Published in final edited form as:

Pediatr Infect Dis J. 2018 November ; 37(11): 1176–1183. doi:10.1097/INF.0000000000002174.

Pneumococcal Colonization and the Nasopharyngeal Microbiota of Children in Botswana

Matthew S. Kelly, MD, MPH^{1,2}, Michael G. Surette, PhD³, Marek Smieja, MD, PhD^{3,6}, Laura Rossi, BSc³, Kathy Luinstra, BSc⁶, Andrew P. Steenhoff, MBBCh, FCPaed^{1,7,8}, David M. Goldfarb, MD^{1,9}, Jeffrey M. Pernica, MD⁵, Tonya Arscott-Mills, MD, MPH^{1,7}, Sefelani Boiditswe, BNSc¹, Tiny Mazhani, MD¹⁰, John F. Rawls, PhD¹¹, Coleen K. Cunningham, MD², Samir S. Shah, MD, MSCE^{12,13}, Kristen A. Feemster, MD, MPH, MSHPR^{#7,8}, and Patrick C. Seed, MD, PhD^{#14,15}

¹Botswana-UPenn Partnership, Gaborone, Botswana

²Division of Pediatric Infectious Diseases, Duke University Medical Center, Durham, NC, USA

³Department of Medicine, McMaster University, Hamilton, Ontario, Canada

⁴Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Ontario, Canada

⁵Department of Pediatrics, McMaster University, Hamilton, Ontario, Canada

⁶St. Joseph's Healthcare, Hamilton, Ontario, Canada

⁷Global Health Center, Children's Hospital of Philadelphia, Philadelphia, PA, USA

⁸Division of Pediatric Infectious Diseases, Children's Hospital of Philadelphia, Philadelphia, PA, USA

⁹Department of Pathology and Laboratory Medicine, BC Children's Hospital, Vancouver, British Columbia, Canada

¹⁰University of Botswana School of Medicine, Gaborone, Botswana

¹¹Duke Microbiome Center, Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC

¹²Divisions of Hospital Medicine and Infectious Diseases, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

¹³University of Cincinnati College of Medicine, Cincinnati, OH, USA

¹⁴Ann and Robert H. Lurie Children's Hospital, Chicago, IL

¹⁵Department of Pediatrics, Feinberg School of Medicine, Northwestern University, Chicago, IL

[#] These authors contributed equally to this work.

Corresponding author: Matthew S. Kelly, MD, MPH, Box 3499, Duke University Medical Center, Durham, NC 27710; Matthew.kelly@duke.edu; (919)-684-6335 Fax: (919)-668-4859.

Conflicts of Interest: The authors have no conflicts of interest to declare.

Abstract

Background: Nasopharyngeal colonization precedes infections caused by *Streptococcus pneumoniae*. A more detailed understanding of interactions between *S. pneumoniae* and the nasopharyngeal microbiota of children could inform strategies to prevent pneumococcal infections.

Methods: We collected nasopharyngeal swabs from children 1 to 23 months of age in Botswana between August 2012 and June 2016. We tested samples for *S. pneumoniae* and common respiratory viruses using PCR. We sequenced the V3 region of the bacterial 16S ribosomal RNA gene and used random forest models to identify clinical variables and bacterial genera that were associated with pneumococcal colonization.

Results: Mean age of the 170 children included in this study was 8.3 months. Ninety-six (56%) children were colonized with *S. pneumoniae*. Pneumococcal colonization was associated with older age ($P=0.0001$), a lack of electricity in the home ($P=0.02$), and household use of wood as a cooking fuel ($P=0.002$). Upper respiratory symptoms were more frequent in children with *S. pneumoniae* colonization (60% vs. 32%; $P=0.001$). Adjusting for age, nasopharyngeal microbiota composition differed in colonized and noncolonized children ($P=0.001$). *S. pneumoniae* colonization was associated with a higher relative abundance of *Moraxella* ($P=0.001$) and lower relative abundances of *Corynebacterium* ($P=0.001$) and *Staphylococcus* ($P=0.03$). A decision tree model containing the relative abundances of bacterial genera had 81% sensitivity and 85% specificity for the determination of *S. pneumoniae* colonization status.

Conclusions: *S. pneumoniae* colonization is associated with characteristic alterations of the nasopharyngeal microbiota of children. Prospective studies should determine if nasopharyngeal microbial composition alters the risk of pneumococcal colonization and thus could be modified as a novel pneumonia prevention strategy.

Keywords

microbiota; microbial communities; *Streptococcus pneumoniae*; respiratory viruses; children

INTRODUCTION

Streptococcus pneumoniae (pneumococcus) is the predominant bacterial respiratory pathogen during childhood. Despite the increasing use of pneumococcal conjugate vaccines, *S. pneumoniae* is still responsible for more than 500,000 child deaths each year.¹ Colonization of the nasopharynx precedes respiratory infections and invasive disease caused by *S. pneumoniae*.² Variations in the local environment of the nasopharynx affect whether pneumococcal colonization occurs and, once colonization is established, whether infection occurs.^{3,4} The nasopharyngeal microbiota may provide a barrier to colonization by *S. pneumoniae* through competition for nutrients, production of antimicrobial substances, or modulation of the local immune response.⁵ However, the bacteria within the nasopharyngeal microbiota that modify the risk of pneumococcal colonization have not been systematically identified.

In this study, we used nasopharyngeal samples from children in Botswana to identify clinical risk factors for pneumococcal colonization and to identify associations between the bacterial composition of the nasopharyngeal microbiota and *S. pneumoniae* colonization.

MATERIALS AND METHODS

Setting

Botswana is a landlocked country in southern Africa that has a semi-arid climate, with a short rainy season that typically occurs from November to March. The country's under-five child mortality rate was estimated to be 41 per 1,000 live births in 2016.⁶ *Haemophilus influenzae* type B (Hib) and 13-valent pneumococcal conjugate (PCV-13) vaccinations were introduced in November 2010 and July 2012, respectively. Gaborone, the capital and largest city of Botswana, is located in the country's South-East District.

Study population

We conducted a case-control study of pneumonia between April 2012 and June 2016, as described elsewhere in detail.⁷ For the purposes of these analyses, we utilized the control group comprised of children one to 23 months of age receiving well child or acute care services at one of 18 public clinics in the Gaborone area. Controls were matched to pneumonia cases 1:1 by primary care clinic and date (2 weeks from the case enrollment). Exclusion criteria included World Health Organization clinical pneumonia ("cough or difficult breathing with lower chest wall indrawing"), a chronic medical condition predisposing to pneumonia (other than HIV), hospitalization in the prior 14 days, and asthma.⁸ Children with upper respiratory symptoms – defined as "cough or nasal congestion or discharge" – were eligible for inclusion. Sociodemographic and clinical data were recorded from a face-to-face interview with the child's caregivers and review of infant and maternal medical records.

Laboratory Methods

Nasopharyngeal specimens were collected using flocked swabs and universal transport medium (Copan Italia, Brescia, Italy). These samples were placed into a cooler with ice packs immediately after collection and transported within six hours to the National Health Laboratory in Gaborone where they were transferred to cryovials for storage at –80°C. Samples were shipped on dry ice at six-month intervals to St. Joseph's Healthcare (Hamilton, ON, Canada). Total genomic DNA was extracted from nasopharyngeal specimens using a protocol involving mechanical and enzymatic lysis.⁹ Polymerase chain reaction (PCR) was used to amplify the variable V3 region of the bacterial 16S ribosomal RNA (rRNA) gene using broad-range primers and an Illumina MiSeq instrument (BioProject accession number PRJNA423191).^{9,10} These same nasopharyngeal specimens were tested for respiratory syncytial virus, influenza viruses A and B, parainfluenza virus types 1–4, human metapneumovirus, adenovirus, coronaviruses (NL63, 229E, OC43, HKU1), bocavirus, and rhinovirus/enterovirus using nucleic acid extracted on the NucliSens easyMAG platform (bioMérieux, Marcy l'Etoile, France) and amplified using a Rotor-Gene 6500 real-time PCR thermocycler (Qiagen, Hilden, Germany). This nucleic acid extract was

also tested for *S. pneumoniae* using a quantitative PCR assay targeting the autolysin (*lytA*) gene.⁹

Bioinformatics

Raw sequences were analyzed through a pipeline that included Cutadapt and PANDAseq tools.⁹ Sequences were clustered at 97% identity using AbundantOTU+, and given taxonomic assignments using the Ribosomal Database Project classifier against the Greengenes 2011 database.⁹ Each read was grouped into an operational taxonomic unit (OTU), a cluster of similar 16S rRNA sequences corresponding roughly to a species or a group of closely related species. We excluded three nasopharyngeal samples with less than 5,000 reads. Using the R package *phyloseq*, OTUs with fewer than ten counts across all samples or that did not appear in at least 10% of samples were removed.¹¹ Sample reads were then normalized to the cumulative 75th percentile using the *metagenomeSeq* package in R.¹² Because a single OTU accounted for >75% of reads for 82% of genera, we aggregated OTUs at the genus level for compositional analyses.

Statistical Analyses

Characteristics of children colonized and noncolonized with *S. pneumoniae* were compared using Chi square or Fisher's exact tests (for categorical variables) and two-sample t tests (for continuous variables). Species richness was measured by the Chao1 index, and microbial diversity was assessed using the Shannon diversity index.^{13,14} After verifying that these measures approximated normal distributions in the study population, we used multivariable linear regression to compare the Chao1 and Shannon diversity indices in nasopharyngeal samples from colonized and noncolonized children. We considered genera with a mean relative abundance of 0.5% in the overall population to be highly abundant. We included these genera in logistic regression models to evaluate whether the proportion of each genus was associated with *S. pneumoniae* colonization. These analyses were adjusted for age given the substantial changes in the nasopharyngeal microbiota during infancy that we previously reported in this study population.¹⁵ We used random forest models implemented with the R package *caret* to identify the bacterial genera that were most important in discriminating *S. pneumoniae*-colonized children from noncolonized children.¹⁶ For these analyses, the study population was divided into training (50%) and test (50%) sets and the genera *Streptococcus* and unclassified *Streptococcaceae* were excluded to ensure that *S. pneumoniae* was not assigned to genera within the model. The model was initially fit on the training set and the sensitivity and specificity of the model was then calculated based on fitting of test set data. To further evaluate the association between *S. pneumoniae* colonization and upper respiratory symptoms, we included the density of *S. pneumoniae* colonization (measured as copies/mL by quantitative PCR) in a multivariable logistic regression model adjusting for the following potential confounders: age, sex, season, enrollment year, maternal HIV infection, current breastfeeding, electricity in the home, household use of wood as a cooking fuel, number of household members, PCV-13 doses, receipt of antibiotics within the prior 7 days, and detection of one or more respiratory viruses. Statistical analyses were conducted using R version 3.4.2.

RESULTS

Patient Characteristics

One hundred seventy children were included in these analyses. Mean age was 8.3 months and 51% were female. Nasopharyngeal colonization by *S. pneumoniae* was detected in 96 (56%) children. Sixty (35%) children had one or more respiratory viruses, the majority (76%) of which were rhinovirus/enterovirus infections. Characteristics of the study population are shown in Table 1. Age, electricity in the home, household use of wood as a cooking fuel, and the prevalence of upper respiratory symptoms differed between *S. pneumoniae*-colonized and noncolonized children.

Nasopharyngeal Microbiota Diversity

A total of 9,016,655 high-quality 16S rRNA sequences (mean of 53,039 sequences per sample) were obtained from the 170 nasopharyngeal samples included in this study. Rarefaction curves were constructed to ensure coverage of the bacterial diversity present (Supplemental Figure 1). All samples had sufficient sequencing depth to obtain a high degree of sequence coverage. The mean (standard deviation) Chao1 and Shannon diversity indices in the overall population were 250.5 (106.5) and 1.30 (0.61). Species richness, as measured by the Chao1 index, and microbial diversity, assessed using the Shannon diversity index, did not differ in *S. pneumoniae*-colonized and noncolonized children after adjustment for age (Supplemental Figure 2).

Nasopharyngeal Microbiota Composition

Sequences were clustered into 336 bacterial OTUs (after filtering), representing 155 genera from ten phyla. Principal component analysis using Bray-Curtis distances demonstrated that overall nasopharyngeal samples clustered by detection of *S. pneumoniae* and that *S. pneumoniae*-colonized and noncolonized microbial communities differed significantly [Figure 1; $P<0.001$ by PERMANOVA, adjusting for age]. The abundances of highly abundant genera in children by *S. pneumoniae* colonization status are shown in Figures 2 and 3. Adjusting for age, *S. pneumoniae* colonization was associated with higher relative abundances of *Moraxella* ($P=0.001$) and *Streptococcus* ($P<0.0001$) and lower relative abundances of *Corynebacterium* ($P=0.001$) and *Staphylococcus* ($P=0.03$). Respiratory virus infection was associated with a higher relative abundance of *Streptococcus* ($P=0.01$). The most abundant genera in nasopharyngeal samples contained typical commensal bacteria (*Corynebacterium*, *Dolosigranulum*) and common potential respiratory pathogens (*Haemophilus*, *Moraxella*, *Staphylococcus*, and *Streptococcus*). Each of the highly abundant genera containing a common respiratory pathogen was comprised almost entirely of a single dominant OTU (Supplemental Figure 3).

To further evaluate for confounding by age, we performed analyses in a subset of *S. pneumoniae*-colonized and noncolonized children matched on age. Mean (standard deviation) ages of the 57 noncolonized and 57 colonized children in these matched analyses were 7.9 (4.8) and 8.2 (4.8) months, respectively. Analyses comparing nasopharyngeal microbial diversity and composition in *S. pneumoniae*-colonized and noncolonized children were substantively unchanged in this matched cohort.

Bacterial Genera Predictive of Pneumococcal Colonization

A random forest analysis including non-streptococcal bacterial genera was performed to identify the genera that were most useful in discriminating *S. pneumoniae*-colonized and noncolonized children (Figure 4). We calculated a sensitivity of 81% and a specificity of 85% for the determination of *S. pneumoniae* colonization status by this decision tree model. The most discriminative genera were *Moraxella*, unclassified *Lactobacillales*, *Haemophilus*, unclassified *Bacillaceae*, and *Klebsiella*. Among the top 20 predictive genera that were not highly abundant, *S. pneumoniae* colonization was associated with a higher relative abundance of unclassified *Lactobacillales* ($P=0.005$) and lower relative abundances of *Chryseobacterium* ($P=0.03$), *Granulicatella* ($P=0.04$), *Halomonas* ($P=0.04$), *Rothia* ($P=0.03$), and unclassified *Bacillaceae* ($P=0.046$) after adjustment for age.

Upper Respiratory Symptoms

We performed additional analyses to investigate the association between *S. pneumoniae* colonization and upper respiratory symptoms. First, even when children with respiratory viruses were excluded, we found that the prevalence of upper respiratory symptoms was higher in children colonized with *S. pneumoniae* than in noncolonized children (49% vs. 20%; $P=0.002$). Moreover, we found that the density of *S. pneumoniae* colonization was an independent predictor of the presence of upper respiratory symptoms ($P=0.001$) in a multivariable logistic regression model. The composition of the nasopharyngeal microbiota of children with upper respiratory symptoms also differed significantly from that of children without upper respiratory symptoms ($P<0.001$ by PERMANOVA, adjusting for age). The presence of upper respiratory symptoms was associated with a higher relative abundance of *Streptococcus* ($P=0.0004$) and a lower relative abundance of *Corynebacterium* ($P=0.04$) in analyses adjusted for age.

DISCUSSION

In this study of infants and young children in Botswana, we observed marked alterations in nasopharyngeal microbial composition in children colonized with *S. pneumoniae*. These alterations included changes in the relative abundances of genera containing potential respiratory pathogens (*Moraxella*, *Staphylococcus*) and a genera that contains common nasopharyngeal commensals (*Corynebacterium*).

The nasopharynx is the site of dynamic interactions between diverse communities of commensal bacteria and potential respiratory pathogens. Our findings indicate that colonization of the nasopharynx by *S. pneumoniae* is associated with distinct alterations in these microbial communities. Specifically, we observed a higher relative abundance of *Moraxella* and lower relative abundances of *Corynebacterium* and *Staphylococcus* in children colonized with *S. pneumoniae*. Although the sequencing methodology that we used did not enable determination of bacterial composition at the species level, the genera containing common respiratory pathogens (*Moraxella*, *Staphylococcus*, *Streptococcus*) were dominated by a single OTU, and prior culture-based studies suggest that *Moraxella catarrhalis*, *Staphylococcus aureus*, and *S. pneumoniae* are the predominant species from these genera within the nasopharynxes of children.^{17,18} The observed associations between

the relative abundances of *Moraxella* and *Staphylococcus* and detection of *S. pneumoniae* by species-specific PCR are consistent with previous studies that examined density interactions between these bacteria in the upper respiratory tract.^{19–22} These studies generally reported a synergistic relationship between *S. pneumoniae* and *M. catarrhalis* and an antagonistic relationship between *S. pneumoniae* and *S. aureus*.^{19–23} A prior study that used high-throughput sequencing methods also reported an inverse association between *S. pneumoniae* and *Corynebacterium*, as was observed among children in our cohort.²⁴ Using a decision tree model, we also identify several other bacterial genera that accurately discriminate *S. pneumoniae*-colonized from noncolonized children. While these findings should be interpreted with caution given the cross-sectional study design, the associations suggest that these genera may modify the risk of pneumococcal colonization. The identification of bacteria that inhibit colonization by *S. pneumoniae* and other respiratory pathogens could form the basis for novel pneumonia prevention strategies, including use of a probiotic cocktail to selectively modify the nasopharyngeal microbiota such that it more effectively resists pathogen colonization.

We found that the biodiversity of nasopharyngeal microbial communities was similar in children with and without *S. pneumoniae* colonization. Given the cross-sectional design of this study, we cannot exclude the possibility that nasopharyngeal microbial diversity influences *S. pneumoniae* colonization risk, or that nasopharyngeal colonization by *S. pneumoniae* alters microbial diversity. Laufer et al. reported that culture-identified *S. pneumoniae* colonization was associated with lower microbial diversity among children with and without acute otitis media in the United States.²⁴ Notably, nasopharyngeal microbial diversity in this study was substantially higher than in our cohort, which could be related to differences in age, socioeconomic status, or genetic and environmental factors between these populations.²⁴ Conversely, higher nasopharyngeal microbial diversity was associated with the establishment of colonization in a pneumococcal challenge study conducted in adults, breaking with the traditional dogma that a higher diversity microbial community is more effective in resisting colonization by exogenous pathogens.²⁵ We may not have found an association between *S. pneumoniae* colonization and nasopharyngeal microbial diversity because colonization is associated with only a transient loss of diversity, or because this effect is strain-dependent. In a mouse model, pneumococcal challenge led to a reduction in microbial diversity, but diversity had returned to baseline within one week.²⁶ Moreover, in the previously described pneumococcal challenge study conducted in adults, microbial diversity did not change following challenge with a serotype 6B strain, while challenge with a serotype 23F strain actually transiently increased diversity.²⁵

S. pneumoniae colonization was associated with upper respiratory symptoms in our cohort, even after adjusting for respiratory virus infection and other measured confounding factors. This raises the possibility that *S. pneumoniae* colonization may be a symptomatic event in some infants and young children. Although historically upper respiratory symptoms have been attributed to respiratory viruses, two prior studies in children also suggest a potential role for *S. pneumoniae*. In a longitudinal study of 234 Australian children, a nasopharyngeal microbiota profile dominated by *Streptococcus* was associated with acute respiratory infection symptoms independent of respiratory virus infection.²⁷ Moreover, in a study of 2,840 children in Kenya, *S. pneumoniae* colonization was associated with upper respiratory

symptoms in the preceding two weeks, although testing for respiratory viruses was not performed.²⁸ The potential mechanisms for this association have not been defined, but mild nasal mucosal inflammation is observed following exposure to *S. pneumoniae*. In a study conducted among adults in the United States, experimental pneumococcal challenge resulted in recruitment of neutrophils and mononuclear cells to the nasal mucosa in some subjects.²⁹ Moreover, it is possible that *S. pneumoniae* colonization results in upper respiratory symptoms through alterations of the nasopharyngeal microbiota. In particular, we found that the presence of upper respiratory symptoms was associated not only with enrichment of the nasopharyngeal microbiota with *Streptococcus* but also with a lower relative abundance of *Corynebacterium*, a genus that includes common nasopharyngeal commensals.

Respiratory viruses are postulated to play a key role in the pathogenesis of *S. pneumoniae* infections.²⁰ Although the synergistic interactions between influenza viruses and *S. pneumoniae* are perhaps the most well described, *S. pneumoniae* colonization has been associated with infection by most common respiratory viruses, including rhinoviruses, respiratory syncytial virus, and parainfluenza viruses.²⁰ One mechanism by which respiratory viruses have been proposed to influence the risk of *S. pneumoniae* infection is through alteration of the composition of the nasopharyngeal microbiota. In our cohort, respiratory virus infection was associated with a higher relative abundance of *Streptococcus* but was not associated with the relative abundances of other highly abundant genera. Prior studies evaluating the impact of respiratory virus infection on nasopharyngeal microbial composition suggest that this effect may be virus-specific. In a study of 740 infants hospitalized for bronchiolitis, RSV and rhinovirus infections were associated with specific alterations of the nasopharyngeal microbiota.³⁰ Similarly, higher relative abundances of *Streptococcus* and *Haemophilus* and a lower relative abundance of *Staphylococcus* were observed in a cohort of previously healthy children with RSV infection.³¹ Finally, among pneumonia patients, influenza A (H1N1) virus infection was associated with enrichment of the oropharyngeal microbiota by *Pseudomonas*, *Bacillus*, and *Ralstonia* species possessing certain signaling and motility genes.³²

Our study has several limitations. First, given the cross-sectional design, we were unable to assess whether the alterations of the nasopharyngeal microbiota observed in children with *S. pneumoniae* predisposed to or were the result of pneumococcal colonization. Moreover, because the median duration of pneumococcal colonization may exceed one month in infants and young children, we are unable to determine the timing of *S. pneumoniae* acquisition relative to nasopharyngeal sampling in the study population.³³ In addition, the method for 16S rRNA gene sequencing that we used did not permit determination of bacterial composition at the species level. Alternative approaches, such as shotgun sequencing, could provide a more detailed understanding of bacterial density interactions in the upper respiratory tract. Although we tested for the most common respiratory viruses in children, it remains possible that the association between *S. pneumoniae* colonization and upper respiratory symptoms is related to undetected respiratory virus infection or other unmeasured factors. Moreover, given that the majority of viral infections in our cohort were rhinoviruses, we were unable to evaluate for associations between infection with other respiratory viruses and the composition of the nasopharyngeal microbiota. Finally, we excluded *Streptococcus* and unclassified *Streptococcaceae* from the microbial decision tree

model. Although this was necessary to ensure that *S. pneumoniae* was not assigned to genera within the model, the exclusion of non-pneumococcal streptococci may have reduced the accuracy of this microbial decision tree model. In particular, previous studies observed an inverse relationship between *S. pneumoniae* colonization and the density of viridans group streptococci in the nasopharynx.^{18,38}

In conclusion, pneumococcal colonization was associated with characteristic alterations of the nasopharyngeal microbiota of children that reflect synergistic and antagonist interactions of *S. pneumoniae* with commensal bacteria and other potential respiratory pathogens. Prospective studies are needed to determine if the nasopharyngeal microbiota provides a barrier to *S. pneumoniae* colonization and thus could be modified as a novel pneumonia prevention strategy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We would like to thank Copan Italia (Brescia, Italy) for their generous donation of the universal transport media and flocked swabs used for the collection of nasopharyngeal specimens. We offer our sincere thanks to the children and families who participated in this study.

Source of Funding: This research was supported by an Early Career Award from the Thrasher Research Fund (to MSK), a Burroughs Wellcome / American Society of Tropical Medicine and Hygiene Postdoctoral Fellowship in Tropical Infectious Diseases (to MSK), by Children's Hospital of Philadelphia (to APS, KAF) and Pincus Family Foundation, and through core services from the Penn Center for AIDS Research, a National Institutes of Health (NIH)-funded program (P30-AI045008). Funding for this project was also made possible in part by a CIPHER grant (to MSK) from the International AIDS Society, supported by ViiV Healthcare. The views expressed in this publication do not necessarily reflect the official policies of the International AIDS Society or ViiV Healthcare. MSK and CKC received financial support from the NIH through the Duke Center for AIDS Research (P30-AI064518). MSK was supported by a NIH Career Development Award (K23-AI135090) and a research grant from the Society for Pediatric Research (2018–2). JMP was supported by a Hamilton Health Sciences Early Career Award. PCS received funding from the NIH through a Research Project Grant (7R01-GM108494).

REFERENCES

1. World Health Organization. Estimated Hib and pneumococcal deaths for children under 5 years of age, 2008 Available at: http://www.who.int/immunization/monitoring_surveillance/burden/estimates. Accessed April 26, 2017.
2. Bogaert D, de Groot R, Hermans P. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. *Lancet Infect Dis* 2004;4(3):144–154. [PubMed: 14998500]
3. Wolter N, Tempia S, Cohen C, et al. High nasopharyngeal pneumococcal density, increased by viral coinfection, is associated with invasive pneumococcal pneumonia. *J Infect Dis* 2014;210(10):1649–1657. [PubMed: 24907383]
4. Garcia-Rodriguez JA, Fresnadillo Martinez MJ. Dynamics of nasopharyngeal colonization by potential respiratory pathogens. *J Antimicrob Chemother* 2002;50 Suppl S2:59–73.
5. de Steenhuijsen Piers WA, Sanders EA, Bogaert D. The role of the local microbial ecosystem in respiratory health and disease. *Philos Trans R Soc Lond B Biol Sci* 2015;370(1675).
6. United Nations Children's Fund. Under-five mortality.. Available at: <https://data.unicef.org/topic/child-survival/under-five-mortality>. Accessed June 1, 2018.
7. Kelly MS, Smieja M, Luinstra K, et al. Association of respiratory viruses with outcomes of severe childhood pneumonia in Botswana. *PLoS One* 2015;10(5):e0126593. [PubMed: 25973924]

8. World Health Organization. Pocket book of hospital care for children: guidelines for the management of common childhood illnesses Geneva: WHO; 2013.
9. Whelan FJ, Verschoor CP, Stearns JC, et al. The loss of topography in the microbial communities of the upper respiratory tract in the elderly. *Ann Am Thorac Soc* 2014;11(4):513–521. [PubMed: 24601676]
10. Bartram AK, Lynch MD, Stearns JC, et al. Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end Illumina reads. *Appl Environ Microbiol* 2011;77(11):3846–3852. [PubMed: 21460107]
11. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS One* 2013;8(4):e61217. [PubMed: 23630581]
12. Paulson JN. metagenomeSeq: Statistical analysis for sparse high-throughput sequencing. Bioconductor package 2014.
13. Gotelli NJ, Colwell RK. Estimating species richness. Magurran AE, McGill BJ. *Biological Diversity: Frontiers in Measurement and Assessment* Oxford, U.K.: Oxford University Press; 2011: 39–54.
14. Magurran AE. *Measuring Biological Diversity* 2nd ed. Oxford, U.K: Blackwell Science Ltd; 2004.
15. Kelly MS, Surette MG, Smieja M, et al. The nasopharyngeal microbiota of children with respiratory infections in Botswana. *Ped Infect Dis J* 2017;36(9):e211–e218.
16. Kuhn M, Wing J, Weston S, et al. caret: Classification and regression training. R package version 6.0–21. CRAN, VIENNA, Austria 2015.
17. Gehanno P, Lenoir G, Barry B, et al. Evaluation of nasopharyngeal cultures for bacteriologic assessment of acute otitis media in children. *Pediatr Infect Dis J* 1996;15(4):329–332. [PubMed: 8866802]
18. Faden H, Stanievich J, Brodsky L, et al. Changes in nasopharyngeal flora during otitis media of childhood. *Pediatr Infect Dis J* 1990;9(9):623–626. [PubMed: 2122410]
19. Chien YW, Vidal JE, Grijalva CG, et al. Density interactions among *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Staphylococcus aureus* in the nasopharynx of young Peruvian children. *Pediatr Infect Dis J* 2013;32(1):72–77. [PubMed: 22935873]
20. Bosch AA, Biesbroek G, Trzcinski K, et al. Viral and bacterial interactions in the upper respiratory tract. *PLoS Pathog* 2013;9(1):e1003057. [PubMed: 23326226]
21. Jacoby P, Watson K, Bowman J, et al. Modelling the co-occurrence of *Streptococcus pneumoniae* with other bacterial and viral pathogens in the upper respiratory tract. *Vaccine* 2007;25(13):2458–2464. [PubMed: 17030494]
22. Bogaert D, van Belkum A, Sluijter M, et al. Colonisation by *Streptococcus pneumoniae* and *Staphylococcus aureus* in healthy children. *Lancet* 2004;363(9424):1871–1872. [PubMed: 15183627]
23. Pericone CD, Overweg K, Hermans PW, et al. Inhibitory and bactericidal effects of hydrogen peroxide production by *Streptococcus pneumoniae* on other inhabitants of the upper respiratory tract. *Infect Immun* 2000;68(7):3990–3997. [PubMed: 10858213]
24. Laufer AS, Metlay JP, Gent JF, et al. Microbial communities of the upper respiratory tract and otitis media in children. *MBio* 2011;2(1):e00245–10.
25. Cremers AJ, Zomer AL, Gritzfeld JF, et al. The adult nasopharyngeal microbiome as a determinant of pneumococcal acquisition. *Microbiome* 2014;2:44. [PubMed: 25671106]
26. Krone CL, Biesbroek G, Trzci ski K, et al. Respiratory microbiota dynamics following *Streptococcus pneumoniae* acquisition in young and elderly mice. *Infect Immun* 2014;82(4):1725–1731. [PubMed: 24516113]
27. Teo SM, Mok D, Pham K, et al. The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development. *Cell Host Microbe* 2015;17(5):704–715. [PubMed: 25865368]
28. Abdullahi O, Karani A, Tigoi CC, et al. The prevalence and risk factors for pneumococcal colonization of the nasopharynx among children in Kilifi District, Kenya. *PLoS One* 2012;7(2):e30787. [PubMed: 22363489]

29. Wright AK, Ferreira DM, Gritzfeld JF, et al. Human nasal challenge with *Streptococcus pneumoniae* is immunising in the absence of carriage. *PLoS Pathog* 2012;8(4):e1002622. [PubMed: 22496648]
30. Mansbach JM, Hasegawa K, Henke DM, et al. Respiratory syncytial virus and rhinovirus severe bronchiolitis are associated with distinct nasopharyngeal microbiota. *J Allergy Clin Immunol* 2016;137(6):1909–1913. [PubMed: 27061249]
31. de Steenhuijsen Piters WA, Heinonen S, Hasrat R, et al. Nasopharyngeal microbiota, host transcriptome, and disease severity in children with respiratory syncytial virus infection. *Am J Respir Crit Care Med* 2016;194(9):1104–1115. [PubMed: 27135599]
32. Leung RK, Zhou JW, Guan W, et al. Modulation of potential respiratory pathogens by pH1N1 viral infection. *Clin Microbiol Infect* 2013;19(10):930–935. [PubMed: 23167452]
33. Heinsbroek E, Tafatatha T, Chisambo C, et al. Pneumococcal acquisition among infants exposed to HIV in rural Malawi: a longitudinal household study. *Am J Epidemiol* 2016;183(1):70–78. [PubMed: 26628514]
34. Johanson WG, Jr., Blackstock R, Pierce AK, et al. The role of bacterial antagonism in pneumococcal colonization of the human pharynx. *J Lab Clin Med* 1970;75(6):946–952. [PubMed: 4393022]

Figure 1. Principal Component Analysis plot of Bray-Curtis distances.

Each dot corresponds to a single nasopharyngeal sample. The ellipses define the regions containing 90% of all samples that can be drawn from the underlying multivariate distribution. Clustering of nasopharyngeal samples by *S. pneumoniae* colonization status is observed, and the clusters are significantly different in analyses adjusting for age (PERMANOVA; $P < 0.001$).

Figure 2. Relative abundances of highly abundant genera by detection of *S. pneumoniae* and respiratory viruses.

Bar plots depicting the mean relative abundances of genera in nasopharyngeal swab samples by *S. pneumoniae* colonization status and respiratory virus detection. Genera with a mean relative abundance of 0.5% in the overall population are shown.

Figure 3. Relative abundances of highly abundant genera by *S. pneumoniae* colonization status. Each dot represents the relative abundance of the genus in a single nasopharyngeal sample. The rectangles depict the lower quartile, median, and upper quartile for these relative abundances in children testing negative or positive for *S. pneumoniae* by PCR. Genera with a mean relative abundance of 0.5% in the overall population are shown. *P* values were estimated using multivariable logistic regression to evaluate for associations between the relative abundance of each genus and *S. pneumoniae* colonization adjusting for age.

Figure 4. Decision tree model for the prediction of *S. pneumoniae* colonization status. Variables of importance for the prediction of *S. pneumoniae* colonization are shown from a random forest model containing non-streptococcal bacterial genera. This microbial decision tree model had high sensitivity and specificity for the identification of *S. pneumoniae* colonization.

Table 1.Characteristics of the study population by detection of *Streptococcus pneumoniae* by PCR.

	<i>S. pneumoniae</i> PCR				
	Negative (n=74)		Positive (n=96)		<i>P</i> ^a
Demographics					
Mean (SD) age, months	6.6	4.8	9.6	5.1	0.0001
Sex					>0.99
Female	37	50%	49	51%	
Male	37	50%	47	49%	
Season					>0.99
Dry (April – October)	47	64%	62	65%	
Rainy (November – March)	27	36%	34	35%	
Enrollment year					0.21
2012	5	7%	11	11%	
2013	39	53%	45	47%	
2014	22	30%	19	20%	
2015	5	7%	14	15%	
2016	3	4%	7	7%	
Low birthweight (<2500 grams)	14	19%	11	11%	0.25
Maternal HIV infection ^b	13	19%	19	22%	0.79
Current breastfeeding	42	57%	51	53%	0.75
Socioeconomic factors					
Maternal educational level					0.35
None or primary	1	1%	5	5%	
Secondary	48	65%	64	67%	
Tertiary	25	34%	27	28%	
Electricity in the home	57	77%	57	59%	0.02
Household use of wood as a cooking fuel	7	9%	29	30%	0.002
Mean (SD) household members	5.3	2.9	5.4	2.4	0.85
Clinical factors					
Upper respiratory symptoms ^c	24	32%	58	60%	0.001
Pneumococcal conjugate vaccine (PCV-13)					0.90
0 doses	23	31%	30	31%	
1 dose	11	15%	12	13%	
2 doses	40	54%	54	56%	
Receipt of antibiotics within the prior 7 days	4	5%	5	5%	>0.99
Detection of one or more respiratory viruses	23	31%	37	39%	0.40
Rhinoviruses/enteroviruses	19		27		

	<i>S. pneumoniae</i> PCR				
	Negative (n=74)		Positive (n=96)		<i>P</i> ^a
Coronaviruses	3		5		
Other viruses	1		4		
>1 virus	0		1		

PCR, polymerase chain reaction; SD, standard deviation

^a *P* values were calculated using Chi-square or Fisher's exact tests for categorical variables and two-sample t tests for continuous variables

^b Data on maternal HIV infection status were missing from n=14 children

^c Defined as "cough or nasal congestion or discharge"