. Strains at Sciensano were incubated directly from a $-80°C$ stock cryotube on Columbia agar base medium plus 5% sheep blood (CB + 5% SB) (Oxoid) and grown for 24 h at 36°C in 5% CO_2 (Binder C170 incubator; VWR) before being measured.

There were some small differences in the incubation protocols of the other sites (Bruker and University Hospital of Heidelberg). At Bruker, isolates were incubated from a $-80°C$ Microbank (Thermo Fisher Scientific) on CB + 5% SB (Becton Dickinson [BD]) at 35°C to 37°C with a 2.5-L CO_2 gen pack (Thermo Fisher Scientific). The next day, isolates were subcultured on fresh CB medium and incubated for 24 h under the same conditions, after which FT-IR spectroscopy measurements were performed. At the University Hospital of Heidelberg (UHH), strains were incubated on CB + 5% SB (BD) for 24 h at 36°C in 5% CO_2 (53) and subsequently subcultured for another 24 h on fresh CB, after which FT-IR spectroscopy measurements were performed.

FT-IR spectroscopy. Pneumococcal strains were prepared for FT-IR spectroscopy using the direct smear method. Briefly, a 1-μL plastic loop was used to carefully remove the biomass from the plate and apply it evenly to a silicon plate containing 96 spots (Bruker, Germany). Apart from the pneumococcal strains, two *E. coli* reference strains (IRTS 1 and IRTS2; Bruker) were applied to the silicon plate to validate the run. After the samples were dried, the silicon plate was placed under a UV lamp (25 W; Philips) at a distance of 5 cm for 10 min to inactivate the bacterial cells. Finally, the plate was inserted into the IR Biotyper (Bruker, Germany) and the measurements were started.

Analysis of FT-IR spectroscopy spectra. (i) HCA: the dendrogram. Analysis of the spectra was partly performed with the IR Biotyper software (version 3.0) and partly using third party software (Excel, RapidMiner, and PAST). First, average spectra were created from all spectra of each individual strain using the IR Biotyper software and, in the case of the training set strains, further labeled with their respective serotypes. Next, all individual and average spectra were exported as a spreadsheet in XLS format, creating a large data set containing 521 data points (between wavenumbers 800 and 1,300 cm^{-1}) for each spectrum. All individual spectra were subsequently used to build a principal-component analysis (PCA) model with the RapidMiner Studio 9.8 software using a variance threshold of 0.999, after which the freshly built model was applied to all average spectra. The resulting simplified data set now contained all average spectra of the 233 strains with 86 principal components (PCs) and was subjected to linear discriminant analysis (LDA) using the open-source PAST4.03 software (54). As a final analysis step, hierarchical cluster analysis (HCA) was applied to all 233 strains and their respective 34 linear discriminants using Ward's algorithm as the linkage type, generating a large dendrogram containing all 233 training set strains. For the analysis of the validation set strains, the analysis procedure was the same, except that these strains, labeled "SAMPLE" of unknown serotype ("?"), did not contribute to building the PCA model and did not influence the supervised LDA model, as they were unlabeled. The validation set strains appear after HCA in the dendrogram in the cluster of strains most closely resembling them.

(ii) Artificial neural network model: the PneumoClassifier. The IR Biotyper software offers the possibility to use machine-learning algorithms on the measured spectra. Two PneumoClassifiers were developed (version 1.0 and version 2.1) and varied in the numbers of spectra used for training and/or the numbers of training cycles. The accuracy of the classifiers was determined in a stratified 4-fold cross-validation test with a confusion matrix as the output.
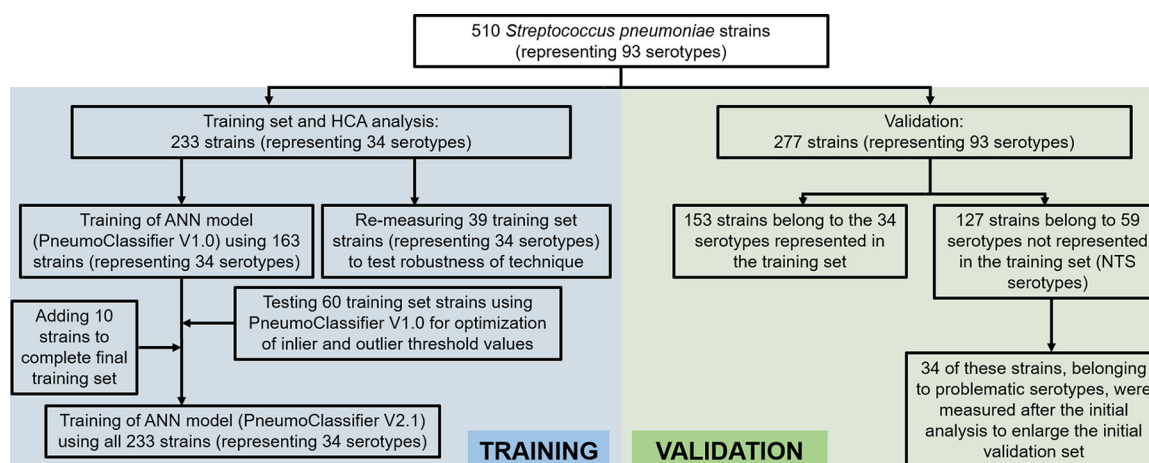
FIG 1 Schematic overview of the *S. pneumoniae* strains measured with FT-IR spectroscopy and their purpose throughout the study.

Finally, validation strains were analyzed in technical pentaplicates and the resulting average spectrum was used by PneumoClassifier version 2.1 to predict the serotype. A green-colored serotyping output (outlier value of <1.4) signified that the result was within the boundaries of the training set and could be trusted, while an orange-colored serotyping output (outlier value between 1.4 and 2.0) signified that the result was at the boundaries of the training set and should be interpreted with care. When the result was colored red (outlier value of >2), the result was outside the boundaries of the training set and thus could not be trusted.

Whole-genome sequencing (WGS). Overnight *S. pneumoniae* cultures of all serogroup 6 (SG6) training set strains (Table S2) were scraped from the plates and dissolved in BHI broth (BD), and DNA was extracted using an MgC Bacterial DNA Kit (Atrida, Amersfoort, The Netherlands), following the manufacturer's instructions. Sequencing libraries were prepared using the Illumina Nextera XT DNA sample preparation kit according to the manufacturer's instructions (Illumina, San Diego, CA, USA), and DNA sequencing was performed on an Illumina MiSeq instrument (MiSeq version 3 chemistry). Trimming, clipping, filtering out non-confident bases, correcting sequencing errors, reducing redundant reads, and merging overlapping paired reads were performed using fqCleanER version 3.01 (55–57). All FASTQ files were uploaded to Pathogenwatch (https://pathogen.watch/), which uses the SeroBA (34) algorithm to predict the pneumococcal serotype. In addition to this, a more detailed analysis was performed by looking into the *wciP* and *wciN* genes, which are known to be the genetic determinants of the different SG6 capsule structures (58–61). In short, FASTQ reads were imported into CLC Genomics Workbench 21.0.4 (Qiagen) and mapped in parallel to the *cps* clusters of both an SG6 subtype B (6B) (NCBI accession number CR931639) and a 6C (NCBI accession number EF538714) reference strain. Strains with the *wciP* (584A) single-nucleotide polymorphism (SNP) were either 6B or 6D, while strains with the *wciP* (584G) SNP were either 6A or 6C. Information on which *wciN* variant is present (*wciNα* in 6A and 6B and *wciNβ* in 6C and 6D) in these strains completes the subtyping process.

Data availability. All sequences have been made publicly available in the NCBI Sequence Read Archive and can be retrieved using the accession number PRJNA836581.

RESULTS

Construction of FT-IR spectroscopy training set for *S. pneumoniae* serotypes. A collection of *S. pneumoniae* isolates from Belgium (173 strains) and The Netherlands (60 strains) from patients with invasive pneumococcal disease (IPD) (178 strains) or noninvasive pneumococcal disease (NIPD) (55 strains) was analyzed to create a robust database (Fig. 1 and Table S1). These 233 strains represented 34 different serotypes (24 VTs and 10 NVTs), with each serotype being represented by at least 4 different strains (Table 1). All strains were serotyped using the quellung reaction (and/or the latex agglutination test) and were representative of the circulating serotypes causing IPD in Belgium/Europe (51, 62, 63).

All 233 pneumococcal strains were measured using FT-IR spectroscopy on three independent occasions, yielding a minimum of 12 spectra per strain. This large and complex data set was analyzed using a combination of dimensionality reduction techniques (PCA, grouped by isolate, and LDA, grouped by serotype) and hierarchical cluster analysis (Euclidean distance with Ward's algorithm linkage type). The exact workflow is described in detail in Materials and Methods. The chosen output for this analysis procedure was a dendrogram (Fig. 2A), a tree-shaped figure with one stem splitting into two branches, eventually ending in 233 branch tips, with all strains represented and forming clusters. The vertical scale represents the distance between strains, so that the more similar they are, the later they split off, eventually

TABLE 1 Serotypes represented in the training set and the respective numbers of strains used

Serotype (*n* = 34)	No. of strains in training set (*n* = 233)
1	4
2	5
3	9
4	5
5	4
6A	8
6B	6
6C	11
7F	7
8	5
9V	7
9N	6
10A	5
11A	9
12F	5
14	10
15A	9
15B	10
16F	6
17F	6
18C	5
19A	9
19F	11
20	5
22F	6
23A	7
23B	10
23F	5
24A	4
24B	5
24F	8
33F	6
35B	10
38	5

ending up in clusters. The cutoff value (COV) for the dendrogram was chosen manually, with the goals being (i) to separate as many of the different serotypes as possible into different clusters and (ii) to group as many individual strains of the same serotype as possible within one cluster.

When interpreting the dendrogram for the strains in our training set strains, it is clear that almost all isolates of the same serotype cluster together and different serotypes form distinct clusters (a detailed view is presented in Fig. 2B). The exceptions to the rule are the strains from SG24 (subtypes 24A, 24B, and 24F) that form one distinct serogroup cluster but cannot be accurately separated into their respective subtype clusters (Fig. 2B). Together, these results are highly encouraging and show that FT-IR spectroscopy can accurately separate 31 different pneumococcal serotypes (encompassing all VTs) and group 24A, 24B, and 24F strains at the serogroup level.

Training an ANN model to classify pneumococcal serotypes. Using the IR Biotyper (version 3.0) software, an artificial neural network (ANN) model was trained using all individual training set spectra. First, an ANN model was built to use the first 20 principal components (PCs) on a subset of the data set (2,246 spectra; PneumoClassifier version 1.0) and trained for 200 cycles. Using this classifier, 60 training set strains that were not yet incorporated in the model (Fig. 1 and Table S1) were utilized as mock serotyping strains to tweak the inlier and outlier thresholds of the model. It was found that increasing the inlier threshold from 1.0 (default setting) to 1.4 provided a more realistic result in confidently typing training set strains. Next, the final ANN model (PneumoClassifier version 2.1) was created with the full training set (3,276 spectra) and trained for 500 cycles. The PneumoClassifier had a total accuracy of 92% in correctly predicting the serotypes of strains that were in the training set. A
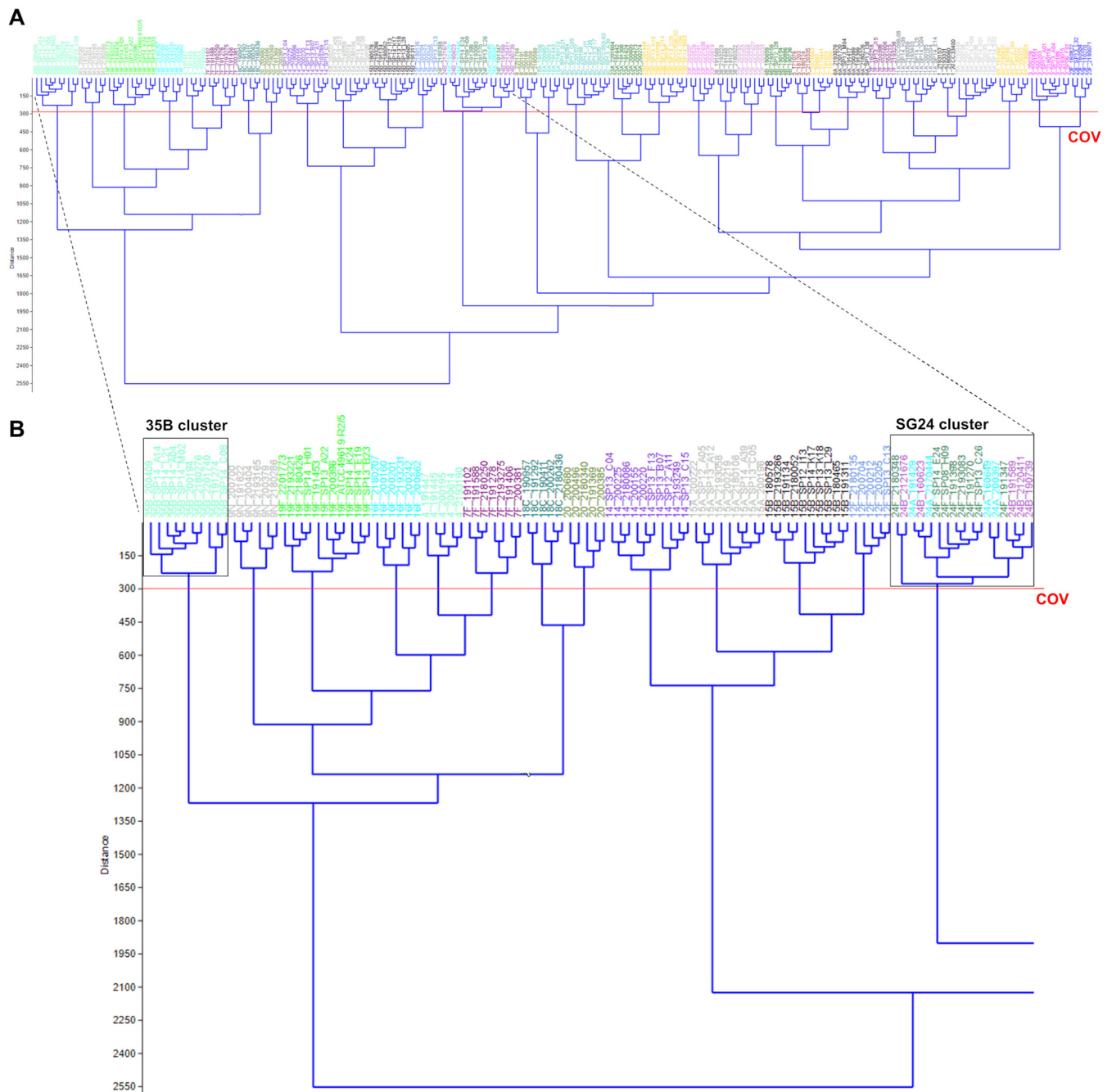
FIG 2 (A) Dendrogram visualizing the 233 pneumococcal training set strains representing 34 serotypes. The red line depicts the cutoff value (COV), separating strains from different serotypes into distinct clusters but at the same time grouping strains from the same serotype in one cluster. The linkage type used for the dendrogram was Euclidean distance with Ward's algorithm. (B) Zoomed-in image of the outer left part of the dendrogram shown in panel A (dashed lines). Different clusters grouping strains from the same serotypes can be observed (e.g., 35B cluster). The SG24 cluster is shown as a clearly distinct cluster, but within this cluster, the different subtypes (24A, 24B, and 24F) are not accurately separated.

confusion matrix was obtained as visual output to get a detailed view of the classifier's performance for every serotype (Fig. 3). The large majority of errors was found within SG24 (and to a lesser extent within SG6), thus encountering a similar problem as with the dendrogram output. Because SG6 is an important serogroup consisting of both vaccine (6A and 6B) and nonvaccine (6C and 6D) types, all SG6 training set strains were subjected to WGS and the SeroBA algorithm was used for serotype prediction. This resulted in three strains for which a different serotype was predicted by WGS, and after repeating the quellung reaction on these strains, one strain was found to be incorrectly categorized as 6A instead of 6C (Table S2).

Predicted class

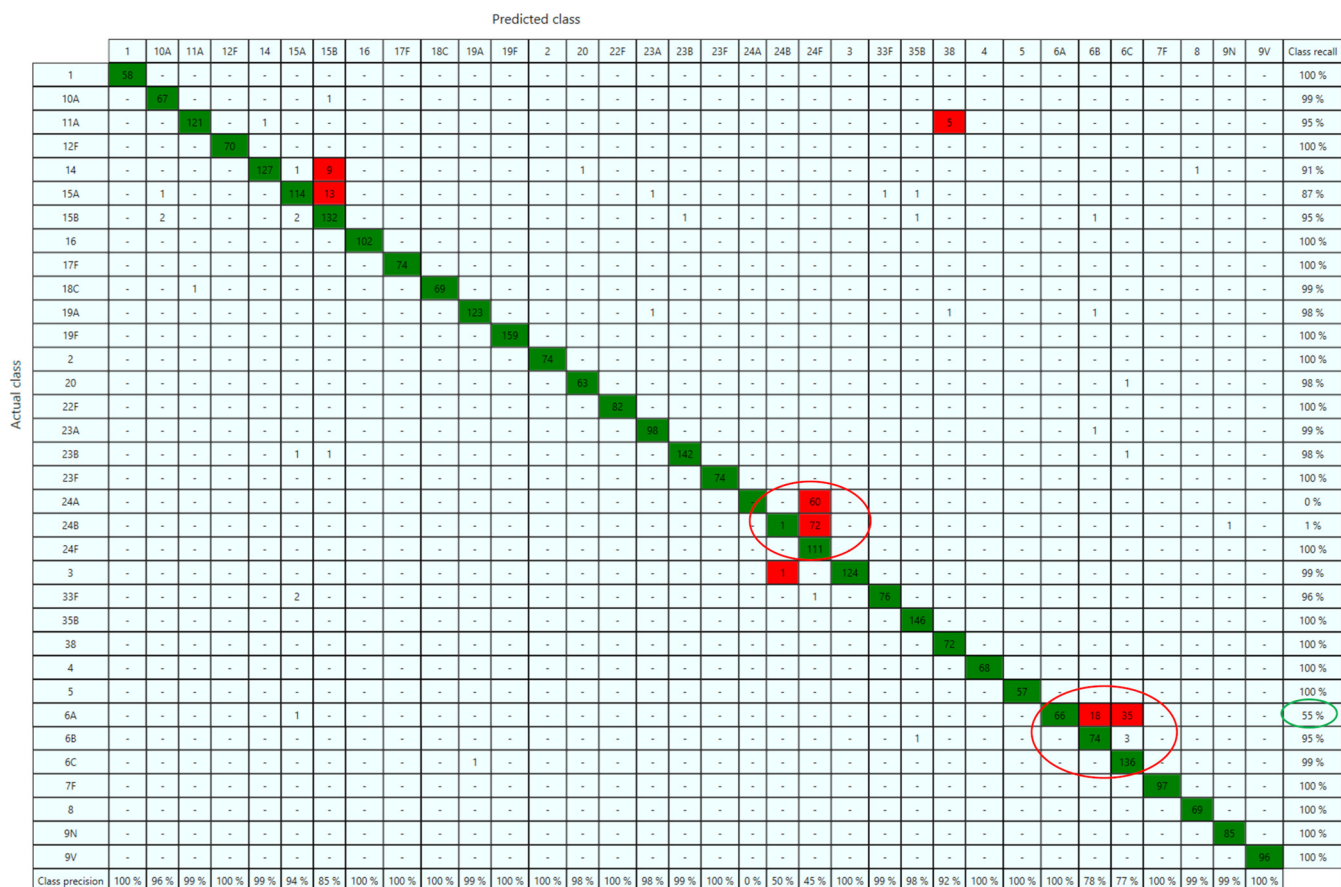| Actual class	1	10A	11A	12F	14	15A	15B	16	17F	18C	19A	19F	2	20	22F	23A	23B	23F	24A	24B	24F	3	33F	35B	38	4	5	6A	6B	6C	7F	8	9N	9V	Class recall
1	58	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	100 %
10A	·	67	·	·	·	1	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	99 %
11A	·	·	121	·	1	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	5	·	·	·	·	·	·	·	·	·	95 %
12F	·	·	·	70	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	100 %
14	·	·	·	·	127	1	9	·	·	·	·	·	1	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1	·	91 %
15A	·	1	·	·	·	114	13	·	·	·	·	·	·	·	·	1	·	·	·	1	1	·	·	·	·	·	·	·	·	·	·	·	·	·	87 %
15B	·	2	·	·	·	2	132	·	·	·	·	·	·	·	·	·	·	1	·	·	·	·	·	1	·	·	·	·	1	·	·	·	·	·	95 %
16	·	·	·	·	·	·	·	102	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	100 %
17F	·	·	·	·	·	·	·	·	74	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	100 %
18C	·	·	1	·	·	·	·	·	·	69	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	99 %
19A	·	·	·	·	·	·	·	·	·	·	123	·	·	·	·	1	·	·	·	·	·	·	·	1	·	·	·	1	·	·	·	·	·	·	98 %
19F	·	·	·	·	·	·	·	·	·	·	·	159	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	100 %
2	·	·	·	·	·	·	·	·	·	·	·	·	74	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	100 %
20	·	·	·	·	·	·	·	·	·	·	·	·	·	63	·	·	·	·	·	·	·	·	·	·	·	·	·	1	·	·	·	·	·	·	98 %
22F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	82	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	100 %
23A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	98	·	·	·	·	·	·	·	·	·	·	·	1	·	·	·	·	·	·	99 %
23B	·	·	·	·	·	1	1	·	·	·	·	·	·	·	·	·	142	·	·	·	·	·	·	·	·	·	·	·	1	·	·	·	·	·	98 %
23F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	74	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	100 %
24A	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	60	·	·	·	·	·	·	·	·	·	·	·	·	·	·	0 %
24B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1	72	·	·	·	·	·	·	·	·	·	·	·	·	1	·	1 %
24F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	111	·	·	·	·	·	·	·	·	·	·	·	·	·	100 %
3	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1	·	124	·	·	·	·	·	·	·	·	·	·	·	·	99 %
33F	·	·	·	·	·	2	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1	·	76	·	·	·	·	·	·	·	·	·	·	·	96 %
35B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	146	·	·	·	·	·	·	·	·	·	·	100 %
38	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	72	·	·	·	·	·	·	·	·	·	100 %
4	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	68	·	·	·	·	·	·	·	·	100 %
5	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	57	·	·	·	·	·	·	·	100 %
6A	·	·	·	·	·	1	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	66	18	35	·	·	·	·	55 %
6B	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1	·	·	·	·	·	·	74	3	·	·	·	·	95 %
6C	·	·	·	·	·	·	·	·	1	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	136	·	·	·	·	99 %
7F	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	97	·	·	·	100 %
8	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	69	·	·	100 %
9N	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	85	·	100 %
9V	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	96	100 %
Class precision	100 %	96 %	99 %	100 %	99 %	94 %	85 %	100 %	100 %	100 %	99 %	100 %	100 %	98 %	100 %	98 %	99 %	100 %	0 %	50 %	45 %	100 %	99 %	98 %	92 %	100 %	100 %	100 %	78 %	77 %	100 %	99 %	99 %	100 %	

FIG 3 Confusion matrix showing the output of the trained PneumoClassifier version 2.1 (500 training cycles). The highlighted SG24 and SG6 subtypes in the red-encircled areas are spectra that were often misclassified within the serogroup and led to low sensitivities for these serogroups. The green-encircled value is the sensitivity for predicting 6A spectra before the classifier was retrained with the corrected SG6 serotypes.

When this error was corrected and the classifier retrained, the sensitivity for predicting 6A spectra increased from 55% to 85%, thus yielding satisfying accuracies for all SG6 subtypes.

Validation of the serotyping workflow. To validate the FT-IR spectroscopy approach as an *S. pneumoniae* serotyping method, a serotyping workflow was established to analyze pneumococcal strains that were not part of the training set (Fig. 4A), as follows. (i) The serotype of an analyzed strain was called when the serotypes obtained via the dendrogram and the PneumoClassifier were concordant and the result was within both the dendrogram cluster demarcated by the COV and the training set of the PneumoClassifier (green or orange code). (ii) A strain was classified as a non-training set serotype (NTS) when the serotypes obtained via the dendrogram and the PneumoClassifier were discordant. (iii) The measurement was repeated with more technical replicates when the serotypes obtained via the dendrogram and the PneumoClassifier were concordant but the result was outside the dendrogram cluster demarcated by the COV or outside the training set of the PneumoClassifier. If there was still any doubt after this, the strain was classified as an NTS serotype.

Using this workflow, a validation set of 277 strains belonging to 93 different serotypes and different from the strains used to train the model were measured with FT-IR spectroscopy (Fig. 1 and Table S1). A total of 153 strains belonged to the 34 serotypes included in the training set, and all of these serotypes were challenged at least once (Table S3). In total, 150 of 153 strains were found to be correctly serotyped (or grouped, in the case of SG24), yielding an accuracy of 98.0% in predicting serotypes included in the training set. Only three strains, belonging to three different serotypes (17F, 9V, and 2), were mistyped as NTS serotypes (Fig. 4B).

Furthermore, a set of 90 validation strains belonging to 59 NTS serotypes (Table S4) were measured initially and analyzed using the serotyping workflow. For the majority of NTS serotypes, one strain was tested. Only serotypes 7D, 10D, 10E, 11E, and 33E could not be
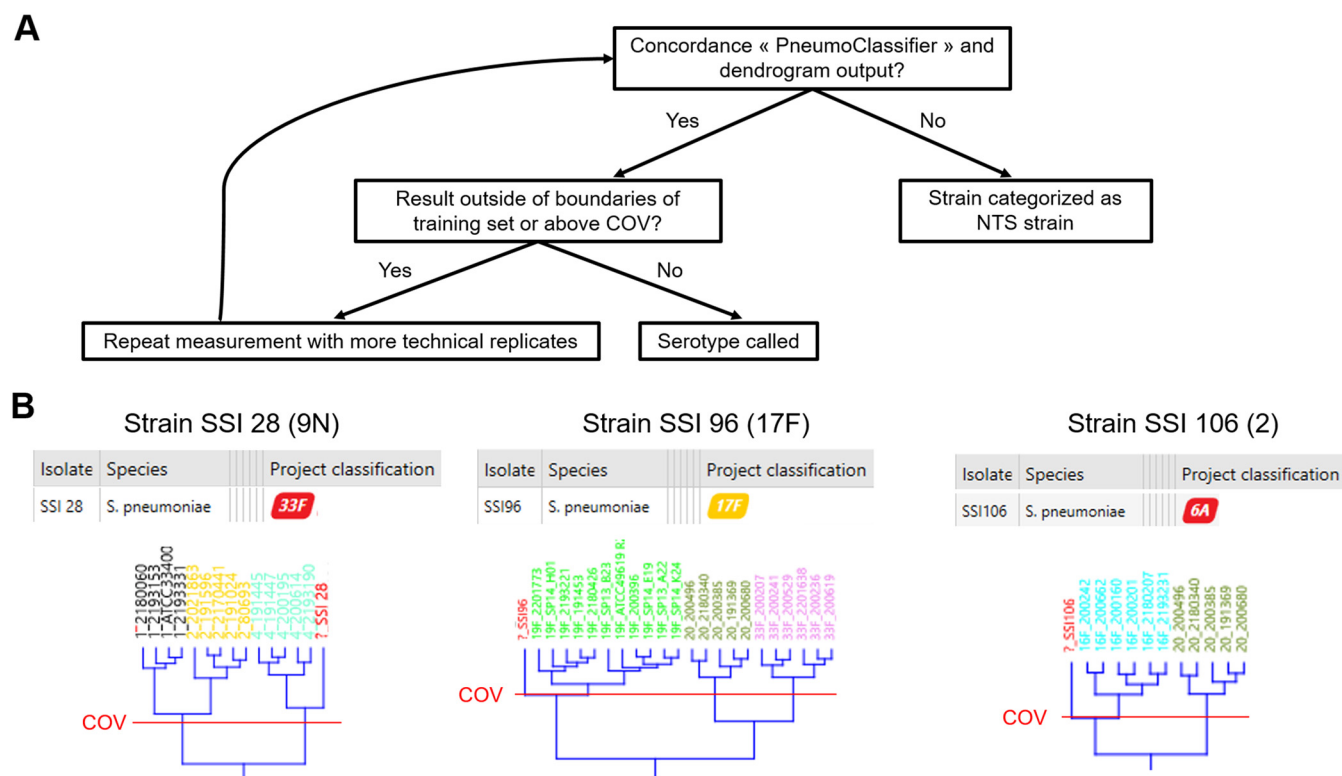
FIG 4 (A) Schematic overview of the serotyping workflow adopted throughout the study. (B) Serotyping output for three validation set strains that were mistyped as NTS strains using the serotyping workflow shown in panel A. Top, the tables show the output of the PneumoClassifier version 2.1; bottom, the relevant part of the dendrogram for each validation set strain is shown. The red lines depict the cutoff value (COV).

tested, as they were lacking in the collections of the reference centers of Belgium, The Netherlands, Denmark, and Germany. Applying the serotyping workflow to the NTS validation strains resulted in 42 of 59 serotypes being correctly classified as NTS serotypes, thus attaining an accuracy of 71.2% in correctly predicting NTS serotypes. The NTS serotypes that were confounded for training set serotypes are listed in Table 2. Of the 17 problematic serotypes, 14 serotypes were only represented by one validation strain, and an extra set of validation strains ($n = 34$, making a total of 124 NTS validation strains) was acquired for 12 of these problematic serotypes (7A, 9L, 9A, 11D, 12B, 15F, 18A, 19B, 29, 33A, 39, and 40). After measuring and analyzing this extra set of strains, all 12 serotypes were confirmed to be confounded with training set serotypes.

Robustness of the FT-IR spectroscopy technique for pneumococcal serotyping. FT-IR spectroscopy is a highly sensitive technique, picking up very small chemical differences. This is a big advantage when serotyping strains with very similar capsule structures but can also be potentially troublesome when environmental conditions change (e.g., type of medium and composition, temperature, humidity, and incubation time), sometimes in ways that are beyond the control of the operator, and cause alterations in biomass composition.

To test the robustness of the technique, two approaches were adopted. First, 39 training set strains (with every training set serotype represented at least once) were remeasured 1 year after they had been initially measured to construct the training set (Table S1). Of all the retested strains, 22 were measured in technical pentaplicates and 17 in technical triplicates. Of these 39 strains, 36 strains were serotyped correctly, while 3 strains were incorrectly categorized as NTS serotypes, indicating that a degree of variation (of thus-far-unknown origin) had emerged that could lead to mistyping. The three strains belonged to three different serotypes (23A, 23B, and 6A) and were measured in technical triplicates, suggesting that measuring fewer technical replicates could lead to more mistyping. To evaluate this, the three problematic strains were measured once more in technical pentaplicates, and all were then found to be serotyped correctly.

TABLE 2 Summary of the 17 NTS serotypes that were mistyped as training set serotypes using FT-IR spectroscopy

Serotype from FT-IR spectroscopy[a]	NTS serotype from quellung reaction[b]	Differences in chemical structures of capsule[c]
6C	6D	Difference in positions of glycosidic linkages between adjacent sugar moieties
7F	7A	7F harbors an extra branching sugar
9N	9L	One sugar difference in sugar backbone (glucose vs galactose)
9V	9A	Different amounts of acetylation
10A	39	Difference in acetylation and positions of glycosidic linkages
11A	11D	11D is part 11A and part different sugar backbone (glucose vs *N*-acetylglucosamine) and different acetylation amt
12F	12B	No structural information available on 12B
15B	15C	15B contains acetylated sugar moiety
18C	18A	Difference in acetylation and one sugar moiety (glucose vs *N*-acetylglucosamine)
19A	19B	19B contains an extra sugar in its backbone and additional branching sugars
33F	33A and 33B	33A and 33B differ in amounts of acetylation and 33B contains different sugars in its backbone
15A	15F	Difference in side chains (glycerol vs choline) and acetylation
24	40	No structural information available on 40
35B	29 and 35D	Different sugars in backbones of 35B and 29 and difference in acetylation between 35B and 35D
38	25A	No structural information available on 25A

[a]Vaccine types (VTs) are shaded in gray.
[b]NTS, non-training set.
[c]The main chemical structural differences of the capsules of these serotypes are listed.

A second approach to testing the robustness of the technique was to compare spectra acquired at Sciensano (Brussels, Belgium), at Bruker (Bremen, Germany), and the UHH (Heidelberg, Germany). The conditions under which the strains were grown at the three sites were very similar in incubation parameters (36°C and 5% CO_2), incubation time (24 h), and growth medium (Columbia agar base medium plus 5% sheep blood [CB + 5% SB]). Remarkably, it was observed that when comparing all the different spectra, they were not clustering per serotype as expected but, rather, formed two distinct clouds on a three-dimensional (3-D) PCA plot, clearly separating Sciensano spectra from UHH spectra (Fig. 5A). This pointed to the existence of one or more variables that had a larger contribution to the total variance of the data set than the individual chemical differences of the different serotypes. After close examination of the exact incubation conditions, it was noticed that the incubation medium originated from two different suppliers (Oxoid versus BD). An experiment was designed to test the alleged influence of the medium supplier. Strains from 8 different serotypes were grown on CB + 5% SB from BD, Oxoid, or bioMérieux and analyzed using FT-IR spectroscopy. When interpreting the 2-D PCA plot, it was observed that the spectra of the strains measured on the Oxoid medium clustered far away from the spectra of the same strains measured on BD and bioMérieux media, regardless of the site at which the spectra were acquired (Fig. 5B). Moreover, measurements of the same reference strain grown at different locations (UHH and Sciensano) and on CB from different suppliers (BD and Oxoid) showed that the influence of the medium supplier is far greater than other interlaboratory differences that might exist (Fig. S1). This indicates that the Oxoid medium contains a component(s) that, presumably indirectly, changes the chemical composition of the sugar-rich outer surface of the pneumococci and that these changes obfuscate the serotype-dependent capsule differences, ultimately complicating the serotyping process when comparing spectra from strains grown on CB + 5% SB plates from different suppliers.

DISCUSSION

S. pneumoniae is still among the leading causes of infection-induced mortality, especially in developing countries (7). The introduction of PCVs in the early 2000s and their implementation in childhood vaccination programs has greatly decreased the incidence of IPD caused by VT serotypes (19–22). This effect was also initially observed in the unvaccinated adult population in the United States and was attributed to herd immunity (64, 65). However, in recent years, the incidence of IPD worldwide has plateaued or even slightly increased as the result of (i) serotype replacement, since NVT serotypes fill the ecological niche that is left by VT serotypes, and (ii) limitations to herd protection (29, 66–69). As long as there
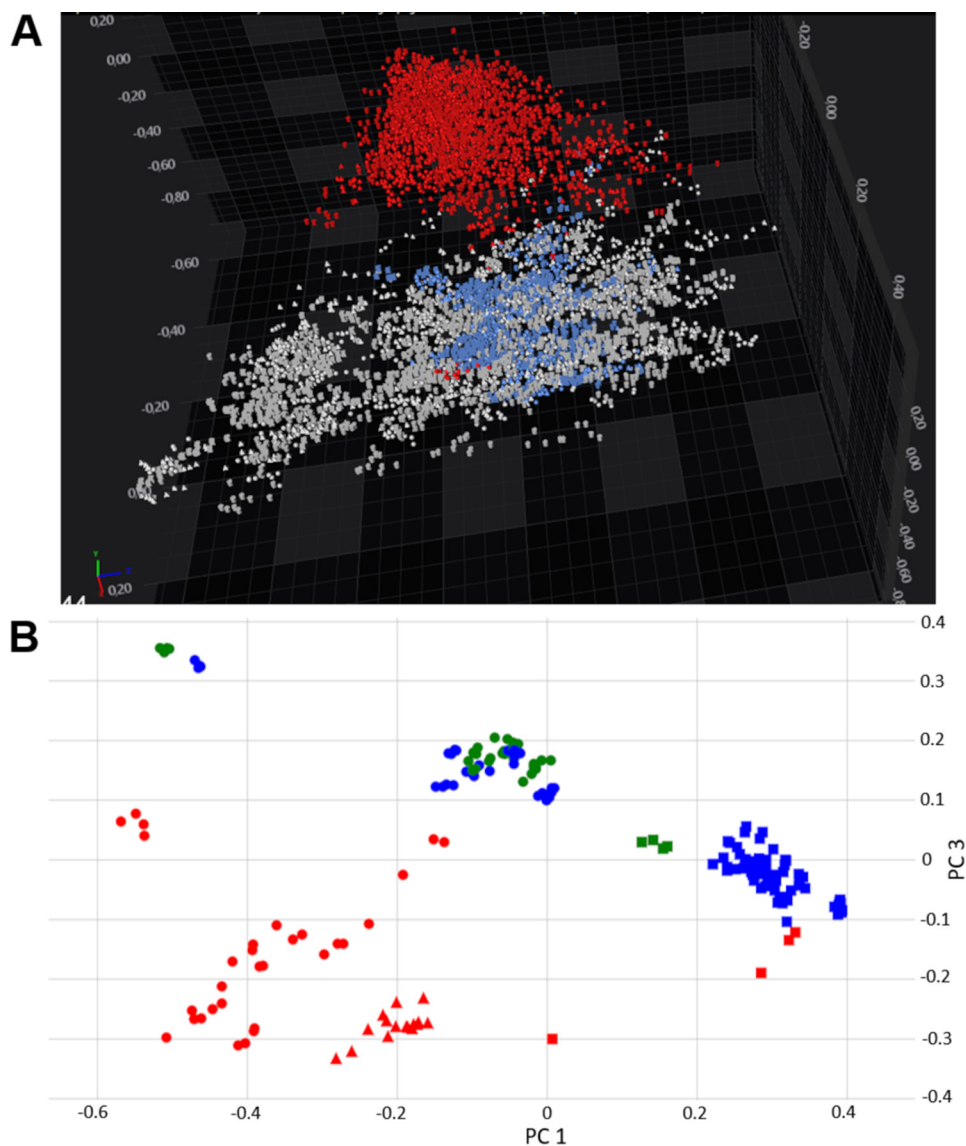
FIG 5 (A) Three-dimensional PCA plot constructed using around 6,000 spectra acquired from three different locations (Sciensano, Brussels; Bruker, Bremen; and UHH, Heidelberg). Data points are colored according to the location where they were measured: red data points at Sciensano, blue data points at UHH, and gray data points at Bruker. The spectra acquired at Sciensano cluster together strongly but far away from the spectra acquired at Bremen and Heidelberg. Together, PC1, PC2, and PC3 account for 60% of the total variation. (B) Two-dimensional PCA plot (PC1, 46.2% of the total variance explained, and PC3, 15.9% of the total variance explained) depicting 184 spectra from strains (representing 9 different serotypes) that were measured on CB + 5% SB from three different suppliers (Oxoid, BD, and bioMérieux). The spectra acquired on Oxoid medium (red data points) cluster away from the spectra acquired on BD (blue data points) and bioMérieux (green data points) media, regardless of the location at which the spectra were measured. Data points with different shapes depict different locations: triangles correspond to Sciensano spectra, squares to UHH spectra, and dots to Bruker spectra.

is no vaccine protecting against every pneumococcal strain, independent of serotype, surveillance of circulating serotypes will remain essential in the future and serotyping techniques will continue to be an important tool to monitor pneumococcal seroepidemiology (62, 70).

FT-IR spectroscopy uses infrared light to create absorption spectra of the biological material being probed, yielding information on the vibration and rotation of covalent bonds of the biological structure (41, 42). This technique can detect and discriminate very small chemical differences and was recently tested for serotyping *S. pneumoniae* (37). Burckhardt et al. (37) used a training set containing the spectra of 120 strains (representing the 24 VT serotypes) to type a set of 168 challenge strains (representing 48 different serotypes, both VT and NVT) and achieved excellent concordance between the results of FT-IR spectroscopy and the quellung reaction. In this study, extensive serotyping validation results for *S. pneumoniae*

were generated using two different approaches (HCA and ANN), surpassing the initial validation study of Burckhardt et al., and promising accuracies were obtained, 98% for predicting training set serotypes and 71.2% for predicting non-training set serotypes.

Interestingly, FT-IR spectroscopy was capable of accurately categorizing all training set strains at the serotype level except for SG24, which could only be classified on the serogroup level. At the time of strain collection, the existence of the new 24C variant (14), which contains a mixture of 24B and 24F repeating units, was not yet known. This finding is relevant because our analysis technique relies on feeding in correct *a priori* information (supervised learning is applied when building both the LDA, grouped by serotype, or the ANN model), and when this condition is not met, the models become highly unreliable. However, it is highly unlikely that the 24B and 24F strains used in our study were "hidden" 24C strains, as 24C strains already show an atypical serological profile when exposed to the routinely used SG24-typing antisera. In this regard, it should be noted that subtyping of SG24 strains using the reference phenotypic methods, such as the quellung reaction and the latex agglutination assay, proves difficult too because of low reactivity with the antisera. Future work will include the incorporation of more SG24 training set strains to increase the robustness of the training set, followed by an in-depth analysis of the measured spectra.

For the prediction of training set serotypes (24 VT and 10 NVT), FT-IR spectroscopy performed very well, with 98% accuracy, and problematic strains seemed to randomly spread out over different serotypes. There could be various reasons explaining the few discordances observed between the results of FT-IR spectroscopy and the quellung reaction. First, although both techniques measure the capsule structure itself, they do this differently. While the quellung reaction uses the 3-D specificity that polyclonal antibodies possess with the pneumococcal capsule, FT-IR spectroscopy probes the vibration and rotation of covalent bonds (typically C-O stretching and O-H bending) both within the pneumococcal outer surface and inside the cell. This implies that some differences between serotypes might be picked up more easily when using the quellung reaction rather than FT-IR spectroscopy and vice versa. Furthermore, FT-IR spectroscopy is not limited to measuring the chemical bonds of the capsule structure, as all C-O and O-H bonds are detected and it can consequently also pick up differences in glycoproteins and/or glycolipids between related strains.

Another reason for the observed discordances could be the limited number of *S. pneumoniae* strains per serotype ($n \geq 4$) used for training, in combination with large genetic variations for particular serotypes (71–74), possibly translating into small differences between pneumococcal strains belonging to the same serotype. Although a minimum of four strains per serotype was used to build the training set and care was taken to incorporate strains originating from both IPD and NIPD specimens and from two geographical locations (Belgium and The Netherlands), it is possible that the training set was not yet robust enough for certain serotypes. In this regard, it should be noted that there was no WGS information on most of the training set strains and it was thus unknown how much (genetic) variation was represented within the different serotypes in our training set. Currently, work is ongoing to increase the number of strains per serotype in our training set to create a higher degree of robustness.

Finally, we sought to minimize variations in the environmental conditions when growing the pneumococci or measuring them with FT-IR spectroscopy but found them to be unavoidable. It was shown that some degree of variability between measurements of the same strain on different occasions exists and that repeating the measurement with more technical replicates can improve the serotyping accuracy. Furthermore, care was taken to use the growth medium within the mentioned expiration date. However, this still leaves a window of a couple of weeks where the unexpired medium could change slightly and alter the growing conditions, possibly influencing the outcome of the measurement. In this regard, it was found that using CB + 5% SB past the expiration date did alter the serotyping accuracy for certain serotypes (data not shown). Together, the existence of this environmental variation may result in less accurate serotyping and be (partly) responsible for the observed discordances between FT-IR spectroscopy and the quellung reaction.

Given the inherent limitations of the trained models, predicting serotypes not present in the training set (NTS serotypes) and categorizing them as such proved to be challenging, reaching an accuracy of 71.2%. The majority of NTS serotypes that were wrongly typed as training set serotypes had capsule structures that were very similar to some training set serotypes (e.g., 6C versus 6D, 7F versus 7A, 9N versus 9L, and 19A versus 19B), and this appears to be a challenge in our serotyping workflow. It seems unlikely that all these problematic serotypes lack enough discriminatory features in their capsule for the FT-IR spectroscopy to separate them accurately. For example, 19A and 19F are both training set serotypes and can be discriminated accurately using FT-IR spectroscopy, even though their capsule structure is more alike than the 19B capsule structure (8). The nature of the supervised serotyping technique and HCA is to force the categorization of unknown samples within a known data set, and it seems that whenever a serotype is not part of the training set (and our supervised models have not had the opportunity to be trained using that particular serotype), this serotype tends to end up in another serotype cluster based on capsule structure resemblance. This problem can be solved by building extra serogroup-specific ANN models trained exclusively on serotypes with similar capsule structures, thus creating a hierarchical order of ANN models that can be used sequentially to yield the correct serotype. Currently, efforts are being made to incorporate more strains of these problematic serotypes in an extended training set to (i) test whether these serotypes can in fact be discriminated when feeding the supervised models more information and (ii) create more specific ANN models.

Comparison of spectra acquired from strains grown under similar conditions and originating from the same type of IR device (IR Biotyper, Bruker) but measured at different sites was found to be challenging. It was shown that the supplier of CB + 5% SB impacted the comparability of IR spectra between different institutions. More specifically, spectra acquired from strains grown on CB from BD or bioMérieux compared well, as opposed to the spectra measured from strains grown on CB from Oxoid. This observation was independent of the site at which the spectra were measured, suggesting that interlaboratory comparison of spectra was realistic as long as strains were grown on CB from the same supplier. The anticipated variation in composition between the CB media of the different suppliers is currently unknown, and work is ongoing to pinpoint the cause of the observed differences between the different medium suppliers. Future supervised models will be developed to account for this "supplier variation," enabling the straightforward comparison of IR spectra between laboratories. Interestingly, different broths or the same broth from different suppliers also affect the accuracy of the latex agglutination test (75), suggesting that the challenging comparability between spectra of different laboratories is not solely the result of the FT-IR spectroscopy measuring technique *per se*. At present, FT-IR spectroscopy yields excellent results for accuracy but relies on the construction of a robust in-house training set, a labor-intensive endeavor that is not achievable for every clinical laboratory. However, with the knowledge acquired in this study, the application of a strictly standardized workflow, including defined medium suppliers, should be feasible in the near future.

The results presented herein underline the potential of FT-IR spectroscopy as a serotyping technique for *S. pneumoniae*. Concordance between FT-IR spectroscopy and the quellung reaction is very high for serotypes that are part of the training set, and future work aims at an all-encompassing, reliable training set. Moreover, FT-IR spectroscopy is a promising technique because of its ease of use, cost effectiveness, and medium-throughput potential, thus making it the ideal candidate to complement the existing phenotypic methods. The workflow of the technique is straightforward and requires a minimal amount of training: Pneumococcal strains grown on blood agar plates are directly applied to a silicon plate using a sterile loop, and the plate is then inserted into the IR Biotyper. The hands-on time is limited to 1 h to 1 h 30 min, and the FT-IR spectroscopy measurement itself takes around 1 h 30 min for a full plate (96 spots). Up to 30 strains can be measured in one measurement run, and the subsequent serotyping analysis workflow described herein should not take longer than 30 min. Consequently, the labor time for serotyping one strain can be as low as 4 min, which is considerably faster than the average time needed to serotype a pneumococcal strain using the quellung reaction (between 5 and 30 min per strain, depending on

the specific serotype). Performing an accurate cost comparison between the different phenotypic methods is tricky, mainly because (i) the cost of performing the quellung reaction very much depends on the exact serotype that is being probed and whether a latex agglutination assay is performed too and (ii) the cost of the IR Biotyper machine needs to be added into the equation as a yearly depreciation cost, and this depends on how heavily the machine is used for serotyping pneumococci. Nonetheless, based on the information obtained from the national reference centers involved in this study, the average cost per sample for the quellung reaction was estimated to be around €60. The cost per sample using the IR Biotyper was estimated based on a yearly use of serotyping 1,500 pneumococcal strains (rough average of the yearly number of invasive pneumococci serotyped at the NRC Belgium) and was €20 per sample when using three technical replicates per strain or €60 when using five technical replicates per sample. The majority of strains are accurately serotyped using three technical replicates, and only a subset of strains need to be measured again using five technical replicates, yielding an average cost per sample of around €35. These calculations show that investing in the IR Biotyper machine and using it to serotype pneumococci is cost effective compared to the costs of other phenotypic methods. In this regard, it should be noted that the IR Biotyper has the potential to be used for serotyping other pathogens too, or for outbreak analysis (43–48).

Surprisingly, it was shown that the origin of the growth medium impacts the comparability of IR spectra between different institutions, and more work is required to fully understand the influence of environmental conditions on the IR spectra of pneumococci. In the future, FT-IR spectroscopy has the potential to be used for first-line serotyping (quick, cost effective, and accurate), after which strains that prove more difficult to serotype could be typed using the classical phenotypic or genotypic techniques. In this regard, it should be noted that FT-IR spectroscopy is not dependent on the development of new type-specific antisera and has the potential to quickly detect unknown serotypes. Finally, because certain serotypes are well known to be linked to high antimicrobial resistance, FT-IR spectroscopy could prove useful in clinical laboratory settings in the future, when quick serotyping results could assist clinicians in their decision-making process when considering different treatment options.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.
SUPPLEMENTAL FILE 1, PDF file, 0.9 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Janoff EN, Musher DM. 2015. Chapter 201. Streptococcus pneumoniae, p 2310–2327.e5. *In* Bennett JE, Dolin R, Blaser MJ (ed), Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases, 8th ed, vol 2. Saunders, Philadelphia, PA.
2. Oligbu G, Fry NK, Ladhani SN. 2019. The epidemiology and biostatistics of Pneumococcus. Methods Mol Biol 1968:215–224. https://doi.org/10.1007/978-1-4939-9199-0_18.
3. Oligbu G, Collins S, Djennad A, Sheppard CL, Fry NK, Andrews NJ, Borrow R, Ramsay ME, Ladhani SN. 2019. Effect of pneumococcal conjugate vaccines on pneumococcal meningitis, England and Wales, July 1, 2000-June 30, 2016. Emerg Infect Dis 25:1708–1718. https://doi.org/10.3201/eid2509.180747.
4. Ferreira-Coimbra J, Sarda C, Rello J. 2020. Burden of community-acquired pneumonia and unmet clinical needs. Adv Ther 37:1302–1318. https://doi.org/10.1007/s12325-020-01248-7.
5. Dagan R, Pelton S, Bakaletz L, Cohen R. 2016. Prevention of early episodes of otitis media by pneumococcal vaccines might reduce progression to complex disease. Lancet Infect Dis 16:480–492. https://doi.org/10.1016/S1473-3099(15)00549-6.
6. Desmet S. 2020. Report National Reference Centre Streptococcus pneumoniae 2020. National Reference Centre *Streptococcus pneumoniae*, UZ Leuven, Belgium. https://nrchm.wiv-isp.be/nl/ref_centra_labo/streptococcus_pneumoniae_invasive/Rapporten/Rapport%20Streptococcus%20pneumoniae%202020.pdf.
7. Troeger C, Forouzanfar M, Rao PC, Khalil I, Brown A, Swartz S, Fullman N, Mosser J, Thompson RL, Reiner RC, Abajobir A, Alam N, Alemayohu MA, Amare AT, Antonio CA, Asayesh H, Avokpaho E, Barac A, Beshir MA, Boneya DJ, Brauer M, Dandona L, Dandona R, Fitchett JRA, Gebrehiwot TT, Hailu GB, Hotez PJ, Kasaeian A, Khoja T, Kissoon N, Knibbs L, Kumar GA, Rai RK, Razek HMAE, Mohammed MSK, Nielson K, Oren E, Osman A, Patton G, Qorbani M, Roba HS, Sartorius B, Savic M, Shigematsu M, Sykes B, Swaminathan S, Topor-Madry R, Ukwaja K, Werdecker A, Yonemoto N, Zaki MES, Lim SS, et al. 2017. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Infect Dis 17:1133–1161. https://doi.org/10.1016/S1473-3099(17)30396-1.
8. Geno KA, Gilbert GL, Song JY, Skovsted IC, Klugman KP, Jones C, Konradsen HB, Nahm MH. 2015. Pneumococcal capsules and their types: past, present, and future. Clin Microbiol Rev 28:871–899. https://doi.org/10.1128/CMR.00024-15.
9. Paton JC, Trappetti C. 2019. Streptococcus pneumoniae capsular polysaccharide. Microbiol Spectr 7:7.2.33. https://doi.org/10.1128/microbiolspec.GPP3-0019-2018.

10. Weiser JN, Ferreira DM, Paton JC. 2018. Streptococcus pneumoniae: transmission, colonization and invasion. Nat Rev Microbiol 16:355–367. https://doi.org/10.1038/s41579-018-0001-8.

11. Hyams C, Camberlein E, Cohen JM, Bax K, Brown JS. 2010. The Streptococcus pneumoniae capsule inhibits complement activity and neutrophil phagocytosis by multiple mechanisms. Infect Immun 78:704–715. https://doi.org/10.1128/IAI.00881-09.

12. Abeyta M, Hardy GG, Yother J. 2003. Genetic alteration of capsule type but not PspA type affects accessibility of surface-bound complement and surface antigens of Streptococcus pneumoniae. Infect Immun 71:218–225. https://doi.org/10.1128/IAI.71.1.218-225.2003.

13. Nelson AL, Roche AM, Gould JM, Chim K, Ratner AJ, Weiser JN. 2007. Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. Infect Immun 75:83–90. https://doi.org/10.1128/IAI.01475-06.

14. Ganaie F, Maruhn K, Li C, Porambo RJ, Elverdal PL, Abeygunwardana C, van der Linden M, Duus JØ, Sheppard CL, Nahm MH. 2021. Structural, genetic, and serological elucidation of Streptococcus pneumoniae serogroup 24 serotypes: discovery of a new serotype, 24C, with a variable capsule structure. J Clin Microbiol 59:e0054-21. https://doi.org/10.1128/JCM.00540-21.

15. Izurieta P, Bahety P, Adegbola R, Clarke C, Hoet B. 2018. Public health impact of pneumococcal conjugate vaccine infant immunization programs: assessment of invasive pneumococcal disease burden and serotype distribution. Expert Rev Vaccines 17:479–493. https://doi.org/10.1080/14760584.2018.1413354.

16. Lawrence H, Pick H, Baskaran V, Daniel P, Rodrigo C, Ashton D, Edwards-Pritchard RC, Sheppard C, Eletu SD, Litt D, Fry NK, Rose S, Trotter C, McKeever TM, Lim WS. 2020. Effectiveness of the 23-valent pneumococcal polysaccharide vaccine against vaccine serotype pneumococcal pneumonia in adults: a case-control test-negative design study. PLoS Med 17: e1003326. https://doi.org/10.1371/journal.pmed.1003326.

17. Berild JD, Winje BA, Vestrheim DF, Slotved H-C, Valentiner-Branth P, Roth A, Storsäter J. 2020. A systematic review of studies published between 2016 and 2019 on the effectiveness and efficacy of pneumococcal vaccination on pneumonia and invasive pneumococcal disease in an elderly population. Pathogens 9:259. https://doi.org/10.3390/pathogens9040259.

18. Djennad A, Ramsay ME, Pebody R, Fry NK, Sheppard C, Ladhani SN, Andrews NJ. 2019. Effectiveness of 23-valent polysaccharide pneumococcal vaccine and changes in invasive pneumococcal disease incidence from 2000 to 2017 in those aged 65 and over in England and Wales. EClinicalMedicine 6:42–50. https://doi.org/10.1016/j.eclinm.2018.12.007.

19. Murdoch DR. 2020. Assessing the impact of pneumococcal conjugate vaccines. Clin Infect Dis 70:1589–1590. https://doi.org/10.1093/cid/ciz484.

20. Chapman R, Sutton K, Dillon-Murphy D, Patel S, Hilton B, Farkouh R, Wasserman M. 2020. Ten year public health impact of 13-valent pneumococcal conjugate vaccination in infants: a modelling analysis. Vaccine 38:7138–7145. https://doi.org/10.1016/j.vaccine.2020.08.068.

21. Jayasinghe S, Chiu C, Quinn H, Menzies R, Gilmour R, McIntyre P. 2018. Effectiveness of 7- and 13-valent pneumococcal conjugate vaccines in a schedule without a booster dose: a 10-year observational study. Clin Infect Dis 67:367–374. https://doi.org/10.1093/cid/ciy129.

22. Wasserman M, Chapman R, Lapidot R, Sutton K, Dillon-Murphy D, Patel S, Chilson E, Snow V, Farkouh R, Pelton S. 2021. Twenty-year public health impact of 7- and 13-valent pneumococcal conjugate vaccines in US children. Emerg Infect Dis 27:1627–1636. https://doi.org/10.3201/eid2706.204238.

23. Song J-Y, Chang C-J, Andrews C, Diez-Domingo J, Oh M-D, Dagan R, Hartzel J, Pedley A, Li J, Sterling T, Tamms G, Chiarappa JA, Lutkiewicz J, Musey L, Tu Y, Buchwald UK, V114-016 (PNEU-PATH) study group. 2021. Safety, tolerability, and immunogenicity of V114, a 15-valent pneumococcal conjugate vaccine, followed by sequential PPSV23 vaccination in healthy adults aged ≥50 years: a randomized phase III trial (PNEU-PATH). Vaccine 39:6422–6436. https://doi.org/10.1016/j.vaccine.2021.08.038.

24. Platt HL, Cardona JF, Haranaka M, Schwartz HI, Narejos Perez S, Dowell A, Chang C-J, Dagan R, Tamms GM, Sterling T, Morgan L, Shi Y, Pedley A, Musey LK, Buchwald UK. 2022. A phase 3 trial of safety, tolerability, and immunogenicity of V114, 15-valent pneumococcal conjugate vaccine, compared with 13-valent pneumococcal conjugate vaccine in adults 50 years of age and older (PNEU-AGE). Vaccine 40:162–172. https://doi.org/10.1016/j.vaccine.2021.08.049.

25. Severance R, Schwartz H, Dagan R, Connor L, Li J, Pedley A, Hartzel J, Sterling TM, Nolan KM, Tamms GM, Musey LK, Buchwald UK. 2022. Safety, tolerability, and immunogenicity of V114, a 15-valent pneumococcal conjugate vaccine, administered concomitantly with influenza vaccine in healthy

adults aged ≥50 years: a randomized phase 3 trial (PNEU-FLU). Hum Vaccin Immunother 18:1–14. https://doi.org/10.1080/21645515.2021.1976581.

26. Klein NP, Peyrani P, Yacisin K, Caldwell N, Xu X, Scully IL, Scott DA, Jansen KU, Gruber WC, Watson W. 2021. A phase 3, randomized, double-blind study to evaluate the immunogenicity and safety of 3 lots of 20-valent pneumococcal conjugate vaccine in pneumococcal vaccine-naive adults 18 through 49 years of age. Vaccine 39:5428–5435. https://doi.org/10.1016/j.vaccine.2021.07.004.

27. Senders S, Klein NP, Lamberth E, Thompson A, Drozd J, Trammel J, Peng Y, Giardina PC, Jansen KU, Gruber WC, Scott DA, Watson W. 2021. Safety and Immunogenicity of a 20-valent pneumococcal conjugate vaccine in healthy infants in the United States. Pediatr Infect Dis J 40:944–951. https://doi.org/10.1097/INF.0000000000003277.

28. Oligbu G. 2020. Higher valent pneumococcal conjugate vaccines: is it a roller coaster? AIMS Public Health 7:29–32. https://doi.org/10.3934/publichealth.2020004.

29. Dobay O. 2019. The complexity of serotype replacement of pneumococci. Hum Vaccin Immunother 15:2725–2728. https://doi.org/10.1080/21645515.2019.1621068.

30. Beckler E, Macleod P. 1934. The Neufeld method of Pneumococcus type determination as carried out in a public health laboratory: a study of 760 typings. J Clin Invest 13:901–907. https://doi.org/10.1172/JCI100634.

31. Austrian R. 1976. The quellung reaction, a neglected microbiologic technique. Mt Sinai J Med 43:699–709.

32. Slotved H-C, Kaltoft M, Skovsted IC, Kerrn MB, Espersen F. 2004. Simple, rapid latex agglutination test for serotyping of pneumococci (Pneumotest-latex). J Clin Microbiol 42:2518–2522. https://doi.org/10.1128/JCM.42.6.2518-2522.2004.

33. Jauneikaite E, Tocheva AS, Jefferies JMC, Gladstone RA, Faust SN, Christodoulides M, Hibberd ML, Clarke SC. 2015. Current methods for capsular typing of Streptococcus pneumoniae. J Microbiol Methods 113:41–49. https://doi.org/10.1016/j.mimet.2015.03.006.

34. Epping L, van Tonder AJ, Gladstone RA, Bentley SD, Page AJ, Keane JA, The Global Pneumococcal Sequencing Consortium. 2018. SeroBA: rapid high-throughput serotyping of Streptococcus pneumoniae from whole genome sequence data. Microb Genom 4:e000186. https://doi.org/10.1099/mgen.0.000186.

35. Kapatai G, Sheppard CL, Al-Shahib A, Litt DJ, Underwood AP, Harrison TG, Fry NK. 2016. 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.

36. Slotved H-C, Sheppard CL, Dalby T, van der Ende A, Fry NK, Morfeldt E, Nyholm O, Rokney A, Ron M, Siira L, Scott KJ, Smith A, Thom L, Toropainen M, Vestrheim DF. 2017. External quality assurance for laboratory identification and capsular typing of Streptococcus pneumoniae. Sci Rep 7:13280. https://doi.org/10.1038/s41598-017-13605-8.

37. Burckhardt I, Sebastian K, Mauder N, Kostrzewa M, Burckhardt F, Zimmermann S. 2019. Analysis of Streptococcus pneumoniae using Fourier-transformed infrared spectroscopy allows prediction of capsular serotype. Eur J Clin Microbiol Infect Dis 38:1883–1890. https://doi.org/10.1007/s10096-019-03622-y.

38. Leung MH, Bryson K, Freystatter K, Pichon B, Edwards G, Charalambous BM, Gillespie SH. 2012. Sequetyping: serotyping Streptococcus pneumoniae by a single PCR sequencing strategy. J Clin Microbiol 50:2419–2427. https://doi.org/10.1128/JCM.06384-11.

39. Velusamy S, Tran T, Mongkolrattanothai T, Walker H, McGee L, Beall B. 2020. Expanded sequential quadriplex real-time polymerase chain reaction (PCR) for identifying pneumococcal serotypes, penicillin susceptibility, and resistance markers. Diagn Microbiol Infect Dis 97:115037. https://doi.org/10.1016/j.diagmicrobio.2020.115037.

40. Knight JR, Dunne EM, Mulholland EK, Saha S, Satzke C, Tothpal A, Weinberger DM. 2021. Determining the serotype composition of mixed samples of pneumococcus using whole-genome sequencing. Microb Genom 7:mgen000494. https://doi.org/10.1099/mgen.0.000494.

41. Novais Â, Freitas AR, Rodrigues C, Peixe L. 2019. Fourier transform infrared spectroscopy: unlocking fundamentals and prospects for bacterial strain typing. Eur J Clin Microbiol Infect Dis 38:427–448. https://doi.org/10.1007/s10096-018-3431-3.

42. Naumann D, Helm D, Labischinski H. 1991. Microbiological characterizations by FT-IR spectroscopy. Nature 351:81–82. https://doi.org/10.1038/351081a0.

43. Preisner OE, Menezes JC, Guiomar R, Machado J, Lopes JA. 2012. Discrimination of Salmonella enterica serotypes by Fourier transform infrared spectroscopy. Food Res International 45:1058–1064. https://doi.org/10.1016/j.foodres.2011.02.029.

44. Campos J, Sousa C, Mourão J, Lopes J, Antunes P, Peixe L. 2018. Discrimination of non-typhoid Salmonella serogroups and serotypes by Fourier transform infrared spectroscopy: a comprehensive analysis. Int J Food Microbiol 285:34–41. https://doi.org/10.1016/j.ijfoodmicro.2018.07.005.

45. Cordovana M, Mauder N, Kostrzewa M, Wille A, Rojak S, Hagen RM, Ambretti S, Pongolini S, Soliani L, Justesen US, Holt HM, Join-Lambert O, Hello SL, Auzou M, Veloo AC, May J, Frickmann H, Dekker D. 2021. Classification of Salmonella enterica of the (para-)typhoid fever group by Fourier-transform infrared (FTIR) spectroscopy. Microorganisms 9:853. https://doi.org/10.3390/microorganisms9040853.

46. Dinkelacker AG, Vogt S, Oberhettinger P, Mauder N, Rau J, Kostrzewa M, Rossen JWA, Autenrieth IB, Peter S, Liese J. 2018. Typing and species identification of clinical Klebsiella isolates by Fourier transform infrared spectroscopy and matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol 56:e00843-18. https://doi.org/10.1128/JCM.00843-18.

47. Martak D, Valot B, Sauget M, Cholley P, Thouverez M, Bertrand X, Hocquet D. 2019. Fourier-transform infrared spectroscopy can quickly type gram-negative bacilli responsible for hospital outbreaks. Front Microbiol 10:1440. https://doi.org/10.3389/fmicb.2019.01440.

48. Vogt S, Löffler K, Dinkelacker AG, Bader B, Autenrieth IB, Peter S, Liese J. 2019. Fourier-transform infrared (FTIR) spectroscopy for typing of clinical Enterobacter cloacae complex isolates. Front Microbiol 10:2582. https://doi.org/10.3389/fmicb.2019.02582.

49. Ceyssens P-J, Bambeke FV, Mattheus W, Bertrand S, Fux F, Bossuyt EV, Damée S, Nyssen H-J, Craeye SD, Verhaegen J, Pneumoniae S, Tulkens PM, Vanhoof R, The Belgian Streptococcus pneumoniae Study Group. 2016. Molecular analysis of rising fluoroquinolone resistance in Belgian non-invasive Streptococcus pneumoniae isolates (1995–2014). PLoS One 11:e0154816. https://doi.org/10.1371/journal.pone.0154816.

50. Vanhoof R, Camps K, Carpentier M, De Craeye S, Frans J, Glupczynski Y, Goffinet P, Gordts B, Govaerts D, Ide L, Lefèvre P, Lontie M, Cartuyvels R, Meunier F, Mulongo B, Philippart I, Surmont I, Van Bossuyt E, Van Eldere J, Verhaegen J. 2010. 10th survey of antimicrobial resistance in noninvasive clinical isolates of Streptococcus pneumoniae collected in Belgium during winter 2007–2008. Pathol Biol (Paris) 58:147–151. https://doi.org/10.1016/j.patbio.2009.07.018.

51. Desmet S, Lagrou K, Wyndham-Thomas C, Braeye T, Verhaegen J, Maes P, Fieuws S, Peetermans WE, Blumental S. 2021. Dynamic changes in paediatric invasive pneumococcal disease after sequential switches of conjugate vaccine in Belgium: a national retrospective observational study. Lancet Infect Dis 21:127–136. https://doi.org/10.1016/S1473-3099(20)30173-0.

52. Peckeu L, van der Ende A, de Melker HE, Sanders EAM, Knol MJ. 2021. Impact and effectiveness of the 10-valent pneumococcal conjugate vaccine on invasive pneumococcal disease among children under 5 years of age in the Netherlands. Vaccine 39:431–437. https://doi.org/10.1016/j.vaccine.2020.11.018.

53. Burckhardt I, Panitz J, Burckhardt F, Zimmermann S. 2017. Identification of Streptococcus pneumoniae: development of a standardized protocol for optochin susceptibility testing using total lab automation. Biomed Res Int 2017:4174168. https://doi.org/10.1155/2017/4174168.

54. Hammer O, Harper DAT, Ryan PD. 2001. PAST: paleontological statistics software package for education and data analysis. Palaeontolog Electron 4:art 4.

55. Criscuolo A, Brisse S. 2013. AlienTrimmer: a tool to quickly and accurately trim off multiple short contaminant sequences from high-throughput sequencing reads. Genomics 102:500–506. https://doi.org/10.1016/j.ygeno.2013.07.011.

56. Crusoe MR, Alameldin HF, Awad S, Boucher E, Caldwell A, Cartwright R, Charbonneau A, Constantinides B, Edvenson G, Fay S, Fenton J, Fenzl T, Fish J, Garcia-Gutierrez L, Garland P, Gluck J, González I, Guermond S, Guo J, Gupta A, Herr JR, Howe A, Hyer A, Härpfer A, Irber L, Kidd R, Lin D, Lippi J, Mansour T, McA'Nulty P, McDonald E, Mizzi J, Murray KD, Nahum JR, Nanlohy K, Nederbragt AJ, Ortiz-Zuazaga H, Ory J, Pell J, Pepe-Ranney C, Russ ZN, Schwarz E, Scott C, Seaman J, Sievert S, Simpson J, Skennerton CT, Spencer J, Srinivasan R, Standage D, et al. 2015. The khmer software package: enabling efficient nucleotide sequence analysis. F1000Res 4:900. https://doi.org/10.12688/f1000research.6924.1.

57. Liu Y, Schröder J, Schmidt B. 2013. Musket: a multistage k-mer spectrum-based error corrector for Illumina sequence data. Bioinformatics 29:308–315. https://doi.org/10.1093/bioinformatics/bts690.

58. Mavroidi A, Aanensen DM, Godoy D, Skovsted IC, Kaltoft MS, Reeves PR, Bentley SD, Spratt BG. 2007. Genetic relatedness of the Streptococcus pneumoniae capsular biosynthetic loci. J Bacteriol 189:7841–7855. https://doi.org/10.1128/JB.00836-07.

59. Sheppard CL, Pichon B, George RC, Hall LMC. 2010. Streptococcus pneumoniae isolates expressing a capsule with epitopes of both serotypes 6A and 6B. Clin Vaccine Immunol 17:1820–1822. https://doi.org/10.1128/CVI.00335-10.

60. Park IH, Pritchard DG, Cartee R, Brandao A, Brandileone MCC, Nahm MH. 2007. Discovery of a new capsular serotype (6C) within serogroup 6 of Streptococcus pneumoniae. J Clin Microbiol 45:1225–1233. https://doi.org/10.1128/JCM.02199-06.

61. Jin P, Kong F, Xiao M, Oftadeh S, Zhou F, Liu C, Russell F, Gilbert GL. 2009. First report of putative Streptococcus pneumoniae serotype 6D among nasopharyngeal isolates from Fijian children. J Infect Dis 200:1375–1380. https://doi.org/10.1086/606118.

62. Deloria Knoll M, Bennett JC, Garcia Quesada M, Kagucia EW, Peterson ME, Feikin DR, Cohen AL, Hetrich MK, Yang Y, Sinkevitch JN, Ampofo K, Aukes L, Bacci S, Bigogo G, Brandileone M-CC, Bruce MG, Camilli R, Castilla J, Chan G, Chanto Chacón G, Ciruela P, Cook H, Corcoran M, Dagan R, Danis K, de Miguel S, De Wals P, Desmet S, Galloway Y, Georgakopoulou T, Hammitt LL, Hilty M, Ho P-L, Jayasinghe S, Kellner JD, Kleynhans J, Knol MJ, Kozakova J, Kristinsson KG, Ladhani SN, Lara CS, León ME, Lepp T, Mackenzie GA, Mad'arová L, McGeer A, Mungun T, Mwenda JM, Nuorti JP, Nzoyikorera N, the Pserenade Team, et al. 2021. Global landscape review of serotype-specific invasive pneumococcal disease surveillance among countries using PCV10/13: the Pneumococcal Serotype Replacement and Distribution Estimation (PSERENADE) Project. Microorganisms 9:742. https://doi.org/10.3390/microorganisms9040742.

63. Hanquet G, Krizova P, Valentiner-Branth P, Ladhani SN, Nuorti JP, Lepoutre A, Mereckiene J, Knol M, Winje BA, Ciruela P, Ordobas M, Guevara M, McDonald E, Morfeldt E, Kozakova J, Slotved H-C, Fry NK, Rinta-Kokko H, Varon E, Corcoran M, van der Ende A, Vestrheim DF, Munoz-Almagro C, Latasa P, Castilla J, Smith A, Henriques-Normark B, Whittaker R, Pastore Celentano L, Savulescu C, SpIDnet/I-MOVE+ Pneumo Group. 2019. Effect of childhood pneumococcal conjugate vaccination on invasive disease in older adults of 10 European countries: implications for adult vaccination. Thorax 74:473–482. https://doi.org/10.1136/thoraxjnl-2018-211767.

64. Miller E, Andrews NJ, Waight PA, Slack MP, George RC. 2011. 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.

65. Corcoran M, Vickers I, Mereckiene J, Murchan S, Cotter S, Fitzgerald M, Mcelligott M, Cafferkey M, O'Flanagan D, Cunney R, Humphreys H. 2017. The epidemiology of invasive pneumococcal disease in older adults in the post-PCV era. Has there been a herd effect? Epidemiol Infect 145:2390–2399. https://doi.org/10.1017/S0950268817001194.

66. Amin-Chowdhury Z, Collins S, Sheppard C, Litt D, Fry NK, Andrews N, Ladhani SN. 2020. Characteristics of invasive pneumococcal disease caused by emerging serotypes after the introduction of the 13-valent pneumococcal conjugate vaccine in England: a prospective observational cohort study, 2014–2018. Clin Infect Dis 71:e235–e243. https://doi.org/10.1093/cid/ciaa043.

67. van der Linden M, Imöhl M, Perniciaro S. 2019. Limited indirect effects of an infant pneumococcal vaccination program in an aging population. PLoS One 14:e0220453. https://doi.org/10.1371/journal.pone.0220453.

68. Vestjens SMT, Sanders EAM, Vlaminckx BJ, de Melker HE, van der Ende A, Knol MJ. 2019. Twelve years of pneumococcal conjugate vaccination in the Netherlands: impact on incidence and clinical outcomes of invasive pneumococcal disease. Vaccine 37:6558–6565. https://doi.org/10.1016/j.vaccine.2019.08.025.

69. Lewnard JA, Hanage WP. 2019. Making sense of differences in pneumococcal serotype replacement. Lancet Infect Dis 19:e213–e220. https://doi.org/10.1016/S1473-3099(18)30660-1.

70. Converso TR, Assoni L, André GO, Darrieux M, Leite LCC. 2020. The long search for a serotype independent pneumococcal vaccine. Expert Rev Vaccines 19:57–70. https://doi.org/10.1080/14760584.2020.1711055.

71. Arends DW, Miellet WR, Langereis JD, Ederveen THA, van der Gaast-de Jongh CE, van Scherpenzeel M, Knol MJ, van Sorge NM, Lefeber DJ, Trzciński K, Sanders EAM, Dorfmueller HC, Bootsma HJ, de Jonge MI. 2021. Examining the distribution and impact of single-nucleotide polymorphisms in the capsular locus of Streptococcus pneumoniae serotype 19A. Infect Immun 89:e00246-21. https://doi.org/10.1128/IAI.00246-21.

72. Ruiz García Y, Nieto Guevara J, Izurieta P, Vojtek I, Ortega-Barría E, Guzman-Holst A. 2021. Circulating clonal complexes and sequence types of Streptococcus pneumoniae serotype 19A worldwide: the importance of multidrug resistance: a systematic literature review. Expert Rev Vaccines 20:45–57. https://doi.org/10.1080/14760584.2021.1873136.

73. Oliveira LMA, Souza ARV, Pinto TCA, Teixeira LM. 2021. Characterization of Streptococcus pneumoniae serotype 19F-variants occurring in Brazil uncovers a predominant lineage that can lead to misinterpretation in capsular typing. Int J Infect Dis 104:580–583. https://doi.org/10.1016/j.ijid.2021.01.030.

74. Lo SW, Gladstone RA, van Tonder AJ, Lees JA, Du Plessis M, Benisty R, Givon-Lavi N, Hawkins PA, Cornick JE, Kwambana-Adams B, Law PY, Ho PL, Antonio M, Everett DB, Dagan R, von Gottberg A, Klugman KP, McGee L, Breiman RF, Bentley SD, Brooks AW, Corso A, Davydov A, Maguire A, Pollard A, Kiran A, Skoczynska A, Moiane B, Beall B, Sigauque B, Aanensen D, Lehmann D, Faccone D, Foster-Nyarko E, Bojang E, Egorova E, Voropaeva E, Sampane-Donkor E, Sadowy E, Bigogo G, Mucavele H, Belabbès H, Diawara I, Moïsi J, Verani J, Keenan J, Bhai JNNT, Ndlangisa KM, Zerouali K, Ravikumar KL, et al. 2019. Pneumococcal lineages associated with serotype replacement and antibiotic resistance in childhood invasive pneumococcal disease in the post-PCV13 era: an international whole-genome sequencing study. Lancet Infect Dis 19:759–769. https://doi.org/10.1016/S1473-3099(19)30297-X.

75. Slotved H-C, Kerrn MB. 2005. The effect of broth media on pneumococcal growth and the latex serotyping result. J Microbiol Methods 61:181–186. https://doi.org/10.1016/j.mimet.2004.11.013.