

https://dspace.org/

# Identification of virulence genes from non-typeable haemophilus influenzae isolates recovered from pneumonia cases and community controls

**Boitshoko Mahlangu, Boitshoko**

**2020**
https://hdl.handle.net/10539/29999

Downloaded from a DSpace repository. https://dspace.org/

# Identification of virulence genes from non-typeable *Haemophilus influenzae* isolates recovered from pneumonia cases and community controls

Boitshoko Mahlangu

Student no: 1408533

A Dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Medicine.

Johannesburg, 2020

## Declaration

I Boitshoko Mahlangu declare that this Dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science in Medicine at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

B. Mahlangu

Boitshoko Mahlangu

12th day of June 2020 in Johannesburg

## Abstract

Non-typeable *Haemophilus influenzae* (NTHi) causes severe upper (URTI) and lower respiratory tract infections (LRTI). Characterising virulence factors of NTHi associated with LRTI, could contribute to identifying potential vaccine epitopes. Our objective was to characterise the prevalence of select NTHi putative virulence factor genes in cultured strains obtained from induced sputum samples of children hospitalised with pneumonia compared to NTHi strains colonising the nasopharynx (NP) of otherwise healthy age group matched community controls (including children with URTI).

NTHi samples were collected as part of the Pneumonia Etiology Research for Child Health study conducted between 2011 and 2013 and retrospectively analysed. Cases were defined as children aged 1-59 months hospitalised with World Health Organisation defined severe or very severe pneumonia, and controls were children living within the same study catchment area as the cases without signs and symptoms of severe or very severe pneumonia. *H. influenzae* isolates underwent standard microbiological tests and were classified as NTHi by PCR using the *bexB* assay. Additionally, a novel high-throughput genotyping assay was developed and optimised (BioMark-HD system) in order to detect twelve NTHi putative virulence factors.

Overall, the study included 113 and 298 NTHi isolates from cases and controls, respectively. NTHi cases were younger than controls (10.1 months vs. 14.2 months; P=0.021) and out of the 12 analysed genes for NTHi proteins, two were significantly higher in cases than controls; i.e. *hmwC (*58% vs. 39%; P=0.002) and *hmw2A* (18% vs. 5%; P<0.001). We identified two combinations of genes that were more common in cases than controls, namely *hmwC+/infA+/hmw2A+/tehB+/lic2C+* and *hmwC+/infA+/hmw2A-/tehB+/lic2C+*.

The potential of *hmwC* and *hmw2A* with or without other putative virulence proteins of NTHi warrant further investigation as potential vaccine targets for the prevention of severe pneumonia due to NTHi.

## Acknowledgements

Pneumonia Etiology Research for Child Health (PERCH) teams and study participants.

Respiratory and Meningeal Pathogens Research Unit (RMPRU) students and staff.

Prof Shabir Madhi, Dr Elize de Bruyn, Prof Marta Nunes, Prof Jeffrey Dorfman and Dr Courtney Olwagen – Thank you for the academic support that you provided throughout my studies.

Dr Vicky Baillie – Thank you for your patience, understanding and support. I know it was as difficult for you as it was for me.

## Research Funding

This work is based upon research supported by the Department of Science and Technology/National research Foundation: South African Research Chair Initiative: Vaccine Preventable Diseases.

## Table of Contents


Declaration ..... ii
Abstract ..... iii
Acknowledgements ..... iv
Research Funding ..... iv
Table of Contents ..... v
List of Figures ..... vii
List of Tables ..... viii
Nomenclature ..... x
1. Introduction ..... 1
1.1. Pneumonia ..... 1
*1.2. Haemophilus influenzae* ..... 7
*1.3.* Structure and classification of *H. influenzae* ..... 7
1.3.1. Typeable *H. influenzae* ..... 8
1.3.2. NTHi ..... 9
1.4. Pathogenesis ..... 10
1.5. Virulence factors of NTHi as potential targets for vaccine development ..... 11
1.6. Treatment for *H. influenzae* ..... 13
1.7. Hib Vaccine ..... 15
1.7.1. NTHi vaccine development ..... 22
1.8. Nanofluidic real-time polymerase chain reaction (PCR) (Fluidigm®) ..... 26
1.9. Pneumonia Etiology Research for Child Health (PERCH) ..... 26
1.10. Study Justification ..... 27
1.11. Objectives ..... 28
2. METHODS ..... 29
2.1. Study population ..... 29
2.2. Study Samples ..... 29
2.3. Real-time PCR detection for *H. influenzae* and other common respiratory pathogens ..... 30
2.4. DNA extraction ..... 31
2.5. Culture identification of NTHi ..... 31
2.6. Classification of *H. influenzae* ..... 31
2.6.1. Primer optimisation, standard curves and quantification of the real-time PCR assays ..... 32
2.7. Nanofluidic real-time PCR (Fluidigm®) ..... 33
2.7.1. Pre-amplification of DNA ..... 37

2.7.2. Real-time qPCR using the Biomark$^{TM}$ HD System (Fluidigm) .....37
2.7.3. Primer optimisation, Standard curves and quantification in the nanofluidic real-time PCR system. .....38
2.8. Statistics .....38
2.9. Ethics .....39
3. Results .....40
3.1. Optimisation of the *bexB* and *hpd* PCR primers .....40
3.2. Optimisation for the Fluidigm PCR .....40
3.3. Demographics and clinical characteristics of controls by *H. influenzae* status 43
3.4. Demographics and clinical features of the cases according to NTHi status by IS Culture .....48
3.5. Demographics and clinical characteristics of NTHi colonised cases and controls by culture .....55
3.6. Detection of virulence genes among NTHi positive cases and controls .....57
4. Discussion .....60
5. References .....63
6. Apendices ..... **Error! Bookmark not defined.**
6.1. Apendix A: HREC certificate .....76

## List of Figures


Figure 1.1 The basic structure of capsulated *H. influenzae* ..... 7

Figure 1.2 H. influenzae capsule locus ..... 8

Figure 1.3 Attachment of NTHi to epithelial cells and invasion of NTHi to host cells ..... 9

Figure 1.4 Graphical representation of the β-lactam action on Gram-negative bacterium ..... 10

Figure 1.5 β-lactamase action on the β-lactam ring ..... 11

Figure 1.6 Graphical representation of countries using Hib vaccines in 1997 and 2016 ..... 13

Figure 1.7 Flex Six microfluidic chip ..... 23

Figure 3.1 Real-time PCR analysis of the *hmwC* primers ..... 40

Figure 3.2 Real-time PCR analysis of the *igA1* primers ..... 40

Figure 3.3 Real-time PCR analysis of the *infA* primers ..... 40

Figure 3.4 Real-time PCR analysis of the *lic2C* primers ..... 40

Figure 3.5 Real-time PCR analysis of the *oapA* primers ..... 41

Figure 3.6 Real-time PCR analysis of the *pilA* primers ..... 41

Figure 3.7 Real-time PCR analysis of the *hmw1A* primers ..... 41

Figure 3.8 Real-time PCR analysis of the *hmw2A* primers ..... 41

Figure 3.9 Real-time PCR analysis of the *losB* primers ..... 41

Figure 3.10 Real-time PCR analysis of the *msfA1* primers ..... 42

Figure 3.11 Real-time PCR analysis of the *tehB* primers ..... 42

Figure 3.12 Real-time PCR analysis of the *hia* primers ..... 42

Figure 3.13 Virulence factors detected among NTHi positive cases (n=83) and controls (n=273) ..... 56

## List of Tables


Table 1.1 Review of studies reporting on the aetiology of bacterial pneumonia in African Countries since 2000 ..... 3

Table 1.2 Summary of potential virulence factors for *H. influenzae* ..... 11

Table 1.3 A review of articles on NTHi as a replacement for Hib ..... 20

Table 1.4 Studies on *H. influenzae* virulence factors ..... 14

Table 2.1 Primers used for Fluidigm ..... 29

Table 3.1 Performance of the Real-Time PCR ..... 39

Table 3.2 Demographics and clinical features of the community controls according to *H. influenzae* status and samples that did not have enough volume ..... 44

Table 3.3 Demographics and clinical features of the community controls according to *H. influenzae* capsule status ..... 45

Table 3.4 Demographics and clinical features of the community controls according to NTHi NP Culture status ..... 46

Table 3.5 Demographics and clinical features of the cases according to *H. influenzae* status and samples that did not have enough volume ..... 48

Table 3.6 Demographics and clinical features of the cases according to *H. influenzae* capsule status ..... 50

Table 3.7 Demographics and clinical features of the cases according to NTHi status by IS Culture ..... 52

Table 3.8 Prevalence of cases and controls in whom NTHi was cultured on IS and NP/OP swabs, respectively according to age ..... 54

Table 3.9 Demographics and clinical characteristics of cases and controls in whom NTHi was cultured on IS and NP/OP swabs, respectively ..... 55

Table 3.10 Different types of combinations detected among NTHi

cases and NTHi controls..... 60

## Nomenclature

% - Percentage

°C – degrees Celsius

BLAST – Basic Local Alignment Search Tool

CHBAH – Chris Hani Baragwanath Academic Hospital

DNA – Deoxyribonucleic acid

Hi – *Haemophilus influenzae*

HibCV – *H. influenzae* type B conjugate vaccine

HIV – Human immunodeficiency virus

HREC – Human Research Ethics Committee

IFC – Integrated fluidic circuit

IS – Induced sputum

LRTI – Lower respiratory tract infection

NTHi – Nontypeable *H. influenzae*

NP – Nasopharyngeal

PCR – Polymerase chain reaction

PCV – Pneumococcal conjugate vaccine

PERCH – Pneumonia Etiology Research for Child Health

qPCR – Quantitative polymerase chain reaction

RMPRU – Respiratory and Meningeal Pathogens Research Unit

STGG – Skim-milk, tryptone, glucose and glycerine broth

URTI – Upper respiratory tract infection

# 1. Introduction

## 1.1. Pneumonia

Pneumonia is one of the most common, life-threatening childhood illnesses with approximately 120 million cases annually, the majority of which occur in developing countries (1–3). In 2015, pneumonia was the second leading cause of death in children under-5 years of age, accounting for 15.5 % (Uncertainty Range [UR]: 13.9-17.6) of the 5, 942 million total deaths in this age group (4). The global estimates for pneumonia incidence in low- to middle-income countries was 0.23 episodes per child-year, 11.5% of which were severe pneumonia (5). In a study conducted on a South African cohort, the pneumonia incidence is estimated at 0.32 episodes per child-year during infancy (6). Another study found a pneumonia incidence of 0.27 episodes per child-year and severe pneumonia incidence of 0.05 per child-year and a case fatality ratio of 1% (7) In 2016, pneumonia was attributed as the cause of 19,638 deaths in South African children (8).

Determining the aetiology of pneumonia is challenging as direct sampling of the site of infection, i.e. lung, is difficult with potential risk to the patient (9). Furthermore, organisms associated with causing pneumonia commonly colonise the upper airway of healthy individuals, thus limiting the use of sampling of the nasopharynx as a proxy for inferring causality of pneumonia (10,11). Additionally, blood culture which is often used for identifying bacterial causes of pneumonia is not sensitive for diagnosing bacterial pneumonia (1-27%) (12). Another challenge in the etiologic diagnosis of pneumonia in children is the difficulty in obtaining adequate specimen for pathogen detection (13).

The leading bacteria attributed as causes of pneumonia in children are *Streptococcus pneumoniae, Haemophilus influenzae* (encapsulated and non-encapsulated), *Moraxella catarrhalis* and *Staphylococcus aureus.* Between 2000 and 2018, there were 11 studies that investigated the aetiology of childhood pneumonia (0-10 years) in African countries (detailed in Table 1.1) which included three from South Africa (14–16), one each from Mali (17), Kenya (18), Malawi (19), Morocco (20), Niger (21), Senegal (22), Tanzania (23), and The Gambia (24). Seven studies collected nasopharyngeal (NP) swabs and used either blood culture or real-time

polymerase chain reaction (PCR) or both to infer pneumonia aetiology (7,14,17,18,20,21,23). Furthermore, two studies undertook lung aspirates (including analysing by culture and real-time PCR), and blood culture (15,19).

*S. pneumoniae* was the most prevalent bacteria (range: 26.7%-87.5%) detected in these studies (14,15,17–22,24) followed by *H. influenzae* (range: 1% - 54%) (14,19–22,24). *Moraxella catarrhalis* was the most prevalent bacteria (range: 21.7% - 75%) detected in two of the 11 studies (7,23). The *H. influenzae* type b conjugate vaccine (HibCV) was in routine use in 6/11 countries (7,17–19,21,24) at the time of the study, whereas the pneumococcal conjugate vaccines (PCV) were only in routine use in three of the eleven countries (7,17,20). Countries in which HibCV had not been introduced in the public immunisation program reported a high prevalence of *H. influenzae* (46% to 83.6%) as the cause of pneumonia (22,23). In countries where HibCV was included in the public immunisation program, the prevalence of *H. influenzae* as the attributed cause of pneumonia range from 1.5% to 6.8% (7,14,15,17–21,24).

Table 1.1: Review of studies reporting on the aetiology of bacterial pneumonia in African Countries since 2000

| **Country** | **Design** | **Case definition** | **Number of participants** | **Study population** | **Organism isolated** | **Types of samples (techniques)** | **Use of HibCV and PCV** | **Study duration** |
|---|---|---|---|---|---|---|---|---|
| South Africa (15) | Prospective | Children meeting WHO criteria for severe and very severe pneumonia | 358 | 1 to 59 months (median age: 4.8 months) | *S. pneumoniae* (n=26, 7%)<br>*S. aureus* (n=16, 5%)<br>Other streptococci (n=12, 3%)<br>*H. influenzae* (n=2, 1%)<br>*K. pneumoniae* (n=2,1%) | Bronchoalveolar lavage fluid (culture) | HibCV | January 2001 to December 2002 |
| South Africa (14) | Prospective | HIV-infected children aged 8 weeks or older. | 203 | More than 8 weeks (median age: 21.6 months) | *S. pneumoniae* (n=48, 22%)<br>*H. influenzae* (n=47, 21.7%)<br>*M. catarrhalis (*n=47, 21.7%)<br>*S. aureus* (n=44,20.4%)<br>Enterobacteriaceae (n=32, 14.1%) | Nasopharyngeal swabs (culture) | HibCV | December 2002 |
| South Africa (7) | Nested case-control | Any episode of pneumonia regardless of severity, excluding congenital pneumonia (defined as presentation before postnatal discharge) | 967 (284 pneumonia cases 412 controls) | Less than 5 years (Median age: 11 months)<br>(Male: 53%) | ***Cases (n=967)***<br>*S. pneumoniae* (n=168, 60%)<br>*H. influenzae* (n=152, 54%)<br>Hib (n=4, 1%)<br>*M. catarrhalis* (n=214, 75%)<br>*S. aureus* (n=81, 28%)<br>**Controls (n=412)**<br>*S. pneumoniae* (n=237, 58%)<br>*H. influenzae* (n=164, 40%)<br>Hib (n=5, 1%)<br>*M. catarrhalis* (n=292, 71%)<br>*S. aureus* (n=142, 35%) | Nasopharyngeal swabs (culture and real-time PCR) | HibCV PCV13 | May 2012 to Dec 2014 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Mali (17) | Prospective case-control | Cases – children with radiologically-confirmed pneumonia. Controls – hospitalised children without respiratory features. Cases and controls matched for age and period. | 216 patients (118 cases and 98 controls) | Less than 5 years (median age:11 months) (Male: 55.9%) | **Cases (n=118)** *S. pneumoniae* (n=85, 72%) *S. aureus* (n=23, 19.5%) *H. influenzae* (n=8, 6.8%) **Controls (n=98)** *S. pneumoniae* (n=47, 48%) *S. aureus* (n=19, 19.4%) *H. influenzae* (n=6, 6.2%) | Nasopharyngeal swabs (culture and real-time PCR) | HibCV PCV13 | July 2011 – December 2012 |
| Kenya (18) | Prospective | Children meeting the WHO syndromic definition of severe pneumonia or very severe pneumonia. | 568 | 2 to 59 months (median age: 11 months) (male: 57%) | *S. pneumoniae* (n=15, 46.9%) *E. coli* (n=5, 15.6%) *S. aureus* (n=3, 9.4%) *H. influenzae* (n=1, 3.1%) *Acinetobacter spp* (1,3.1%) | Nasopharyngeal swabs (culture) | HibCV | May 2007 to May 2008 |
| Malawi (19) | Prospective observational | Children with a clinical diagnosis of pneumonia only, and only those with radiologically confirmed pneumonia with focal, segmental or lobar pneumonia. | 95 | Less than 10 years (median age 31.2 months) (male: 61%) | *S. pneumoniae* (n=37, 39%) *H. influenzae* (n=6, 6.3%) *S. typhimurium* (n=2, 2.1%) Viruses (n=24, 25.3%) | Lung aspirates (culture and real-time PCR) | HibCV | April 2004 to October 2006 |
| Morocco (20) | Prospective | Children meeting WHO criteria for severe pneumonia. | 700 (195 pneumonia cases) | 2 to 59 months (mean age: 21.5 months) (male: 64.1%) | *S. pneumoniae* (n=10, 5.2%) *H. influenzae* (n=3, 1.5%) Viruses (n=353, 51.5%) | Nasopharyngeal swabs (culture and real-time PCR) | HibCV PCV | November 2010 to December 2011 |
| Niger (21) | Retrospective | Cases - defined as a hospitalized child <5 years of age with onset of cough | 160 random stored NP samples | Less than 5 years (median age: 9 months) (Male: 54.4%) | *S. pneumoniae* (n=90, 56.3%) *H. influenzae* (n=20, 12.5%) | Nasopharyngeal swabs (real-time PCR) | HibCV | January 2010 to December 2012 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | or difficulty breathing within 7 days prior to admission, and at least one of the following danger signs: inability to drink or breastfeed, lethargy, vomiting everything, convulsions, nasal flaring, chest indrawing, stridor in a calm child or tachypnoea | | | *C. pneumoniae* (n=4, 2.5%)<br>*S. aureus* (n=18, 11.3%)<br>Any bacteria (n=102, 63.7%) | | | |
| Senegal (22) | Prospective | Children presenting to Ndioum's hospital paediatric service that did not have antimicrobial therapy in the past 2 weeks. | 114 | 2 to 59 months (mean age: 17.8 months) (male:54.4%) | *S. pneumoniae* (n=64, 56%)<br>*H. influenzae* (n=53, 46%) | Nasopharyngeal aspirates (culture) | None | March to May 2000 |
| Tanzania (25) | Prospective | Acute respiratory infection – any acute (≤1 week) infection manifested by at least one respiratory sign or symptom localized to the upper or lower respiratory tract and divided into two categories: clinical pneumonia or upper respiratory tract infection (URTI). | 177 clinical pneumonia cases | 2 months to 10 years (Male: 51.1%) | *S. pneumoniae* (n=155, 87.5%)<br>*H. influenzae* (n=148, 83.6%)<br>*M. catarrhalis* (n=170, 96%)<br>*N. meningitidis* (n= 82, 46.3%) | Nasopharyngeal swabs (real-time PCR) | None | April to August 2008<br>June to December 2008 |
| The Gambia (24) | | Severe pneumonia - cough or difficulty in breathing, plus any of the following: | 56 participants | 2 months to 5 years | *S. pneumoniae* (n=48, 91%)<br>*H. influenzae* (n=12, 23%) | | HibCV | 2007 to 2009 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | lower chest wall indrawing, nasal flaring, or an oxygen saturation of <90% on pulse oximetry. | | | *Klebsiella spp* (n=3, 6%)<br>*Streptococcus spp* (non-*pneumoniae*) (n=2, 4%)<br>Viruses (n=10, 17%) | | | |

Pneumonia remains a major cause of morbidity and mortality in children (3). Tuberculosis is increasingly recognised as occurring in acute pneumonia in high burden areas, including other factors such as malnutrition, HIV infection or exposure, tobacco smoke exposure and attending daycare (26).

## *1.2. Haemophilus influenzae*

*H. influenzae* is a Gram-negative coccobacillus bacterium identified in 1892 during the Influenza pandemic of 1889 to 1892 (27–29). Haemophilus means blood-loving, which explains the absolute nutritional requirement for haem (known as X-factor) during growth (30). *H. influenzae* also requires nicotinamide adenine dinucleotide (known as V-factor) for growth (30).

*H. influenzae* is an opportunistic organism that forms part of the human nasopharyngeal microbiome (10,27). Intercurrent respiratory viral infections could lead to new acquisition of bacteria and an increase in density in the nasopharynx, which could predispose to subsequent translocation into the lung and cause pneumonia (31). Encapsulated *H. influenzae*, and particularly Hib, also cause bacteraemia (32), epiglottitis (33) and meningitis in children (16). Transmission of *H. influenzae* occurs through respiratory droplet spread (34), and susceptibility to acquisition is determined by living conditions such as larger family size, household crowding, poor household ventilation and environmental smoke exposure (35).

## *1.3.* Structure and classification of *H. influenzae*

The *H. influenzae* cell envelope consists of the outer membrane (OM), the periplasmic space (PS) and the inner membrane (IM) (36). The inner layer of the outer membrane is made up of phospholipids while the outer layer is made up of lipo-oligosaccharides (LOS). The PS is made up of long polymers of repeating disaccharides with insertions of prostaglandins (PG). The IM is made up of a phospholipid bilayer. The cell envelope consists of soluble proteins, transmembrane proteins or anchored proteins (Figure 1.1) (37).

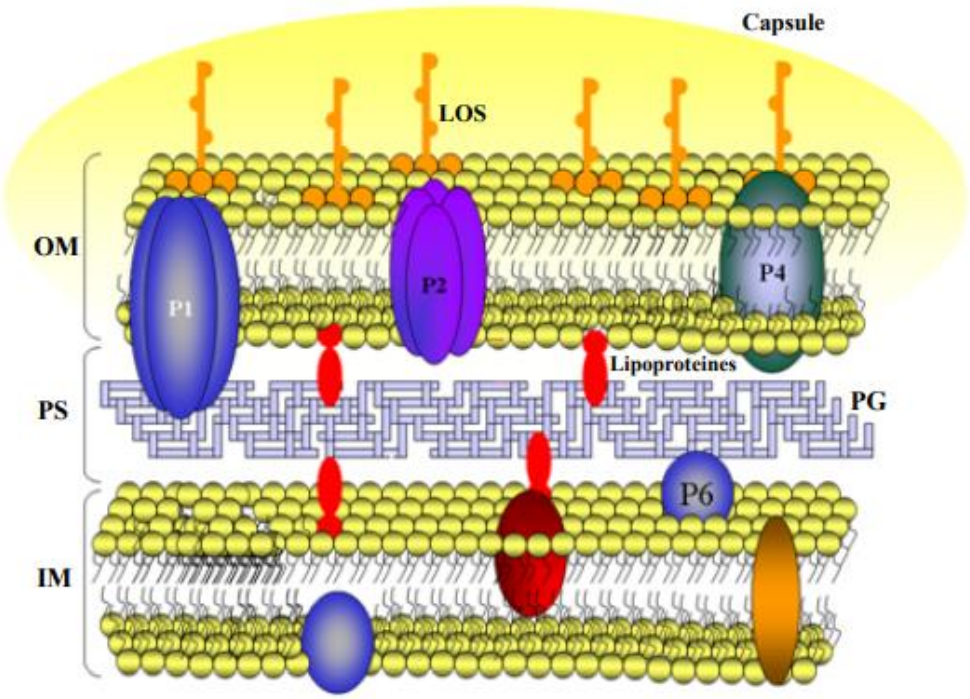



Figure 1.1: The basic structure of capsulated *H. influenzae* from Varela et al, (2012)(36).

*H. influenzae* is classified by the presence (encapsulated/typeable strains) or absence (unencapsulated/nontypeable strains (NTHi)) of the polysaccharide capsule (indicated in Figure 1.1) (27). The encapsulated strains are further classified by their unique antigenically distinct capsular polysaccharides (a-f) (38). Slide agglutination was the preferred method for classifying *H. influenzae* according to the different serotypes (a-f) until it was noted that some of the isolates were misidentified after using real-time PCR (39). A more reliable method for subtyping is real-time PCR which targets the capsule locus, particularly the genes *bexA* and *bexB* (Figure 1.2) (40).

### 1.3.1. Typeable *H. influenzae*

The polysaccharide capsule is encoded by the capsular (cap) gene which can be divided into three functionally distinct regions, designated I, II and III (Figure 1.2) (41). Regions I and III are common to all capsular types and contain highly conserved genes necessary for processing and transport of capsular material to the

cell surface (42). Region II contains type-specific genes involved in capsule biosynthesis which gives rise to the different antigenically distinct capsules (42). The most pathogenic strain is the encapsulated type b (Hib) strain which uses the capsule to evade the host's immune system by amplifying the capsule gene locus leading to decreased complement-mediated killing (43). The Hib capsule is made up of the polyribosyl-ribitol-phosphate (PRP) which makes Hib the most pathogenic serotype (44). Some strains of *H. influenzae,* including Hib, contain multiple copies of the cap locus which increases the virulence, and it has been shown that as the number of copies increase, the susceptibility to complement-mediated lysis decreases (45–47).

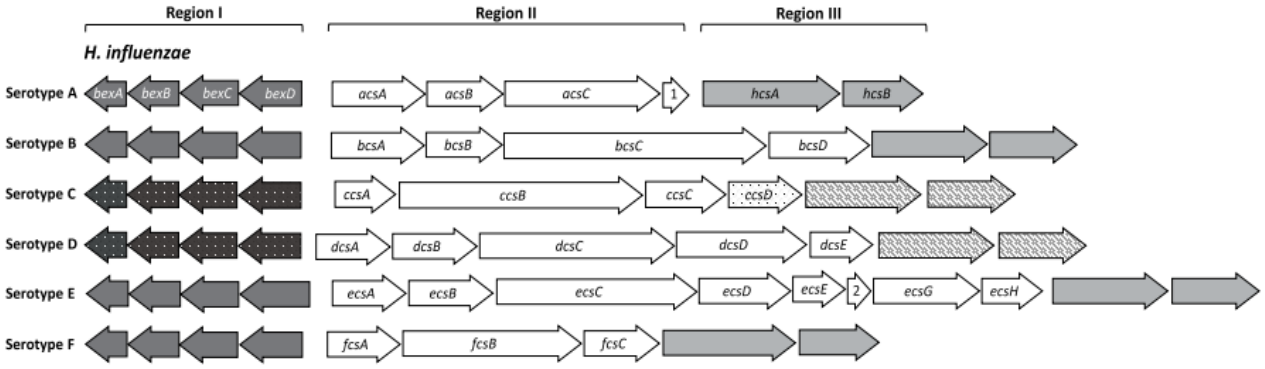



Figure 1.2: *H. influenzae* capsule locus from González-Díaz et al. (2019)(48).
Genes in dark grey belong to region 1. Genes in light grey belong to region III. The white genes indicate genes found in region II.

### 1.3.2. NTHi

The NTHi strains have a greater genetic diversity but are generally less invasive than the typeable strains, because of the absence of the capsule, NTHi are intracellular organisms to avoid complement-mediated lysis (49). Gilsdorf et al., (2004), proposed several mechanisms that might be responsible for the lack of the capsular polysaccharide by the NTHi strains, these include phase variation (slipped-strand mispairing, mediated by short DNA repeats); unstable mRNA (which results in transcription but not translation); point mutations (non-synonymous rather than synonymous (silent) mutations encountered) and insertions and deletions (addition or

removal of new DNA). All of these mechanisms allow NTHi to survive under various selective pressures, and to adapt to changing microenvironments (50).

## 1.4. Pathogenesis

*H. influenzae* has a number of surface exposed proteins that can be used to attach to epithelial cells (51) including the high molecular weight 1 and 2 (*hmw1/2*), type IV pilus (*pilA), H. influenzae* adhesin (*hia*) and the opacity associated protein A (*oapA)* (Figure 1.3) (52–55). These outer membrane proteins attach to host cell receptors (Figure 1.3) (53). Once inside the cell, *H. influenzae* uses proteins such as protein-D and macrophage survival factor (*msf)* to survive within host cells (56–58). In the host cells, *H.* influenzae escapes the phagosomes and live in the cytoplasm of the host cells or release factors that inhibit phagocytosis (59–61).

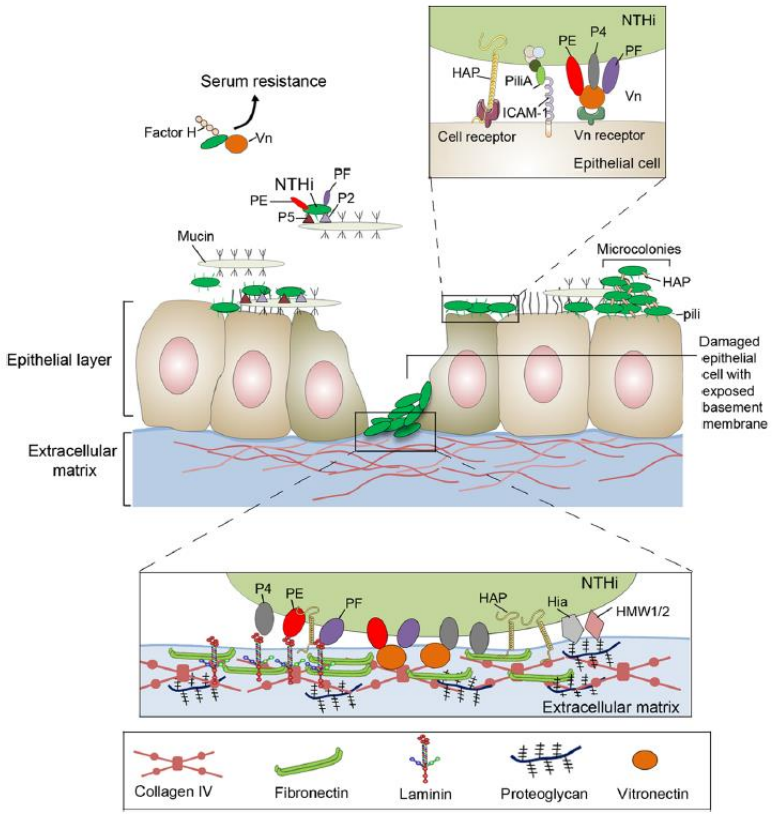



Figure 1.3: Attachment of NTHi to epithelial cells and invasion of NTHi to host cells from Duell et al, (2016)(53).

NTHi has the ability to bind to a variety of host proteins. NTHi attaches to the epithelial cell surfaces and the underlying extracellular matrix layer. The binding of NTHi to the cell surfaces allows NTHi to

strongly attach to the host cells to mediate colonisation or entry, or to provide defensive mechanisms such as evading the host immune response (53).

## 1.5. Virulence factors of NTHi as potential targets for vaccine development

NTHi lacks the main virulence factor of typeable strains, namely the capsule, but may have different combinations of virulence-related genes that allow the bacteria to evade the host's immune defences and cause disease (62). Analyses of strains recovered from children with pneumonia and from adults with chronic obstructive pulmonary disease (COPD) indicated that certain NTHi strains possess different repertoire or prevalence of virulence genes compared to colonising strains (61). Identification of the virulence genes associated with NTHi strains that cause disease could assist in identifying potential vaccine epitopes (63). Some of the genes which have been associated with NTHi virulence are summarised in Table 1.2 below:

Table 1.2: Summary of potential virulence factors for *H. influenzae*

| **Target gene** | **Functions** |
|---|---|
| High Molecular Weight 1/2 (*hmw1/2*) | HMW1 and HMW2 mediate the attachment of the bacterium to the human nasopharynx for colonisation which is an essential first step during infection (64). HMW1 and HMW2 are found in 75% of *H. influenzae* strains whereas the remaining mostly contain Hia which has a similar function to HMW1/2 (54). Factors that determine whether HMW1/2 or Hia will be present have not been identified yet but there is an assumption that HMW-deficient strains might have evolved from encapsulated isolates (65). There is no proof to suggest that strains that possess HMW1/2 are more virulent than Hia containing strains (65). |
| Adhesin auto-transporter protein *(Hia)* | |
| *Lipo-oligosaccharide (LOS)* | LOS forms a part of the cell wall (66). LOS is phase variable which is the reversible, high-frequency switching of gene expression thus allowing for the generation of a diverse population of phenotypically distinct bacteria (67). LOS contributes to antigenic diversity (allows the bacteria to alter its surface protein in order to avoid the host immune response), molecular mimicry |

| | |
|---|---|
| | (sequence similarities between bacterial and self-proteins allows the bacteria to evade host immune response), host-cell adherence (NTHi bind to receptors on host cells in order to gain entry into those cells) and resistance to complement-mediated killing (some configurations of LOS lead to resistance of complement-mediated killing) (53,67–72). |
| Protein D (*hpd)* | Protein D is an immunoglobulin D-binding membrane protein exposed on the bacterial surface. Plays a role in pathogenesis due to absence antigenic drift and surface localisation (73). |
| Immunoglobulin A1 protease (*igA1*) | An enzyme that cleaves the human IgA1 at its hinge region into its $F_c$ and V regions so that the immunoglobulin is unable to carry out clearance of the bacteria. Four different cleavage patterns have been discovered thus far. It is unknown whether any pattern is more pathogenic than the other (59); however, levels of IgA1 protease are higher in invasive strains compared to colonizing strains (60). *IgA1* also mediates invasion and trafficking in human respiratory epithelial cells, facilitating the persistence of *H. influenzae* (74). |
| Type IV pilus (*pilA*) | *pilA* is a major subunit for type IV pilus (Tfp) (filamentous polymers composed of helically arranged pilin subunits, that are assembled and expressed on the surface of numerous gram-negative pathogens) which has the ability to facilitate the uptake of DNA across the bacterial cell membrane and also mediate the adherence to mammalian epithelial cells (69,75). *PilA* has also been associated with the possibility of providing structural stability for biofilm formation which is believed to play a role in chronic infections (75). The whole pilus gene cluster consists of *pilA-D* and *ComA-F* (52). |
| Macrophage survival factor | Plays a role in the intracellular survival of *H. influenzae*, although the mechanism is not understood, research has shown that strains lacking SLR have shorter survival time compared to strains possessing SLR (57). Any *SlrVA* is present in 79% of isolates associated with disease (Kress-Bennett, *et al*., 2016). |

| | |
|---|---|
| Tellurite resistance protein | Plays a role in the utilization of low levels of exogenous haem to support aerobic growth (76). |
| Opacity associated protein A (*oapA*) | Plays a role in phase variation and adhesion to epithelial cells during initial colonization and cell invasion (77). oapA is also responsible for the transparent-colony phenotype and mediates attachment to epithelial cells (55). |

## 1.6. Treatment for *H. influenzae*

First line therapy involves the use of oral antibiotics such as beta (β)-lactams (27), which are a class of broad-spectrum antibiotics containing a β-lactam ring (Figure 1.5) (78). These antibiotics act by inhibiting the synthesis of the bacterial cell wall (79). β-lactams bind to the penicillin-binding proteins making up the bacterial cell wall thus inhibiting the formation of the cross-linking polymers thereby inactivating the inhibitor of autolytic enzymes that are responsible for bacterial lysis (Figure 1.4) (80,81).

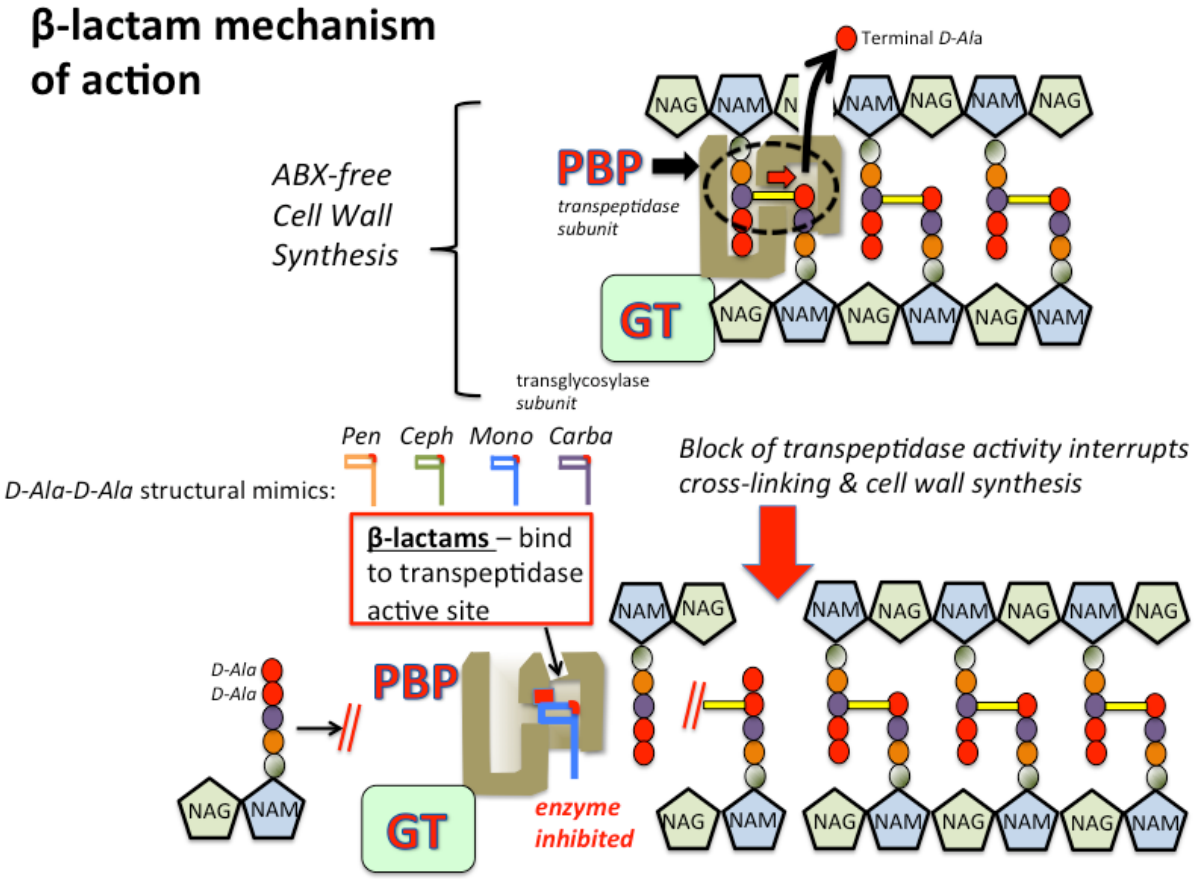



Figure 1.4: Graphical representation of the β-lactam action from Drusano et al, (2004)(80).

Top: In the absence of drug, transpeptidase enzymes (also known as Penicillin Binding Proteins; PBP) in the cell wall catalyse cross-links between adjacent glycan chains, which involves the removal of a terminal D-alanine residue from one of the peptidoglycan precursors (highlighted by the broken oval). Glycosyltransferases (GT), which exist as either separate subunits, or tightly associated with transpeptidases (e.g. as is the case for PBP-2) create covalent bonds between adjacent sugar molecules NAM & NAG. The net result of covalent bonds between both the peptide and sugar chains creates a rigid cell wall that protects the bacterial cell from osmotic forces that would otherwise result in cell rupture. Bottom: Beta-lactam antibiotics, bear a structural resemblance to the natural D-Ala-D-Ala substrate for the transpeptidase, and exert their inhibitory effects on cell wall synthesis by tightly binding to the active site of the transpeptidase (PBP) (80).

There is however, the emergence of drug-resistant typeable *H. influenzae*, and NTHi strains that produce a β-lactamase enzyme that breaks down the β-lactam ring of the antibiotic. Currently, between 10-25% of NTHi strains are β-lactamase-positive (42).

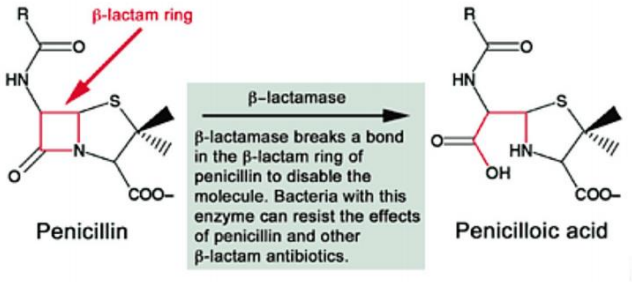



Figure 1.5: β-lactamase action on the β-lactam ring from Urumova et al, (2015)(82).

Antibiotics are commonly prescribed for respiratory tract infections, but the prevalence of resistance and their mechanisms keep increasing and becoming more complex (83). Thus alternative therapies need to be evaluated including extended-spectrum cephalosporins, tetracyclines, amoxicillin-clavulanic acid, trimethoprim-sulfamethoxazole, quinolones and macrolide antibiotics (27). Furthermore, the development of vaccines may play a role in the prevention of NTHi illness (63).

## 1.7. Hib Vaccine

The Hib vaccine was first licensed for use in 1985 in the United States of America for children under-5 years of age (84–86). This vaccine used PRP as an epitope and is only effective against Hib (87). This initial polysaccharide-only Hib vaccine was discontinued in 1988 due to eliciting only weak B-cell response and no memory responses; resulting in vaccine effectiveness of 35% to 42% after the third dose (87,88). The next generation of Hib vaccine was a conjugate of PRP to protein carriers, which elicited Th1-cell dominant immune responses (89), that also induced memory responses (90). In 2016, 190 countries had already introduced the Hib vaccine to their national immunisation programmes (Figure 1.6). In 2017, the coverage of the Hib conjugate vaccine (HibCV) in South Africa was 82% (91).

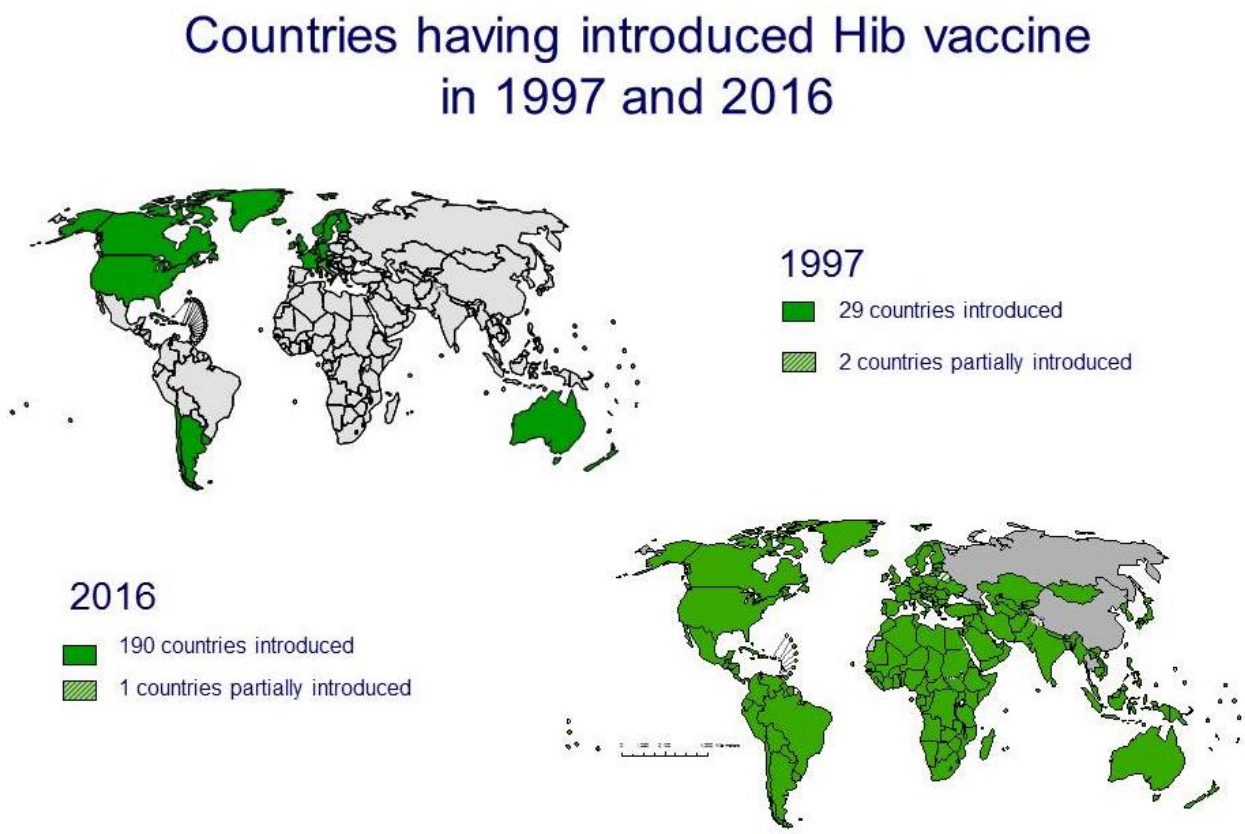



Figure 1.6: Graphical representation of countries using Hib vaccines in 1997 and 2016 by Loharikar et al, (2016) (91).

In 1997, 29 countries had already introduced the Hib vaccine into their national immunisation programmes and only two were in the process of introducing it. In 2016 a total of 190 countries have introduced the Hib vaccine into their national immunisation programmes.

In 1999, South Africa became the first African country to include the HibCV into its routine childhood immunisation programme (92). A study done in Soweto from 1996 to 2000 reported a reduction of invasive Hib disease from 25 to 1 per 100,000 children following the introduction of the HibCV (92). Furthermore, since the introduction of the HibCV in South Africa, Hib cases have decreased by 65% from 55 (1999-2000) to 19 cases (2003-2004) in children less than 1 year of age (86).

Langereis et al (2015) (93), tried to sum up the possible explanations for the replacement of Hib by NTHi which include (1) vaccine-mediated strain replacement, (2) improved bacterial detection and serotyping (culture takes long before an organism can be isolated but techniques such as PCR are rapid and more sensitive), (3) increased virulence of NTHi strains (possibly NTHi might have acquired new

virulence genes) and (4) demographic changes (*H. influenzae* disease was mainly a paediatric disease caused by Hib but now is caused by other strains such as Hia, Hif and NTHi and affects the elderly).

Sixteen studies from low-middle income countries that introduced the HibCV investigated the strains of *H. influenzae* identified post-vaccine introduction; Table 1.3. There were three studies from Italy (94–96), two from Japan (97,98), and Australia (99,100) and one from Turkey (101), The Gambia (102), Kenya (103), Central Asia (104), Korea (105), China (106), Brazil (107), Spain (108) and Israel (109). Twelve studies collected NP swabs and employed either blood culture or real-time PCR or both (94–96,101–109). While four studies collected blood and cerebrospinal fluid (CSF) and employed slide agglutination serotyping (SAST) and real-time PCR (97–100). Studies that enrolled only healthy children (6/16) found that NTHi NP colonisation was prevalent in 70% to 100% of participants compared to <1% prevalence of Hib colonisation (94,94,102,103,107,109). In children with respiratory tract infections (10/16), NTHi colonisation was prevalent in 50% to 100% of individuals, whereas Hib colonisation prevalence range from 0% to 11% (95,97–101,104–106,108).

Table 1.3: A review of articles on NTHi as a replacement for Hib

| Country | Design | Case definition | Number of participants | Study participants | Prevalence of Hi species identified | Types of samples (technique) | Study period |
|---|---|---|---|---|---|---|---|
| Italy (96) | Prospective | Healthy children | 717 | Less than 6 years<br>(median age: 11.6 months)<br>(male:55.3%) | *H. influenzae* (n=101, 14.1%)<br>NTHi (n=98, 97%)<br>Hie (n=2, 1.98%)<br>Hif (n=1, 0.99%) | OP swabs (real-time PCR) | June 2012 to July 2013 |
| Italy (94) | Prospective | Healthy children | 301 | Less than 6 years<br>(median age: 31.61 months)<br>(male: did not specify) | *H. influenzae* (n=130, 43.18%)<br>NTHi (100%) | NP swabs (culture and real-time PCR) | January to April 2012 |
| Italy (95) | Prospective surveillance | Children with invasive infections | 523 cases | 0-94 years (median age: 64 years)<br>(male: 51.6%) | NTHi (n=230, 84.2%)<br>Hib (n=24, 8.79%)<br>Hif (n=12, 4.40%)<br>Hie (n=5, 1.83%)<br>Hia (n=2, 0.7%) | NP swabs (real-time PCR and *mlst*) | January 2012 to December 2016 |
| Japan (98) | Prospective | Paediatric patients with invasive infections | 566 cases | Less than 5 years<br>2008 to 2012<br>(median age: not specified)<br>(male: 51.6%) | 2008 to 2012<br>NTHi (n=29, 11.1%)<br>Hib (n=234, 88.9%)<br>2013 to 2017<br>NTHi (n=262, 96.8%)<br>Hib (n=1, 3.2%) | Blood and CSF (SAST and real-time PCR) | January 2008 to 2012 then 2013 to December 2017 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | 2013 to 2017<br>(media age: specified)<br>(male: 64.9%) | | | |
| Japan (97) | Retrospective | Paediatric patients with invasive infections | 53 (septic pneumonia: 3) | 0-6 years<br>(median age: not specified)<br>(male: 43.5%) | *H. influenzae* (n=3)<br>NTHi (n=3, 100%) | Blood and CSF (SAST and real-time PCR | Jan 2000 to May 2003 |
| Australia (99) | Retrospective review | Paediatric patients with invasive infections | 148 cases | 72 % Less than 5 years<br>(median age: not specified)<br>(male: 56%) | NTHi (n=77, 52%)<br>Encapsulated (n=77, 48%) | Blood and CSF (culture) | January 2002 to December 2011 |
| Australia (100) | Retrospective | Paediatric patients with invasive infections | 737 cases<br>Pneumonia (n=360) | 0-10 years<br>(media age: not specified)<br>(males: n=345, 46.8%) | *H. influenzae*<br>Hib (n=33, 9.17%)<br>Non-b (n=40, 11.1%)<br>NTHi (n=198, 55%)<br>Untyped (n=89, 24.7%) | Blood and CSF (SAST and real-time PCR) | 2001-2013 |
| Turkey (101) | Prospective | Clinically diagnosed as having URTI or sinusitis | 1753 | 0-10 years<br>(median age: not specified)<br>(male: not specified) | **Controls (n=1382)**<br>*H. influenzae* (n=315, 22.7%)<br>Hib (n=98, 7.1%)<br>**Cases (n=371)** | NP swabs (culture) | Oct 1999 to Aug 2001 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | *H. influenzae* (n=84, 22.6%)<br>Hib (n=18, 21.4%) | | |
| The Gambia (102) | Prospective | Only healthy infants were enrolled. | 498 swabs | Birth to 12 months<br>(mean age: not specified)<br>(male: not specified) | NTHi (n=348, 70%)<br>Hib (n=3, 0.7%)<br>Encapsulated (n=7%, 35) | NP swab (real-time PCR) | Not specified |
| Kenya (103) | Prospective survey | Nasopharyngeal carriage prevalence among healthy subjects | 450 | 0-59 months<br>(median age: not specified)<br>(males: n=225, 50%) | 90 *H. influenzae*<br>NTHi (n=78, 86.7%)<br>Hib (n=6, 0.07%)<br>Hia (n=3, 0.03%)<br>Hie (n=2, 0.02%)<br>Hic (n=1, 0.01%) | NP swabs (culture and real-time PCR) | 2004 |
| Central Asia (104) | Prospective | Ill and well children visiting outpatient clinics | 630 | 2-59 months<br>(median age: not specified)<br>(male: n=331, 52.5%) | *H. influenzae* (n=357, 57%)<br>Hib (n=34, 11%) | NP swabs (culture) | Jan 1997 |
| Korea (105) | Prospective surveillance | Participants with respiratory tract infections | 100 | Participant age not specified<br>(male: not specified) | Sputum *H. influenzae* (n=58, 58%)<br>Hia (n=2, 3.45%)<br>Hif (n=2, 3.45%)<br>NTHi (n=54, 93.1%)<br>Nasopharynx *H. influenzae* (n=6, 6%)<br>NTHi (n=6, 100%) | NP swabs (culture and real-time PCR) | 2000 to 2004 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| China (106) | Prospective | Children with acute respiratory infections | 1664 cases of ARI | Less than 5 years<br>(median age: not specified)<br>(males: not specified) | *H. influenzae* (n=348, 21%) | NP swabs (culture) | February to May 2000, 2002, 2010 and 2012 |
| Brazil (107) | Prospective | Children attending day-care centres | 1192 | 2-59 months<br>(median age: 39 months)<br>Male (n=645, 54.1%) | *H. influenzae* (n=383, 32.1%)<br>Encapsulated (n=33.7, 8.8%)<br>Hif (n=18, 4.6%)<br>Hia (n=8, 2.0%)<br>Hib (n=3, 0.7%)<br>NTHi (n=278, 72.6%)<br>$Hib^-$ (n=17, 1.4%) | NP swabs (culture) | Aug to Dec 2005 |
| Spain (108) | Cross-sectional | Clinical suspicion of chronic LRTI - persistent or recurrent respiratory signs/symptoms | 197 | 6 months to 6 years<br>(median age:39.84 months)<br>(Male:51.8%) | *H. influenzae* (92)<br>NTHi n=90, 95.7%)<br>Hif (n=2, 4.3%) | NP swabs (culture and real-time PCR) | September 2013 to September 2015 |
| Israel (109) | Prospective pre-implementation randomized controlled trial. | Only healthy children were enrolled. | 769 (183 NTHi positive) | 2 to 30 months<br>(Median age: not specified)<br>(Male: 53.5%) | NTHi (n=165, 90%) | NP swabs (culture) | Not specified |

Abbreviations: NTHi – Nontypeable *H. influenzae;* OP – Oropharyngeal; NP – Nasopharyngeal; SAST – Slide Agglutination Serotyping; CSF – Cerebrospinal Fluid; URTI – Upper Respiratory Tract Infection.

The studies show that Hib vaccination can reduce infective exacerbations and antibiotic use in children and those vaccines may have a role in preventing chronic disease progression and reducing antibiotic reliance.

Finding a vaccine against NTHi will not just be beneficial towards the fight against pneumonia and other respiratory infections but will cover other types of infections. NTHi has been associated with early pregnancy loss and it has been termed as emerging neonatal and maternal pathogen (110).

### 1.7.1. NTHi vaccine development

The development of vaccines against NTHi poses a challenge in the identification of suitable epitopes. The antigen should be conserved across the genus, induce protective immune mechanisms against infection which do not cause tissue damage at the site of infection, and immune-modulatory and evasion capabilities of the bacteria must be overcome by the immune response induced (42). The focus on NTHi vaccines has mainly been based on outer membrane proteins and other genes that play a role in virulence (63).

One of the potential vaccines against NTHi is protein-D (which is used as a conjugating protein to the 10-valent pneumococcal conjugate vaccine; PHiD-CV) (73). Protein-D is a 42-kDa immunoglobulin D (IgD) binding lipoprotein found on the outer membrane of *H. influenzae* (56). Protein-D contains the phosphodiesterase activity which catalyses the hydrolysis of glycerophosphodiesters to glycerol 3-phosphate and alcohol (111,112). PHiD-CV is made up of ten purified polysaccharide capsules of *S. pneumoniae* conjugated to *H. influenzae* protein D (113), thus conferring protection to 10 *S. pneumoniae* serotypes (4, 6B, 9V, 14, 18C, 19F, 23F, 1, 5 and 7F) and NTHi (114).

The PHiD-CV vaccine induced high levels of protein-D specific antibodies, but only showed a small reduction in NTHi bacterial load in the airways of the mice (73), possibly because the protein-D used in PHiD-CV might not be genetically similar to the protein-D found in the whole organism (56,115). Regardless of this, there are

countries that use the PHiD-CV vaccines including New Zealand, however, PHiD-CV is being used against *S. pneumoniae* and not NTHi (116). Similarly, in an Australian study, they found that PHiD-CV did not select for protein-D PCR negative isolates (117).

Since the failure of the PHiD-CV to confer protection against NTHi infection, there have been studies focusing on virulence genes for vaccine formulations against NTHi (63).

Nine selected studies investigating different virulence factors for NTHi are summarised in Table 1.4. These included studies that investigated the following virulence factors; *hmw1A, hmw2A, hmwC, igA1, tehB, pilA, msfA1* and the LOS associated genes (*losB*, *infA* and *lic2C*) and identified the roles of these virulence factors in the pathogenesis of NTHi illness (57,60,66,118) and as possible vaccine targets (54,76,119,120). All the studies indicate that NTHi and *H. influenzae* have an array of virulence factors at their disposal. These virulence factors differ in prevalence and sequences hence making them difficult to target with one vaccine. They are all in agreement that vaccine candidates for NTHi still need to be explored since no vaccine is capable yet of preventing NTHi disease (54,57,60,66,76,118–121).

Table 1.4: Studies on *H. influenzae* virulence factors

| Reference | Aim | Types of isolates (number – year) | Virulence target | Findings |
|---|---|---|---|---|
| Mistry et al, (2011)(60) | To characterise the cleavage characteristics of the IgA1 protease of an isolate of NTHi obtained from the respiratory tract of a bronchiectasis patient with chronic colonisation. | NTHi 6988 | Immunoglobulin A1 (*igA1)* protease | *igA1* protease is produced by many strains of *H. influenzae* which results in different cleavage specificities of the human immunoglobulin A1. |
| Pettigrew et al, (2002)(118) | To identify virulence genes that might play a role in otitis media. | Only ill children (NTHi, 90 Hi 40 Hib – 1998 to 2001) | Lipooligosaccharide (LOS) associated genes – *lic2A* and *lic2B.* | *lic2B* was present in more than half of NTHi isolates and 88% of Hib. *lic2B* is important for otitis media pathogenesis. |
| Morey et al, (2013)(66) | To provide context for the relative importance of NTHi LOS moieties to a range of virulence traits. | NTHi 375 with 12 different mutants. 23 different plasmids with different gene inserts. | LOS associated genes – *lic1, lic2A, siaB, siaT, hap, lpsA, igtF, opsX* and *igtF* | The interactions between NTHi and the host have complex and multifactorial relations. LOS plays an important role in NTHi pathogenesis. |
| Ecevit et al, (2004)(54) | To investigate the prevalence of the *hifBC, hmw1A, hmw2A, hmwC* and *hia* genes in a collection of Hib and NTHi invasive and respiratory mucosal isolates. To assess the potential importance of these adhesins in mediating adherence to various body tissues. | Only ill children (170 Hib and 97 NTHi – 1982 to 1995) (26 NTHi – 1979 to 1982) (16 NTHi – 1996 to 1997) (5 NTHi – 2002) (18 NTHi – 1998) | *hifBC, hmw1A, hmw2A, hmwC* and *hia* | There is vast genetic variability between NTHi strains compared to Hib strains. The genetic variability allows *H. influenzae* to survive inside the host. *hifBC* plays an important role in nasopharyngeal colonisation while the HMW adhesins play a role in otitis media. |
| Winter et al, (2014)(120) | To assess the ability of antisera raised against purified HMW1/HMW2 proteins or recombinant Hia proteins to mediate opsonophagocytic killing | 89 NTHi strains (65 expressing HMW1/HMW2 24 expressing Hia – 1975 to 2011) | *hmw1/hmw2* and *hia* | Antisera raised against *hmw1/hmw2* killed 48 of 65 *hmw1/hmw2* expressing strains, while antisera raised against *hia* killed 15 of 24 *hia* expressing strains. |

| | | | | |
|---|---|---|---|---|
| | of a large panel of unrelated NTHi strains. | | | A vaccine formulated with *hmw1/hmw2* and *hia* proteins might be effective against disease caused by most NTHi strains. |
| Shahini et al, (2016)(119) | To determine the prevalence of *hmwA* which encodes *hmw1/hmw2* | 32 NTHi isolates | *hmwA* | *hmwA* was found in 62.5% (n=20/32) NTHi isolates.<br>Sequences of *hmwA* differed according to the site of isolation. |
| Whitby et al, (2010)(76) | To determine the prevalence of tehB across the species *H. influenzae* and the impact of *tehB* on tellurite resistance as well as other biologically relevant phenotypes, including iron and haem utilisation and resistance to oxidative stress. | 3 NTHi<br>1 Hib<br>1 Hid | *tehB* | *tehB* plays a role in both resistance to oxidative damage and haem uptake/utilisation.<br>*tehB* protects *H. influenzae* from tellurite exposure. |
| Mokrzan et al, (2018)(121) | To explore the potential of *pilA* as a candidate vaccine antigen against polymicrobial OM due to NTHi and *M. Catarrhalis.* | 5 NTHi<br>1 *M. Catarrhalis* | *pilA* | Immunisation with recombinant, soluble form of *pilA* (rsPilA) with reduced dosage of appropriate antibiotic can prevent or resolve complex infections. |
| Kress-Bennett et al, (2016)(57) | To present initial characterisation of Msf, a novel distributed NTHi virulence factor with a role in macrophage survival and disease. | 210 *H. influenzae* | *msfA1-4* | Deletion of all four genes resulted in a highly significant decrease in phagocytosis and survival in macrophages, which was fully complemented by a single copy the msfA1 gene. |

## 1.8. Nanofluidic real-time polymerase chain reaction (PCR) (Fluidigm®)

This PCR system allows for the reaction of assays using nano-volumes of samples and reagents thus reducing the cost of the experiment and increasing the number of assays that can be tested in limited clinical sample volumes. The Fluidigm uses what is known as a microfluidic chip which has separate inlets for samples and assays. The samples and assays are mixed within the inter-connecting tubes and valves which means less pipetting. The microfluidic chips come in different formats. One such chip is known as the Flex Six format (Figure 1.7) and it contains 12 inlets for assays and 12 inlets for samples (122). The Flex Six can be used in different setups allowing for a maximum of 144 reactions (12 assays X 12 inlets) on six different occasions and a maximum of 864 reactions ((12 assays X 12 samples) X 6) in a single run (123).

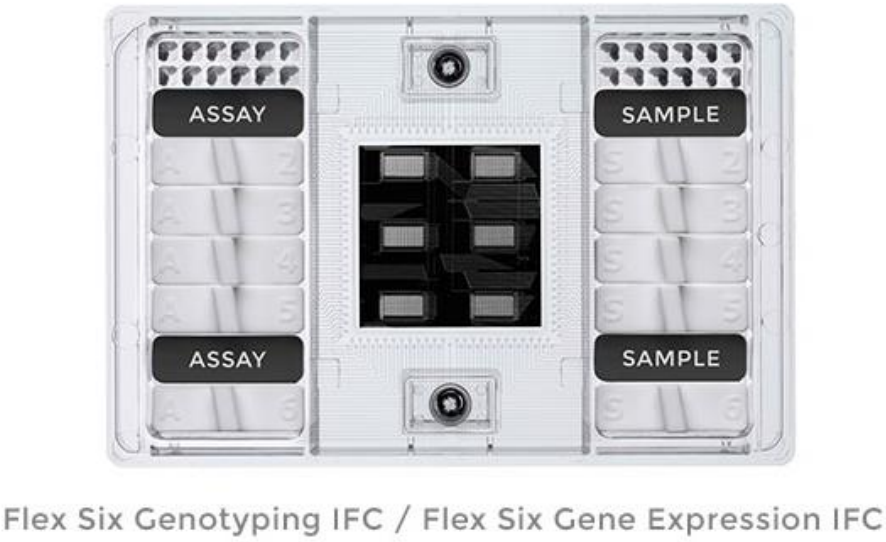



Figure 1.7: Flex six microfluidic chip from www.fluidigm.com (123)

## 1.9. Pneumonia Etiology Research for Child Health (PERCH)

The PERCH project aimed at providing a comprehensive evaluation of the etiologic agents causing WHO-defined severe and very severe pneumonia among children (1-59 months of age) in developing countries (108). In addition to the cases, age-frequency matched community controls were enrolled. The controls could have signs and symptoms of a respiratory tract infection (RTI) provided they did not have severe

pneumonia. NP swabs were collected from all cases and controls. Additionally, IS was collected from cases.

The PERCH site in South Africa was based at Chris Hani-Baragwanath Academic Hospital, in Soweto, which is a low-income community with a diversity of ethnic backgrounds (109). Enrolment started in August 2011 and was completed by August 2013; during this time 805 and 822 HIV-uninfected cases and controls were enrolled in the project.

## 1.10. Study Justification

Since NTHi is now more likely to be causative agents of *Haemophilus* disease, it would be beneficial to characterise the bacterial virulence factors of South African strains associated with the disease. Knowledge regarding the virulence potential of prevalent South African NTHi strains will also support current research towards finding the best NTHi vaccine targets.

Real-time PCR target gene comparison between putative pathogens and non-pathogenic colonising strains within a species could assist in identifying candidate genes important in bacterial pathogenesis and which could be potential vaccine epitopes.

The aim of this study was to undertake molecular characterisation of *H. influenzae* isolates identified in the induced sputum (IS) of HIV-uninfected children hospitalised with severe or very severe pneumonia, and from NP swabs in age group matched community controls enrolled into the PERCH study in South Africa between 2011 and 2013.

## 1.11. Objectives

To characterise the prevalence of select NTHi putative virulence factor genes in cultured strains obtained from IS samples of children hospitalised with severe or very severe pneumonia compared to *H. influenzae* strains colonising the NP of age group matched community controls (that could have URTI) and to determine the association of certain NTHi virulence genes with pneumonia in children

# 2. METHODS

## 2.1. Study population

This study retrospectively analysed samples collected from participants in the PERCH study (protocol M101129). Soweto is home to approximately 1.5 million people living in a low-middle income urban setting (124). HibCV and Pneumococcal Conjugate Vaccines (PCV) were introduced into the national vaccine schedule in 1999 and 2009, respectively and vaccine coverage in 2016 for three doses of each vaccine was estimated at 74% and 65%, respectively (125–127).

Cases were defined as children aged 1-59 months hospitalised with World Health Organisation (WHO) defined severe or very severe pneumonia. Severe pneumonia was defined as having a cough or difficulty breathing and lower chest wall indrawing, and very severe pneumonia as having a cough or difficulty breathing and at least one of the following: central cyanosis, difficulty breastfeeding/drinking, vomiting everything, convulsions, lethargy, unconsciousness or head nodding (124). Cases were excluded if the children presented with wheeze which resolved following bronchodilator therapy, children living outside of the designated study zone or if the child was enrolled as a PERCH case in the last 30 days or was admitted to hospital within the last 2 weeks. Controls were defined as children living within the same study catchment area as the cases without signs and symptoms of severe or very severe pneumonia. Cases and controls were age-frequency matched and to avoid bias participants with non-severe pneumonia or lower respiratory tract infections were allowed as controls provided none of the conditions for pneumonia or severe pneumonia diagnosis was met upon enrolment. The inclusion and exclusion criteria are detailed elsewhere (128).

## 2.2. Study Samples

On enrolment into the study, flocked NP swabs (Flexible minitip, Copan®) and rayon OP swabs (Flexible minitip, Copan®) were collected from all cases and controls and placed in 3mL of Universal transport media (Copan®). In addition, NP swabs were collected and placed in STGG (Respiratory and Meningeal Pathogens Research Unit, RMPRU, South Africa). In cases only, IS samples were collected and standard culture diagnostic techniques were performed for all common respiratory pathogens

including *H. influenzae.* All samples, as well as pathogens cultured from the IS samples stored in STGG, were archived at -70°C until subsequent testing. The characterisation of NTHi factors was analysed for all available NTHi isolates obtained from IS culture in the cases.

## 2.3. Real-time PCR detection for *H. influenzae* and other common respiratory pathogens

Extracted nucleic acid was amplified using the Fast Track Diagnostics Respiratory 33 kit (FTD-33) according to manufacturer's instructions (Fast-track Diagnostics, Sliema, Malta). FTD-33 uses real-time PCR with assays detecting 33 respiratory pathogens (viruses, bacteria and fungi) including *H. influenzae* and Hib together with: parainfluenza viruses 1, 2, 3 and 4 (PIV); coronaviruses NL63, 229E, OC43, and HKU1 (HCoV); human metapneumovirus A/B (HMPV); human rhinovirus (HRV); respiratory syncytial virus A/B (RSV); adenovirus (AdV); enterovirus, parechovirus; bocavirus (HBoV); cytomegalovirus; *Pneumocystis jirovecii*; *Mycoplasma pneumoniae*; *Chlamydophila pneumoniae*; *Streptococcus pneumoniae*; *Staphylococcus aureus*; *Moraxella catarrhalis*; *Bordetella pertussis*; *Klebsiella pneumoniae*; *Legionella* species; and *Salmonella* species (129,130). PCR was performed on the Applied Biosystems 7500 instrument (Foster City, California, United States). Cycling conditions were 50°C for 15 minutes, 95°C for 10 minutes and 40 cycles of 95°C for 8 seconds followed by 60°C for 34 seconds.

Fast Track Diagnostics provided plasmid standards as controls. Every 3 months, the plasmids were serially diluted 10-fold to generate standard curves that were required for the calculation of pathogen densities (copies/ml) from the sample cycle threshold (Ct) values (129). The linear range of detection of the assay provided by the standards was from 4.0 to 8.0 $\log_{10}$ copies/ml.

NP/OP isolates from the controls and NP/OP and IS from the cases that were *H. influenzae* positive were archived. As a result, the upcoming methods presented below were necessary for re-growth and confirmation of NTHi isolates.

## 2.4. DNA extraction

Total nucleic acid extraction was performed using the NucliSENS® easyMAG® extraction as per manufacturer's instructions (BioMerieux, Marcy l'Etoile, France). Nucleic acids were extracted from 400µL of the study samples, namely IS, NP/OP swabs in VTM and *H. influenzae* cultures, and eluted into a final volume of 100µL. Extraction products were archived at -70ºC until further testing.

## 2.5. Culture identification of NTHi

The NP swabs samples collected from the control participants that tested positive for *H. influenzae* using the FTD-33 assay were cultured for isolation of *H. influenzae* using methods described by Ren et al, (2012) (131). IS samples from cases were previously cultured and stored, as a result they were re-cultured using methods described by Ren et al, (2012) (131). Briefly, samples were cultured for 24 hours at 37°C in 5% $CO_2$ on chocolate agar with bacitracin (Media Mage, South Africa) which is selective for *Haemophilus* species. X and V factor dependence were assessed by growing the suspected *H. influenzae* isolates for 24 hours at 37°C in 5% $CO_2$ on XV sucrose media (Diagnostic Media Products, South Africa) together with XV factor disks from Sigma-Aldrich (St. Louis, Missouri, United States). Cultures were confirmed to be *H. influenzae* if growth was observed around the X+V factor disk only. *H. influenzae* confirmed colonies were stored in skim milk, tryptone, glucose and glycerine (STGG) at -70ºC until extraction. *H. influenzae* isolates were re-cultured if no growth was observed.

## 2.6. Classification of *H. influenzae*

It was important to verify that the isolates that were stored were indeed *H. influenzae,* therefore, it was important to classify those isolates as *H. influenzae* using the *hpd* and *bexB* PCR assays. Positive control strains for typeable and NTHi subtyping were ATCC 9422 (*H. influenzae* type *f)*, ATCC 19418 (NTHi) and ATCC 51456 (Hib). The colony forming units per millilitre (CFU/mL) were determined by picking 5 – 10 colonies grown on chocolate agar and inoculating into brain heart infusion broth (BHI) supplemented with NAD and haemin at 37ºC in 5%$CO_2$. Cultures were grown until an optical density measurement of 1 was reached at a wavelength of 600nm. Bacterial

isolates were serially diluted 10-fold with three 20µL drops of each being placed on chocolate agar plates and incubated for 24 hours in 5%$CO_2$. The number of colonies was counted and the following equation was used to calculate the CFU/ml:

$$\frac{drop\ \#1+drop\ \#2+drop\ \#3}{3}\ X\ 50\ X\ dilution\ factor = \text{CFU/ml}$$

All primers and dye-labelled minor groove binding (MGB) probes were designed using the ABI primer express software 3.0.1. from Applied Biosystems (Foster City, California, United States) based on previously published *H. influenzae* sequences obtained from GenBank (40,132). Quantitative PCR was performed using the Applied Biosystems 7500 (ABI-7500) platform. Cycling conditions were incubation at 50°C for 2 minutes, hold at 95°C for 20 seconds then 40 cycles of denaturation at 95°C for 3 seconds and annealing/extension at 60°C for 30 seconds.

The culture positive cases (cultured from the IS samples) and controls isolates (cultured from the NP swabs) were confirmed to be *H. influenzae* using single round PCR assays targeting a 137bp region in the lipoprotein D gene (*hpd*), which is present in 99% of *H. influenzae* strains (132) and capable of differentiating *H. influenzae* from other closely related *Haemophilus* species. The primer sequences were *hpd*#3 forward (5'-GGTTAAATATGCCGATGGTGTTG-3'), hdp#3 reverse (5'-CGGTGTAAGGATGCACTTCCA-3') and hdp#3 probe (5'-(FAM)-CCAGGTTGGTATATGTTAG-(MGB)-3'). All isolates testing positive on the *hpd*#3 PCR underwent further testing using primers specific for the *bexB* gene, required for capsular export, which is capable of differentiating NTHi from capsular isolates (40). The primer sequences were *bexB* forward (5'- TCACCGAAATGTTCGAGTGCTA-3'), *bexB* reverse (5'- AATAGAAGCACCAGCAACCTCAA-3') and *bexB* probe (5'- (FAM)-TACCATTTTTACTCGCGTGTTG-(MGB)-3').

### 2.6.1. Primer optimisation, standard curves and quantification of the real-time PCR assays

Primer optimisation for *bexB* and *hpd* primers was performed as follows: a primer matrix was set up using 25µL reaction volumes containing 12.5µL Taqman® fast

advanced master mix (Applied Biosystems, ABI, Foster City, USA) and 0.025mM MGB dye labelled probe (Applied Biosystems, ABI, Foster City, USA) in order to obtain the optimal primer concentration. Reactions were amplified using the ABI 7500 Real-Time PCR system (Applied Biosystems, ABI, Foster City, USA) as per manufacturer's instructions with fast cycling times which were incubation at 50°C for 2 minutes, followed by activation at 95°C for 20 seconds then 40 cycles which consist of denaturing at 95°C for 3 seconds and annealing/extension at 60°C for 30 seconds.

Standard curves were determined by serially diluting the reference strains genomic DNA using a linear dynamic range between $10^2$ and $10^8$ copies/ml based on the CFU/mL concentrations. The standard curves were also used to further identify the lower limit of detection (LOD) which is the minimum number of copies that can be detected by the primer pairs. Early exponential phase cultures of the reference strain DNA were used to set up the standard curves and the Cq values relative to the standard curves were used to calculate the CFU/swab.

Using the gradient determined from the standard curve, the amplification efficiency (e) of each primer pair was calculated by $e=-1+10^{(-1/m)}$. The intra-assay (repeatability) variation was determined by amplifying four aliquots of the same control DNA at $10^5$ copies/ml in the same run and then calculating the standard deviation (SD). The inter-assay variation (reproducibility) was determined by amplifying the same control DNA during different runs by calculating the SD. Accuracy was determined by calculating the difference between the observed copy number and the expected copy number of the samples.

## 2.7. Nanofluidic real-time PCR (Fluidigm®)

Targets to be detected using Fluidigm® are specified in Table 2.1 and were either designed from this study or taken from published sequences. Targets were selected based on published gene distribution and NTHi disease association. Published primers sequences were used to find the corresponding gene sequences on NCBI (https://blast.ncbi.nlm.nih.gov). Primers for *hmw1A, hmw2A, losB, infA* and *lic2C*

were redesigned through the Fluidigm® D3 assay design portal (https://d3.fluidigm.com/account/login) where a gene sequence is inserted and primers are designed from the most conserved region of that gene. Primers for *pilA, msfA1, tehB, oapA* and *hmwC* were taken from published sequences. The blast algorithm was used through the NCBI website to confirm that each set of primers bind to the gene it was designed for only. Resulting genes from NCBI Blast were compared to the primer sequences to verify the similarities.

Table 2.1: Primers used for Fluidigm

| Target name | Gene | Forward (5'-3') | Reverse (5'-3') | Functions | Reference |
|---|---|---|---|---|---|
| High Molecular Weight 1A | *Hmw1A* | CAAAGGACGCACTTACTGGAA | ATCGCTTCCTCTGGAGTCAA | Mediate the attachment of the bacterium to the human nasopharynx for colonisation which is an essential first step during infection | (54) |
| High Molecular weight 2A | *Hmw2A* | GCGGCGATGACGCTTTTAA | TTTCGTCTGTCTGAGGCTGAAA | | (54) |
| High Molecular Weight C | *HmwC* | TTATGGGCAGGGAATCAACAACTTT | CACTGCCCACATAATCATCTTCTACGA | | (54) |
| *Haemophilus influenzae* adhesin | *Hia* | CGGAATGCGAACGAAGTGAA | AATTTCACGCCTACCGTTGAC | | (54) |
| Immunoglobulin A1 protease | *igA1* | ACGCCGTGAAGACTACTATATG | CTCGTTGTTGATATGGTTCAT | Cleaves the human IgA1 at its hinge region into its $F_c$ and V regions so that the immunoglobulin is unable to carry out clearance of the bacteria. | (133) |
| Tellurite-resistance determinants | *tehB* | CAGCAAAAATTATCTCGCCTTGT | TATTAGGCTATGATGTGACCTCTTGG | Plays a role in the utilization of low levels of exogenous haem to support aerobic growth. | (134) |
| Type V pili | *pilA* | CTATATACACATAATTCCACATCAGCCTTA | CCACCATCGCAATTCCTTCTT | Facilitates the uptake of DNA across the bacterial cell membrane and also mediate the adherence to mammalian epithelial cells. | (75) |

| | | | | | |
|---|---|---|---|---|---|
| Sel-1-Like Repeat | *msfA1* | TTTCAATCCACCGCTTGG | GGCTTCAGAACCTTGTTGGA | Plays a role in the intracellular survival. | (57) |
| Lipooligosaccharide-associated | *losB* | AGCTGGTGCTATTAAGCTATCC | TATGCGGATAATCGGCTGAC | Genes are phase variable which is the reversible, high-frequency switching of gene expression thus allowing for the generation of a diverse population of phenotypically distinct bacteria. | (135) |
| Lipooligosaccharide-associated | *infA* | GCATTGAGATGCAAGGCACAA | CACGTGACCGTTTTCTAACTCC | | (135) |
| Lipooligosaccharide-associated | *Lic2C* | AATTTCAATACGGTGGAGGTATGGAACG | GGATTGAGAATACGGTGTAGCTGTTCAGATGC | | (72) |
| Opacity-associated protein A | *oapA* | GCACGAGAAATTGCGGG | GAGACAGATTGCGTTGC | Plays a role in phase variation and adhesion to epithelial cells during initial colonization and cell invasion. | (55) |

### 2.7.1. Pre-amplification of DNA

The initial step of Fluidigm involves specific target amplification (STA) of the nucleic acids and was performed on the T100 Thermal Cycler (Bio-Rad Laboratories, CA, USA) according to the manufacturer's instructions. A pool of primers was prepared by mixing each of the 12 assays (forward and reverse primer mixed in the same tube) at 100µM and DNA suspension buffer to a final volume of 200µL. PCR reactions containing 0.5µL of the pooled assays, 2.25µL of nuclease-free water, 1µL of PreAmp master mix (Fluidigm, CA, USA) and 1.25µL of the extracted DNA were carried out with the T100 Thermal Cycler (Bio-Rad Laboratories, CA, USA) and cycling conditions included an initial activation at 95ºC for 2 minutes followed by 10 cycles (denaturation at 95ºC for 15 seconds and annealing/extension at 60ºC for 4 minutes). Following pre-amplification, the samples were treated with exonuclease I (Inqaba Biotechnical Industries, South Africa) to remove unincorporated primers using the T100 Thermal Cycler (Bio-Rad Laboratories, CA, USA) with cycling conditions digest at 37°C for 30 minutes and inactivation at 80°C for 15 minutes.

### 2.7.2. Real-time qPCR using the Biomark™ HD System (Fluidigm)

A Flex six microfluidic chip (Fluidigm, CA, USA) was used. The chip is made up of interconnecting tubes and valves, the NanoFlex™ Integrated Fluidic Circuit (IFC) applies pressure which opens the valves and pushes the samples and the assays into their reaction chambers which are located in the middle of the plate. It has 6 independent partitions with 12 inlets for samples and 12 inlets assays. This setup allows for 12 samples to be run with 12 assays at any given time. Before using the Flex Six microfluidic chip, the chip is primed by injecting 150µL of control line fluid (Fluidigm, CA, USA), the chip is placed into the NanoFlex™ integrated fluidic circuit (IFC) controller and the prime script is then run. Samples and assays are added to their respective inlets. The chip is then placed back in the NanoFlex™ IFC controller for loading and mixing. After loading, the valves are closed and the chip is ready for thermal cycling.

The assays were labelled with SYBR® Green (Eugene, Oregon, USA), which is a fluorescent dye that binds double-stranded DNA (136). The dye absorbs blue light

and emits green light. The amount of DNA amplified can determine by measuring the fluorescence at the end of each cycle.

Thermal cycling was performed using the Biomark™ instrument at 95ºC for 10 minutes followed by 40 cycles at 95ºC for 15 seconds and 60 ºC for 1 minute. Real-time PCR Analysis Software in the BioMark-HD instrument (Fluidigm Corporation, CA, USA) was used to analyse the data.

### 2.7.3. Primer optimisation, Standard curves and quantification in the nanofluidic real-time PCR system.

The positive controls used for the 12 virulence genes were synthetic gBlocks® gene fragments from Integrated DNA Technologies (Coralville, USA). gBlocks® gene fragments are sequence-verified, double-stranded DNA fragments specifically designed for target sequences. The gBlocks® were used to optimise the real-time PCR assays. The gBlocks® were delivered as pellets normalised to 500ng. Resuspension was performed by adding 50µL TE making the concentration 10ng/µL which can be stored at -20ºC for up to 24 months.

## 2.8. Statistics

Statistical analysis was performed using STATA Version 13.1 (College Station, TX, USA) and a two-sided P-value <0.05 was considered statistically significant. For the demographics data analysis, Student's t-test was used on normally distributed continuous variables while the chi-square test was used for categorical variables. Logistic regression was used to obtain odds ratios (ORs) which were used to model the prevalence of *H. influenzae* with the cases and controls. Univariate analysis was performed on a number of population characteristics and clinical features. All the variables that had a univariate P-value of <0.2 were used to generate the multivariate model. The variables included in each of the multivariate models are listed in the footnotes below each of the tables in the results section. For controls, the multivariate analysis was adjusted for age (months), attending daycare, preterm birth and respiratory tract infection. For cases, the multivariate analysis was adjusted for age

(months), sex distribution, very severe pneumonia, tachypnoea, fever and head nodding. For the case/control multivariate analysis, the model was adjusted for age (months). Frequency and proportion of each of the genes in both cases and controls were calculated. Analysis was limited to genes that showed a higher prevalence in the cases than the controls.

## 2.9. Ethics

The PERCH study in South Africa was approved by the Wits Human Research Ethics Committee (HREC) (certificate no: M10M101129). On the 24th January 2017 approval to conduct the analysis of the *H. influenzae* virulence factors in the PERCH project was obtained from the PERCH executive committee and ethical approval was granted by the HREC board (certificate no: M170282 - available in Appendix A).

## 3. Results

A total of 802 HIV-uninfected children with clinically diagnosed severe/very-severe pneumonia and 822 HIV-uninfected community-controls were enrolled between August 2011 and August 2013. Four-percent (n=25/822) of controls had signs and symptoms of RTI.

### 3.1. Optimisation of the *bexB* and *hpd* PCR primers

Primers for *hpd* and for the capsule locus *bexB* were able to amplify their respective targets. Both had a limit of detection of 100copies/PCR and the average coefficient of variation ($r^2$) was 0.993; Table 3.1.

Table 3.1: Performance of the Real-Time PCR

| Target | Slope | Y-intercept | Efficiency (%) | Coefficient of correlation ($r^2$) | Limit of detection (copies/PCR) | Intra-assay variation (Repeatability) (%) | Inter-assay variation (Reproducibility) (%) |
|---|---|---|---|---|---|---|---|
| *hpd* | -3.444 | 46.378 | 95.131 | 0.995 | 100 | 0.73 | 3.79 |
| *bexB* | -3.759 | 55.606 | 84.511 | 0.991 | 100 | 0.77 | 1.12 |

### 3.2. Optimisation for the Fluidigm PCR

As described in the methods section, ATCC 9422 (*H. influenzae* type *f)*, ATCC 19418 (NTHi) and ATCC 51456 (Hib) were used as *H. influenzae* positive control isolates. Below are images from the Fluidigm real-time analysis software showing the amplification of the positive control isolates. Following the pre-amplification step which includes 10 cycles, isolates were confirmed as positive if threshold cycle was 21 and above.

All the primer pairs were able amplify their respective targets, except for *hia.* Thus, there was a need to troubleshoot the *hia* assay. Initially, we checked for operational error by restarting the assay and following the protocol while making sure that all the

steps in the protocol are adhered to. Then we checked the master mix and used fresh stock. Then we checked the primers buy verifying the design and confirming the required concentration. Then we checked and verified that the correct template was used and fresh stock was as well. The PCR instrument and programs were also checked for optimal operations and correct setting. After multiple failures optimisation runs for *hia* primers, they were removed and the results were excluded.

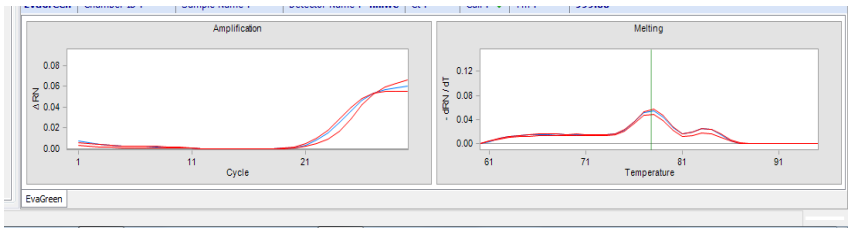



Figure 3.1: Real-time PCR analysis of the *hmwC* primers.

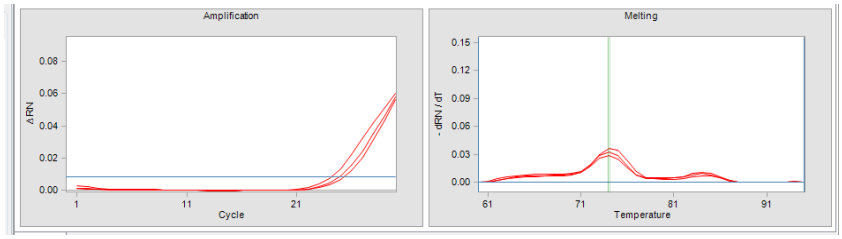



Figure 3.2: Real-time PCR analysis of the *igA1* primers.

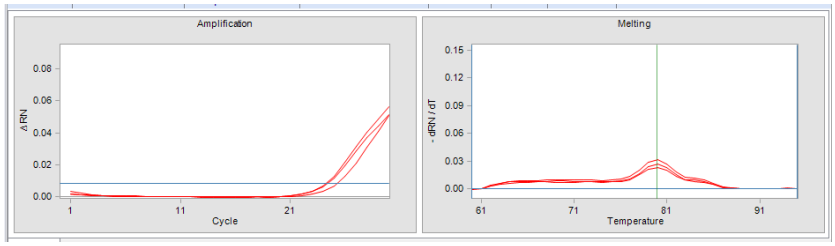



Figure 3.3: Real-time PCR analysis of the *infA* primers.

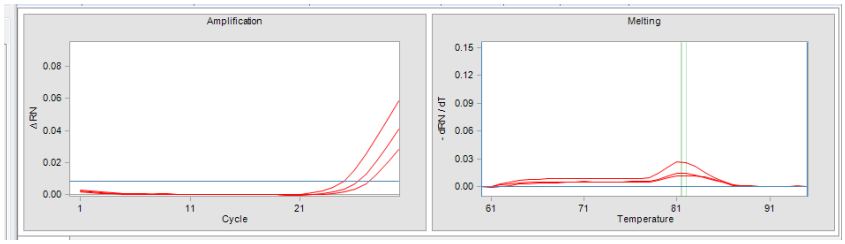



Figure 3.4: Real-time PCR analysis of the *lic2C* primers.

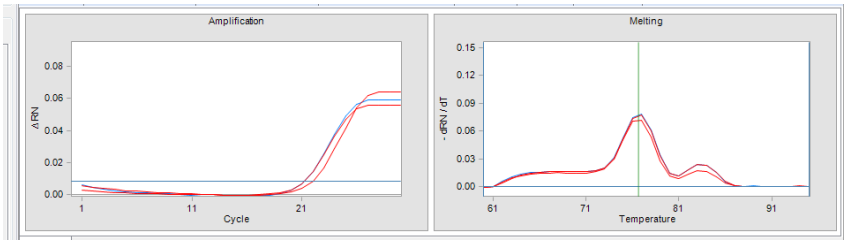



Figure 3.5: Real-time PCR analysis of the *oapA* primers.

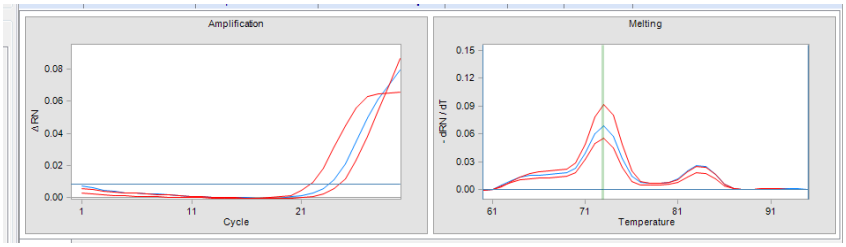



Figure 3.6: Real-time PCR analysis of the *pilA* primers.

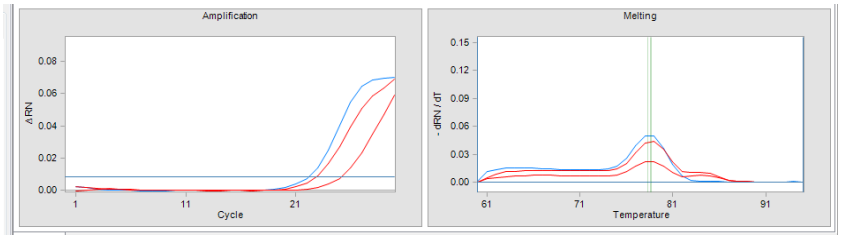



Figure 3.7: Real-time PCR analysis of the *hmw1A* primers.

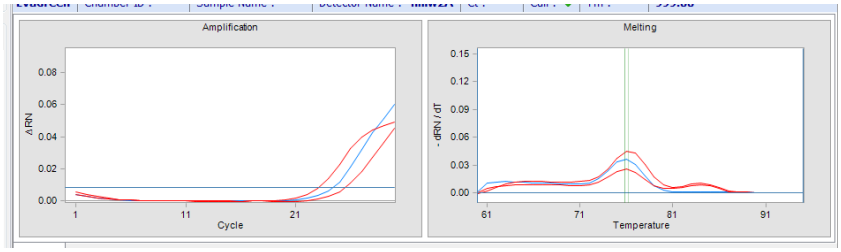



Figure 3.8: Real-time PCR analysis of the *hmw2A* primers.

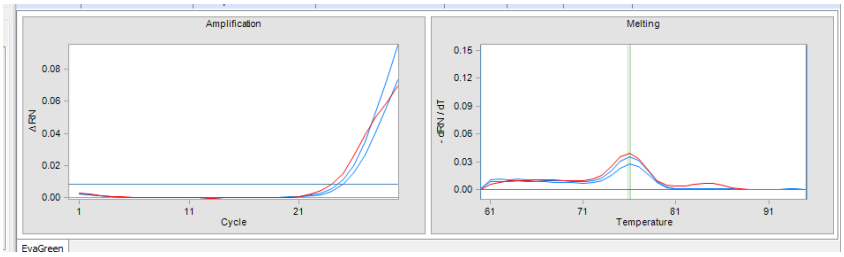



Figure 3.9: Real-time PCR analysis of the *losB* primers.

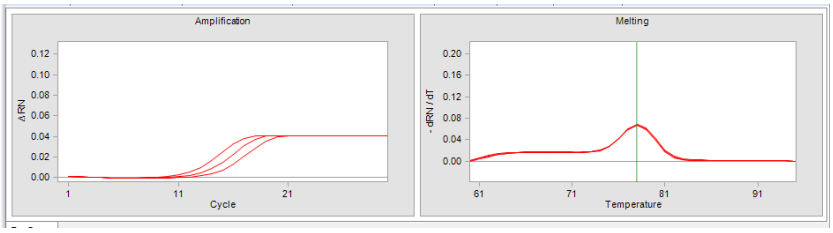



Figure 3.10: Real-time PCR analysis of the *msfA1* primers.

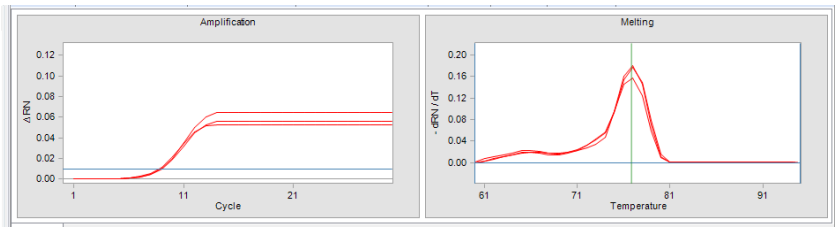



Figure 3.11: Real-time PCR analysis of the *tehB* primers.

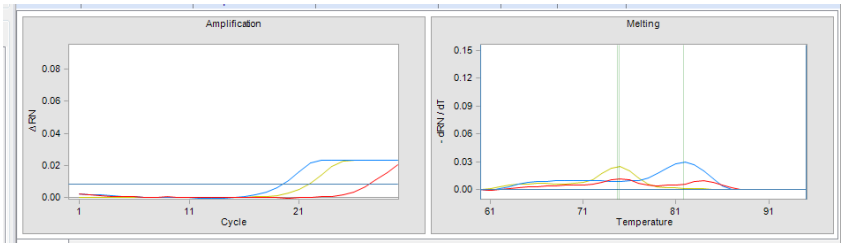



Figure 3.12: Real-time PCR analysis of the *hia* primers.

## 3.3. Demographics and clinical characteristics of controls by *H. influenzae* status

Forty-seven percent (n=387/817) of the controls were positive for *H. influenzae* colonisation on the FTD assay, of which 39 had inadequate archived sample volume for further culture. Of the 348 FTD positive samples on which culture was attempted, *H. influenzae* was grown in 86.8% (n=302/348). The proportion of FTD positive and culture negative was 13.2% (n=46/348). The culture negative samples all had very low copy numbers (Ct >35) on the FTD assay. There were no differences in the demographic and clinical characteristics of the participants that were culture-positive (n=302) compared to controls that were either culture-negative or in whom samples were unavailable for culture (n=85), except for a higher prevalence of there being a smoker in the household (37.1% vs. 22.6%; aOR: 2.0; 95% CI: 1.1-3.5); Table 3.2.

Of the *H. influenzae* culture-positive among the controls, 8.4% (n=25/298) were encapsulated (serotype characterisation not done further) and 91.6% (n=273/298) were NTHi. There were no differences in the demographic and clinical characteristics of controls in which the *H. influenzae* was of capsulated serotype compared to NTHi; Table 3.3.

On comparison of the NTHi positive and *H. influenzae* negative community controls (FTD negative), children colonised with NTHi were older (mean age: 13.9 vs. 10.7 months; P<0.001), and 2.8-fold (aOR 95% CI: 1.4-5.6), and more likely to have signs and symptoms of RTI (9.2% vs. 3.3%). There were no other differences in characteristics including sex, breastfeeding history, daycare centre attendance, preterm birth and low birth weight between the NTHi positive and negative controls (FTD negative); Table 3.4.

Table 3.2: Demographics and clinical features of the community controls according to *H. influenzae* status and samples that did not have enough volume

| | | *H. influenzae* identified on FTD | | | | | |
|---|---|---|---|---|---|---|---|
| Characteristics n(%) | N=387 | *H. influenzae* culture positive (N=302) | *H. influenzae* culture negative or un-available sample (N=85) | Unadjusted P-value | OR (95% CI) | Adjusted P-value | aOR (95% CI) |
| Age (months), mean (IQR) | 387 | 13.9 (5-19) | 14.8 (5-22) | 0.720 | Not done | 0.409 | Not done |
| Male | 387 | 143 (47.4) | 43 (50.6) | 0.598 | 0.9 (0.5-1.4) | 0.487 | 0.8 (0.5-1.4) |
| Never breast fed | 387 | 186 (61.6) | 52 (61.2) | 0.945 | 1.0 (0.6-1.7) | 0.890 | 1.0 (0.6-1.7) |
| Under weight[a] | 383 | 0 | 1 (1.2) | 0.055 | Not calculable | | Not calculable |
| HEU[b] | 387 | 84 (27.8) | 27 (31.8) | 0.477 | 0.8 (0.5-1.4) | 0.462 | 0.8 (0.5-1.4) |
| Daycare attendance | 386 | 66 (21.9) | 12 (14.3) | 0.099 | 1.7 (0.9-3.3) | 0.160 | 1.6 (0.8-2.9) |
| Smoker in household | 386 | 112 (37.1) | 19 (22.6) | 0.013 | 2.0 (1.2-3.6) | 0.022 | 2.0 (1.1-3.5) |
| Premature birth[c] | 386 | 95 (31.5) | 19 (22.6) | 0.814 | 1.0 (0.9-1.2) | 0.903 | 1.0 (0.8-1.2) |
| Birth weight, mean (IQ) | | 3.0 (2.8-3.3) | 3.0 (2.6-3.4) | <0.001 | Not done | 0.253 | Not done |
| Clinical features | | | | | | | |
| Respiratory tract infection[f] | 387 | 26 (8.6) | 5 (5.9) | 0.413 | 1.5 (0.6-4.1) | 0.640 | 1.3 (0.5-3.5) |
| Tachypnoea[d] | 380 | 32 (10.8) | 10 (11.9) | 0.778 | 0.9 (0.4-1.9) | 0.746 | 0.9 (0.4-2.0) |
| Cough | 387 | 9 (3.0) | 1 (1.2) | 0.254 | 0.7 (0.4-1.3) | 0.328 | 0.8 (0.4-1.3) |
| Fever[e] | 386 | 1 (0.3) | 0 | 0.597 | Not calculable | Not calculable | Not calculable |
| Diarrhoea | 386 | 2 (0.7) | 0 | 0.474 | Not calculable | Not calculable | Not calculable |
| Rhinorrhoea | 386 | 9 (3.0) | 2 (2.4) | 0.601 | 0.9 (0.6-1.2) | 0.554 | 0.9 (0.7-1.2) |

Abbreviations - SD: Standard deviation; HIV: Human immunodeficiency virus; HEU: HIV-expose but HIV uninfected; IQ: Inter-quartile range.
*P*-values from Chi-squared models for NTHi-positive compared to all *H. influenzae*-negative controls, logistic regression models adjusted for confounding variates (<0.2 in univariate analysis) including attending daycare and smoker in household. Odds ratios could not be calculated for continuous variables or variables with 0 values, thus cells not calculable.
[a] - Underweight defined as weight for age <-2SD of the median age-sex specific WHO reference; [b] - HEU defined as undetectable viral load, HIV seronegative in the child with a positive maternal history. Positive maternal status based on self-report was accepted, except for seronegative children < 7 months of age where documented positive maternal status was required; [c] - premature birth defined as <37 weeks gestational age; [d] - Tachypnoea defined as respiratory rate >60 breaths/minute if aged <2 months, respiratory rate >50 breaths/minute if aged 2-12 months, respiration rate >40 breaths/minute if aged >12 months; [e]- Fever defined as temperature >38$^{o}$C; [f] – Respiratory tract infection defined as cough, runny nose, ear discharge, wheeze, difficulty breathing, fever or sore throat.

Table 3.3: Demographics and clinical features of the community controls according to *H. influenzae* capsule status

| Characteristics n(%) | N=298 | Encapsulated *H. influenzae* (N=25) | NTHi (N=273) | Unadjusted P-value[a] | OR (95% CI) | Adjusted P-value | aOR (95% CI) |
|---|---|---|---|---|---|---|---|
| Age (months), mean (IQR) | 298 | 8.4 (2-11) | 14.2 (5-19) | 0.033 | Not done | 0.027 | Not done |
| Male | 298 | 15 (60.0) | 126 (46.2) | 0.189 | 1.8 (0.8-4.0) | 0.182 | 1.8 (0.8-4.3) |
| Never breast fed | 298 | 15 (60.0) | 169 (61.9) | 0.851 | 0.9 (0.4-2.1) | 0.752 | 0.9 (0.4-2.1) |
| Under weight[a] | 297 | 0 (0) | 0 (0) | Not calculable | Not calculable | Not calculable | Not calculable |
| HEU[b] | 298 | 5 (20.0) | 78 (28.6) | 0.364 | 0.6 (0.2-1.7) | 0.421 | 0.6 (0.2-1.9) |
| Daycare attendance | 298 | 4 (16.0) | 60 (22.0) | 0.410 | 0.7 (0.2-1.8) | 0.591 | 0.8 (0.4-1.7) |
| Smoker in household | 298 | 10 (40.0) | 100 (36.6) | 0.974 | 1.0 (0.5-1.9) | 0.706 | 1.1 (0.6-2.3) |
| Premature birth[c] | 298 | 7 (28.0) | 86 (31.5) | 0.186 | 1.2 (0.9-1.4) | 0.122 | 1.2 (1.0-1.5) |
| Birth weight, mean (IQ) | 296 | 3.0 (2.8-3.3) | 2.9 (2.7-3.3) | 0.289 | Not done | 0.181 | Not done |
| Clinical features | | | | | | | |
| Respiratory tract infection[f] | 298 | 1 (4.0) | 25 (9.2) | 0.396 | 0.4 (0.1-3.2) | 0.611 | 0.6 (0.1-4.7) |
| Tachypnoea[d] | 292 | 2 (8.3) | 28 (10.5) | 0.744 | 0.8 (0.2-3.5) | 0.220 | 0.4 (0.1-1.8) |
| Cough | 298 | 0 (0) | 9 (3.3) | 0.357 | Not calculable | Not calculable | Not calculable |
| Fever[e] | 298 | 0 (0) | 1 (0.4) | 0.762 | Not calculable | Not calculable | Not calculable |
| Diarrhoea | 298 | 0 (0) | 2 (0.7) | 0.870 | Not calculable | Not calculable | Not calculable |
| Rhinorrhoea | 298 | 1 (4.0) | 8 (2.9) | 0.849 | 0.9 (0.3-2.6) | 0.781 | 0.9 (0.4-2.0) |

HIV uninfected; IQ: Inter-quartile range.

*P*-values from Chi-squared models for NTHi-positive compared to all *H. influenzae*-negative controls, logistic regression models adjusted for confounding variates (<0.2 in univariate analysis) including age, sex and premature birth. Odds ratios could not be calculated for continuous variables or variables with 0 values, thus cells not calculable.

[a] - Underweight defined as weight for age <-2SD of the median age-sex specific WHO reference; [b] - HEU defined as undetectable viral load, HIV seronegative in the child with a positive maternal history. Positive maternal status based on self-report was accepted, except for seronegative children < 7 months of age where documented positive maternal status was required; [c] - premature birth defined as <37 weeks gestational age; [d] - Tachypnoea defined as respiratory rate >60 breaths/minute if aged <2 months, respiratory rate >50 breaths/minute if aged 2-12 months, respiration rate >40 breaths/minute if aged >12 months; [e]- Fever defined as temperature >38$^{o}$C; [f] – Respiratory tract infection defined as cough, runny nose, ear discharge, wheeze, difficulty breathing, fever or sore throat.

Table 3.4: Demographics and clinical features of the community controls according to NTHi NP Culture status

| Characteristics, n(%) | N | NTHi positive (N=273) | *H. influenzae* negative (N=430) | Unadjusted P-value | OR (95% CI) | Adjusted P-value | aOR (95% CI) |
|---|---|---|---|---|---|---|---|
| Age (months), mean (IQR) | 703 | 13.9 (5-18) | 10.7 (3-13) | 0.000 | Not done | 0.001 | Not done |
| Male | 703 | 126 (46.2) | 212 (49.3) | 0.416 | 0.9 (0.7-1.2) | 0.510 | 0.9 (0.7-1.2) |
| Never breast fed | 703 | 169 (61.9) | 281 (65.4) | 0.354 | 0.9 (0.6-1.2) | 0.523 | 0.9 (0.6-1.2) |
| Under weight[a] | 699 | 0 | 9 (2.1) | 0.016 | Not calculable | Not calculable | Not calculable |
| HEU[b] | 703 | 78 (28.6) | 113 (26.3) | 0.506 | 1.1 (0.8-1.5) | 0.380 | 1.2 (0.8-1.7) |
| Daycare attendance | 699 | 60 (22.0) | 43 (10.1) | 0.122 | 1.1 (1.0-1.3) | 0.295 | 1.1 (0.9-1.3) |
| Smoker in household | 699 | 100 (36.6) | 123 (28.9) | 0.386 | 1.1 (0.9-1.3) | 0.695 | 1.0 (0.8-1.3) |
| Premature birth[c] | 699 | 86 (31.5) | 136 (31.9) | 0.118 | 0.9 (0.8-1.0) | 0.143 | 0.9 (0.8-1.0) |
| Birth weight, mean (IQR) | 703 | 3.0 (2.8-3.3) | 3.0 (2.6-3.3) | 0.077 | Not done | 0.086 | Not done |
| Clinical features | | | | | | | |
| Respiratory infection[f] | 703 | 25 (9.2) | 14 (3.3) | 0.001 | 3.0 (1.5-5.9) | 0.004 | 2.8 (1.4-5.6) |
| Tachypnoea[d] | 690 | 28 (10.5) | 33 (7.8) | 0.237 | 1.4 (0.8-2.3) | 0.078 | 1.7 (0.9-2.9) |
| Cough | 702 | 9 (3.3) | 5 (1.2) | 0.918 | 1.0 (0.7-1.6) | 0.589 | 0.8 (0.3-1.8) |
| Fever[e] | 703 | 1 (0.4) | 4 (0.9) | 0.403 | 0.4 (0.0-3.5) | 0.142 | 1.0 (0.0-1.8) |
| Diarrhoea | 700 | 2 (0.7) | 0 | 0.615 | Not calculable | Not calculable | Not calculable |
| Rhinorrhoea | 700 | 8 (2.9) | 3 (0.7) | 0.401 | 1.1 (0.8-1.6) | 0.344 | 1.9 (0.5-6.7) |

Abbreviations - SD: Standard deviation; HIV: Human immunodeficiency virus; HEU: HIV-expose but HIV uninfected; IQR: Inter-quartile range; OR: Odds ratio; aOR: Adjusted odds ratio.

*H. influenzae* negative refers to all the children that are negative for all *H. influenzae* isolates.

*P*-values from Chi-squared models for NTHi-positive compared to all *H. influenzae*-negative controls, logistic regression models adjusted for confounding variates (<0.2 in univariate analysis) including age (months), attending daycare, preterm birth, respiratory tract infection.

Odds ratios could not be calculated for continuous variables or variables with 0 values, thus cells not calculable.

[a] - Underweight defined as weight for age <-2SD of the median age-sex specific WHO reference; [b] - HEU defined as undetectable viral load, HIV seronegative in the child with a positive maternal history. Positive maternal status based on self-report was accepted, except for seronegative children < 7 months of age where documented positive maternal status was required; [c] - premature birth defined as <37 weeks gestational age; [d] - Tachypnoea defined as respiratory rate >60 breaths/minute if aged <2 months, respiratory rate >50 breaths/minute if aged 2-12 months, respiration rate >40 breaths/minute if aged >12 months; [e]- Fever defined as temperature >38$^{o}$C; [f] – Respiratory tract infection defined as cough, runny nose, ear discharge, wheeze, difficulty breathing, fever or sore throat.

## 3.4. Demographics and clinical features of the cases according to NTHi status by IS Culture

Twenty-two percent (n=164/722) of IS samples from cases were positive for *H. influenzae* after standard microbiological culture, of which in 36 the retained sample had inadequate volume for subtyping and virulence factor detection. Of the remaining 149 samples tested, 10.1% failed to sub-culture (n=15) and culture was successful in 113 (75.8%). There were no differences in the demographic and clinical characteristics of the participants that were culture-positive (n=113) versus the participants that were culture-negative or in whom samples were unavailable for culture (n=51); Table 3.5.

Of the *H. influenzae* culture-positive cases, 26.5% (n=30/113) were characterised as encapsulated *H. influenzae* (further serotype characterisation not done) and 73.5% (n=83/113) were NTHi. There were no differences in the demographic and clinical characteristics of cases in whom encapsulated *H. influenzae* compared to NTHi was cultured; Table 3.6.

Cases with NTHi cultured from IS were older than those in whom NTHi was negative on IS culture (mean age: 10.8 vs 8.4 months; P=0.026); but did not differ in any other demographic characteristic; Table 3.6. Further, there were no differences in prevalence of clinical signs and symptoms between NTHi cultured on IS and non-cultured NTHi on IS cases; Table 3.6. Cases with NTHi cultured on IS, however, were more likely to have fever (73.5%) than those without NTHi on IS (62.2%; aOR1.6, 95% CI: 1.0-2.6); Table 3.7.

Table 3.5: Demographics and clinical features of the cases according to *H. influenzae* status and samples that did not have enough volume

| Characteristics n(%) | N=413 | *H. influenzae* culture positive (N=365) | *H. influenzae* culture negative or un-available sample (N=36) | Unadjusted P-value[a] | OR (95% CI) | Adjusted P-value | aOR (95% CI) |
|---|---|---|---|---|---|---|---|
| Age (months), mean (IQR) | | 9.6 (3-13) | 9.4 (3-12) | 0.782 | Not done | 0.411 | Not done |
| Male | 413 | 190 (50.4) | 17 (47.2) | 0.716 | 1.1 (0.6-2.3) | 0.550 | 1.3 (0.6-2.7) |
| Never breast fed | 413 | 252 (66.8) | 24 (66.7) | 0.983 | 1.0 (0.5-2.1) | 0.742 | 1.1 (0.5-2.6) |
| Under weight[a] | 412 | 50 (13.3) | 5 (13.9) | 0.921 | 1.0 (0.4-2.6) | 0.506 | 0.6 (0.2-2.4) |
| HEU[b] | 413 | 138 (36.6) | 10 (27.8) | 0.291 | 1.5 (0.7-3.2) | 0.848 | 1.1 (0.5-2.5) |
| Daycare attendance | 413 | 68 (18.0) | 6 (16.7) | 0.566 | 0.9 (0.7-1.2) | 0.481 | 1.3 (0.6-3.0) |
| Smoker in household | 413 | 150 (39.8) | 11 (30.6) | 0.245 | 0.8 (0.5-1.2) | 0.250 | 0.8 (0.5-1.2) |
| Premature birth[c] | 413 | 69 (18.3) | 7 (19.4) | 0.038 | 0.9 (0.8-1.0) | 0.374 | 0.9 (0.8-1.1) |
| Birth weight, mean(IQ) | | 2.9 (2.5-3.3) | 2.7 (2.3-3.3) | <0.001 | Not done | 0.316 | Not done |
| Clinical features | | | | | | | |
| Very severe pneumonia | 413 | 125 (33.2) | 10 (27.8) | 0.511 | 1.3 (0.6-2.8) | 0.285 | 1.6 (0.7-4.0) |
| Tachypnoea[g] | 407 | 300 80.7) | 29 (82.9 | 0.751 | 0.9 (0.3-2.2) | 0.905 | 1.1 (0.4-2.8) |
| Tachycardia[f] | 412 | 192 (51.1) | 19 (52.8) | 0.844 | 0.9 (0.5-1.9) | 0.802 | 0.9 (0.4-1.9) |
| Cough | 413 | 327 (86.7) | 33 (91.7) | 0.780 | 0.9 (0.5-1.6) | 0.536 | 0.7 (0.2-2.4) |
| Fever[h] | 413 | 260 (69.0) | 27 (75.0) | 0.452 | 0.7 (0.3-1.6) | 0.674 | 0.8 (0.3-2.0) |
| Diarrhoea | 413 | 63 (16.7) | 5 (13.9) | 0.663 | 1.2 (0.5-3.3) | 0.611 | 1.3 (0.5-3.7) |
| Rhinorrhoea | 413 | 114 (30.2) | 12 (33.3) | 0.057 | 0.7 (0.4-1.0) | 0.860 | 1.1 (0.5-2.4) |
| Wheezing | 404 | 122 (33.2) | 12 (33.3) | 0.982 | 1.0 (0.5-2.1) | 0.961 | 1.0 (0.4-2.3) |
| Chest X-ray abnormal[d] | 413 | 368 (96.0) | 35 (97.2) | 0.839 | 1.0 (0.7-1.6) | 0.840 | 1.0 (0.7-1.6) |
| Supplementary O2 therapy | 413 | 340 (90.2) | 33 (91.7) | 0.774 | 0.8 (0.2-2.9) | 0.452 | 0.6 (0.2-2.2) |
| Hypoxia[e] | 409 | 290 (77.8) | 26 (72.2) | 0.450 | 1.3 (0.6-2.9) | 0.357 | 1.5 (0.6-3.4) |
| Lethargic | 413 | 21 (5.6) | 1 (2.8) | 0.485 | 2.1 (0.3-15.8) | 0.598 | 1.7 (0.2-14.0) |
| Convulsions | 410 | 7 (1.9) | 0 | 0.415 | Not calculable | Not calculable | Not calculable |
| Head nodding | 413 | 108 (28.7) | 10 (27.8) | 0.912 | 1.0 (0.5-2.2) | 0.503 | 1.4 (0.6-3.3) |
| Central cyanosis | 413 | 5 (1.3) | 0 | 0.487 | Not calculable | Not calculable | Not calculable |
| Vomiting everything | 413 | 6 (1.6) | 0 | 0.446 | Not calculable | Not calculable | Not calculable |
| Lower chest wall indrawing | 413 | 359 (95.2) | 35 (97.2) | 0.590 | 0.6 (0.1-4.4) | 0.647 | 0.6 (0.1-5.1) |
| Stridor | 413 | 20 (5.3) | 1 (2.8) | 0.517 | 2.0 (0.3-15.1) | 0.580 | 1.8 (0.2-15.4) |
| Grunting | 413 | 28 (7.4) | 2 (5.6) | 0.625 | 1.4 (0.4-5.4) | 0.430 | 0.5 (0.1-2.6) |
| Nasal flaring | 413 | 318 (84.4) | 27 (75.0) | 0.153 | 1.8 (0.8-4.0) | 0.268 | 1.7 (0.6-4.5) |
| Hospital stay >5 Days | 413 | 222 (58.9) | 16 (44.4) | 0.094 | 1.8 (0.9-3.6) | 0.189 | 1.7 (0.8-3.5) |

| | | | | |
|---|---|---|---|---|
| Case fatality ratio | 356 | 10 (3.1) | 0 | 0.322 |

Abbreviations - SD: Standard deviation; HIV: Human immunodeficiency virus; HEU: HIV-expose but HIV uninfected; IQ: Inter-quartile range.
*P*-values from Chi-squared models for NTHi-positive compared to all *H. influenzae*-negative cases, logistic regression models adjusted for confounding variates (<0.2 in univariate analysis) including preterm birth, rhinorrhoea and staying in hospital for more than 5 days. Odds ratios could not be calculated for continuous variables or variables with 0 values, thus cells not calculable.
[a] - Underweight defined as weight for age <-2SD of the median age-sex specific WHO reference; [b] - HEU defined as undetectable viral load, HIV seronegative in the child with a positive maternal history. Positive maternal status based on self-report was accepted, except for seronegative children < 7 months of age where documented positive maternal status was required; [c] - Premature birth defined as <37 weeks gestational age; [d] - Defined as primary end point pneumonia or any infiltrates observed on the chest X-rays; [e] - A child was considered to be hypoxic if 1) a room air pulse-oximetry reading indicated oxygen saturation <90% at the two sites at elevation (Zambia and South Africa) or <92% at all other sites, or 2) a room air oxygen saturation reading was not available and the child was on oxygen; [f] - Tachycardia defined as heart rate >160 beats per minute (bpm) if aged <11 months, heart rate >150 bpm if aged 12-35 months, heart rate >140 bpm if aged 36-59 months; [g] - Tachypnoea defined as respiratory rate >60 breaths/minute if aged <2 months, respiratory rate >50 breaths/minute if aged 2-12 months, respiration rate >40 breaths/minute if aged >12 months; [h] - Fever defined as temperature >38$^{o}$C

Table 3.6: Demographics and clinical features of the cases according to *H. influenzae* capsule status

| Characteristics n(%) | N=113 | Encapsulated *H. influenzae* (N=30) | NTHi (N=83) | Unadjusted P-value | OR (95% CI) | Adjusted P-value | aOR (95% CI) |
|---|---|---|---|---|---|---|---|
| Age (months), mean (IQR) | 113 | 10.1 (3-13) | 12.7 (4-19) | 0.121 | Not done | 0.271 | Not done |
| Male | 113 | 16 (53.3) | 38 (45.8) | 0.479 | 1.4 (0.6-3.1) | 0.238 | 1.8 (0.7-4.7) |
| Never breast fed | 113 | 20 (66.7) | 58 (69.9) | 0.744 | 0.9 (0.4-2.1) | 0.444 | 0.7 (0.2-1.9) |
| Under weight[a] | 112 | 3 (10.0) | 13 (15.9) | 0.437 | 0.6 (0.2-2.2) | 0.945 | 1.1 (0.2-4.5) |
| HEU[b] | 113 | 10 (33.3) | 29 (34.9) | 0.874 | 0.9 (0.4-2.3) | 0.588 | 1.4 (0.5-4.0) |
| Daycare attendance | 113 | 9 (30.0) | 15 (18.1) | 0.113 | 2.1 (0.8-5.1) | 0.128 | 2.3 (0.8-6.5) |
| Smoker in household | 113 | 8 (26.7) | 33 (39.8) | 0.205 | 0.6 (0.2-1.4) | 0.065 | 0.3 (0.1-1.1) |
| Premature birth[c] | 113 | 4 (13.3) | 17 (20.5) | 0.961 | 1.0 (0.8-1.2) | 0.544 | 0.9 (0.7-1.2) |
| Birth weight, mean(IQ) | 105 | 2.8 (2.5-3.2) | 2.9 (2.7-3.4) | 0.488 | Not done | 0.342 | Not done |
| Clinical features | | | | | | | |
| Very severe pneumonia | 113 | 8 (26.7) | 18 (21.7) | 0.579 | 1.3 (0.5-3.4) | 0.924 | 1.1 (0.3-4.0) |
| Tachypnoea[g] | 113 | 24 (80.0) | 71 (85.5) | 0.477 | 0.7 (0.2-2.0) | 0.425 | 0.6 (0.2-2.2) |
| Tachycardia[f] | 113 | 17 (56.7) | 39 (47.0) | 0.365 | 1.5 (0.6-3.4) | 0.328 | 1.6 (0.6-4.3) |
| Cough | 113 | 25 (83.3) | 76 (91.6) | 0.150 | 0.4 (0.1-1.4) | 0.102 | 0.3 (0.1-1.3) |
| Fever[h] | 113 | 23 (76.7) | 60 (72.3) | 0.642 | 1.3 (0.5-3.3) | 0.689 | 1.3 (0.4-3.9) |
| Diarrhoea | 113 | 6 (20.0) | 10 (12.1) | 0.284 | 1.8 (0.6-5.6) | 0.189 | 2.5 (0.6-10.1) |
| Rhinorrhoea | 113 | 11 (36.7) | 19 (22.9) | 0.147 | 2.0 (0.8-4.8) | 0.168 | 2.0 (0.7-5.7) |
| Wheezing | 112 | 13 (44.8) | 24 (28.9) | 0.120 | 2.0 (0.8-4.8) | 0.057 | 2.8 (1.0-7.8) |
| Chest X-ray abnormal[d] | 113 | 0 (0) | 1 (1.2) | 0.573 | Not calculable | Not calculable | Not calculable |
| Supplementary O2 therapy | 113 | 26 (86.7) | 74 (89.2) | 0.714 | 0.8 (0.2-2.8) | 0.413 | 0.5 (0.1-2.3) |
| Hypoxia[e] | 112 | 21 (70.0) | 58 (70.7) | 0.940 | 1.0 (0.4-2.4) | 0.506 | 0.7 (0.2-2.1) |
| Lethargic | 113 | 2 (6.7) | 2 (2.4) | 0.299 | 2.9 (0.4-21.5) | 0.854 | 0.8 (0.1-9.1) |
| Convulsions | 113 | 1 (3.3) | 1 (1.2) | 0.467 | 2.8 (0.2-46.7) | 0.478 | 0.2 (0.0-12.3) |
| Head nodding | 113 | 5 (16.7) | 15 (18.1) | 0.863 | 0.9 (0.3-2.8) | 0.938 | 0.9 (0.2-3.7) |
| Central cyanosis | 113 | 0 (0) | 1 (1.2) | 0.546 | Not calculable | Not calculable | Not calculable |
| Vomiting everything | 113 | 1 (3.3) | 1 (1.2) | 0.467 | 2.8 (0.2-46.7) | 0.846 | 0.7 (0.0-16.1) |
| Lower chest wall indrawing | 113 | 27 (90.0) | 81 (97.6) | 0.109 | 0.2 (0.0-1.4) | 0.058 | 0.1 (0.0-1.1) |
| Stridor | 113 | 0 (0) | 5 (6.0) | 0.169 | Not calculable | Not calculable | Not calculable |
| Grunting | 113 | 0 (0) | 2 (2.4) | 0.573 | Not calculable | Not calculable | Not calculable |
| Nasal flaring | 113 | 24 (80.0) | 74 (89.2) | 0.212 | 0.5 (0.2-1.5) | 0.213 | 0.4 (0.1-1.6) |
| Hospital stay >5 Days | 113 | 14 (46.7) | 44 (53.0) | 0.551 | 0.8 (0.3-1.8) | 0.765 | 0.9 (0.3-2.3) |
| Case fatality ratio | 99 | 0 (0) | 0 (0) | | Not calculable | Not calculable | Not calculable |

Abbreviations - SD: Standard deviation; HIV: Human immunodeficiency virus; HEU: HIV-expose but HIV uninfected; IQ: Inter-quartile range.

*P*-values from Chi-squared models for NTHi-positive compared to all *H. influenzae*-negative cases, logistic regression models adjusted for confounding variates (<0.2 in univariate analysis) including age, attending daycare, cough, wheezing, rhinorrhoea, lower chest wall indrawing and stridor. Odds ratios could not be calculated for continuous variables or variables with 0 values, thus cells not calculable.

[a] - Underweight defined as weight for age <-2SD of the median age-sex specific WHO reference; [b] - HEU defined as undetectable viral load, HIV seronegative in the child with a positive maternal history. Positive maternal status based on self-report was accepted, except for seronegative children < 7 months of age where documented positive maternal status was required; [c] - Premature birth defined as <37 weeks gestational age; [d] - Defined as primary end point pneumonia or any infiltrates observed on the chest X-rays; [e] - A child was considered to be hypoxic if 1) a room air pulse-oximetry reading indicated oxygen saturation <90% at the two sites at elevation (Zambia and South Africa) or <92% at all other sites, or 2) a room air oxygen saturation reading was not available and the child was on oxygen; [f] - Tachycardia defined as heart rate >160 beats per minute (bpm) if aged <11 months, heart rate >150 bpm if aged 12-35 months, heart rate >140 bpm if aged 36-59 months; [g] - Tachypnoea defined as respiratory rate >60 breaths/minute if aged <2 months, respiratory rate >50 breaths/minute if aged 2-12 months, respiration rate >40 breaths/minute if aged >12 months; [h] - Fever defined as temperature >38$^{o}$C

Table 3.7: Demographics and clinical features of the cases according to NTHi status by IS Culture

| Characteristics n(%) | N=673 | NTHi positive (N=83) | *H. influenzae* Negative (N=560) | Unadjusted P-value | OR (95% CI) | Adjusted P-value | aOR (95% CI) |
|---|---|---|---|---|---|---|---|
| Age (months), mean (IQR) | 673 | 10.8 (4-14) | 8.4 (2-11) | 0.001 | Not done | 0.013 | Not done |
| Male | 673 | 54 (47.8) | 311 (55.5) | 0.132 | 0.7 (0.5-1.1) | 0.151 | 0.7 (0.5-1.1) |
| Never breast fed | 673 | 78 (69.0) | 367 (65.5) | 0.474 | 1.2 (0.8-1.8) | 0.530 | 1.2 (0.7-1.8) |
| Under weight[a] | 669 | 16 (14.3) | 70 (12.6) | 0.620 | 1.2 (0.6-2.1) | 0.499 | 1.2 (0.7-2.3) |
| HEU[b] | 673 | 39 (34.5) | 208 (37.1) | 0.597 | 0.9 (0.6-1.4) | 0.905 | 1.0 (0.6-1.5) |
| Daycare attendance | 673 | 24 (21.2) | 77 (13.8) | 0.633 | 1.0 (0.9-1.3) | 0.965 | 0.1 (0.8-1.3) |
| Smoker in household | 673 | 41 (26.3) | 201 (35.9) | 0.601 | 0.9 (0.6-1.2) | 0.299 | 0.8 (0.6-1.2) |
| Premature birth[c] | 673 | 21 (18.6) | 100 (17.9) | 0.872 | 1.0 (0.9-1.1) | 0.814 | 1.0 (0.9-1.1) |
| Birth weight, mean(IQR) | 639 | 2.8 (2.5-3.2) | 2.9 (2.6-3.3) | 0.321 | Not done | 0.420 | Not done |
| Clinical features | | | | | | | |
| Very severe pneumonia | 673 | 26 (23.0) | 173 (30.9) | 0.094 | 0.7 (0.4-1.1) | 0.283 | 1.7 (0.6-4.7) |
| Tachypnoea[g] | 662 | 95 (84.1) | 424 (77.2) | 0.110 | 1.6 (0.9-2.7) | 0.201 | 1.5 (0.8-2.6) |
| Tachycardia[f] | 673 | 56 (49.6) | 265 (47.3) | 0.664 | 1.1 (0.7-1.6) | 0.617 | 1.1 (0.7-1.7) |
| Cough | 673 | 101 (89.4) | 490 (87.5) | 0.270 | 1.2 (0.9-1.6) | 0.297 | 1.2 (0.9-1.6) |
| Fever[h] | 673 | 83 (73.5) | 337 (62.2) | 0.008 | 1.8 (1.2-2.9) | 0.035 | 1.6 (1.0-2.6) |
| Diarrhoea | 673 | 16 (14.2) | 85 (15.2) | 0.782 | 0.9 (0.5-1.6) | 0.834 | 0.9 (0.5-1.7) |
| Rhinorrhoea | 673 | 30 (26.6) | 146 (26.1) | 0.916 | 1.0 (0.6-1.6) | 0.687 | 0.9 (0.6-1.5) |
| Wheezing | 661 | 37 (33.0) | 190 (34.6) | 0.749 | 0.9 (0.6-1.4) | 0.580 | 0.9 (0.6-1.4) |
| Chest X-ray abnormal[d] | 673 | 110 (97.4) | 534 (95.4) | 0.453 | 0.9 (0.7-1.2) | 0.225 | 0.8 (0.6-1.1) |
| Supplementary $O_2$ therapy | 673 | 100 (88.5) | 673 490 (87.5) | 0.922 | 1.0 (0.6-1.6) | 0.871 | 1.0 (0.6-1.8) |
| Hypoxia[e] | 669 | 79 (70.5) | 416 (74.7) | 0.361 | 0.8 (0.5-1.3) | 0.490 | 0.8 (0.5-1.4) |
| Lethargic | 673 | 4 (3.5) | 17 (3.0) | 0.779 | 1.2 (0.4-3.6) | 0.829 | 1.2 (0.3-4.4) |
| Convulsions | 671 | 2 (1.8) | 7 (1.3) | 0.664 | 1.4 (0.3-6.9) | 0.799 | 1.3 (0.2-8.5) |
| Head nodding | 673 | 20 (17.7) | 155 (27.7) | 0.027 | 0.6 (0.3-0.9) | 0.049 | 0.3 (0.1-1.0) |
| Central cyanosis | 673 | 1 (0.9) | 7 (1.3) | 0.744 | 0.7 (0.1-5.8) | 0.503 | 0.4 (0.0-4.8) |
| Vomiting everything | 673 | 2 (1.8) | 10 (1.8) | 0.991 | 1.0 (0.2-4.6) | 0.622 | 0.6 (0.1-3.9) |
| Lower chest wall indrawing | 0.947 | 108 (95.6) | 536 (95.7) | 0.947 | 1.0 (0.4-2.6) | 0.483 | 0.7 (0.2-2.0) |
| Stridor | 673 | 5 (4.4) | 28 (5.0) | 0.510 | 0.8 (0.4-1.6) | 0.500 | 0.7 (0.3-1.7) |
| Grunting | 673 | 2 (1.8) | 38 (6.8) | 0.988 | 1.0 (0.7-1.4) | 0.939 | 1.0 (0.7-1.4) |
| Nasal flaring | 673 | 98 (86.7) | 473 (84.5) | 0.541 | 1.2 (0.7-2.2) | 0.412 | 1.3 (0.7-2.4) |
| Hospital stay >5 Days | 673 | 58 (51.3) | 291 (52.0) | 0.902 | 1.0 (0.7-1.5) | 0.678 | 1.1 (0.7-1.7) |
| Case fatality ratio | 575 | 0 | 12 (2.5) | 0.110 | Not calculable | Not calculable | Not calculable |

Abbreviations - SD: Standard deviation; HIV: Human immunodeficiency virus; HEU: HIV-expose but HIV uninfected; IQR: Inter-quartile range; OR: Odds ratio; aOR: Adjusted odds ratio.
*H. influenzae* negative refers to all the children that are negative for all *H. influenzae* isolates.
*P*-values from Chi-squared models for NTHi-positive compared to all *H. influenzae*-negative cases, logistic regression models adjusted for confounding variates (<0.2 in univariate analysis) including age (months), sex distribution, very severe pneumonia, tachypnoea, fever and head nodding. Odds ratios could not be calculated for continuous variables or variables with 0 values, thus cells not calculable.
[a] - Underweight defined as weight for age <-2SD of the median age-sex specific WHO reference; [b] - HEU defined as undetectable viral load, HIV seronegative in the child with a positive maternal history. Positive maternal status based on self-report was accepted, except for seronegative children < 7 months of age where documented positive maternal status was required; [c] - Premature birth defined as <37 weeks gestational age; [d] - Defined as primary end point pneumonia or any infiltrates observed on the chest X-rays; [e] - A child was considered to be hypoxic if 1) a room air pulse-oximetry reading indicated oxygen saturation <90% at the two sites at elevation (Zambia and South Africa) or <92% at all other sites, or 2) a room air oxygen saturation reading was not available and the child was on oxygen; [f] - Tachycardia defined as heart rate >160 beats per minute (bpm) if aged <11 months, heart rate >150 bpm if aged 12-35 months, heart rate >140 bpm if aged 36-59 months; [g] - Tachypnoea defined as respiratory rate >60 breaths/minute if aged <2 months, respiratory rate >50 breaths/minute if aged 2-12 months, respiration rate >40 breaths/minute if aged >12 months; [h] - Fever defined as temperature >38$^{o}$C

As the aim of our study was to investigate the virulence factors of NTHi, further analyses was limited to cases and controls identified with NTHi on induced sputum and nasopharyngeal swabs, respectively.

## 3.5. Demographics and clinical characteristics of NTHi colonised cases and controls by culture

Among those cases and controls in whom *H. influenzae* was cultured on IS (n=113) and NP swabs (n=298), respectively, the isolate was less likely to be NTHi in cases than controls (73.5% vs 91.6%; P<0.001), including in infants 6 to 11 months (74.2% vs. 96.3%; P<0.001), and the 12 to 59 months age (65.9% vs 93.6%; adjusted P<0.001); Table 3.8.

Table 3.8: Prevalence of cases and controls in whom NTHi was cultured on IS and NP/OP swabs, respectively according to age

| Age groups | Cases | Controls | P-value* |
|---|---|---|---|
| All | 83/113 (73.5) | 273/298 (91.6) | <0.001 |
| 1-5 months | 33/41 (80.5) | 78/92 (84.8) | 0.538 |
| 6-11 months | 23/31 (74.2) | 79/82(96.3) | <0.001 |
| 12-59 months | 27/41 (65.9) | 116/124 (93.6) | <0.001 |

Abbreviations: NTHi – Nontypeable *Haemophilus influenzae*
"*" - P-vales calculated using chi-square tests

Among cases and controls with NTHi, cases were older than controls (mean age: 14.2 vs. 10.1 months; P=0.021), were more likely to be underweight (15.9% vs 0%; P<0.001) and were born at lower birth weight (mean: 2.8 vs 3.0 kg; P<0.001); Table 3.9.

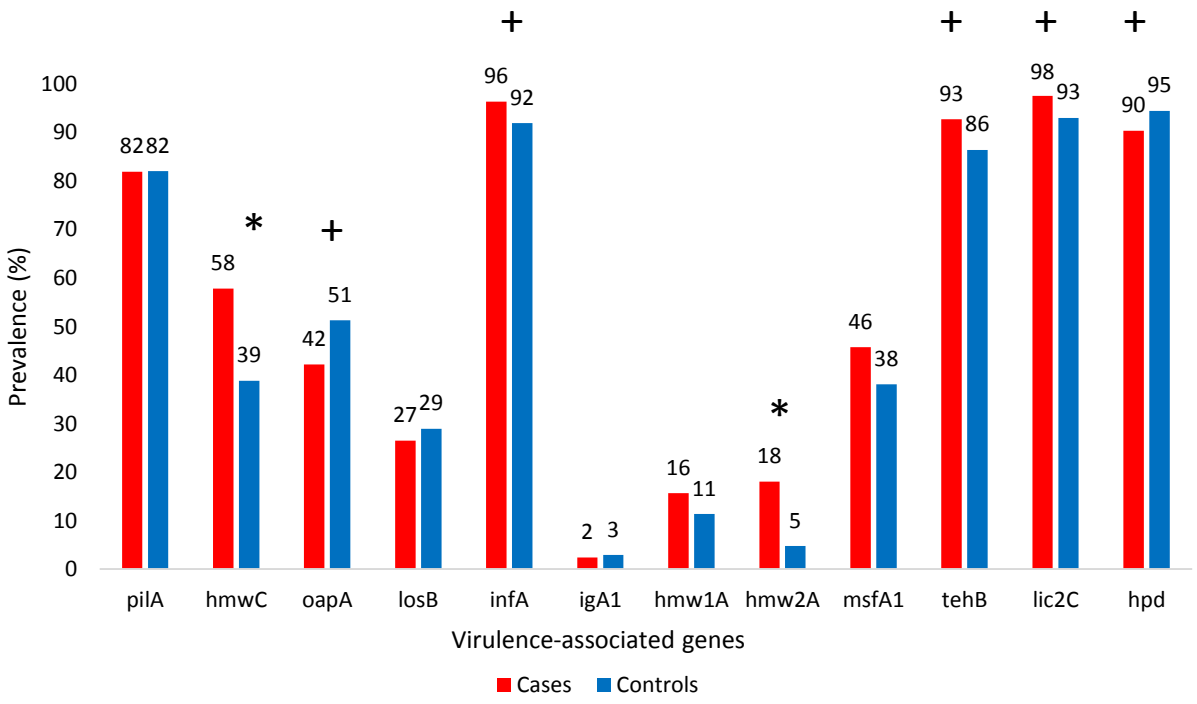
Table 3.9: Demographics and clinical characteristics of cases and controls in whom NTHi was cultured on IS and NP/OP swabs, respectively

| Characteristics n(%) | N=356 | Cases (n=83) | Controls (n=273) | Unadjusted P-value | OR (95% CI) | Adjusted P-value | aOR (95% CI) |
|---|---|---|---|---|---|---|---|
| Age (months), mean (IQR) | 356 | 10.1 (3-13) | 14.2 (5-19) | 0.021 | Not done | 0.027 | Not done |
| Male | 356 | 38 (45.8) | 126 (46.2) | 0.953 | 1.0 (0.6-1.6) | 0.857 | 1.0 (0.6-1.7) |
| Never breast fed | 356 | 58 (69.9) | 169 (61.9) | 0.186 | 1.4 (0.8-2.4) | 0.225 | 1.4 (0.8-2.4) |
| Underweight[a] | 354 | 13 (15.9) | 0 | <0.001 | Not calculable | Not calculable | Not calculable |
| HEU[b] | 356 | 29 (34.9) | 78 (28.6) | 0.268 | 1.3 (0.8-2.3) | 0.098 | 1.6 (0.9-3.0) |
| Daycare attendance | 356 | 15 (18.1) | 60 (22.0) | 0.234 | 0.7 (0.4-1.2) | 0.499 | 0.9 (0.6-1.3) |
| Smoker in household | 356 | 33 (39.8) | 100 (36.6) | 0.980 | 1.0 (0.7-1.5) | 0.886 | 1.0 (0.7-1.5) |
| Premature birth[c] | 356 | 17 (20.5) | 86 (31.5) | 0.368 | 1.1 (0.9-1.2) | 0.364 | 1.1 (0.9-1.3) |
| Birth weight, mean (IQ) | 349 | 2.8 (2.5-3.2) | 3.0 (2.8-3.3) | <0.001 | Not done | 0.000 | Not done |

Abbreviations - SD: Standard deviation; HIV: Human immunodeficiency virus; HEU: HIV-expose but HIV uninfected; IQ: Inter-quartile range.
*P*-values from Chi-squared models for *H. influenzae*-positive cases compared to all *H. influenzae*-positive controls, logistic regression models adjusted for confounding variates (<0.2 in univariate analysis) including age (months). Odds ratios could not be calculated for continuous variables or variables with 0 values, thus cells not calculable.
[a] - Underweight defined as weight for age <-2SD of the median age-sex specific WHO reference; [b] - HEU defined as undetectable viral load, HIV seronegative in the child with a positive maternal history. Positive maternal status based on self-report was accepted, except for seronegative children < 7 months of age where documented positive maternal status was required; [c] - Premature birth defined as <37 weeks gestational age.

## 3.6. Detection of virulence genes among NTHi positive cases and controls

The most commonly detected *H. influenzae* virulence factor amongst NTHi cases was *lic2C* (98%) followed by *infA* (96%) and *tehB* (93%). Amongst controls the most common putative virulence factors were *hpd* (95%), *lic2C* (93%) and *infA* (92%). On comparison of the prevalence of the 12 virulence genes in NTHi cases and controls, *hmwC* (57.8% vs 38.8%; P<0.001) and *hmw2A* (18.1% vs 4.8%; P<0.001) were more prevalent in isolates from cases than controls, Figure 3.13.




"*" Indicates P-values < 0.05
"+" Indicates P-values < 0.2

Figure 3.13: Virulence factors detected among NTHi positive cases (n=83) and controls (n=273)

Among the 12 genes detected, five (*hmwC, infA, hmw2A, tehB,* and *lic2C*) had a P-value < 0.2 and were included in analyses aimed at computing different genetic combinations that were more prevalent in isolates from cases than controls.

Out of 120 possible combinations of virulence factors, only 10 were possible; Table 3.9. Two combination of virulence factors were detected more commonly in NTHi from cases than controls, which were: i-*hmwC+/infA+/hmw2A+/tehB+/lic2C+* (12.4%

vs. 3.7%; OR: 3.7; 95% CI: 1.6-8.4) and ii-*hmwC+/infA+/hmw2A-/tehB+/lic2C+* (45.1% vs. 34.2%; OR: 1.6; 95% CI: 1.0-2.5); Table 3.10.

Table 3.10: Different types of combinations detected among NTHi cases and NTHi controls

| *hmwC* | *infA* | *hmw2A* | *tehB* | *lic2C* | N | Cases | Controls | P-value | 95% CI |
|---|---|---|---|---|---|---|---|---|---|
| - | + | + | + | + | 5 | 3 (2.7) | 2 (0.7) | 0.101 | 4.0 (0.7-24.5) |
| + | + | + | + | + | 25 | 14 (12.4) | 11 (3.7) | 0.001 | 3.7 (1.6-8.4) |
| + | + | - | + | + | 153 | 51 (45.1) | 102 (34.2) | 0.041 | 1.6 (1.0-2.5) |
| - | + | - | - | + | 22 | 4 (3.5) | 18 (6.0) | 0.315 | 0.6 (0.2-1.7) |
| - | + | - | + | + | 178 | 37 (32.7) | 141 (47.3) | 0.008 | 0.5 (0.3-0.9) |
| - | - | - | - | - | 22 | 3 (2.7) | 19 (6.4) | 0.147 | 0.4 (0.1-1.4) |
| - | - | - | - | + | 3 | 0 (0) | 3 (1.0) | 0.147 | Not calculable |
| - | - | - | + | + | 1 | 1 (0.9) | 0 (0) | 0.104 | Not calculable |
| - | - | + | - | + | 1 | 0 (0) | 1 (0.3) | 0.538 | Not calculable |
| - | + | - | - | - | 1 | 0 (0) | 1 (0.3) | 0.538 | Not calculable |

Odds ratio could not be calculated for continuous variables or variables with 0 values, thus cells not calculable.

## 4. Discussion

To test whether NTHI detected in children with pneumonia express different virulence factors, we used a novel nanofluidic real-time PCR system to compare the proportion and combination of virulence genes in children from South Africa with pneumonia and community controls. This successfully established and reliable nanofluidic PCR assay was able to simultaneously detect 12 virulence factors of NTHi in 60 different samples (58 samples together with a positive and negative controls) in a single PCR run.

We found significantly higher proportions of *hmwC* and *hmw2A* genes in children with clinical pneumonia compared with those in the community controls. These genes play a role in adherence to epithelial cells (54,65,137). This finding supports the suggestion made by Ecevit et al. (2004) which states that adherence to epithelial cells is the first step to pathogenesis which is facilitated by *hmw1/hmw2* genes (54). However, the prevalence of these proteins was low in isolates from cases (57.8% for *hmwC* and 18.1% for *hmw2A*), hence limiting their potential as individual epitopes for the construct of NTHi vaccine. Our findings are consistent with that of Ecevit et al. (54), who identified the prevalence of genes for these proteins in 46% for *hmwC* and 31% for *hmw2A* in isolates cultured from IS samples while in the isolates cultured from NP samples they found 37% for *hmwC* and 15% for *hmw2A*. Other studies have however found the prevalence of *hmwC, hmw2A* and *hmw1A* to be significantly higher (70%-80%) in NTHi isolates cultured from NP swabs (64,65,137). Furthermore, in an exploratory analysis, we found that two combinations of virulence factors; namely *hmwC+/infA+/hmw2A+/tehB+/lic2C+* and *hmwC+/infA+/hmw2A-/tehB+/lic2C+* were present in significantly higher proportions in the cases compared to the controls; albeit, only in low numbers; 12.4% vs. 3.7% to 45.1% vs. 34.2 respectively. This too would be a limitation in them being considered as a cocktail of proteins as a potential vaccine candidate. Nevertheless, *hmwC* and *hmw2A* together with other identified virulence proteins (e.g. *tehB, infA* and *lic2C*), warrant further consideration for the construct of NTHi vaccines that will likely consist of a cocktail of proteins.

The *tehB* gene which is involved in growth of NTHi when haem levels are low, was highly prevalent in both cases and controls; furthermore, it was present in both virulence combinations found to be associated with the cases. Thus potentially indicating that it is an important virulence factor and a possible vaccine target. NTHi has the ability to generate LOS which has many benefits including the ability to resist complement-mediated bactericidal killing (67,72,138). LOS is phase variable because its formation depends on many different genes, thus in this study we chose *infA, lic2C* and *losB* as these form the core of the LOS structure*. losB* did not show any significant prevalence between cases and controls, however *infA* and *lic2C* showed higher prevalence in the cases than the controls which is consistent with findings from other studies (138,139). Both *infA* and *lic2C* are located adjacent to each other on the bacterial genome and were detected in both virulent combinations signifying their importance in disease. Thus the *infA* and *lic2C* virulence factors are potentially good vaccine targets.

Furthermore, the genes *pilA, oapA* and *hpd* were found in similar prevalence amongst the cases and controls and in similar proportions as reported previously (55,75,132). Whereas, *msfA1* and *igA1* were found in lower proportions than expected when compared to other studies (57,133). The *msf* gene can be divided into four different regions, namely *msfA1, msfA2, msfA3* and *msfA4*. In this study, only *msfA1* was detected (45.8% of the cases and 38.1% of the controls); whereas, other studies looked at all four sub-sets of the *msfA* gene which could account for the differences between our study and other published data*.* Similarly, *igA1* was found at the lowest prevalence in this study which was unexpected as previous studies have reported its prevalence to be > 90% (133,140). These studies, investigated *igA1* in invasive *H. influenzae* isolates whereas, we were investigating isolates causing non-bacteraemic pneumonia.

Children colonised with NTHi were older in both cases and controls compared to non-colonised children. Furthermore, cases were significantly younger than the controls. This observation supports other studies that state that pneumonia affects neonates (6), which is supported by the high number of participants between the ages of 1 and 5 months for the cases. Other studies have suggested that age has an

effect on the development of lungs and therefore a younger child is susceptible to various lung infections (131,132).

Study limitations include that ten percent of the *H. influenzae* positive cases and 14 percent of the *H. influenzae* positive controls failed to culture viable *H. influenzae* isolates which is most likely due to the loss of viability associated with prolonged storage intervals (143–145) and repeated freezing and thawing of samples. Thus we were unable to determine the virulence characteristics of the full data set. Further, this study made use of gene expression assays which are only able to determine whether or not a gene is present in the bacterial genome; further studies are required to determine whether in fact these genes are expressed. Understanding the expression of the virulence genes may provide insights into the pathogenesis of NTHi and the best possible vaccine candidate to interrupt the pathogenesis.

In conclusion, we developed a nanofluidic assay that can simultaneously detect 12 virulence factors; however, future studies need to expand on this assay to look at more virulence genes. NTHi has a large genome and a vast array of virulence genes, therefore improving knowledge on the known virulence genes and finding new ones is important. Emphasis should be placed on vaccines that eradicate NTHi from the airways completely. Furthermore, NTHi has been known to be affected by environmental pressures. Therefore, changes in genomes due to mutations on various NTHi isolates should be monitored as acquisition of new genes might result in new virulence factors. The potential of *hmwC* and *hmw2A* with or without other putative virulence proteins of NTHi warrant further investigation as potential vaccine targets for the prevention of severe pneumonia due to NTHi.

## 5. References


1. Madhi SA. Introduction of the pneumococcal conjugate vaccine into the South African public immunisation programme: dawn of a new era? South Afr J Infect Dis. 2009 Jan 28;23(4):5–9.

2. Walker CLF, Rudan I, Liu L, Nair H, Theodoratou E, Bhutta ZA, et al. Global burden of childhood pneumonia and diarrhoea. Lancet Lond Engl. 2013 Apr 20;381(9875):1405–16.

3. Grousd JA, Rich HE, Alcorn JF. Host-Pathogen Interactions in Gram-Positive Bacterial Pneumonia. Clin Microbiol Rev [Internet]. 2019 Jun 19 [cited 2020 Jun 7];32(3). Available from: https://cmr.asm.org/content/32/3/e00107-18

4. Liu L, Oza S, Hogan D, Chu Y, Perin J, Zhu J, et al. Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the Sustainable Development Goals. The Lancet. 2016 Dec 17;388(10063):3027–35.

5. McAllister DA, Liu L, Shi T, Chu Y, Reed C, Burrows J, et al. Global, regional, and national estimates of pneumonia morbidity and mortality in children younger than 5 years between 2000 and 2015: a systematic analysis. Lancet Glob Health. 2019 Jan;7(1):e47–57.

6. le Roux DM, Myer L, Nicol MP, Zar HJ. Incidence and severity of childhood pneumonia in the first year of life in a South African birth cohort: the Drakenstein Child Health Study. Lancet Glob Health. 2015 Feb 1;3(2):e95–103.

7. Zar H, Barnett W, Stadler A, Gardner-Lubbe S, Myer L, Nicol M. Aetiology of childhood pneumonia in a well vaccinated South African birth cohort: a nested case-control study. Lancet Respir Med. 2016 Jun;4(6):463–72.

8. Statistics SA. Mortality and causes of death in southAfrica, 2016: Findings from death notification [Internet]. 2019. Available from: https://www.statssa.gov.za/publications/P03093/P030932016.pdf

9. Nunes MC, Soofie N, Downs S, Tebeila N, Mudau A, de Gouveia L, et al. Comparing the Yield of Nasopharyngeal Swabs, Nasal Aspirates, and Induced Sputum for Detection of Bordetella pertussis in Hospitalized Infants. Clin Infect Dis Off Publ Infect Dis Soc Am. 2016 Dec 1;63(Suppl 4):S181–6.

10. Xu Q, Pichichero ME. Co-colonization by Haemophilus influenzae with Streptococcus pneumoniae enhances pneumococcal-specific antibody response in young children. Vaccine. 2014 Feb 3;32(6):706–11.

11. De Schutter I, De Wachter E, Crokaert F, Verhaegen J, Soetens O, Piérard D, et al. Microbiology of Bronchoalveolar Lavage Fluid in Children With Acute Nonresponding or Recurrent Community-Acquired Pneumonia: Identification of Nontypeable Haemophilus influenzae as a Major Pathogen. Clin Infect Dis. 2011 Jun 15;52(12):1437–44.

12. Nascimento-Carvalho C. Etiology of childhood community acquired pneumonia and its implications for vaccination. Braz J Infect Dis. 2001;5(2):87–97.

13. Yun KW, Wallihan R, Juergensen A, Mejias A, Ramilo O. Community-Acquired Pneumonia in Children: Myths and Facts. Am J Perinatol. 2019 Jul;36(S 02):S54–7.

14. Cotton MF, Wasserman E, Smit J, Whitelaw A, Zar HJ. High incidence of antimicrobial resistant organisms including extended spectrum beta-lactamase producing Enterobacteriaceae and methicillin-resistant Staphylococcus aureusin nasopharyngeal and blood isolates of HIV-infected children from Cape Town, South Africa. BMC Infect Dis. 2008 Apr 1;8(1):40.

15. McNally LM, Jeena PM, Gajee K, Thula SA, Sturm AW, Cassol S, et al. Effect of age, polymicrobial disease, and maternal HIV status on treatment response and cause of severe pneumonia in South African children: a prospective descriptive study. The Lancet. 2007 Apr;369(9571):1440–51.

16. Zarei AE, Almehdar HA, Redwan EM. Hib Vaccines: Past, Present, and Future Perspectives. J Immunol Res [Internet]. 2016;2016. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4745871/

17. Bénet T, Sylla M, Messaoudi M, Sánchez Picot V, Telles J-N, Diakite A-A, et al. Etiology and Factors Associated with Pneumonia in Children under 5 Years of Age in Mali: A Prospective Case-Control Study. PLoS ONE [Internet]. 2015 Dec 22;10(12). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4687909/

18. Webb C, Ngama M, Ngatia A, Shebbe M, Morpeth S, Mwarumba S, et al. Treatment Failure among Kenyan Children with Severe Pneumonia – a Cohort Study. Pediatr Infect Dis J. 2012 Sep;31(9):e152–7.

19. Carrol ED, Mankhambo LA, Guiver M, Banda DL, The IPD Study Group, Denis B, et al. PCR Improves Diagnostic Yield from Lung Aspiration in Malawian Children with Radiologically Confirmed Pneumonia. Nizami Q, editor. PLoS ONE. 2011 Jun 14;6(6):e21042.

20. Jroundi I, Mahraoui C, Benmessaoud R, Moraleda C, Tligui H, Seffar M, et al. The Epidemiology and Aetiology of Infections in Children Admitted with Clinical Severe Pneumonia to a University Hospital in Rabat, Morocco. J Trop Pediatr. 2014 Aug 1;60(4):270–8.

21. Lagare A, Maïnassara HB, Issaka B, Sidiki A, Tempia S. Viral and bacterial etiology of severe acute respiratory illness among children < 5 years of age without influenza in Niger. BMC Infect Dis [Internet]. 2015 Dec [cited 2019 Feb 25];15(1). Available from: http://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-015-1251-y

22. Echave P. Percentage, Bacterial Etiology and Antibiotic Susceptibility of Acute Respiratory Infection and Pneumonia among Children in Rural Senegal. J Trop Pediatr. 2003 Feb 1;49(1):28–32.

23. Chochua S, D'Acremont V, Hanke C, Alfa D, Shak J, Kilowoko M, et al. Increased Nasopharyngeal Density and Concurrent Carriage of Streptococcus pneumoniae, Haemophilus influenzae, and Moraxella catarrhalis Are Associated with Pneumonia in Febrile Children. PLoS ONE [Internet]. 2016 Dec

1;11(12). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5132320/

24. Howie SRC, Morris GAJ, Tokarz R, Ebruke BE, Machuka EM, Ideh RC, et al. Etiology of Severe Childhood Pneumonia in The Gambia, West Africa, Determined by Conventional and Molecular Microbiological Analyses of Lung and Pleural Aspirate Samples. Clin Infect Dis. 2014 Sep 1;59(5):682–5.

25. Chochua S, D'Acremont V, Hanke C, Alfa D, Shak J, Kilowoko M, et al. Increased Nasopharyngeal Density and Concurrent Carriage of Streptococcus pneumoniae, Haemophilus influenzae, and Moraxella catarrhalis Are Associated with Pneumonia in Febrile Children. PLoS ONE [Internet]. 2016 Dec 1;11(12). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5132320/

26. Marangu D, Zar HJ. Childhood pneumonia in low-and-middle-income countries: An update | Elsevier Enhanced Reader [Internet]. 2019 [cited 2020 Jun 7]. Available from: https://reader.elsevier.com/reader/sd/pii/S1526054219300594?token=5A12F613AE7D165892AA0129796FD5DB81A74A490BCEF8BCA3CB68E8E6C1FF8ABEF5B0EA527035BFFF55534AC35535FB

27. King P. Haemophilus influenzae and the lung (Haemophilus and the lung). Clin Transl Med. 2012;1(1):10.

28. Murphy TF. Respiratory infections caused by non-typeable Haemophilus influenzae. Curr Opin Infect Dis. 2003 Apr;16(2):129–34.

29. Wollstein M. PFEIFFER'S BACILLUS AND INFLUENZA. J Exp Med. 1919 Nov 30;30(6):555–68.

30. Herriott RM, Meyer EY, Vogt M, Modan M. Defined Medium for Growth of Haemophilus influenzae. 1970;4.

31. Zhang T-G, Li A-H, Lyu M, Chen M, Huang F, Wu J. Detection of respiratory viral and bacterial pathogens causing pediatric community-acquired pneumonia in Beijing using real-time PCR. Chronic Dis Transl Med. 2015 Jun 1;1(2):110–6.

32. Laupland KB, Schønheyder HC, Østergaard C, Knudsen JD, Valiquette L, Galbraith J, et al. Epidemiology of Haemophilus influenzae bacteremia: A multi-national population-based assessment. J Infect. 2011 Feb 1;62(2):142–8.

33. Rogers DJ, Sie KCY, Manning SC. Epiglottitis due to nontypeable Haemophilus influenzae in a vaccinated child. Int J Pediatr Otorhinolaryngol. 2010 Feb;74(2):218–20.

34. Auranen K, Eichner M, Leino T, Takala AK, Mäkelä PH, Takala T. Modelling transmission, immunity and disease of Haemophilus influenzae type b in a structured population. Epidemiol Infect. 2004 Oct;132(5):947–57.

35. Watt JP, Chen S, Santosham M. Haemophilus influenzae type b conjugate vaccine: review of observational data on long term vaccine impact to inform recommendations for vaccine schedules. Rep World Health Organ [Internet].

2012 [cited 2017 Aug 9]; Available from: http://cdrwww.who.int/entity/immunization/sage/meetings/2013/april/4_Hib_Review_observational_data_Watt.pdf

36. Varela, JN. Membrane Protein as Novel Targets for Vaccine Production in Haemophilus influenzae and Neisseria meningitid. J Vaccines Vaccin [Internet]. 2012 [cited 2019 May 9];03(06). Available from: https://www.omicsonline.org/membrane-protein-as-novel-targets-for-vaccine-production-in-haemophilus-influenzae-and-neisseria-meningitid-2157-7560.1000152.php?aid=9336

37. Waxman DJ, Strominger JL. Penicillin-Binding Proteins and the Mechanism of Action of Beta-Lactam Antibiotics. Annu Rev Biochem. 1983;52(1):825–69.

38. King PT, Sharma R. The Lung Immune Response to Nontypeable Haemophilus influenzae (Lung Immunity to NTHi). J Immunol Res. 2015 May 31;2015:e706376.

39. LaClaire LL, Tondella MLC, Beall DS, Noble CA, Raghunathan PL, Rosenstein NE, et al. Identification of Haemophilus influenzae Serotypes by Standard Slide Agglutination Serotyping and PCR-Based Capsule Typing. J Clin Microbiol. 2003 Jan 1;41(1):393–6.

40. Davis GS, Sandstedt SA, Patel M, Marrs CF, Gilsdorf JR. Use of bexB to detect the capsule locus in Haemophilus influenzae. J Clin Microbiol. 2011 Jul;49(7):2594–601.

41. Ohkusu K, Nash KA, Inderlied CB. Molecular characterisation of Haemophilus influenzae type a and untypeable strains isolated simultaneously from cerebrospinal fluid and blood: novel use of quantitative real-time PCR based on the cap copy number to determine virulence. Clin Microbiol Infect. 2005 Aug;11(8):637–43.

42. Van Eldere J, Slack MPE, Ladhani S, Cripps AW. Non-typeable Haemophilus influenzae, an under-recognised pathogen. Lancet Infect Dis. 2014 Dec;14(12):1281–92.

43. Lichtenegger S, Bina I, Durakovic S, Glaser P, Tutz S, Schild S, et al. Serum resistance and phase variation of a nasopharyngeal non-typeable Haemophilus influenzae isolate. Int J Med Microbiol IJMM. 2017 Feb;307(2):139–46.

44. Inzana TJ. Capsules and virulence in the HAP group of bacteria. Can J Vet Res Rev Can Rech Veterinaire. 1990 Apr;54 Suppl:S22-27.

45. Cerquetti M, Cardines R, Giufre M, Sofia T, D'Ambrosio F, Mastrantonio P, et al. Genetic Diversity of Invasive Strains of *Haemophilus influenzae* Type b before and after Introduction of the Conjugate Vaccine in Italy. Clin Infect Dis. 2006 Aug;43(3):317–9.

46. Cerquetti M, Cardines R, Ciofi degli Atti ML, Giufré M, Bella A, Sofia T, et al. Presence of Multiple Copies of the Capsulation b Locus in Invasive *Haemophilus influenzae* Type b (Hib) Strains Isolated from Children with Hib Conjugate Vaccine Failure. J Infect Dis. 2005 Sep;192(5):819–23.

47. Noel GJ, Brittingham A, Granato AA, Mosser DM. Effect of amplification of the Cap b locus on complement-mediated bacteriolysis and opsonization of type b Haemophilus influenzae. Infect Immun. 1996 Nov;64(11):4769–75.

48. González-Díaz A, Tubau F, Pinto M, Sierra Y, Cubero M, Càmara J, et al. Identification of polysaccharide capsules among extensively drug-resistant genitourinary Haemophilus parainfluenzae isolates. Sci Rep [Internet]. 2019 Dec [cited 2019 Apr 15];9(1). Available from: http://www.nature.com/articles/s41598-019-40812-2

49. Gilsdorf JR, Marrs CF, Foxman B. Haemophilus influenzae: Genetic Variability and Natural Selection To Identify Virulence Factors. Infect Immun. 2004 May;72(5):2457–61.

50. Atack JM, Srikhanta YN, Fox KL, Jurcisek JA, Brockman KL, Clark TA, et al. A biphasic epigenetic switch controls immunoevasion, virulence and niche adaptation in non-typeable Haemophilus influenzae. Nat Commun [Internet]. 2015 Jul 28;6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4525171/

51. Foxwell AR, Kyd JM, Cripps AW. Nontypeable Haemophilus influenzae: Pathogenesis and Prevention. Microbiol Mol Biol Rev. 1998 Jun;62(2):294–308.

52. Carruthers MD, Tracy EN, Dickson AC, Ganser KB, Munson RS, Bakaletz LO. Biological roles of nontypeable Haemophilus influenzae type IV pilus proteins encoded by the pil and com operons. J Bacteriol. 2012 Apr;194(8):1927–33.

53. Duell BL, Su Y-C, Riesbeck K. Host–pathogen interactions of nontypeable Haemophilus influenzae: from commensal to pathogen. FEBS Lett. 2016;590(21):3840–53.

54. Ecevit IZ, McCrea KW, Pettigrew MM, Sen A, Marrs CF, Gilsdorf JR. Prevalence of the hifBC, hmw1A, hmw2A, hmwC, and hia Genes in Haemophilus influenzae Isolates. J Clin Microbiol. 2004 Jul;42(7):3065–72.

55. Prasadarao NV, Lysenko E, Wass CA, Kim KS, Weiser JN. Opacity-Associated Protein A Contributes to the Binding of Haemophilus influenzae to Chang Epithelial Cells. Infect Immun. 1999 Aug;67(8):4153–60.

56. Janson H, Hedén LO, Forsgren A. Protein D, the immunoglobulin D-binding protein of Haemophilus influenzae, is a lipoprotein. Infect Immun. 1992 Apr;60(4):1336–42.

57. Kress-Bennett JM, Hiller NL, Eutsey RA, Powell E, Longwell MJ, Hillman T, et al. Identification and Characterization of msf, a Novel Virulence Factor in Haemophilus influenzae. PLOS ONE. 2016 Mar 15;11(3):e0149891.

58. von Gottberg A, Cohen C, Whitelaw A, Chhagan M, Flannery B, Cohen AL, et al. Invasive disease due to Haemophilus influenzae serotype b ten years after routine vaccination, South Africa, 2003-2009. Vaccine. 2012 Jan 11;30(3):565–71.

59. Fernaays MM, Lesse AJ, Cai X, Murphy TF. Characterization of igaB, a Second Immunoglobulin A1 Protease Gene in Nontypeable Haemophilus influenzae. Infect Immun. 2006 Oct 1;74(10):5860–70.

60. Mistry DV, Stockley RA. The cleavage specificity of an IgA1 protease from Haemophilus influenzae. Virulence. 2011 Mar 1;2(2):103–10.

61. Murphy TF, Lesse AJ, Kirkham C, Zhong H, Sethi S, Munson RS. A Clonal Group of Nontypeable Haemophilus influenzae with Two IgA Proteases Is Adapted to Infection in Chronic Obstructive Pulmonary Disease. PLoS ONE [Internet]. 2011 Oct 5;6(10). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3187821/

62. Erwin AL, Smith AL. Nontypeable Haemophilus influenzae: understanding virulence and commensal behavior. Trends Microbiol. 2007 Aug 1;15(8):355–62.

63. Murphy TF. Vaccines for Nontypeable Haemophilus influenzae: the Future Is Now. Clin Vaccine Immunol CVI. 2015 May;22(5):459–66.

64. Giufrè M, Muscillo M, Spigaglia P, Cardines R, Mastrantonio P, Cerquetti M. Conservation and Diversity of HMW1 and HMW2 Adhesin Binding Domains among Invasive Nontypeable Haemophilus influenzae Isolates. Infect Immun. 2006 Feb 1;74(2):1161–70.

65. St. Geme JW, Kumar VV, Cutter D, Barenkamp SJ. Prevalence and Distribution of the hmw and hia Genes and the HMW and Hia Adhesins among Genetically Diverse Strains of Nontypeable Haemophilus influenzae. Infect Immun. 1998 Jan;66(1):364–8.

66. Morey P, Viadas C, Euba B, Hood DW, Barberán M, Gil C, et al. Relative Contributions of Lipooligosaccharide Inner and Outer Core Modifications to Nontypeable Haemophilus influenzae Pathogenesis. Infect Immun. 2013 Nov;81(11):4100–11.

67. Fox KL, Atack JM, Srikhanta YN, Eckert A, Novotny LA, Bakaletz LO, et al. Selection for Phase Variation of LOS Biosynthetic Genes Frequently Occurs in Progression of Non-Typeable Haemophilus influenzae Infection from the Nasopharynx to the Middle Ear of Human Patients. PLOS ONE. 2014 Feb 28;9(2):e90505.

68. Campagnari AA, Gupta MR, Dudas KC, Murphy TF, Apicella MA. Antigenic diversity of lipooligosaccharides of nontypable Haemophilus influenzae. Infect Immun. 1987 Apr;55(4):882–7.

69. Garmendia J, Martí-Lliteras P, Moleres J, Puig C, Bengoechea JA. Genotypic and phenotypic diversity of the noncapsulated Haemophilus influenzae: adaptation and pathogenesis in the human airways. Int Microbiol Off J Span Soc Microbiol. 2012 Dec;15(4):159–72.

70. Ikeda M, Enomoto N, Hashimoto D, Fujisawa T, Inui N, Nakamura Y, et al. Nontypeable Haemophilus influenzae exploits the interaction between protein-E and vitronectin for the adherence and invasion to bronchial epithelial cells.

BMC Microbiol [Internet]. 2015 Nov 14;15. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4647820/

71. Langereis JD, Stol K, Schweda EK, Twelkmeyer B, Bootsma HJ, Vries SPW de, et al. Modified Lipooligosaccharide Structure Protects Nontypeable Haemophilus influenzae from IgM-Mediated Complement Killing in Experimental Otitis Media. mBio. 2012 Aug 31;3(4):e00079-12.

72. McCrea KW, Xie J, Daniel D, Ulrich-Lewis JT, Zhang L. Predicted Configurations of Oligosaccharide Extensions in the Lipooligosaccharide of Nontypeable Haemophilus influenzae Isolates. J Clin Microbiol. 2014 Jul;52(7):2659–61.

73. Siggins MK, Gill SK, Langford PR, Li Y, Ladhani SN, Tregoning JS. PHiD-CV induces anti-Protein D antibodies but does not augment pulmonary clearance of nontypeable Haemophilus influenzae in mice. Vaccine. 2015 Sep 11;33(38):4954–61.

74. Murphy TF, Kirkham C, Jones MM, Sethi S, Kong Y, Pettigrew MM. Expression of IgA Proteases by Haemophilus influenzae in the Respiratory Tract of Adults With Chronic Obstructive Pulmonary Disease. J Infect Dis. 2015 Dec 1;212(11):1798–805.

75. Jurcisek JA, Bookwalter JE, Baker BD, Fernandez S, Novotny LA, Munson RS, et al. The PilA protein of non-typeable Haemophilus influenzae plays a role in biofilm formation, adherence to epithelial cells and colonization of the mammalian upper respiratory tract. Mol Microbiol. 2007 Sep;65(5):1288–99.

76. Whitby PW, Seale TW, Morton DJ, VanWagoner TM, Stull TL. Characterization of the Haemophilus influenzae tehB gene and its role in virulence. Microbiology. 2010 Apr;156(Pt 4):1188–200.

77. Lawrence P, Bey R. Map-Based Comparative Genomic Analysis of Virulent *Haemophilus Parasuis* Serovars 4 and 5. J Genomics. 2015;3:59–71.

78. Andersson M, Resman F, Eitrem R, Drobni P, Riesbeck K, Kahlmeter G, et al. Outbreak of a beta-lactam resistant non-typeable Haemophilus influenzae sequence type 14 associated with severe clinical outcomes. BMC Infect Dis [Internet]. 2015 Dec 23;15. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4690285/

79. Williamson R, Collatz E, Gutmann L. [Mechanisms of action of beta-lactam antibiotics and mechanisms of non-enzymatic resistance]. Presse Medicale Paris Fr 1983. 1986 Dec 20;15(46):2282–9.

80. Drusano GL. Antimicrobial pharmacodynamics: critical interactions of “bug and drug.” Nat Rev Microbiol. 2004 Apr;2(4):289–300.

81. Lakshmi R, Nusrin KS, Georgy SA, Sreelakshmi KS. ROLE OF BETA LACTAMASES IN ANTIBIOTIC RESISTANCE: A REVIEW. Int Res J Pharm. 2014 Feb 19;5(2):37–40.

82. Urumova V. Extended spectrum beta lactamase producing animal enterobacteriaceae isolates as potential risk to public health – review. Rev Méd Vét. 2015;16.

83. Ho J, Ip M. Antibiotic-Resistant Community-Acquired Bacterial Pneumonia. Infect Dis Clin. 2019 Dec 1;33(4):1087–103.

84. Hussey G, Hitchcock J, Schaaf H, Coetzee G, Hanslo D, van Schalkwyk E, et al. Epidemiology of invasive Haemophilus influenzae infections in Cape Town, South Africa. Ann Trop Paediatr. 1994;14(2):97–103.

85. Madhi SA, Petersen K, Khoosal M, Huebner RE, Mbelle N, Mothupi R, et al. Reduced effectiveness of Haemophilus influenzae type b conjugate vaccine in children with a high prevalence of human immunodeficiency virus type 1 infection. Pediatr Infect Dis J. 2002 Apr;21(4):315–21.

86. von Gottberg A. Bacterial meningitis in the era of paediatric vaccination against the encapsulated pathogens. Contin Med Educ. 2010 Jun 29;28(6):252.

87. Kelly DF, Moxon ER, Pollard AJ. Haemophilus influenzae type b conjugate vaccines. Immunology. 2004 Oct;113(2):163–74.

88. Ward J, Brenneman G, Letson GW, Heyward WL. Limited efficacy of a Haemophilus influenzae type b conjugate vaccine in Alaska Native infants. The Alaska H. influenzae Vaccine Study Group. N Engl J Med. 1990 Nov 15;323(20):1393–401.

89. Beckenhaupt P, Bryant P, Kroger A, Weaver D, Wolicki J, Wortley P. D-epartment of Health and Human Services. 2012;12.

90. Eskola J, Ward J, Dagan R, Goldblatt D, Zepp F, Siegrist C-A. Combined vaccination of Haemophilus influenzae type b conjugate and diphtheria-tetanus-pertussis containing acellular pertussis. The Lancet. 1999 Dec;354(9195):2063–8.

91. Loharikar A. Status of New Vaccine Introduction — Worldwide, September 2016. MMWR Morb Mortal Wkly Rep [Internet]. 2016 [cited 2020 Jun 7];65. Available from: https://www.cdc.gov/mmwr/volumes/65/wr/mm6541a3.htm

92. Martin M, Casellas JM, Madhi SA, Urquhart TJ, Delport SD, Ferrero F, et al. Impact of haemophilus influenzae type b conjugate vaccine in South Africa and Argentina. Pediatr Infect Dis J. 2004 Sep;23(9):842–7.

93. Langereis, J.D., Jonge, M.I. de. Invasive Disease Caused by Nontypeable Haemophilus influenzae. Emerg Infect Dis. 2015;21:1711–8.

94. Camilli R, Vescio MF, Giufrè M, Daprai L, Garlaschi ML, Cerquetti M, et al. Carriage of Haemophilus influenzae is associated with pneumococcal vaccination in Italian children. Vaccine. 2015 Aug 26;33(36):4559–64.

95. Giufrè M, Fabiani M, Cardines R, Riccardo F, Caporali MG, D'Ancona F, et al. Increasing trend in invasive non-typeable Haemophilus influenzae disease and molecular characterization of the isolates, Italy, 2012–2016. Vaccine. 2018 Oct;36(45):6615–22.

96. Giufrè M, Daprai L, Cardines R, Bernaschi P, Ravà L, Accogli M, et al. Carriage of Haemophilus influenzae in the oropharynx of young children and molecular epidemiology of the isolates after fifteen years of H. influenzae type b vaccination in Italy. Vaccine. 2015 Nov 17;33(46):6227–34.

97. Ishiwada N, Cao Y, Kohno Y. PCR-based capsular serotype determination of Haemophilus influenzae strains recovered from Japanese paediatric patients with invasive infection | Elsevier Enhanced Reader [Internet]. 2004 [cited 2019 May 1]. Available from: https://reader.elsevier.com/reader/sd/pii/S1198743X14633565?token=7114A8339EBB5DD927E904741F31E100F594CCA41C73D633EEB0095BED55DAFFF2BA554BAB3183D1F1B501D8C4CD0903

98. Suga S, Ishiwada N, Sasaki Y, Akeda H, Nishi J, Okada K, et al. A nationwide population-based surveillance of invasive Haemophilus influenzae diseases in children after the introduction of the Haemophilus influenzae type b vaccine in Japan. Vaccine. 2018 Sep 11;36(38):5678–84.

99. Cleland G, Leung C, Wan Sai Cheong J, Francis J, Heney C, Nourse C. Paediatric invasive Haemophilus influenzae in Queensland, Australia, 2002-2011: Young Indigenous children remain at highest risk. J Paediatr Child Health. 2018 Jan;54(1):36–41.

100. Wan Sai Cheong J, Smith H, Heney C, Robson J, Schlebusch S, Fu J, et al. Trends in the epidemiology of invasive *Haemophilus influenzae* disease in Queensland, Australia from 2000 to 2013: what is the impact of an increase in invasive non-typable *H. influenzae* (NTHi)? Epidemiol Infect. 2015 Oct;143(14):2993–3000.

101. Yagci A, Ilki A, Akbenlioglu C, Ulger N, Inanli S, Söyletir G, et al. Surveillance of Haemophilus influenzae among respiratory tract samples of Turkish children. Int J Antimicrob Agents. 2003 Nov 1;22(5):548–50.

102. Kwambana BA, Barer MR, Bottomley C, Adegbola RA, Antonio M. Early acquisition and high nasopharyngeal co-colonisation by Streptococcus pneumoniae and three respiratory pathogens amongst Gambian new-borns and infants. BMC Infect Dis [Internet]. 2011 Dec [cited 2019 Feb 25];11(1). Available from: http://bmcinfectdis.biomedcentral.com/articles/10.1186/1471-2334-11-175

103. Abdullahi O, Nyiro J, Lewa P, Slack M, Scott JAG. The descriptive epidemiology of Streptococcus pneumoniae and Haemophilus influenzae nasopharyngeal carriage in children and adults in Kilifi district, Kenya. Pediatr Infect Dis J. 2008 Jan;27(1):59–64.

104. Factor SH, LaClaire L, Bronsdon M, Suleymanova F, Altynbaeva G, Kadirov BA, et al. type b Carriage, Central Asia. Emerg Infect Dis. 2005;11(9):4.

105. Bae S, Lee J, Kim E, Lee J, Yu J, Kang Y. Serotype distribution and β-lactam resistance in Haemophilus influenzae isolated from patients with respiratory infections in Korea. J Microbiol. 2010 Feb;48(1):84–8.

106. Zhu H, Wang A, Tong J, Yuan L, Gao W, Shi W, et al. Nasopharyngeal carriage and antimicrobial susceptibility of Haemophilus influenzae among children younger than 5 years of age in Beijing, China. BMC Microbiol [Internet]. 2015 Feb 4;15. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4332420/

107. de Carvalho CX, Kipnis A, Thörn L, de Andrade JG, Pimenta F, Brandileone MC de C, et al. Carriage of Haemophilus influenzae among Brazilian children attending day care centers in the era of widespread Hib vaccination. Vaccine. 2011 Feb;29(7):1438–42.

108. Escribano Montaner A, García de Lomas J, Villa Asensi JR, Asensio de la Cruz O, de la Serna Blázquez O, Santiago Burruchaga M, et al. Bacteria from bronchoalveolar lavage fluid from children with suspected chronic lower respiratory tract infection: results from a multi-center, cross-sectional study in Spain. Eur J Pediatr. 2018;177(2):181–92.

109. Lewnard JA, Huppert A, Givon-Lavi N, Pettigrew MM, Regev-Yochay G, Dagan R, et al. Density, Serotype Diversity, and Fitness of Streptococcus pneumoniae in Upper Respiratory Tract Cocolonization With Nontypeable Haemophilus influenzae. J Infect Dis. 2016 Nov 1;214(9):1411–20.

110. Cevik M, Moncayo-Nieto OL, Evans MJ. Non-typeable Haemophilus influenzae-associated early pregnancy loss: an emerging neonatal and maternal pathogen. Infection. 2020 Apr 1;48(2):285–8.

111. Clarke C, Bakaletz LO, Ruiz-Guiñazú J, Borys D, Mrkvan T. Impact of protein D-containing pneumococcal conjugate vaccines on non-typeable *Haemophilus influenzae* acute otitis media and carriage. Expert Rev Vaccines. 2017 Jul 3;16(7):751–64.

112. Toropainen M, Raitolehto A, Henckaerts I, Wauters D, Poolman J, Lestrate P, et al. Pneumococcal Haemophilus influenzae Protein D Conjugate Vaccine Induces Antibodies That Inhibit Glycerophosphodiester Phosphodiesterase Activity of Protein D. Infect Immun. 2008 Oct 1;76(10):4546–53.

113. Forsgren A, Riesbeck K, Janson H. Protein D of Haemophilus influenzae: a protective nontypeable H. influenzae antigen and a carrier for pneumococcal conjugate vaccines. Clin Infect Dis Off Publ Infect Dis Soc Am. 2008 Mar 1;46(5):726–31.

114. Dinleyici EC, Yargic ZA. Pneumococcal conjugated vaccine: PHiD-CV. Expert Rev Anti Infect Ther. 2009 Nov;7(9):1063–74.

115. Janson H, Hedén LO, Grubb A, Ruan MR, Forsgren A. Protein D, an immunoglobulin D-binding protein of Haemophilus influenzae: cloning, nucleotide sequence, and expression in Escherichia coli. Infect Immun. 1991 Jan;59(1):119–25.

116. Varghese L, Talbot L, Govender A, Zhang X-H, Mungall BA. A Cost-Effectiveness Analysis of the 10-Valent Pneumococcal Non-Typeable Haemophilus influenzae Protein D Conjugate Vaccine (PHiD-CV) Compared to the 13-Valent Pneumococcal Conjugate Vaccine (PCV13) for Universal Mass

Vaccination Implementation in New Zealand. Appl Health Econ Health Policy. 2018 Jun;16(3):331–45.

117. Beissbarth J, Smith-Vaughan HC, Harris TM, Binks MJ, Leach AJ. Use of the 10-valent pneumococcal Haemophilus influenzae protein D conjugate vaccine (PHiD-CV10) in an Australian Indigenous paediatric population does not alter the prevalence of nontypeable Haemophilus influenzae without the protein D gene. Vaccine. 2019 Jul 9;37(30):4089–93.

118. Pettigrew MM, Foxman B, Marrs CF, Gilsdorf JR. Identification of the lipooligosaccharide biosynthesis gene lic2B as a putative virulence factor in strains of nontypeable Haemophilus influenzae that cause otitis media. Infect Immun. 2002 Jul;70(7):3551–6.

119. Shahini Shams Abadi M, Siadat SD, Vaziri F, Davari M, Fateh A, Pourazar S, et al. Distribution and Diversity of hmw1A Among Invasive Nontypeable Haemophilus influenzae Isolates in Iran. Avicenna J Med Biotechnol. 2016;8(2):99–102.

120. Winter LE, Barenkamp SJ. Antibodies to the HMW1/HMW2 and Hia Adhesins of Nontypeable Haemophilus influenzae Mediate Broad-Based Opsonophagocytic Killing of Homologous and Heterologous Strains. Clin Vaccine Immunol CVI. 2014 May;21(5):613–21.

121. Mokrzan EM, Novotny LA, Brockman KL, Bakaletz LO. Antibodies against the Majority Subunit (PilA) of the Type IV Pilus of Nontypeable *Haemophilus influenzae* Disperse *Moraxella catarrhalis* from a Dual-Species Biofilm. Rappuoli R, editor. mBio. 2018 Dec 11;9(6):e02423-18, /mbio/9/6/mBio.02423-18.atom.

122. Spurgeon SL, Jones RC, Ramakrishnan R. High Throughput Gene Expression Measurement with Real Time PCR in a Microfluidic Dynamic Array. PLOS ONE. 2008 Feb 27;3(2):e1662.

123. fluidigm.com. Gene Expression with the Flex Six IFC Using Delta Gene Assays [Internet]. Available from: https://www.fluidigm.com/binaries/content/documents/fluidigm/resources/flex-six-ge-delta-gene-fast-std-100-7717/flex-six-ge-delta-gene-fast-std-100-7717/fluidigm%3Afile

124. Park DE, Baggett HC, Howie SRC, Shi Q, Watson NL, Brooks WA, et al. Colonization Density of the Upper Respiratory Tract as a Predictor of Pneumonia—Haemophilus influenzae, Moraxella catarrhalis, Staphylococcus aureus, and Pneumocystis jirovecii. Clin Infect Dis. 2017 Jun 15;64(suppl_3):S328–36.

125. Baker L. The face of South Africa's Expanded Programme on Immunisation (EPI) schedule. SA Pharm J. 2010;77(1):18–49.

126. Cohen C, von Mollendorf C, de Gouveia L, Lengana S, Meiring S, Quan V, et al. Effectiveness of the 13-valent pneumococcal conjugate vaccine against invasive pneumococcal disease in South African children: a case-control study. Lancet Glob Health. 2017;5(3):e359–69.

127. Feldstein LR, Mariat S, Gacic-Dobo M, Diallo MS, Conklin LM, Wallace AS. Global Routine Vaccination Coverage, 2016. MMWR Morb Mortal Wkly Rep. 2017 Nov 17;66(45):1252–5.

128. Deloria-Knoll M, Feikin DR, Scott JAG, O'Brien KL, DeLuca AN, Driscoll AJ, et al. Identification and Selection of Cases and Controls in the Pneumonia Etiology Research for Child Health Project. Clin Infect Dis Off Publ Infect Dis Soc Am. 2012 Apr 1;54(Suppl 2):S117–23.

129. Driscoll AJ, Karron RA, Morpeth SC, Bhat N, Levine OS, Baggett HC, et al. Standardization of Laboratory Methods for the PERCH Study. Clin Infect Dis Off Publ Infect Dis Soc Am. 2017 Jun 15;64(Suppl 3):S245–52.

130. Salez N, Vabret A, Leruez-Ville M, Andreoletti L, Carrat F, Renois F, et al. Evaluation of Four Commercial Multiplex Molecular Tests for the Diagnosis of Acute Respiratory Infections. Chan KH, editor. PLOS ONE. 2015 Jun 24;10(6):e0130378.

131. Ren D, Nelson KL, Uchakin PN, Smith AL, Gu X-X, Daines DA. Characterization of extended co-culture of non-typeable Haemophilus influenzae with primary human respiratory tissues. Exp Biol Med Maywood Nj. 2012 May;237(5):540–7.

132. Wang X, Mair R, Hatcher C, Theodore MJ, Edmond K, Wu HM, et al. Detection of bacterial pathogens in Mongolia meningitis surveillance with a new real-time PCR assay to detect Haemophilus influenzae. Int J Med Microbiol IJMM. 2011 Apr;301(4):303–9.

133. Vitovski S, Dunkin KT, Howard AJ, Sayers JR. Nontypeable Haemophilus influenzae in carriage and disease: a difference in IgA1 protease activity levels. JAMA. 2002 Apr 3;287(13):1699–705.

134. Whitby PW, Seale TW, Morton DJ, VanWagoner TM, Stull TL. Characterization of the Haemophilus influenzae tehB gene and its role in virulence. Microbiology. 2010 Apr;156(Pt 4):1188–200.

135. Erwin AL, Nelson KL, Mhlanga-Mutangadura T, Bonthuis PJ, Geelhood JL, Morlin G, et al. Characterization of Genetic and Phenotypic Diversity of Invasive Nontypeable Haemophilus influenzae. Infect Immun. 2005 Sep 1;73(9):5853–63.

136. Ponchel F, Toomes C, Bransfield K, Leong FT, Douglas SH, Field SL, et al. Real-time PCR based on SYBR-Green I fluorescence: An alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. BMC Biotechnol. 2003 Oct 13;3:18.

137. Rodriguez CA, Avadhanula V, Buscher A, Smith AL, Iii JWSG, Adderson EE. Prevalence and Distribution of Adhesins in Invasive Non-Type b Encapsulated Haemophilus influenzae. Infect Immun. 2003 Apr 1;71(4):1635–42.

138. Hood DW, Deadman ME, Engskog MKR, Vitiazeva V, Makepeace K, Schweda EKH, et al. Genes required for the synthesis of heptose-containing

oligosaccharide outer core extensions in Haemophilus influenzae lipopolysaccharide. Microbiology. 2010 Nov 1;156(11):3421–31.

139. Martí-Lliteras P, López-Gómez A, Mauro S, Hood DW, Viadas C, Calatayud L, et al. Nontypable Haemophilus influenzae Displays a Prevalent Surface Structure Molecular Pattern in Clinical Isolates. PLOS ONE. 2011 Jun 16;6(6):e21133.

140. Staples M, Graham RMA, Jennison AV. Characterisation of invasive clinical Haemophilus influenzae isolates in Queensland, Australia using whole-genome sequencing. Epidemiol Infect. 2017 Jun;145(8):1727–36.

141. Beeton ML, Maxwell NC, Davies PL, Nuttall D, McGreal E, Chakraborty M, et al. Role of pulmonary infection in the development of chronic lung disease of prematurity. Eur Respir J. 2011 Jun 1;37(6):1424–30.

142. Pryhuber GS. Postnatal Infections and Immunology Affecting Chronic Lung Disease of Prematurity. Clin Perinatol. 2015 Dec;42(4):697–718.

143. Aulet de Saab OC, de Castillo MC, de Ruiz Holgado AP, de Nader OM. A comparative study of preservation and storage of Haemophilus influenzae. Mem Inst Oswaldo Cruz. 2001 May;96(4):583–6.

144. Kaijalainen T, Ruokokoski E, Ukkonen P, Herva E. Survival of Streptococcus pneumoniae, Haemophilus influenzae, and Moraxella catarrhalis Frozen in Skim Milk- Tryptone-Glucose-Glycerol Medium. J Clin Microbiol. 2004 Jan;42(1):412–4.

145. Votava M, Strítecká M. Preservation of Haemophilus influenzae and Haemophilus parainfluenzae at -70 degrees C. Cryobiology. 2001 Aug;43(1):85–7.

# 6. Appendices

## 6.1. Appendix A: HREC certificate

R14/49 Mr Boitshoko Mahlangu et al

**HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)**

**CLEARANCE CERTIFICATE NO. M170282**

| | |
|---|---|
| **NAME: (Principal Investigator)** | Mr Boitshoko Mahlangu et al |
| **DEPARTMENT:** | Immunology - School of Pathology<br>Respiratory and Meningeal Pathogens Research Unit |
| **PROJECT TITLE:** | Identification of Virulence Genes from Non-Typeable Haemophilus Influenzae Isolates Recovered from Pneumonia Cases and Community Controls |
| **DATE CONSIDERED:** | 24/02/2017 |
| **DECISION:** | Approved unconditionally |
| **CONDITIONS:** | |
| **SUPERVISOR:** | Prof Shabir Madhi |
| **APPROVED BY:** | Professor P Cleaton-Jones, Chairperson, HREC (Medical) |
| **DATE OF APPROVAL:** | 17/03/2017 |

**This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.**

**DECLARATION OF INVESTIGATORS**

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary in Room 301, Third Floor, Faculty of Health Sciences, Phillip Tobias Building, 29 Princess of Wales Terrace, Parktown, 2193, University of the Witwatersrand. I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report**. The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. In this case, the study was initially reviewed in February and will therefore be due in the month of February each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature Date

**PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES**