Strategies for pneumococcal conjugate vaccine use in humanitarian crises

Kevin van Zandvoort

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy of the University of London

November 2024

Department of Infectious Disease Epidemiology

Faculty of Epidemiology and Population Health

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

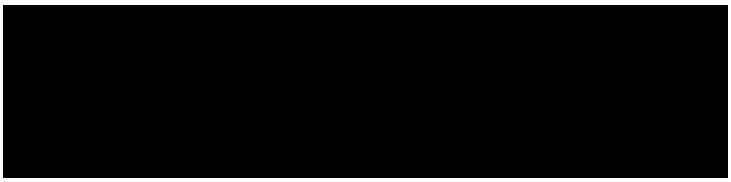
Funded by Elrha's Research for Health in Humanitarian Crises (R2HC) Programme, which aims to improve health outcomes by strengthening the evidence base for public health interventions in humanitarian crises. The R2HC programme is funded by the UK Government (DFID), the Wellcome Trust, and the UK National Institute for Health Research (NIHR).

Research group affiliation(s):
Centre for the Mathematical Modelling of Infectious Diseases
Health in Humanitarian Crises Centre

Declaration of authorship

I, Kevin van Zandvoort, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed,



05/10/2024

Abstract

Humanitarian crises substantially increase the health burden in affected populations. *Streptococcus pneumoniae* is a major cause of morbidity and mortality globally, and likely causes a considerable proportion of the burden in crisis-affected populations. Pneumococcal conjugate vaccines (PCVs) effectively prevent against pneumococcal disease but there is little guidance on their effective and feasible use during humanitarian crises, and they are rarely part of the humanitarian response. I combined primary data collection with mathematical modelling to assess the effect of PCV mass-vaccination campaigns in crisis-affected populations.

I first conducted a cross-sectional survey collecting data on social mixing, demography, pneumococcal carriage and related risk factors in a camp for internally displaced people in Somaliland. Social contact patterns were assortative by age with high rates of physical contacts. Most contacts were made at home and by school-aged children. Pneumococcal carriage prevalence was high, especially in children, and was similar to that in other high-transmission settings. Transmission was driven by children aged 2-5 and 6-14 years.

I then constructed a pneumococcal transmission model parameterized with the collected data, and assessed the effect of a PCV mass-vaccination campaign in different crisis typologies. Single-dose PCV campaigns vaccinating children aged six weeks to one year only have a limited impact. However, extending vaccination to children aged four years or older can partially disrupt the transmission of serotypes targeted by PCV and prevent a substantial proportion of invasive pneumococcal disease over two to three years. Expanded age eligibility may be needed to control transmission in settings with high migration or interaction with unvaccinated populations.

Overall, my research shows that single-dose PCV mass-vaccination campaigns could offer feasible and pragmatic protection against pneumococcal disease to crisis-affected populations where routine immunization may not be possible.

Acknowledgements

I am incredibly grateful for everyone who supported me during this PhD. There are too many people to list. I am especially thankful for the professional and personal support, mentorship, opportunities, motivation, and patience of my supervisors Stefan Flasche, Francesco Checchi, and Rosalind Eggo.

I am also very grateful for my other colleagues, collaborators, and co-authors who have done so much to allow this work to happen, including Catherine McGowan, Anna Carnegie, Bhargavi Rao, and Lucia Cassini at LSHTM, Catherine Satzke, Kim Mulholland, and Casey Pell at MCRI, Mohamed Bobe, Mohammed Saed, Saed Ibrahim, Rachael Cummings, and Mohammed Ismail at Save the Children, and Abdirahman Hassan, Mohamed Hergeye, and all data collectors at the Republic of Somaliland Ministry of Health Development. This work would not have been possible without them and many more. Abdirahman, Dr Bobe, Mohammed, and Saed have always made me feel welcome whenever I visited Hargeisa, for which I am very appreciative.

I also want to acknowledge other staff of the Somaliland TB hospital and MoHD cold-chain facility for storing and safeguarding the collected samples for an extended period, and other staff of the MCRI microbiological laboratory for testing the samples. I also acknowledge the Digaale community leaders for their support in community engagement and the overall community for their participation in the survey.

I also thank my other colleagues at CMMID and LSHTM, both current and former, for creating such an excellent and supportive working environment. I thank my friends, especially Christopher Jarvis, for their encouragement, advice, and distractions.

I thank my family for their love and support. Mum, Roy and Linda, Mr and Mrs Uche, Chidi, and Tohan and all others: thank you. To my dad who has not seen me complete this PhD, I am thankful that he could see me start. Finally, my biggest thanks go to my wife, Onyeka, who has been my rock.

Acronyms and abbreviations

ARI	acute respiratory infection
CCR	case-carrier ratio
CI	confidence interval
CMMID	Centre for Mathematical Modelling of Infectious Diseases
cRCT	cluster-randomised controlled trial
CrI	credible interval
DALY	disability-adjusted life year
DEzs	differential evolution Markov chain with snooker updater
EPI	expanded programme on immunization
FPC	finite population correction
GAM	global acute malnutrition
GE	genome equivalent
HHCC	Health in Humanitarian Crises Centre
HIV	human immunodeficiency virus
IDP	internally displaced people
INGO	international non-governmental organization
IPD	invasive pneumococcal disease
IQR	interquartile range
ISPPD	International Society of Pneumonia and Pneumococcal Diseases
LMIC	lower and middle-income countries
LSHTM	London School of Hygiene and Tropical Medicine
MCMC	Markov chain Monte Carlo
MCRI	Murdoch Children's Research Institute
MoHD	Ministry of Health Development
MSF	Médecins Sans Frontières
MUAC	middle-upper arm circumference
MVC	mass-vaccination campaign
NEps	non-encapsulated
NGM	next generation matrix
NGO	non-governmental organization
NNV	number needed to vaccinate
NVT	non-vaccine serotype
ODK	Open Data Kit
OR	odds ratio
PCM	phase change material
PCR	polymerase chain reaction

PCV	pneumococcal conjugate vaccine
PIRI	periodic intensification of routine immunisation
PoC	protection of civilian
PPSV	pneumococcal polysaccharide vaccine
R0	basic reproduction number
R2HC	Research for Health in Humanitarian Crises
SIS	susceptible-infectious-susceptible
SMART	standardised monitoring and assessment of relief and transitions
STGG	skim milk-tryptone-glucose glycerol
ULT	ultra-low temperature
UN WPP	United Nations World Population Prospects
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund
UNOCHA	United Nations Office for the Coordination of Humanitarian Affairs
UNRWA	United Nations Relief and Works Agency for Palestine Refugees in the Near East
VE	vaccine efficacy
VT	vaccine-targeted serotype
WHO	World Health Organization

Table of Contents

List of Figures

Excludes appendices

List of Tables

Excludes appendices

4. Effectiveness of pneumococcal vaccination campaigns in humanitarian settings: a modelling study

Manuscript

Supplemental material

Structure of the thesis

This thesis describes my research on strategies for pneumococcal conjugate vaccine (PCV) use in humanitarian crises. It is written in a *research paper style*, where every chapter is a published paper, preprint, or soon-to-be-submitted.

An introduction to the topic is provided in chapter 0, which includes background information on humanitarian crises, *Streptococcus pneumoniae*, pneumococcal conjugate vaccines, vaccination in humanitarian crises, and pneumococcal transmission models, as well as the overall aim and objectives of this thesis. Chapter 1 provides a narrative review that summarizes these concepts and proposes the evidence generation pathway used in this thesis: to inform effective PCV strategies by combining primary data collection with mathematical modelling. Chapter 2 presents a research paper describing the study design and results of a cross-sectional survey I designed, conducted, and analysed in a camp for internally displaced people (IDP) in Somaliland in 2019, which estimated the number and patterns of social contacts and prevalence of other risk factors for respiratory infections. Chapter 3 presents a research paper that describes the carriage prevalence and serotype distribution of pneumococci and their association with known risk factors, as estimated in the same cross-sectional survey. Chapter 4 consists of the final research paper introducing an epidemiological model for pneumococcal transmission, stratified by population, age, nutrition, and PCV vaccination, parameterized with the contact and demographic estimates from the Somaliland IDP camp and fitted to pneumococcal prevalence estimates using Bayesian inference. The transmission model is used to assess the effect and efficiency of PCV campaigns in the Somaliland IDP camp and other crisis typologies. The chapter also estimates the global potential demand for PCVs in humanitarian crises if strategies previously identified as optimal were to be implemented in crisis-affected populations worldwide. Finally, chapter 5 discusses the conducted research, and suggests ways forward on PCV vaccination strategies for populations affected by humanitarian crises.

Chapter 0

Introduction

Crisis-affected populations

A humanitarian crisis or emergency is typically described as an event that threatens the health, safety, security, and well-being of a community or large group of people. Crises can be caused by armed conflict including war, natural or human-made disasters including climate emergencies, food insecurity including famine, and infectious disease outbreaks[1–3]. The characteristics of every crisis are unique, but they often result in a breakdown of public health infrastructure, mass displacement, poor sanitation and access to water, and social and political instability[4].

The United Nations estimated 299.4 million people (3.7% of the world population) to be in need of humanitarian assistance worldwide in 2024[5]. At the end of 2022, an estimated 108.4 million people (1.4%) were forcibly displaced globally, including 35.3 million refugees and 62.5 million internally displaced people (IDP)[6]. The number of displaced people has increased from 42.8 million (0.6%) in 2012, and is projected to rise further during the next decades due to climate change and resulting food insecurity.[7]

Crises substantially affect people's lives and can dramatically increase avertable mortality. Estimates vary considerably between settings, but mortality rates are often increased by several folds compared to a pre-crisis baseline, especially in IDPs.[8,9] In most crises, these excess deaths are often attributable to the indirect effects of the crisis, including the breakdown of public health services affecting access to healthcare including vaccination coverage, food insecurity, inadequate water and sanitation, and overcrowding. These factors can increase both the incidence and severity of disease.

Infectious diseases are a particular concern. Epidemics of measles, cholera, shigellosis, meningococcal meningitis, yellow fever, and malaria are commonly reported.[10] The leading causes of under-5 mortality in crisis-affected populations are acute respiratory infections (ARI) and diarrhoeal diseases. A review by Bellos et al. found that 20 to 35% of the under-5 mortality was due to ARIs. They also found that the overall incidence and prevalence of ARIs were much greater compared to stable settings, with the largest excess risk in children 5-14 years old and adults.[11] A second review by Chen et al. found that pneumonia alone was attributable to 20% of child mortality, and that pneumonia incidence rates in children under 5 were 3 to 10 times higher than those in children in stable lower- and middle-income settings. They found that living in crowded settings, unsanitary conditions, being recently displaced, and being malnourished all further increased the risk of pneumonia.[12] There are very limited data on the aetiology of these respiratory infections, but likely common causes include *Streptococcus pneumoniae* (the pneumococcus) and respiratory syncytial virus.

Streptococcus pneumoniae

Streptococcus pneumoniae is a Gram-positive human commensal that commonly resides asymptomatically in the nasopharynx. It occasionally causes disease (e.g. pneumonia, meningitis, and sepsis), especially in young children and people with weakened immune systems.[13] Before the introduction of PCV, pneumococcal pneumonia was estimated to cause 85-95% of the total pneumococcal burden.[14] Although pneumococcal pneumonia has a relatively low case-fatality ratio (3% globally, 6% in Sub-Saharan Africa) compared to pneumococcal meningitis (44% globally, 61% in Sub-Saharan Africa) or other invasive pneumococcal disease (all other IPD; 31% globally, 44% in Sub-Saharan Africa), its incidence is far higher.[13]

As of April 2024, 105 distinct serotypes have been identified, with substantial variability in their spatial distribution[15] and potential to cause invasive disease. Despite regional differences, before the introduction of PCV in the world, only seven serotypes (1, 5, 6A, 6B, 14, 19F, and 23F) were responsible for over 50% of IPD cases.[16] The capsular polysaccharide on the cell surface is believed to be a main determinant of serotype-specific virulence as it modulates tissue invasiveness.[17] This is related to a phase variation process whereby pneumococcal colonies can shift between opaque phenotypes with thin cell walls and thick capsules, and transparent phenotypes with thick cell walls and thin capsules.[18] Different pneumococcal serotypes compete with each other for the same ecological niche. While co-colonization of multiple serotypes often occurs, pre-existing colonisation partially inhibits the acquisition of a new strain.[17]

Pneumococcal colonisation is a prerequisite for both pneumococcal disease and transmission. Natural immunity, a combination of both mucosal and systemic humoral components, is relatively short-lived, as it only protects against reacquisition with the same strain for up to 1 year, with little cross-protection among strains and serotypes.[17] Colonisation prevalence generally peaks in infants (children aged <12 months) and very young children, and decreases in older children as both serotype-specific and nonspecific natural immunity develop with age. For instance, prevalence was 74% in <1y olds, 72% in <5y, 25% in 5-14y, and 25% in 15 y olds in Nigeria.[19] Likewise, the duration of individual carriage episodes decreases with age, though this again is serotype-specific[20,21].

Colonisation is facilitated by inflammation in the upper respiratory tract, particularly following viral infections. Inflammatory periods also increase the density of carriage, which in turn is associated with increased shedding and transmission.[22] Pneumococci are predominantly thought to be transmitted through close contact-dependent exposure to secretions of carriers, though the additional contribution of airborne transmission is unclear.[17]

Besides age and concurrent viral infections, several factors can increase the risk of pneumococcal carriage, and could also shift the age distribution of carriage. These include pre-existing conditions including chronic lung disease and immune-suppressive conditions such as HIV infection, anaemia, and sickle cell disease. Malnutrition likely has a similar effect, and, in terms of its global burden, also occurs disproportionately in children.[23] Environmental factors, such as exposure to cold[24] (poorly insulated housing), outdoor air pollution[25], indoor air pollution[26], and exposure to tobacco smoke[27], may affect susceptibility to acquisition through inflammation and an accumulation of mucus.[28] Crowded living conditions increase the rates at which pneumococci are transmitted. By contrast, personal hygiene and hand washing can be protective against carriage.[28]

Progression to disease and associated case-fatality ratios are predominantly affected by most of the same risk factors[29–31]. The same impaired immune responses that make immunocompromised children susceptible to acquisition lead to high rates of invasion and mortality. HIV-positive children have a 20 times higher risk of IPD compared to HIV-negative children[32], and have a 17-fold higher case-fatality ratio for lower respiratory tract infection.[33] Although older adults generally have low carriage prevalence[34–36], indicating good mucosal immunity against carriage, they experience high rates of invasive disease if colonised.[37]

Pneumococcal carriage prevalence can be assessed by analysing nasal, nasopharyngeal or oropharyngeal swabs or nasopharyngeal aspirates for the presence of pneumococci.[38] Nasopharyngeal swabs are most sensitive in young, healthy children,[38,39] and are often used in studies that aim to assess population-level prevalence.[35,36,40,41] Serotype-specific prevalence is typically of interest, and different microbiological methods can be used to detect the pneumococcal serotypes present. A comparison of various methods found that microarray with a culture amplification step of pneumococcal positive samples is the most accurate method in detecting serotypes, including those carried at lower density.[42]

While the burden of pneumococcal disease, just as the aetiology of ARIs, is largely undocumented in crisis settings, it is likely substantial. Estimates from the Global Burden of Disease Study attribute more than half of all ARI-related deaths worldwide to pneumococci before the introduction of PCVs.[43] In 2000, pneumococcus was responsible for an estimated 11% of all mortality in children under five,[14] and despite marked reductions, a significant pneumococcal burden remains after the introduction of PCVs.[44,45] The generally increased presence of risk factors known to increase pneumococcal carriage, transmission, disease, and mortality may amplify the pneumococcal burden in crisis settings.

Pneumococcal meningitis outbreaks have been reported in populations in the Central African Republic, which have been affected by protracted humanitarian crises for many years[46,47]

These outbreaks both occurred within the African meningitis belt, and specific risk factors may have contributed to these outbreaks. Pneumococcal disease outbreaks have been reported to occur in crowded settings, including long-term care facilities, military settings, childcare settings, homeless shelters, and jails, and crowding is especially common in crisis-affected populations living in camps.[48] Pneumococcal outbreaks may occur regularly in crisis-affected populations, identified as outbreaks of meningitis or more commonly pneumonia, but go unrecognized due to limited surveillance and diagnostic capacity, and other challenges in identifying the causative agent.[49,50]

Pneumococcal vaccines

Two types of pneumococcal vaccines are used to protect against pneumococcal disease, and both only protect against a subset of all serotypes. Pneumococcal polysaccharide vaccines have been used since the 1940s and are currently available as a 23-valent vaccine (PPSV23). PPSV23s are efficacious against disease caused by vaccine-targeted serotypes in adults and older children, but not in children <2 years. They are not considered to be effective against colonization.[51]

Pneumococcal Conjugate Vaccines (PCVs) were developed in the 1990s and have been widely introduced in routine childhood immunisation programmes since the 2000s[13]. In PCVs, the pneumococcal capsular polysaccharides are conjugated to a protein such as a diphtheria or tetanus toxoid, or a *Haemophilus influenzae*-derived protein. This enhances the immune response, particularly in young children, partly accounting for their efficacy in this age group. In addition to the direct protection against pneumococcal disease, PCVs also elicit indirect protection through interrupted acquisition and colonisation of vaccine-targeted serotypes (VT), which reduces their community transmission.[13]

There are currently five PCV products licensed for use in children, with valency against 10 (PCV10: Synflorix and Pneumosil), 13 (PCV13, Prevenar 13), 15 (PCV15, Vaxneuvance), and 20 (PCV20, Prevenar 20) serotypes. All PCVs are highly efficacious against pneumococcal disease. The pooled vaccine efficacy (VE) against VT-IPD is 89% (73 - 96), whilst the pooled efficacy against radiologically confirmed pneumonia (not pneumococcus specific) was 32% (24 - 39).[52] VE against pneumococcal carriage is lower than VE against VT-IPD. The pooled VE against carriage of the seven respective VT serotypes following three infant doses of the PCV7 vaccine (now replaced by PCV13) was estimated as 62% (95%CrI 52 – 72) four months after the final dose, with moderate waning of efficacy thereafter.[53] PCV10, PCV13, PCV15, and PCV20 have been shown to be non-inferior in immunogenicity terms to PCV7 for all seven serotypes, and to increase immunogenicity for the additional serotypes in these vaccines.[54–57]

PCVs are very safe. While relatively mild adverse events, including irritability, pain, fever, drowsiness, tenderness, decreased appetite, and rashes, are common, severe adverse reactions are extremely rare. However, as serotypes compete with one another, the use of PCV frees up an ecological niche for non-vaccine serotypes (NVT), resulting in replacement of VTs by NVTs. This dampens the impact of PCVs, but NVTs are generally less likely to cause severe disease, resulting in a net benefit.[58] The ongoing development of ever-higher valency vaccines is in part an attempt to protect against particularly virulent replacement serotypes.

PCVs have now been introduced in the routine childhood immunisation programmes of the majority of countries, with only 38/194 countries not having rolled out PCV as of 2022.[59] In most places where PCVs are used at high coverage, marked reductions in VT carriage have been observed, including in age groups not eligible for vaccination, indicating strong indirect vaccine effects.[40,60–62] As a result, PCVs have significantly impacted pneumococcal disease, though some of their impact is limited by serotype replacement.[45]

Vaccination in humanitarian crises

Vaccines are an essential tool to reduce the high infectious disease burden in populations affected by humanitarian crises. In stable settings, routine immunisation programmes can effectively deliver a wide range of vaccines using an optimal delivery schedule, usually at different ages. This is highly effective and cost-effective.[63] However, routine immunisation can often be challenging to implement during crises due to disruptions of the cold chain, lack of trained personnel to deliver the vaccines, reduced safety of health care workers, and intermittent access to the affected population.[64,65] As a result, established routine immunisation often breaks down during crises, while displaced people may move to informal camps where no pre-existing health services are available. In the absence of vaccination, vaccination coverage may reduce to levels below the herd immunity threshold, substantially increasing the risk of infectious disease outbreaks. The United Nations Children's Fund (UNICEF) estimates that two-thirds of all zero-dose children live in countries affected by conflict.[66]

Mass-vaccination campaigns (MVC) can be used as an alternative or supplementary strategy to routine immunisation where routine immunisation is not feasible or inadequate, and/or where rapid conferment of immunity is warranted to prevent crisis-attributable excess morbidity and mortality. During an MVC, a large group of people, usually of a wide age range, is vaccinated in a short timeframe. In settings where routine immunisation is not feasible or inadequate, humanitarian actors such as non-governmental organisations

(NGOs) often use MVCs in humanitarian settings either preventatively or reactively to rapidly increase immunity to measles, cholera, and other epidemic-prone infectious disease threats. In 2012, the World Health Organization (WHO) introduced a Framework for Decision-Making on Vaccination in Humanitarian Emergencies, which was updated in 2017.[67] This framework aims to support humanitarian decision-makers in implementing the most appropriate vaccine interventions given the local epidemiology, vaccine characteristics, and other context-specific considerations such as the available cold chain. The framework emphasises expanding the range of vaccines offered to crisis-affected populations, but also recommends the use of adapted vaccination strategies, including expanded age ranges and reduced-dose regimens.

A review by Leach and Checchi on vaccine use in recent humanitarian crises found that the range of vaccines used in humanitarian crises was limited. The vaccines most commonly used in crises were measles and polio, which were primarily provided as MVCs. Other vaccines, however, including PCV and rotavirus vaccines, were predominantly offered using routine immunisation, or not at all.[68]

Grais et al reviewed measles vaccination in humanitarian crises.[69] Measles is prioritized during humanitarian crises as it is highly transmissible and requires high levels of population immunity to control. Case-fatality ratios of measles are often exacerbated in crises due to malnutrition, with estimates as high as 20-30%.[70] The review found that the age-range of affected cases during measles outbreaks was often extended beyond 5 years of age, highlighting the importance of including children up to the age of 14 in the MVCs. There was a clear lack of reports for measles vaccination activities in crises where no outbreak happened. In crises where the outbreaks were reported, the outbreak often ceased shortly after an MVC. However, the impact of measles campaigns is rarely quantified, in part due to security concerns.[69,71] Vaccine efficacy estimates for measles containing vaccines have been estimated at similar levels in crisis-affected settings compared to stable settings.[72] Modelled impact of reactive measles MVCs estimate that these may have prevented 97% of measles cases that would have occurred in the absence of the campaigns in Rohingya refugee camps in 2017.[73]

Eradication efforts of poliomyelitis are threatened by outbreaks of wild and vaccine-derived poliovirus, which have occurred predominantly in crisis-affected countries in recent years.[74] Polio vaccination campaigns are regularly used in countries affected by humanitarian crises, and are often conducted at a national or subnational scale.[68] High vaccination coverage has been achieved in campaigns in multiple countries, including in displaced populations in

Bangladesh, Nigeria, and Kenya with over 80%, 90%, and 95% coverage respectively.[75,76] VE estimates of OPV doses administered in Somalia range between 70 - 95%[77].

Cholera outbreaks occur regularly in crisis-affected populations, particularly those living in large overcrowded camps with poor water, sanitation, and hygiene.[78,79] Mortality due to cholera in humanitarian settings has dramatically reduced following standardized approaches for prevention, especially in refugee camps under UNHCR mandate. Although oral cholera vaccines are affected by stockpile issues, preventative vaccination has been an important element of cholera control strategies.[78,80] Reactive single-dose vaccination campaigns are commonly used to control outbreaks, and the effectiveness of a single-dose has been estimated as 87% in South Sudan.[81] Modelled evidence suggests that a reactive use of single-dose oral cholera vaccines in combination with other targeted interventions would have been able to reduce cholera cases by 81% in Chad, and would have effectively controlled a cholera outbreak in the Democratic Republic of Congo.[82,83]

In general, while vaccines are commonly used in crises, evaluations of vaccination activities in humanitarian settings are rare.[68] A review found only ten percent of reactive campaigns to be evaluated, with evaluations even rarer for preventative vaccination activities.[10] As illustrated by measles, polio, and cholera campaigns, high vaccination coverages can be achieved in populations affected by crises, with vaccine efficacy estimates similar to those in stable populations, and likely substantial impact on the health burden.

Pneumococcal conjugate vaccines in humanitarian crises

Despite the high effectiveness of PCVs, and a likely substantial preventable pneumococcal burden, PCVs are underutilised in populations affected by humanitarian crises. Compared to other vaccines, PCVs are relatively expensive, and their price has long been a barrier to accessing PCV for both national ministries of health and NGOs. In the absence of Gavi support, lower and middle-income countries used to spend 20 and 50 times more for one full PCV10 (Synflorix) or PCV13 regimen compared to a complete regimen of measles and oral polio vaccines, and this used to be indicative of prices paid by humanitarian actors.[84] Since 2017, however, a “Humanitarian Mechanism” sponsored by the WHO, UNICEF, Médecins Sans Frontières, and Save the Children guarantees more affordable PCV procurement of PCV10 (Synflorix) and PCV13, and expedited delivery for humanitarian actors.[85] In addition, the substantially more affordable Pneumosil (PCV10) has been WHO prequalified since 2019.[86,87]

A remaining barrier to using PCV is insufficient evidence on effective and logistically feasible PCV deployment strategies via MCVs, and their expected impact in crises.[88,89] Routinely, most countries use a 3-dose schedule to deliver PCVs as recommended by the WHO, with the first primary dose given at six weeks of age, and at least four weeks between doses. In addition to routine vaccination, catch-up campaigns have been used to accelerate the indirect effect of PCV at the time of vaccine introduction in children aged 1 to 5 years. The WHO recommends that children <2 years of age should be prioritised in such catch-up campaigns because of their higher risk for pneumococcal disease. In addition, they recommend age-appropriate vaccination schedules to be used in humanitarian settings for children <1 year of age and to be considered for children ≤5 years of age. These recommendations are in line with the age groups with the highest burden of pneumococcal disease, but are not necessarily the best way to deliver PCVs. The need for multiple doses in complete PCV schedules also remains a logistical challenge in many humanitarian settings.

There are few known instances of PCVs being administered through MVCs during a crisis.[68] Where PCV MCVs have been used, this has been done using different target age groups[90–93]: in Syria and Niger, PCVs were administered to children up to 5 years of age; In South Sudan, children up to 23 months were eligible, whilst in Niger and the Central African Republic, children aged 11-23 months were vaccinated.[94] The impact and efficiency of these different strategies have not been assessed.

The limited evidence and guidance on the effective use of PCV MVCs is likely a contributing factor to their limited use. Current guidelines mainly focus on the age groups experiencing the highest burden of disease, namely optimising direct protection to those at highest risk. However, it is possible that such an approach would result in providing PCVs to an age group that is too narrow to achieve sufficient levels of herd immunity for substantial impact. Similarly, if the target age group is too broad, many vaccinees would already be protected through indirect effects, and valuable resources may could instead be directed elsewhere. By achieving high vaccination coverage among the main transmitters in a population, a PCV campaign could theoretically rapidly interrupt VT transmission, and sustain these effects for some time, offering indirect protection to those unvaccinated during this period.

Modelling pneumococcal transmission

Transmission models can be used to understand transmission disease dynamics[96–98], including the relative contribution of different age groups to transmission, and thus the direct and indirect impacts of a candidate vaccination strategy[99,100]. Models are a relatively inexpensive method and can be used to simulate and compare many different scenarios,

including alternative vaccination strategies. They also allow combining evidence from different sources, and can thereby be used with minimal but adequate data.[101] Models can thus bridge evidence gaps that may otherwise only be studied with challenging trials that are likely unfeasible to conduct in many humanitarian crises.

The large number of serotypes and the competition among serotypes make modelling pneumococci challenging.[102] Large compartmental models or individual-based models could be used to model individual serotypes, but this is computationally challenging, and there is a lack of serotype-specific data for key epidemiological parameters. Instead, serotypes are often grouped together as VTs and NVTs.[100,103–106]

As pneumococci do not elicit lifelong immunity, adaptations of the susceptible-infectious-susceptible (SIS) model are typically used in compartmental models.[99,100,103,105,106] Pneumococcal serotypes differ in prevalence and transmissibility, and a mechanism for competition between serotypes must be incorporated to model their stable coexistence.[107] Different model structures with different mechanisms have been proposed. Figure 1 shows a non-exhaustive number of model structures reported in the literature, in increasing order of complexity. All are SIS models, where individuals are either Susceptible (S) or carriers of at least one serotype (I) in a specific group of mutually exclusive serotypes. These groups are denoted as A, B, or C in the model structures presented here. Model I is the simplest model structure, which accounts for the carriage of a serotype in either group A or B, but results in competitive exclusion of the serotype with the highest basic reproduction number.[108–110] Model II is an extension that includes competition among serotypes in different groups, allowing for coexistence of both serotypes, but not for co-colonization[34,111–114]. Model III is a further extension with superinfection of heterologous carriage (i.e. simultaneous carriage of both A and B)[99,100,103,112,115,116]. Model IV further allows for superinfection with multiple lineages of the same serotype[117,118]. Models V[112] and VI[106,119] are variations of model III with an additional serotype, where model VI allows for superinfection with all three groups. Individual-based models are often used to model additional complexity in serotypes,[120–122] though a model comparison showed that the incremental benefit of this additional complexity is limited.[102] Additional complexities, such as stratification by age, vaccination status, and other factors, can be added to the model structures as needed.

Pneumococcal serotypes differ in many aspects, including their duration of carriage[21] and propensity to cause invasive disease[123]. It is not known to what extent these and other parameters are responsible for the observed coexistence and difference in frequency of many serotypes. Unlike model structures I, II, and IV, Model III and its variants with increased serotypes (V and VI) are not structurally neutral, meaning that they implicitly

incorporate an unknown mechanism of coexistence.[107] These model structures should, therefore, not be used with the aim of understanding the biological mechanisms causing the coexistence of different strains. However, they have widely been used to successfully model and project the impact of PCVs on pneumococcal transmission.[99,100,103,112,115,116]

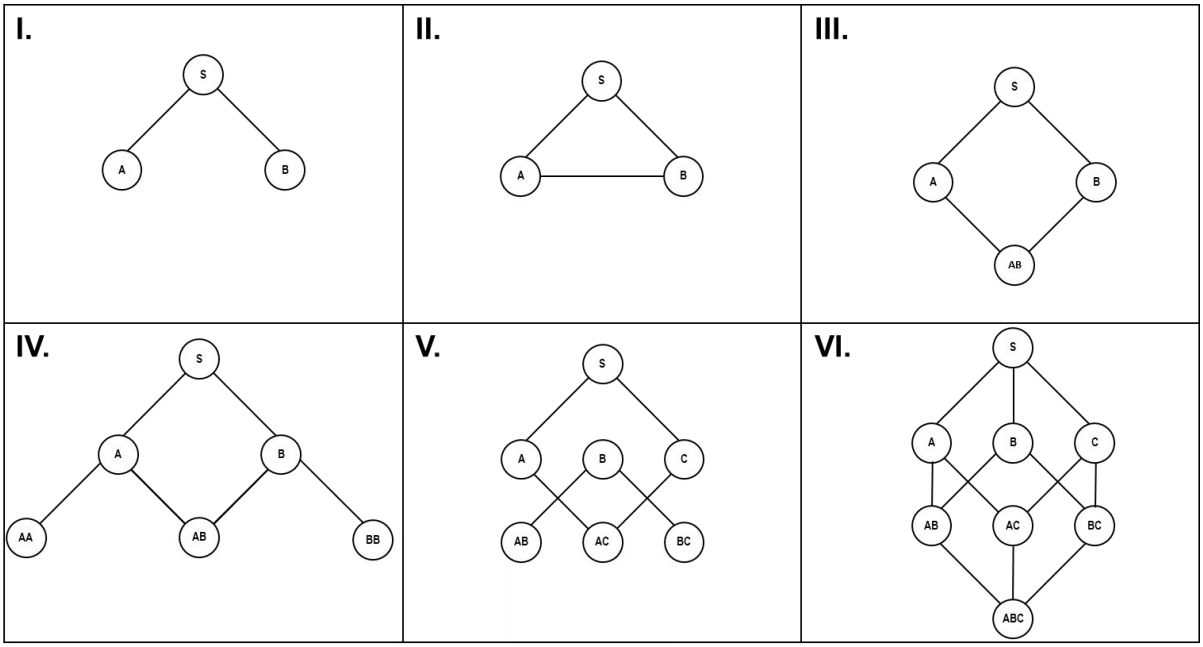


Figure 1. Variants of compartmental susceptible-infectious-susceptible type model structures commonly used to model pneumococcal transmission. In all model structures, people can be susceptible (S) or carriers of serotypes in group A, B, or C. Typically, models group serotypes as vaccine types (VT) or non-vaccine types (NVT) by their inclusion in specific PCV products. Model types I-IV only model two serotype groups, where model type III allows for co-colonization with serotypes in both groups A and B. Models do not explicitly model co-colonization of serotypes within the same group, with the exception of model IV, where a person could be colonized by one or at least two serotypes of a specific group. Models V and VI include a third serotype group C, where model VI allows for cocolonization of all serotype groups. Directional arrows are not shown for clarity, but edges in all model structures represent bidirectional movement where carriage can be acquired (transmission) and lost (clearance).

Model fitting

Model fitting is the process of finding plausible model parameter input values such that resulting model outcomes closely replicate observed data points. Bayesian methods such as Markov Chain Monte Carlo (MCMC) are often used to fit infectious disease models to data and aim to sample values from an unknown posterior distribution of one or several model parameters.[124]

The posterior distribution, or the probability distribution of the fitted parameters, can be calculated using Bayes' Theorem as:

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)} \quad (1)$$

where $\theta = \{a, b, \ldots\}$ is the set of parameters for which the (joint) posterior distribution is calculated, $p(D|\theta)$ is the likelihood, or the probability of the observed data given the model output using parameters θ given the observed data D, $p(\theta)$ is the prior, or the probability of the parameter values based on a prior belief of these parameter values.[125] The normalisation constant $p(D)$ is often ignored as it cancels out when calculating the acceptance ratio, so that:

$$p(\theta|D) \propto p(D|\theta)p(\theta) \quad (2)$$

is used to generate samples from the posterior distribution.

In theory, equation 2 could be solved for all possible values of θ, but that is very inefficient and often impossible due to a large number of fitted variables. Instead, an MCMC sampler can efficiently sample from the posterior distribution. The simplest MCMC sampler is the Metropolis-Hastings algorithm[126], which uses the following process. Here, θ^i denotes the ith sample:

1. Start with a set of parameter values for $\theta = \theta^0$, e.g. by randomly sampling from the prior distribution.
2. Propose a new set of parameter values based on the current values, e.g. $\theta' = \theta + \varepsilon$, where ε may be sampled from random variable Q, and calculate $p(\theta'|D)$.
3. Set $\theta^1 = \begin{cases} \theta', & p(\theta'|D) > p(\theta|D) \\ \theta', & p(\theta'|D) \leq p(\theta|D) \wedge \tau \leq \min(1, p(\theta'|D)/p(\theta|D)) \\ \theta, & p(\theta'|D) \leq p(\theta|D) \wedge \tau > \min(1, p(\theta'|D)/p(\theta|D)) \end{cases}$
   where τ is a sampled value between 0 and 1. I.e. if the proposal values improve the posterior, θ^1 will be set to θ'. If the proposal does not improve the posterior, θ^1 may still be set to θ' with probability $p(\theta'|D)/p(\theta|D)$, and remains at θ otherwise.
4. Repeat steps 2 and 3 to generate a large set of samples.

Due to the stochastic process of sampling new values, an MCMC sampler will, in theory, explore the entire parameter space and always converge to the true posterior distribution when $i \to \infty$.[127]However, the sampler needs to be terminated some point. Several diagnostics can be used to assess whether the sampler may not have converged to the true posterior, but it is not possible to prove whether the sampler has converged to the true posterior.[128] In addition, convergence can be slow, particularly when θ has a high number of dimensions, when the posterior distribution is very wide, and when a single evaluation of the

posterior is relatively expensive, as is the case with many mechanistic models. Several different MCMC samplers have been developed to improve the speed of convergence for those challenges. In this thesis, I have used the Differential Evolution Markov Chain with snooker update, which is an adaptive algorithm that uses a population of multiple MCMC chains ran in parallel.[129] It works well in moderate dimensions with multi-modal distributions. The parallel chains make the sampler computationally efficient to reach convergence, and improve exploration of the global parameter space. Its adaptive nature limits the need to manually tune certain parameters, and it does not require computationally intensive calculation of the gradient of the likelihood as in Hamiltonian Monte Carlo.[130] The sampler also has good integration in existing MCMC frameworks.[131]

Parameterizing transmission model

Transmission models need to be adequately parameterized in order to make useful and realistic projections. While evidence from different sources can be used, these still need to be representative for the local epidemiological characteristics. The structure of a model should also include characteristics important to the settings and research question, such as migration in the case of displaced populations. In general, there is a lack of high-quality data on key parameters from crisis-affected populations, such as prevalence of pneumococcal carriage, incidence of pneumococcal disease and deaths, and transmission patterns. New data could be collected from at least one crisis-affected population to overcome some of these gaps. Estimates of social contacts can be used to understand mixing between age groups, a proxy for transmission, which may be different in camp-like settings compared to a stable host population. Nasopharyngeal swabs can be collected to accurately determine pneumococcal carriage prevalence in different age groups. Ideally, estimates would be collected from different settings representing different crisis typologies, such as large refugee camps, or urban and rural settings where displaced individuals mix with those in a host population, though that is resource intensive. Alternatively, a model that is parameterized well for at least one crisis-affected setting can be used to explore crises of other typologies, as different settings may require different vaccination strategies. Finally, a model able to realistically simulate the effect of PCV campaigns in a range of different crisis-affected settings can be used to quantifying the global need for humanitarian doses, in order to assist efforts to mobilise resources and plan adequate supply as part of the Humanitarian Mechanism or other arrangements.

Overall aims and objectives

This PhD thesis aims to identify strategies for pneumococcal conjugate vaccine use in humanitarian crises. I hypothesised that PCV delivery strategies that rely on indirect protection and can interrupt VT transmission for an extended period would be required in crises where routine PCV delivery is not feasible. Such strategies would be able to provide both direct protection to vaccine recipients and indirect protection to those unable to be vaccinated, including those born after vaccination delivery. I further hypothesised that this would require vaccination of age groups driving transmission in a population, in addition to those experiencing the highest pneumococcal burden.

The following main objectives were pursued to achieve the aim of this PhD:

1. To review the literature on the pneumococcal burden and use of pneumococcal conjugate vaccines in humanitarian crises in order to lay out the evidence generation pathway used in this thesis (chapter 1).
2. To conduct a study in a population affected by a humanitarian crisis with the aim of parameterising a pneumococcal transmission model that is realistic for this setting: this included
    a. Assessing social contact patterns and the prevalence of other risk factors relevant for transmission of pneumococci and other respiratory infections in a crisis-affected population (chapter 2);
    b. Assessing the pneumococcal carriage prevalence and serotype distribution in a crisis-affected population, as well as risk factors for carriage (chapter 3);
3. To develop a pneumococcal transmission model using the collected data to model the effect of different pneumococcal conjugate vaccination campaign strategies in various crisis-affected populations, and quantify the likely global requirement for ‘humanitarian’ PCV doses if these strategies are implemented at scale (chapter 4).

Reference list

1 Ager A, Burnham G, Checchi F, *et al.* Strengthening the evidence base for health programming in humanitarian crises. *Science* 2014; **345**: 1290–2.

2 Humanitarian Coalition. What Is a Humanitarian Emergency? https://www.humanitariancoalition.ca/what-is-a-humanitarian-emergency (accessed April 24, 2024).

3 Bloxham L. What is a humanitarian crisis? Concern Worldwide. 2022; published online Jan 26. https://www.concern.org.uk/news/what-is-a-humanitarian-crisis (accessed April 24, 2024).

4 Inter-Agency Standing Committee. Definition of Complex Emergency | IASC. 1994. https://interagencystandingcommittee.org/content/definition-complex-emergency (accessed April 24, 2024).

5 UN OCHA. Global Humanitarian Overview 2024. 2023; published online Dec 11. https://www.unocha.org/publications/report/world/global-humanitarian-overview-2024-enarfres (accessed April 24, 2024).

6 UNHCR. Global Trends. https://www.unhcr.org/global-trends (accessed April 24, 2024).

7 UNHCR. Climate Action: Focus Area Strategic Plan for 2024 – 2030. https://reporting.unhcr.org/climate-action-focus-area-strategic-plan-2024-2030 (accessed May 28, 2024).

8 Heudtlass P, Speybroeck N, Guha-Sapir D. Excess mortality in refugees, internally displaced persons and resident populations in complex humanitarian emergencies (1998–2012) – insights from operational data. *Conflict and Health* 2016; **10**: 15.

9 Checchi F, Roberts L. Interpreting and using mortality data in humanitarian emergencies. A primer for non-epidemiologists. 2005.

10 Warsame A, Murray J, Gimma A, Checchi F. The practice of evaluating epidemic response in humanitarian and low-income settings: a systematic review. *BMC Medicine* 2020; **18**: 315.

11 Bellos A, Mulholland K, O'Brien KL, Qazi SA, Gayer M, Checchi F. The burden of acute respiratory infections in crisis-affected populations: a systematic review. *Conflict and Health* 2010; **4**: 3.

12 Chen SJ, Walker PJ, Mulholland K, Graham HR. Childhood pneumonia in humanitarian emergencies in low- and middle-income countries: A systematic scoping review. *J Glob Health*; **12**: 10001.

13 Klugman KP, Dagan R, Malley R, Whitney CG. Pneumococcal conjugate vaccine and pneumococcal common protein vaccines. In: Plotkin's Vaccines, 7th edn. Philadelphia: Elsevier, 2018.

14 O'Brien KL, Wolfson LJ, Watt JP, *et al.* Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates. *The Lancet* 2009; **374**: 893–902.

15 Weinberger DM, Trzciński K, Lu Y-J, *et al.* Pneumococcal Capsular Polysaccharide Structure Predicts Serotype Prevalence. *PLOS Pathogens* 2009; **5**: e1000476.

16 Johnson HL, Deloria-Knoll M, Levine OS, *et al.* Systematic Evaluation of Serotypes Causing Invasive Pneumococcal Disease among Children Under Five: The Pneumococcal Global Serotype Project. *PLoS Medicine* 2010; **7**: e1000348.

17 Weiser JN, Ferreira DM, Paton JC. *Streptococcus pneumoniae*: transmission, colonization and invasion. *Nature Reviews Microbiology* 2018; **16**: 355–67.

18 Kim JO, Weiser JN. Association of intrastrain phase variation in quantity of capsular polysaccharide and teichoic acid with the virulence of *Streptococcus pneumoniae*. *J Infect Dis* 1998; **177**: 368–77.

19 Adetifa IMO, Antonio M, Okoromah CAN, *et al.* Pre-Vaccination Nasopharyngeal Pneumococcal Carriage in a Nigerian Population: Epidemiology and Population Biology. *PLoS One* 2012; **7**. DOI:10.1371/journal.pone.0030548.

20 Lees JA, Croucher NJ, Goldblatt D, *et al.* Genome-wide identification of lineage and locus specific variation associated with pneumococcal carriage duration. *eLife* 2017; **6**: e26255.

21 Lipsitch M, Abdullahi O, D'Amour A, *et al.* Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal serotypes in Kenya with a Markov transition model. *Epidemiology* 2012; **23**: 510–9.

22 Siegel SJ, Roche AM, Weiser JN. Influenza promotes pneumococcal growth during coinfection by providing host sialylated substrates as a nutrient source. *Cell Host Microbe* 2014; **16**: 55–67.

23 Rytter MJH, Kolte L, Briend A, Friis H, Christensen VB. The Immune System in Children with Malnutrition—A Systematic Review. *PLOS ONE* 2014; **9**: e105017.

24 Numminen E, Chewapreecha C, Turner C, *et al.* Climate induces seasonality in pneumococcal transmission. *Sci Rep* 2015; **5**. DOI:10.1038/srep11344.

25 Sahuquillo-Arce JM, Ibáñez-Martínez E, Hernández-Cabezas A, *et al.* Influence of environmental conditions and pollution on the incidence of *Streptococcus pneumoniae* infections. *ERJ Open Res* 2017; **3**. DOI:10.1183/23120541.00014-2017.

26 Gordon SB, Bruce NG, Grigg J, *et al.* Respiratory risks from household air pollution in low and middle income countries. *Lancet Respir Med* 2014; **2**: 823–60.

27 Greenberg D, Givon-Lavi N, Broides A, Blancovich I, Peled N, Dagan R. The contribution of smoking and exposure to tobacco smoke to *Streptococcus pneumoniae* and *Haemophilus influenzae* carriage in children and their mothers. *Clin Infect Dis* 2006; **42**: 897–903.

28 Mulholland K, Weber M. Pneumonia in Children: Epidemiology, Prevention and Treatment. 2016.

29 Bello S, Menéndez R, Antoni T, *et al.* Tobacco smoking increases the risk for death from pneumococcal pneumonia. *Chest* 2014; **146**: 1029–37.

30 Balakrishnan I, Crook P, Morris R, Gillespie SH. Early Predictors of Mortality in Pneumococcal Bacteraemia. *Journal of Infection* 2000; **40**: 256–61.

31 Kim PE, Musher DM, Glezen WP, Barradas MCR, Nahm WK, Wright CE. Association of Invasive Pneumococcal Disease with Season, Atmospheric Conditions, Air Pollution, and the Isolation of Respiratory Viruses. *Clin Infect Dis* 1996; **22**: 100–6.

32 von Gottberg A, Cohen C, de Gouveia L, *et al.* Epidemiology of invasive pneumococcal disease in the pre-conjugate vaccine era: South Africa, 2003–2008. *Vaccine* 2013; **31**: 4200–8.

33 Madhi SA, Kuwanda L, Cutland C, Klugman KP. The Impact of a 9-Valent Pneumococcal Conjugate Vaccine on the Public Health Burden of Pneumonia in HIV-Infected and -Uninfected Children. *Clinical Infectious Diseases* 2005; **40**: 1511–8.

34 Melegaro A, Choi Y, Pebody R, Gay N. Pneumococcal carriage in United Kingdom families: estimating serotype-specific transmission parameters from longitudinal data. *Am J Epidemiol* 2007; **166**: 228–35.

35 Nackers F, Cohuet S, le Polain de Waroux O, *et al.* Carriage prevalence and serotype distribution of *Streptococcus pneumoniae* prior to 10-valent pneumococcal vaccine introduction: A population-based cross-sectional study in South Western Uganda, 2014. *Vaccine* 2017; **35**: 5271–7.

36 Hill PC, Akisanya A, Sankareh K, *et al.* Nasopharyngeal Carriage of *Streptococcus pneumoniae* in Gambian Villagers. *Clin Infect Dis* 2006; **43**: 673–9.

37 Adler H, Ferreira DM, Gordon SB, Rylance J. Pneumococcal Capsular Polysaccharide Immunity in the Elderly. *Clin Vaccine Immunol* 2017; **24**: e00004-17.

38 Rapola S, Salo E, Kiiski P, Leinonen M, Takala AK. Comparison of four different sampling methods for detecting pharyngeal carriage of *Streptococcus pneumoniae* and *Haemophilus influenzae* in children. *J Clin Microbiol* 1997; **35**: 1077–9.

39 O'Brien KL, Nohynek H, World Health Organization Pneumococcal Vaccine Trials Carriage Working Group. Report from a WHO Working Group: standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*. *Pediatr Infect Dis J* 2003; **22**: e1-11.

40 Hammitt LL, Akech DO, Morpeth SC, *et al.* Population effect of 10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of *Streptococcus pneumoniae* and non-typeable *Haemophilus influenzae* in Kilifi, Kenya: findings from cross-sectional carriage studies. *The Lancet Global Health* 2014; **2**: e397–405.

41 Heinsbroek E, Tafatatha T, Phiri A, *et al.* Pneumococcal carriage in households in Karonga District, Malawi, before and after introduction of 13-valent pneumococcal conjugate vaccination. *Vaccine* 2018; **36**: 7369–76.

42 Satzke C, Dunne EM, Porter BD, Klugman KP, Mulholland EK, PneuCarriage project group. The PneuCarriage Project: A Multi-Centre Comparative Study to Identify the Best Serotyping Methods for Examining Pneumococcal Carriage in Vaccine Evaluation Studies. *PLOS Medicine* 2015; **12**: e1001903.

43 Troeger C, Blacker B, Khalil IA, *et al.* Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Infectious Diseases* 2018; **18**: 1191–210.

44 O'Brien KL, Baggett HC, Brooks WA, *et al.* Causes of severe pneumonia requiring hospital admission in children without HIV infection from Africa and Asia: the PERCH multi-country case-control study. *The Lancet* 2019; **394**: 757–79.

45 Wahl B, O'Brien KL, Greenbaum A, *et al.* Burden of *Streptococcus pneumoniae* and *Haemophilus influenzae type b* disease in children in the era of conjugate vaccines: global, regional, and national estimates for 2000–15. *The Lancet Global Health* 2018; **6**: e744–57.

46 Coldiron ME, Touré O, Frank T, Bouygues N, Grais RF. Outbreak of Pneumococcal Meningitis, Paoua Subprefecture, Central African Republic, 2016–2017. *Emerg Infect Dis* 2018; **24**: 1720–2.

47 Crellen T, Rao VB, Piening T, Zeydner J, Siddiqui MR. Seasonal upsurge of pneumococcal meningitis in the Central African Republic. *Wellcome Open Research* 2018; **3**: 134.

48 Ihekweazu C, Basarab M, Wilson D, *et al.* Outbreaks of serious pneumococcal disease in closed settings in the post-antibiotic era: A systematic review. *Journal of Infection* 2010; **61**: 21–7.

49 Cliff M, Welaga P, Mohammed N, *et al.* Strategies for controlling pneumococcal disease and outbreaks during humanitarian emergencies. *Nat Med* 2024; **30**: 1515–6.

50 Vernet G, Saha S, Satzke C, *et al.* Laboratory-based diagnosis of pneumococcal pneumonia: state of the art and unmet needs. *Clinical Microbiology and Infection* 2011; **17**: 1–13.

51 Grabenstein JD, Musher DM. Pneumococcal Polysaccharide Vaccines. In: Plotkin's Vaccines, 7th edn. Philadelphia: Elsevier, 2018.

52 Pavia M, Bianco A, Nobile CGA, Marinelli P, Angelillo IF. Efficacy of Pneumococcal Vaccination in Children Younger Than 24 Months: A Meta-Analysis. *Pediatrics* 2009; **123**: e1103–10.

53 Le Polain De Waroux O, Flasche S, Prieto-Merino D, Goldblatt D, Edmunds WJ. The Efficacy and Duration of Protection of Pneumococcal Conjugate Vaccines Against Nasopharyngeal Carriage: A Meta-regression Model. *The Pediatric Infectious Disease Journal* 2015; **34**: 858–64.

54 Dagan R, Juergens C, Trammel J, *et al.* Efficacy of 13-Valent Pneumococcal Conjugate Vaccine (PCV13) Versus That of 7-Valent PCV (PCV7) Against Nasopharyngeal Colonization of Antibiotic-Nonsusceptible *Streptococcus pneumoniae*. *J Infect Dis* 2015; **211**: 1144–53.

55 Cohen O, Knoll M, O'Brian K, *et al.* Pneumococcal Conjugate Vaccine (PCV) Review of Impact Evidence (PRIME) summary of findings from systematic review. 2017. https://terrance.who.int/mediacentre/data/sage/SAGE_Docs_Ppt_Oct2017/9_session_PCV/Oct2019_session9_PCV_PRIMEsummary.pdf (accessed Aug 9, 2019).

56 Chapman TJ, Olarte L, Dbaibo G, *et al.* PCV15, a pneumococcal conjugate vaccine, for the prevention of invasive pneumococcal disease in infants and children. *Expert Rev Vaccines* 2024; **23**: 137–47.

57 Senders S, Klein NP, Lamberth E, *et al.* Safety and Immunogenicity of a 20-valent Pneumococcal Conjugate Vaccine in Healthy Infants in the United States. *Pediatr Infect Dis J* 2021; **40**: 944–51.

58 Feikin DR, Kagucia EW, Loo JD, *et al.* Serotype-specific changes in invasive pneumococcal disease after pneumococcal conjugate vaccine introduction: a pooled analysis of multiple surveillance sites. *PLoS Med* 2013; **10**: e1001517.

59 World Health Organization. Immunization Data portal. https://immunizationdata.who.int/global/wiise-detail-page (accessed April 25, 2024).

60 Klugman KP. Herd protection induced by pneumococcal conjugate vaccine. *The Lancet Global Health* 2014; **2**: e365–6.

61 Loughlin AM, Hsu K, Silverio AL, Marchant CD, Pelton SI. Direct and indirect effects of PCV13 on nasopharyngeal carriage of PCV13 unique pneumococcal serotypes in Massachusetts' children. *Pediatr Infect Dis J* 2014; **33**: 504–10.

62 Simonsen L, Taylor RJ, Young-Xu Y, Haber M, May L, Klugman KP. Impact of Pneumococcal Conjugate Vaccination of Infants on Pneumonia and Influenza Hospitalization and Mortality in All Age Groups in the United States. *mBio* 2011; **2**: e00309-10.

63 Ozawa S, Mirelman A, Stack ML, Walker DG, Levine OS. Cost-effectiveness and economic benefits of vaccines in low- and middle-income countries: A systematic review. *Vaccine* 2012; **31**: 96–108.

64 Close RM, Pearson C, Cohn J. Vaccine-preventable disease and the under-utilization of immunizations in complex humanitarian emergencies. *Vaccine* 2016; **34**: 4649–55.

65 Nnadi C, Etsano A, Uba B, *et al.* Approaches to Vaccination Among Populations in Areas of Conflict. *J Infect Dis* 2017; **216**: S368–72.

66 UNICEF. Two-thirds of unimmunized children live in conflict-affected countries. 2016. https://www.unicef.org/png/press-releases/two-thirds-unimmunized-children-live-conflict-affected-countries-unicef (accessed April 25, 2024).

67 World Health Organization. Vaccination in Acute Humanitarian Emergencies. A framework for decision making. 2017.

68 Leach K, Checchi F. The utilisation of vaccines in humanitarian crises, 2015–2019: A review of practice. *Vaccine* 2022; **40**: 2970–8.

69 Grais RF, Strebel P, Mala P, Watson J, Nandy R, Gayer M. Measles vaccination in humanitarian emergencies: a review of recent practice. *Conflict and Health* 2011; **5**: 21.

70 Wolfson LJ, Grais RF, Luquero FJ, Birmingham ME, Strebel PM. Estimates of measles case fatality ratios: a comprehensive review of community-based studies. *Int J Epidemiol* 2009; **38**: 192–205.

71 Babakura B, Nomhwange T, Jean Baptiste AE, *et al.* The challenges of insecurity on implementing vaccination campaign and its effect on measles elimination and control efforts: A case study of 2017/18 measles campaign in Borno state, Nigeria. *Vaccine* 2021; **39**: C66–75.

72 Jean Baptiste AE, Wagai J, Luce R, *et al.* Measles outbreak in complex emergency: estimating vaccine effectiveness and evaluation of the vaccination campaign in Borno State, Nigeria, 2019. *BMC Public Health* 2021; **21**: 437.

73 Chin T, Buckee CO, Mahmud AS. Quantifying the success of measles vaccination campaigns in the Rohingya refugee camps. *Epidemics* 2020; **30**: 100385.

74 Martini M, Orsini D. Armed conflict in the world threatens the eradication of Poliomyelitis: risks of humanitarian crises. *Pathogens and Global Health* 2022; **116**: 267–8.

75 Khan AI, Islam MdT, Siddique SA, *et al.* Post-vaccination campaign coverage evaluation of oral cholera vaccine, oral polio vaccine and measles–rubella vaccine among Forcibly Displaced Myanmar Nationals in Bangladesh. *Human Vaccines & Immunotherapeutics* 2019; **15**: 2882–6.

76 Sheikh MA, Makokha F, Hussein MA, *et al.* Combined Use of Inactivated and Oral Poliovirus Vaccines in Refugee Camps and Surrounding Communities — Kenya, December 2013. *Morbidity and Mortality Weekly Report* 2014; **63**: 237.

77 Mahamud A, Kamadjeu R, Webeck J, *et al.* Effectiveness of Oral Polio Vaccination Against Paralytic Poliomyelitis: A Matched Case-Control Study in Somalia. *The Journal of Infectious Diseases* 2014; **210**: S187–93.

78 Shannon K, Hast M, Azman AS, Legros D, McKay H, Lessler J. Cholera prevention and control in refugee settings: Successes and continued challenges. *PLOS Neglected Tropical Diseases* 2019; **13**: e0007347.

79 Connolly MA, Gayer M, Ryan MJ, Salama P, Spiegel P, Heymann DL. Communicable diseases in complex emergencies: impact and challenges. *The Lancet* 2004; **364**: 1974–83.

80 Porta MI, Lenglet A, de Weerdt S, *et al.* Feasibility of a preventive mass vaccination campaign with two doses of oral cholera vaccine during a humanitarian emergency in South Sudan. *Trans R Soc Trop Med Hyg* 2014; **108**: 810–5.

81 Azman AS, Parker LA, Rumunu J, *et al.* Effectiveness of one dose of oral cholera vaccine in response to an outbreak: a case-cohort study. *The Lancet Global Health* 2016; **4**: e856–63.

82 Finger F, Bertuzzo E, Luquero FJ, *et al.* The potential impact of case-area targeted interventions in response to cholera outbreaks: A modeling study. *PLOS Medicine* 2018; **15**: e1002509.

83 Ratnayake RC. Case-Area Targeted Intervention for the Control of Cholera Epidemics in Crises: From Spatial Mathematical Modelling to Field Evaluation. 2024; published online April 24. DOI:10.17037/PUBS.04672999.

84 World Health Organization. MI4A: Vaccine Purchase Data. http://www.who.int/immunization/programmes_systems/procurement/v3p/platform/module1/en/ (accessed July 31, 2019).

85 World Health Organization. The Humanitarian Mechanism. 2017. http://www.who.int/immunization/programmes_systems/sustainability/The_Humanitarian_Mechanism_ToRs.pdf?ua=1.

86 World Health Organization. PNEUMOSIL. Prequalification of Medical Products (IVDs, Medicines, Vaccines and Immunization Devices, Vector Control). https://extranet.who.int/prequal/vaccines/p/pneumosilr-0 (accessed April 25, 2024).

87 Gavi. Pneumococcal AMC Supply agreements. 2018; published online June 6. https://www.gavi.org/investing-gavi/innovative-financing/pneumococcal-amc/manufacturers/supply-agreements (accessed April 25, 2024).

88 Blanchet K, Ramesh A, Frison S, *et al.* Evidence on public health interventions in humanitarian crises. *The Lancet* 2017; **390**: 2287–96.

89 Médecins Sans Frontières. Médecins Sans Frontières (MSF) briefing on provisional agenda item 20.1: Global vaccine action plan, Document A71/39. 2018. https://msfaccess.org/sites/default/files/2018-05/WHO-WHA-71_MSF-briefing-GVAP-Agenda-Item-20-1.pdf.

90 Ravelo JL. MSF bends donation policy for pneumonia vaccine. Devex. 2015. https://www.devex.com/news/sponsored/msf-bends-donation-policy-for-pneumonia-vaccine-85317 (accessed Feb 27, 2019).

91 Peyraud N, Quéré M, Duc G, *et al.* A post-conflict vaccination campaign, Central African Republic. *Bull World Health Organ* 2018; **96**: 540–7.

92 Médecins Sans Frontières. Nigeria: 'There were hardly any children under 5 years of age'. 2016. https://www.msf.ie/article/nigeria-there-were-hardly-any-children-under-5-years-age (accessed Feb 27, 2019).

93 Médecins Sans Frontières. Vaccination activities in Médecins Sans Frontières. 2015. https://www.who.int/immunization/sage/meetings/2015/october/MSF_SAGE_October2015_vf.pdf.

94 Médecins Sans Frontières. Humanitarian Mechanism – 1st Monitoring Meeting, Meeting report. 2018.

95 Heymann DL, Aylward RB. Mass vaccination: when and why. *Curr Top Microbiol Immunol* 2006; **304**: 1–16.

96 Vynnycky E, White R. An Introduction to Infectious Disease Modelling. OUP Oxford, 2010.

97 Kucharski AJ, Russell TW, Diamond C, *et al.* Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases* 2020; **20**: 553–8.

98 Sun G-Q, Xie J-H, Huang S-H, Jin Z, Li M-T, Liu L. Transmission dynamics of cholera: Mathematical modeling and control strategies. *Communications in Nonlinear Science and Numerical Simulation* 2017; **45**: 235–44.

99 Le Polain De Waroux O, Edmunds WJ, Takahashi K, *et al.* Predicting the impact of pneumococcal conjugate vaccine programme options in Vietnam. *Hum Vaccin Immunother* 2018; **14**: 1939–47.

100 Flasche S, Ojal J, Le Polain de Waroux O, *et al.* Assessing the efficiency of catch-up campaigns for the introduction of pneumococcal conjugate vaccine: a modelling study based on data from PCV10 introduction in Kilifi, Kenya. *BMC Med* 2017; **15**: 113.

101 Koopman J. Modeling Infection Transmission. *Annual Review of Public Health* 2004; **25**: 303–26.

102 Masala GL, Lipsitch M, Bottomley C, Flasche S. Exploring the role of competition induced by non-vaccine serotypes for herd protection following pneumococcal vaccination. *J R Soc Interface* 2017; **14**. DOI:10.1098/rsif.2017.0620.

103 Melegaro A, Choi YH, George R, Edmunds WJ, Miller E, Gay NJ. Dynamic models of pneumococcal carriage and the impact of the Heptavalent Pneumococcal Conjugate Vaccine on invasive pneumococcal disease. *BMC Infectious Diseases* 2010; **10**: 90.

104 Cai FY, Fussell T, Cobey S, Lipsitch M. Use of an individual-based model of pneumococcal carriage for planning a randomized trial of a whole-cell vaccine. *PLOS Computational Biology* 2018; **14**: e1006333.

105 Ojal J, Griffiths U, Hammitt LL, *et al.* Sustaining pneumococcal vaccination after transitioning from Gavi support: a modelling and cost-effectiveness study in Kenya. *The Lancet Global Health* 2019; **7**: e644–54.

106 Choi YH, Jit M, Flasche S, Gay N, Miller E. Mathematical Modelling Long-Term Effects of Replacing Prevnar7 with Prevnar13 on Invasive Pneumococcal Diseases in England and Wales. *PLOS ONE* 2012; **7**: e39927.

107 Lipsitch M, Colijn C, Cohen T, Hanage WP, Fraser C. No coexistence for free: Neutral null models for multistrain pathogens. *Epidemics* 2009; **1**: 2–13.

108 Davies NG, Flasche S, Jit M, Atkins KE. Within-host dynamics shape antibiotic resistance in commensal bacteria. *Nat Ecol Evol* 2019; **3**: 440–9.

109 Greenhalgh D, Lamb KE, Robertson C. A mathematical model for the spread of *Streptococcus pneumoniae* with transmission dependent on serotype. *J Biol Dyn* 2012; **6 Suppl 1**: 72–87.

110 Trotter CL, Gay NJ, Edmunds WJ. Dynamic models of meningococcal carriage, disease, and the impact of serogroup C conjugate vaccination. *Am J Epidemiol* 2005; **162**: 89–100.

111 Bottomley C, Roca A, Hill PC, Greenwood B, Isham V. A mathematical model of serotype replacement in pneumococcal carriage following vaccination. *J R Soc Interface* 2013; **10**: 20130786.

112 Auranen K, Rinta-Kokko H, Halloran ME. Estimating Strain-Specific and Overall Efficacy of Polyvalent Vaccines Against Recurrent Pathogens From a Cross-Sectional Study. *Biometrics* 2013; **69**: 235–44.

113 Cauchemez S, Temime L, Valleron A-J, *et al. S. pneumoniae* transmission according to inclusion in conjugate vaccines: Bayesian analysis of a longitudinal follow-up in schools. *BMC Infect Dis* 2006; **6**: 14.

114 Obolski U, Lourenço J, Thompson C, Thompson R, Gori A, Gupta S. Vaccination can drive an increase in frequencies of antibiotic resistance among nonvaccine serotypes of *Streptococcus pneumoniae*. *Proc Natl Acad Sci USA* 2018; **115**: 3102–7.

115 Choi YH, Jit M, Gay N, *et al.* 7-Valent Pneumococcal Conjugate Vaccination in England and Wales: Is It Still Beneficial Despite High Levels of Serotype Replacement? *PLOS ONE* 2011; **6**: e26190.

116 Lipsitch M. Vaccination against colonizing bacteria with multiple serotypes. *Proc Natl Acad Sci U S A* 1997; **94**: 6571–6.

117 Gjini E, Valente C, Sá-Leão R, Gomes MGM. How direct competition shapes coexistence and vaccine effects in multi-strain pathogen systems. *J Theor Biol* 2016; **388**: 50–60.

118 Gjini E. Geographic variation in pneumococcal vaccine efficacy estimated from dynamic modeling of epidemiological data post-PCV7. *Sci Rep* 2017; **7**: 3049.

119 van Hoek AJ, Choi YH, Trotter C, Miller E, Jit M. The cost-effectiveness of a 13-valent pneumococcal conjugate vaccination for infants in England. *Vaccine* 2012; **30**: 7205–13.

120 Nurhonen M, Cheng AC, Auranen K. Pneumococcal transmission and disease in silico: a microsimulation model of the indirect effects of vaccination. *PLoS ONE* 2013; **8**: e56079.

121 Flasche S, Edmunds WJ, Miller E, Goldblatt D, Robertson C, Choi YH. The impact of specific and non-specific immunity on the ecology of *Streptococcus pneumoniae* and the implications for vaccination. *Proc Biol Sci* 2013; **280**: 20131939.

122 Cobey S, Lipsitch M. Niche and neutral effects of acquired immunity permit coexistence of pneumococcal serotypes. *Science* 2012; **335**: 1376–80.

123 Løchen A, Truscott JE, Croucher NJ. Analysing pneumococcal invasiveness using Bayesian models of pathogen progression rates. *PLoS Comput Biol* 2022; **18**: e1009389.

124 Dellaportas P, Roberts GO. An Introduction to MCMC. In: Møller J, ed. Spatial Statistics and Computational Methods. New York, NY: Springer, 2003: 1–41.

125 Gelman A. Parameterization and Bayesian Modeling. *Journal of the American Statistical Association* 2004; **99**: 537–45.

126 Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 1970; **57**: 97–109.

127 Brooks S, Gelman A, Jones G, Meng X-L. Handbook of Markov Chain Monte Carlo. CRC Press, 2011.

128 Roy V. Convergence diagnostics for Markov chain Monte Carlo. 2019; published online Oct 16. http://arxiv.org/abs/1909.11827 (accessed April 25, 2024).

129 Ter Braak CJF, Vrugt JA. Differential Evolution Markov Chain with snooker updater and fewer chains. *Stat Comput* 2008; **18**: 435–46.

130 Betancourt M. A Conceptual Introduction to Hamiltonian Monte Carlo. 2018; published online July 15. http://arxiv.org/abs/1701.02434 (accessed May 20, 2024).

131 Hartig F, Minunno F, Paul S, Cameron D, Ott T, Pichler M. BayesianTools: General-Purpose MCMC and SMC Samplers and Tools for Bayesian Statistics. 2023; published online Jan 30. https://cran.r-project.org/web/packages/BayesianTools/index.html (accessed May 20, 2024).

Chapter 1

Pneumococcal conjugate vaccine use during humanitarian crises

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1604011	**Title**	Mr
First Name(s)	Kevin		
Surname/Family Name	van Zandvoort		
Thesis Title	Strategies for pneumococcal conjugate vaccine use in humanitarian crises		
Primary Supervisor	Stefan Flasche		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Vaccine		
When was the work published?	24/09/2019		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I reviewed available literature on the burden of respiratory infections and pneumococcal disease in humanitarian crises, the association between specific risk-factors and pneumococcal transmission, disease, and mortality, and the use of PCV in humanitarian crises. I then wrote the original manuscript draft, edited the manuscript following suggestions and comments from co-authors, and prepared it for submission.

SECTION E

Student Signature	Kevin van Zandvoort
Date	03/10/2024

Supervisor Signature	Stefan Flasche
Date	03/10/2024

Vaccine 37 (2019) 6787–6792

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Review

Pneumococcal conjugate vaccine use during humanitarian crises

Kevin van Zandvoort [a,b,*], Francesco Checchi [a], Emma Diggle [c], Rosalind M. Eggo [a,b], Kartini Gadroen [d,e], Kim Mulholland [a,f], Catherine R. McGowan [c,g], Olivier le Polain de Waroux [a,b,h,j], V. Bhargavi Rao [i], Catherine Satzke [f], Stefan Flasche [a,b]

[a] Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK
[b] Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine, London, UK
[c] Save the Children UK, London, UK
[d] Médecins Sans Frontières, Amsterdam, the Netherlands
[e] Department of Medical Informatics, Erasmus MC, Rotterdam, the Netherlands
[f] Murdoch Children's Research Institute, University of Melbourne, Royal Children's Hospital, Melbourne, Victoria, Australia
[g] Department of Public Health, Environments, and Society, London School of Hygiene & Tropical Medicine. London, UK
[h] UK Public Health Rapid Support Team, London, UK
[i] Manson Unit, Médecins Sans Frontières (MSF UK), London, UK
[j] Public Health England, London, UK

ARTICLE INFO

Article history:
Received 16 May 2019
Received in revised form 16 August 2019
Accepted 9 September 2019
Available online 24 September 2019

Keywords:
Pneumococcal conjugate vaccine
Humanitarian crises
Humanitarian health
Refugees
Internally displaced people
Vaccination strategy
Pneumonia

ABSTRACT

Streptococcus pneumoniae is a common human commensal that causes a sizeable part of the overall childhood mortality in low income settings. Populations affected by humanitarian crises are at especially high risk, because a multitude of risk factors that are enhanced during crises increase pneumococcal transmission and disease severity. Pneumococcal conjugate vaccines (PCVs) provide effective protection and have been introduced into the majority of routine childhood immunisation programmes globally, though several barriers have hitherto limited their uptake during humanitarian crises. When PCV coverage cannot be sustained during crises or when PCV has not been part of routine programmes, mass vaccination campaigns offer a quick acting and programmatically feasible bridging solution until services can be restored. However, we currently face a paucity of evidence on which to base the structure of such campaigns. We believe that, now that PCV can be procured at a substantially reduced price through the Humanitarian Mechanism, this lack of information is a remaining hurdle to PCV use in humanitarian crises. Considering the difficulties in conducting research in crises, we propose an evidence generation pathway consisting of primary data collection in combination with mathematical modelling followed by quasi-experimental evaluation of a PCV intervention, which can inform on optimal vaccination strategies that consider age targeting, dosing regimens and impact duration.

Contents

* Corresponding author at: Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, Keppel Street, WC1E 7HT, London, United Kingdom.
E-mail address: Kevin.Van-Zandvoort@lshtm.ac.uk (K. van Zandvoort).

https://doi.org/10.1016/j.vaccine.2019.09.038
0264-410X/

1. Introduction

Approximately 68.5 million people, nearly 1% of the world's population, were forcibly displaced due to insecurity and war in 2017. In those who are refugees, more than half are under the age of 18, and 17% under the age of five [1,88]. In the same year, hundreds of millions were affected by armed conflicts [2,3], and almost 100 million were impacted by natural disasters [4]. Whether in the acute emergency or the protracted phase, crises substantially affect people's lives, and can dramatically increase premature mortality [5–7]. In most crises, excess deaths are often attributable to the indirect effects of crisis-emergent factors such as the breakdown of public health services, food insecurity, inadequate water and sanitation, and overcrowding; factors that increase both the incidence and severity of disease [8,9].

Infectious diseases are of particular concern, and require specific control measures that include, but are not limited to, vaccines. To date, only a small subset of licensed vaccines that are routinely used in most stable settings is commonly used in humanitarian crises. These usually include measles, polio, and (recently) cholera, with context-specific threats such as meningococcal disease or yellow fever infrequently addressed [10]. However, the prioritisation of pathogens targeted by these vaccines may not comprehensively address the local anticipated preventable disease burden. More recent additions to the vaccine portfolio, such as vaccines protecting against HPV (particularly in settings with high rates of sexual violence), rotavirus, and *Streptococcus pneumoniae*, have rarely been used in humanitarian settings. Using the example of *Streptococcus pneumoniae*, we here propose a framework to overcome some of the barriers for vaccine use in humanitarian settings, and to help prevent the likely substantial disease burden associated with respective pathogens in crises settings.

2. *Streptococcus pneumoniae* in crises

Streptococcus pneumoniae (the pneumococcus) is a human commensal that commonly resides in the nasopharynx, and occasionally causes disease (e.g. pneumonia, meningitis, and sepsis), especially in young children and people with weakened immune systems [11]. The pneumococcal disease burden in crises is largely unknown, but likely substantial. Outbreaks are thought to occur, but often go unnoticed due to non-existent or under resourced surveillance systems and the low specificity of symptoms [12]. Pneumococcal meningitis outbreaks have occasionally been reported in humanitarian settings [13,14], and pneumococcal pneumonia is a major concern. During crises, acute respiratory tract infections (ARI) and diarrhoeal disease make up the top two causes of morbidity in all age groups, with ARIs alone accounting for 20–35% of mortality in children younger than five years of age [15]. The exact aetiology of these ARIs remains unknown, but more than half of all ARI-related deaths worldwide were caused by pneumococci in the pre-pneumococcal conjugate vaccination era [16]. Risk factors that are commonly exacerbated in crises, such as malnutrition, indoor air pollution, and overcrowding, can increase pneumococcal carriage, transmission, disease, and mortality (Table 1). This likely amplifies this burden in crises. Many of these risk factors were also present in pneumococcal outbreaks that have been identified in stable settings [17]. In addition, the displacement and crowding of people from a range of different communities may expose them to a range of circulating serotypes that they have not seen before, increasing the risk of disease and probably extending the risk even more into older age groups.

3. Pneumococcal conjugate vaccines

Pneumococcal conjugate vaccines (PCVs) effectively protect against pneumococcal disease [11]. There are currently two PCV products available, protecting against 10 (PCV10) or 13 (PCV13) of more than 90 known pneumococcal serotypes, and PCVs with increased valency (PCV15 and PCV20) are currently in development [36,37]. In contrast to (unconjugated) pneumococcal polysaccharide vaccines [38], PCVs are recommended for use in children and, in addition to the direct protection against pneumococcal disease, also elicit indirect protection through interrupted transmission of vaccine-targeted serotypes (VT) [11]. Although their impact is dampened by replacement colonisation of the nasopharynx by non-vaccine serotypes, these serotypes are generally less likely to cause severe disease, resulting in a net benefit [39]. PCVs have now been introduced in the routine childhood immunisation programmes of the majority of countries [40]. In most places where PCVs are used at high coverage, the marked reduction in VT transmission has expanded the benefit beyond vaccinees alone [41–44].

4. Vaccination in crises

Vaccination strategies can be categorized into routine immunisation, which aims to reduce the disease burden by sustainable and

Table 1
Crisis-emergent risk factors that can plausibly affect the pneumococcal burden.

Risk factor	Increased transmission (carriage)	Increased probability that carriage leads to disease	Increased case-fatality ratio	Selected references
Acute malnutrition	++*	+++*	+++*	[18,19]
Measles outbreaks and other viral respiratory tract infections	++	++	++	[20–22]
Overcrowding and altered social contact patterns	+++*	–	–	[18,19,24,24]
Disrupted routine pneumococcal conjugate vaccine use	+[i]	+++	–	[25–27]
Low access to curative care	+[ii]	+	+++	[28–31]
Smoke inhalation	–	+	–	[32,33]
Inadequate water and sanitation	++	+	–	[34,35]

– no effect on outcome; + small effect on outcome; ++ medium effect on outcome; +++ large effect on outcome.
* Potential shift in the age-specific risk (younger average acquisition and increased carriage and disease among all age groups).
[i] Increase in carriage of vaccine-targeted serotypes, but not in overall carriage.
[ii] Increased transmission due to reduced bystander effect as a result of limited antibiotic usage in the community.

equitable vaccination of new birth cohorts [45], or mass vaccination campaigns, which aim for a quick but short lived (additional) reduction in disease burden. However, this distinction has become blurred with recent use of 'periodic intensification of routine immunisation' (PIRI) activities [46].

Routine immunisation is highly effective and cost-effective [47], but as a strategy faces a number of challenges during crises, including access to regular timely services, disruption of the cold chain, lack of personnel to deliver vaccines, safety of health care workers, and access of health workers to the affected population [48,49]. Consequently, in the acute phase of a crisis routine immunisation often breaks down and cannot ensure population immunity. Vaccination coverage may drop to levels too low to interrupt transmission in susceptible parts of the population. This is most pronounced in mass displacement scenarios; where overcrowding alone increases the transmission intensity of infections and, in combination with an accumulation of susceptible individuals, increases vaccination requirements to achieve herd immunity.

Accordingly, humanitarian actors including non-governmental organizations (NGOs) emphasise the role of mass vaccination campaigns. These campaigns are regularly used for outbreak control [50], but should in this instance not only aim to quickly control disease but also sustain impact for sufficient time until subsequent campaigns can be performed or routine immunisation can be resumed. The high number of vaccine doses given to extended age groups in a shorter time-frame usually make mass vaccination campaigns more feasible to execute and faster in reducing the disease burden.

Insufficient evidence on the causes underlying the disease burden during crises and limited guidance on vaccine priorities for humanitarian decision-makers may partly explain the hitherto narrow uptake of vaccine interventions. In an attempt to improve this situation the World Health Organization (WHO) introduced a Framework for Decision-Making on Vaccination in Humanitarian Emergencies in 2012, which was updated in 2017 [51]. This three-step framework aims to implement the most appropriate vaccination interventions in each crisis given the local epidemiology, vaccine characteristics, and other context-specific considerations. The framework emphasises expanding the range of vaccines offered to crisis-affected populations, but also recommends adapted vaccination strategies, including expanded age ranges and reduced-dose regimens.

5. PCV use in crises

Although the WHO Framework lists PCVs as one of the vaccines to be considered for use in crises [51], and despite a likely high preventable pneumococcal disease burden, they have rarely been used during crises [53–57]. The rationale for integrated PCV vaccination strategies in crises is clear: mass vaccination campaigns delivered as part of the initial package of interventions in the acute emergency phase of new crises could rapidly establish direct and indirect protection when vulnerability due to malnutrition, congestion of unplanned settlements, and lack of curative health services is likely to peak. These campaigns should ideally be multi-antigen interventions (e.g. bundling measles and cholera) or multi-interventional (e.g. bed nets or micronutrient supplementation).

A PCV-specific barrier to vaccination in crises has long been its price. If not supported by Gavi, lower and middle income countries (LMIC) spend about 20, 50, and 3 times as much for one complete regimen of PCV (50US$) compared to measles containing vaccine, oral polio vaccine, or rotavirus vaccine, which is indicative for prices paid by humanitarian actors until 2017 [57]. While PCVs have been prohibitively expensive, a "Humanitarian Mechanism" sponsored in 2017 by the WHO, Unicef, Médecins Sans Frontières and Save the Children now guarantees more affordable PCV procurement by humanitarian actors and affordable expedited delivery [58]. Although some 600,000 doses of PCV have been delivered through this mechanism to date [55], this only covers a small proportion of crises affected populations at risk. In addition, only multi-dose PCV vials are available through the humanitarian mechanism. Whereas this eases transportation and storage of the vaccine, it also increases wastage and may therefore decrease their cost-effectiveness, especially when used routinely in small populations.

A key barrier that has not yet been addressed is the insufficient evidence on optimal PCV deployment strategies via mass vaccination campaigns and their expected impact in crises [59,60]. In places where they have been used, they have been administered through different strategies targeted at different age groups [61–64]. The impact of those alternative approaches has not been assessed.

The WHO recently updated their recommendations on the use of PCVs in children [65]. These now include a recommendation to use PCV in children under one year of age and consider for children under five years of age during humanitarian crises and other emergencies. This is in line with the aforementioned WHO Framework [51]. However, in the absence of any evidence [66], no further guidance is given to the optimal age range to target in a campaign, the number of doses needed, and the frequency of campaigns.

There is no clear rationale to limit mass vaccination campaigns to those under one, two, or even five years of age. These are the age groups that usually bear the heaviest burden of pneumococcal pneumonia, but in crisis settings where high pneumococcal carriage prevalence likely extends to adulthood, targeting a larger proportion of the transmitting population is probably needed to control VT circulation. This would maximize herd protection, which is crucial in optimising vaccine use, as it protects unvaccinated children and adults. Such control is particularly needed if the effects of a campaign need to sustain protection for months or years until a subsequent campaign is feasible or routine immunisation can be restored. It is also key in settings where high prevalence of acute malnutrition may shift the age spectrum for pneumococcal disease towards older children [67]. Using an extended age range to 14 years of age for example, could be operationally convenient as it may allow co-administration with measles vaccine.

Multi-dose schedules are recommended in routine programmes [65] but may be unfeasible in crisis settings. If, for operational reasons, only a single dose of PCV can be administered, extended age ranges may partially compensate for a lack of optimal direct protection. Single dose strategies only provide moderate direct protection to infants if not followed by a booster dose [68], but this reduced direct protection may be offset by enhanced indirect protection from older age groups, provided that vaccine coverage levels are sufficiently high. Single-dose strategies are being intensively tested in stable settings [68–70], but their exact indirect effects remain unknown.

6. Evaluating optimal vaccination strategies

Vaccination strategies must consider both direct and indirect protection. The former will require estimation of age specific pneumonia burden, which is likely to vary considerably between crisis settings depending on malnutrition rates and other factors. The best evidence of vaccine impact comes from cluster-randomised controlled trials (cRCT). However, these are resource-intensive and exceptionally challenging to conduct during crises, with additional ethical concerns related to randomisation of vulnerable

populations to potentially less protected trial arms [71]. Moreover, only a small subset of many possible combinations of potentially viable dosing strategies and age ranges can be investigated.

We propose instead a sequential evidence generation pathway, consisting of primary data collection in combination with mathematical modelling followed by quasi-experimental evaluation of PCV intervention. Mathematical models are increasingly used to synthesize a multitude of evidence for vaccine decision making, particularly if indirect vaccine effects form a key part of the desired impact [72–74]. If adequately parameterised, these models are useful to simulate the pneumococcal epidemiology of a specific setting and predict PCV impact under various vaccination strategies, as has been done in stable settings such as Kenya [25,52] and Vietnam [75]. However, the use of modelling to inform and evaluate vaccine decision making in crises is limited. It has predominantly been used to assess reactive strategies for outbreaks [76–78], e.g. the potential of ring-vaccination strategies for Ebola control [79], but has for instance also been used for pre-emptive strategies for Hepatitis E in displaced populations [80].

PCV vaccination strategies have, to our knowledge, only been explored in stable settings. A limitation to the use of modelling to inform PCV use in crises-affected populations is the lack of context-specific data for model parameterisation. The key drivers of pneumococcal transmission are social contact behaviour (a proxy for disease transmission routes) and the pre-PCV prevalence of nasopharyngeal carriage that helps identifying pockets of the population driving pneumococcal transmission. Consequently studies have measured both in a multitude of settings [81,82], but few have been done in LMICs and evidence from crisis-affected populations is entirely absent. The main drivers of transmission are often children, due to the nature and frequency of their contacts in combination with high prevalence of pneumococcal carriage [83,84]. However, in displaced populations, both social contact patterns and pneumococcal carriage may be considerably altered from their pre-crisis baseline (see Table 1). As this may significantly affect the appropriate strategy, primary data is needed to construct meaningful models for hypothesis generation.

Specifically, we argue that a seemingly natural assessment of age targeting through PCV use in the age groups with highest incidence of pneumococcal disease is unlikely to make best use of PCV. Whereas this strategy would indeed provide direct protection to those at highest risk, it may lead to either under or over use of PCV. Without an assessment of transmission dynamics, such strategy could end up providing PCV to an age group that is too narrow so that no herd immunity is achieved. This would leave the rest of the population vulnerable, and upcoming generations who are at exceptionally high risk unprotected. Alternatively, the age group may be too broad, and many who would have been protected through herd effects will receive PCV without much added benefit.

Mathematical modelling can be used to study transmission dynamics, needed to predict vaccine impact. Specifically, it can formally integrate available evidence and their associated uncertainty into a prediction framework that can explore and propose vaccination strategies to potentially optimize impact, namely: (i) PCV target age groups for mass vaccination in crises; (ii) minimum vaccination coverage needed; (iii) single vs. multi-dose vaccination options; and (iv) the frequency with which campaigns should be implemented to sustain PCV effects until routine immunisation can be re-established. It can also be used to extrapolate to different crises settings such as overcrowded acute displacement camps or slow-onset food security crises in rural areas.

Although modelling can narrow down the range of potential strategies, pilot implementation of these strategies should be accompanied by impact measures. At a minimum this should include cross-sectional nasopharyngeal carriage studies in the target population before and after PCV use, though ideally extend to measures of impact on morbidity or even mortality. Quasi-experimental designs can be used to evaluate their impact with relatively low resources [85], as has been done in multiple post-licensure PCV studies where no or only a limited number of control sites is available [43,44,86,87]. In addition, such results can feed back into mathematical models [52], leading to more robust predictions of vaccine strategies and impact.

7. Conclusions

Vaccines that are most commonly used in humanitarian crises settings have not necessarily been prioritised based on the current or expected local preventable disease burden. More recent additions to the vaccine portfolio that could potentially prevent a disproportionally large burden, such as PCVs, are infrequently deployed. The high costs of PCVs are now largely mitigated by the availability of PCV through the Humanitarian Mechanism, but the lack of specific PCV usage recommendations is among the key factors that hinder uptake as a routine part of humanitarian responses. Evidence on practical, effective, and cost-effective ways to use PCV is critical for humanitarian actors to better evaluate the role of PCV in the vaccine portfolio for crises use.

Preventing a large proportion of the pneumococcal disease burden through PCV use would contribute to the overarching aim of humanitarian action: to save lives. We propose that a combination of targeted data collection in combination with mathematical modelling can be used to generate evidence-based hypotheses on optimal vaccination strategies for PCVs in crises, and ultimately pave the way for rational PCV use in crises. This evidence pathway could similarly be applied to other vaccine-preventable diseases, for which indirect effects are a key part of their overall effects, to eventually achieve an evidence based prioritisation strategy for optimal vaccine use in humanitarian crises.

Contributors

FC conceived the idea for the manuscript. KvZ wrote the first manuscript draft with support from SF and FC. All authors contributed to, and approved, the final draft. All authors attest they meet the ICMJE criteria for authorship.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

KvZ, FC, CS and KM are supported by Elrha's Research for Health in Humanitarian Crises (R2HC) Programme, which aims to improve health outcomes by strengthening the evidence base for public health interventions in humanitarian crises. The R2HC programme is funded by the UK Government (DFID), the Wellcome Trust, and the UK National Institute for Health Research (NIHR).

SF was supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and Royal Society (grant number 208812/Z/17/Z).

RME acknowledges funding from an HDR UK Innovation Fellowship (grant MR/S003975/1).

CS is supported by an Australian NHMRC Career Development Fellowship (1087957) and a Veski Inspiring Women Fellowship. MCRI is supported by the Victorian Government's Operational Infrastructure Support Program.

The UK Public Health Rapid Support Team is funded by the National Institute for Health Research and the Department of Health and Social Care. The views expressed are those of the authors and not necessarily those of the NIHR or DHSC.

The funders had no involvement in the conceptualization, research, or writing of the manuscript. The corresponding author had final responsibility for the decision to submit for publication.

References

[1] United Nations High Commissioner for Refugees. Population statistics – Data – Demographics; 2018. http://popstats.unhcr.org/en/demographics [accessed July 29, 2019].
[2] Sundberg R, Melander E. Introducing the UCDP georeferenced event dataset. J Peace Res 2013;50:523–32.
[3] Croicu M, Sundberg R. UCDP GED codebook version 18.1; 2017.
[4] Université catholique de Louvain CRED. EM-DAT: The emergency events database; 2019. http://www.emdat.be.
[5] Toole M, Waldman R. The public health aspects of complex emergencies and refugee situations. Annu Rev Public Health 1997;18:283–312.
[6] Heudtlass P, Speybroeck N, Guha-Sapir D. Excess mortality in refugees, internally displaced persons and resident populations in complex humanitarian emergencies (1998–2012) – insights from operational data. Confl Health 2016;10:15.
[7] Thomas SL, Thomas SD. Displacement and health. Br Med Bull 2004;69:115–27.
[8] Checchi F, Warsame A, Treacy-Wong V, Polonsky J, van Ommeren M, Prudhon C. Public health information in crisis-affected populations: a review of methods and their use for advocacy and action. The Lancet 2017;390:2297–313.
[9] Checchi F, Gayer M, Grais RF, Mills EJ. Public health in crisis-affected populations: a practical guide for decision-makers. London: Humanitarian Practice Network, Overseas Development Institute; 2007.
[10] Bellos A. Vaccination in humanitarian emergencies: literature review and case studies; 2012. http://www.who.int/immunization/sage/meetings/2012/april/2_SAGE_WGVHE_SG1__Lit_Review_CaseStudies.pdf [accessed Nov 21, 2018].
[11] Klugman KP, Dagan R, Malley R, Whitney CG. Pneumococcal conjugate vaccine and pneumococcal common protein vaccines. Plotkin's vaccine. Philadelphia: Elsevier; 2018.
[12] Andrejko K, Hosangadi D, Cohen O, et al. WHO technical expert consultation report on optimization of PCV impact: review of evidence and programmatic considerations to inform policy; 2017.
[13] Coldiron ME, Touré O, Frank T, Bouygues N, Grais RF. Outbreak of pneumococcal meningitis, Paoua Subprefecture, Central African Republic, 2016–2017. Emerg Infect Dis 2018;24:1720–2.
[14] Crellen T, Rao VB, Piening T, Zeydner J, Siddiqui MR. Seasonal upsurge of pneumococcal meningitis in the Central African Republic. Wellcome Open Res 2018;3:134.
[15] Bellos A, Mulholland K, O'Brien KL, Qazi SA, Gayer M, Checchi F. The burden of acute respiratory infections in crisis-affected populations: a systematic review. Confl Health 2010;4:3.
[16] Troeger C, Blacker B, Khalil IA, et al. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Infect Dis 2018;18:1191–210.
[17] Zivich PN, Grabenstein JD, Becker-Dreps SI, Weber DJ. *Streptococcus pneumoniae* outbreaks and implications for transmission and control: a systematic review. Pneumonia 2018;10:11.
[18] Fonseca Lima E, Mello MJG, Albuquerque MFPM, et al. Risk factors for community-acquired pneumonia in children under five years of age in the post-pneumococcal conjugate vaccine era in Brazil: a case control study. BMC Pediatr 2016;16. https://doi.org/10.1186/s12887-016-0695-6.
[19] Selwyn BJ. The epidemiology of acute respiratory tract infection in young children: comparison of findings from several developing countries. Coordinated Data Group of BOSTID Researchers. Rev Infect Dis 1990;12 (Suppl 8):S870–88.
[20] Small C, Shaler CR, McCormick S, et al. Influenza infection leads to increased susceptibility to subsequent bacterial superinfection by impairing NK cell responses in the lung. J Immunol Baltim Md 1950;2010(184):2048–56.
[21] Zhou H, Haber M, Ray S, Farley MM, Panozzo CA, Klugman KP. Invasive pneumococcal pneumonia and respiratory virus co-infections. Emerg Infect Dis 2012;18:294–7.
[22] Akramuzzaman SM, Cutts FT, Wheeler JG, Hossain MJ. Increased childhood morbidity after measles is short-term in urban Bangladesh. Am J Epidemiol 2000;151:723–35.
[23] Dueger EL, Asturias EJ, Matheu J, Gordillo R, Torres O, Halsey N. Increasing penicillin and trimethoprim-sulfamethoxazole resistance in nasopharyngeal *Streptococcus pneumoniae* isolates from Guatemalan children, 2001–2006. Int J Infect Dis IJID Off Publ Int Soc Infect Dis 2008;12:289–97.
[24] Petrosillo N, Pantosti A, Bordi E, et al. Prevalence, determinants, and molecular epidemiology of *Streptococcus pneumoniae* isolates colonizing the nasopharynx of healthy children in Rome. Eur J Clin Microbiol Infect Dis 2002;21:181–8.
[25] Ojal J, Griffiths U, Hammitt LL, et al. Sustaining pneumococcal vaccination after transitioning from Gavi support: a modelling and cost-effectiveness study in Kenya. Lancet Glob Health 2019;7:e644–54.
[26] Choi YH, Jit M, Flasche S, Gay N, Miller E. Mathematical modelling long-term effects of replacing Prevnar7 with Prevnar13 on invasive pneumococcal diseases in England and Wales. PLoS ONE 2012;7:e39927.
[27] Le Polain De Waroux O, Flasche S, Prieto-Merino D, Goldblatt D, Edmunds WJ. The efficacy and duration of protection of pneumococcal conjugate vaccines against nasopharyngeal carriage: a meta-regression model. Pediatr Infect Dis J 2015;34(8):858–64.
[28] Singer M, Nambiar S, Valappil T, Higgins K, Gitterman S. Historical and regulatory perspectives on the treatment effect of antibacterial drugs for community-acquired pneumonia. Clin Infect Dis 2008;47:S216–24.
[29] File TM, Schentag JJ. What can we learn from the time course of untreated and partially treated community-onset *Streptococcus pneumoniae* pneumonia? A clinical perspective on superiority and noninferiority trial designs for mild community-acquired pneumonia. Clin Infect Dis 2008;47:S157–65.
[30] Tedijanto C, Olesen SW, Grad YH, Lipsitch M. Estimating the proportion of bystander selection for antibiotic resistance among potentially pathogenic bacterial flora. Proc Natl Acad Sci 2018;115:E11988–95.
[31] Varon E, Levy C, De La Rocque F, et al. Impact of antimicrobial therapy on nasopharyngeal carriage of *Streptococcus pneumoniae*, *Haemophilus influenzae*, and *Branhamella catarrhalis* in children with respiratory tract infections. Clin Infect Dis 2000;31:477–81.
[32] Muthumbi E, Lowe BS, Muyodi C, Getambu E, Gleeson F, Scott JAG. Risk factors for community-acquired pneumonia among adults in Kenya: a case-control study. Pneumonia Nathan Qld 2017;9:17.
[33] Dherani M, Pope D, Mascarenhas M, Smith KR, Weber M, Bruce N. Indoor air pollution from unprocessed solid fuel use and pneumonia risk in children aged under five years: a systematic review and meta-analysis. Bull World Health Organ 2008;86:390–398C.
[34] Reisman J, Rudolph K, Bruden D, Hurlburt D, Bruce MG, Hennessy T. Risk factors for pneumococcal colonization of the nasopharynx in Alaska native adults and children. J Pediatr Infect Dis Soc 2014;3:104–11.
[35] Collins DA, Hoskins A, Snelling T, et al. Predictors of pneumococcal carriage and the effect of the 13-valent pneumococcal conjugate vaccination in the Western Australian Aboriginal population. Pneumonia 2017;9. https://doi.org/10.1186/s41479-017-0038-x.
[36] Stacey HL, Rosen J, Peterson JT, et al. Safety and immunogenicity of 15-valent pneumococcal conjugate vaccine (PCV-15) compared to PCV-13 in healthy older adults. Hum Vaccines Immunother 2019;15:530–9.
[37] Trial to evaluate the safety and immunogenicity of a 20-valent pneumococcal conjugate vaccine in pneumococcal vaccine-naïve adults – full text view - ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/NCT03760146 [accessed July 31, 2019].
[38] Grabenstein JD, Musher DM. Pneumococcal polysaccharide vaccines. Plotkin's Vaccines. Philadelphia: Elsevier; 2018.
[39] Feikin DR, Kagucia EW, Loo JD, et al. Serotype-specific changes in invasive pneumococcal disease after pneumococcal conjugate vaccine introduction: a pooled analysis of multiple surveillance sites. PLoS Med 2013;10:e1001517.
[40] World Health Organization. Official country reported coverage estimates time series; 2018. http://apps.who.int/immunization_monitoring/globalsummary/timeseries/tscoveragebcg.html [accessed Feb 27, 2019].
[41] Klugman KP. Herd protection induced by pneumococcal conjugate vaccine. Lancet Glob Health 2014;2:e365–6.
[42] Loughlin AM, Hsu K, Silverio AL, Marchant CD, Pelton SI. Direct and indirect effects of PCV13 on nasopharyngeal carriage of PCV13 unique pneumococcal serotypes in Massachusetts' children. Pediatr Infect Dis J 2014;33:504–10.
[43] Hammitt LL, Akech DO, Morpeth SC, et al. Population effect of 10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of *Streptococcus pneumoniae* and non-typeable *Haemophilus influenzae* in Kilifi, Kenya: findings from cross-sectional carriage studies. Lancet Glob Health 2014;2:e397–405.
[44] Simonsen L, Taylor RJ, Young-Xu Y, Haber M, May L, Klugman KP. Impact of pneumococcal conjugate vaccination of infants on pneumonia and influenza hospitalization and mortality in all age groups in the United States. mBio 2011;2(1).
[45] World Health Organization. Global Routine Immunization Strategies and Practices (GRISP). World Health Origination; 2019 [accessed Feb 18, 2019] http://www.who.int/immunization/programmes_systems/policies_strategies/GRISP/en/.
[46] World Health Organization. Periodic intensification of routine immunization. World Health Origination; 2009 [accessed Feb 18, 2019] https://www.who.int/immunization/programmes_systems/policies_ strategies/piri_020909.pdf?ua=1&ua=1.
[47] Ozawa S, Mirelman A, Stack ML, Walker DG, Levine OS. Cost-effectiveness and economic benefits of vaccines in low- and middle-income countries: a systematic review. Vaccine 2012;31:96–108.
[48] Close RM, Pearson C, Cohn J. Vaccine-preventable disease and the under-utilization of immunizations in complex humanitarian emergencies. Vaccine 2016;34:4649–55.
[49] Nnadi C, Etsano A, Uba B, et al. Approaches to vaccination among populations in areas of conflict. J Infect Dis 2017;216:S368–72.
[50] Lam E, McCarthy A, Brennan M. Vaccine-preventable diseases in humanitarian emergencies among refugee and internally-displaced populations. Hum Vaccines Immunother 2015;11:2627–36.

[51] World Health Organization. Vaccination in acute humanitarian emergencies. A framework for decision making. World Health Origination; 2017.
[52] Flasche S, Ojal J, Le Polain de Waroux O, et al. Assessing the efficiency of catch-up campaigns for the introduction of pneumococcal conjugate vaccine: a modelling study based on data from PCV10 introduction in Kilifi, Kenya. BMC Med 2017;15. https://doi.org/10.1186/s12916-017-0882-9.
[53] Gargano LM, Hajjeh R, Cookson ST. Pneumonia prevention during a humanitarian emergency: cost-effectiveness of *Haemophilus influenzae* type B conjugate vaccine and pneumococcal conjugate vaccine in Somalia. Prehospital Disaster Med 2015;30:402–11.
[54] Gargano LM, Hajjeh R, Cookson ST. Pneumonia prevention: Cost-effectiveness analyses of two vaccines among refugee children aged under two years, *Haemophilus influenzae* type b-containing and pneumococcal conjugate vaccines, during a humanitarian emergency, Yida camp, South Sudan. Vaccine 2017;35:435–42.
[55] Unicef. Update on humanitarian mechanism; 2018. https://www.unicef. org/supply/files/SESSION_7_2018_VIC_Update_on_Humanitarian_Mechanism. pdf.
[56] de Pereira A de L, Southgate R, Ahmed H, O'Connor P, Cramond V, Lenglet A. Infectious disease risk and vaccination in Northern Syria after 5 years of civil war: the MSF experience. PLOS Curr Disasters 2018. https://doi.org/10.1371/currents.dis.bb5f22928e631dff9a80377309381feb [published online Feb 2].
[57] WHO|MI4A: Vaccine purchase data. WHO. http://www.who.int/immunization/programmes_systems/procurement/v3p/platform/module1/en/ [accessed July 31, 2019].
[58] World Health Organization. The humanitarian mechanism. World Health Organization; 2017. http://www.who.int/immunization/programmes_systems/sustainability/The_Humanitarian_Mechanism_ToRs.pdf?ua=1.
[59] Blanchet K, Ramesh A, Frison S, et al. Evidence on public health interventions in humanitarian crises. The Lancet 2017;390:2287–96.
[60] Médecins Sans Frontières. Médecins Sans Frontières (MSF) briefing on provisional agenda item 20.1: Global vaccine action plan, Document A71/39; 2018. https://msfaccess.org/sites/default/files/2018-05/WHO-WHA-71_MSF-briefing-GVAP-Agenda-Item-20-1.pdf.
[61] Ravelo JL. MSF bends donation policy for pneumonia vaccine. Devex; 2015 [accessed Feb 27, 2019] https://www.devex.com/news/sponsored/msf-bends-donation-policy-for-pneumonia-vaccine-85317.
[62] Peyraud N, Quéré M, Duc G, et al. A post-conflict vaccination campaign, Central African Republic. Bull World Health Organ 2018;96:540–7.
[63] Médecins Sans Frontières. Nigeria: 'There were hardly any children under 5 years of age'; 2016. https://www.msf.ie/article/nigeria-there-were-hardly-any-children-under-5-years-age [accessed Feb 27, 2019].
[64] Médecins Sans Frontières. Vaccination activities in Médecins Sans Frontières; 2015. https://www.who.int/immunization/sage/meetings/2015/october/MSF_SAGE_October2015_vf.pdf.
[65] World Health Organization. Pneumococcal conjugate vaccines in infants and children under 5 years of age: WHO position paper – February 2019. Wkly Epidemiol Rec 2019;94:85–104.
[66] Ager A, Burnham G, Checchi F, et al. Strengthening the evidence base for health programming in humanitarian crises. Science 2014;345:1290–2.
[67] Heymann DL, Aylward RB. Mass vaccination: when and why. Curr Top Microbiol Immunol 2006;304:1–16.
[68] Goldblatt D, Southern J, Andrews NJ, et al. Pneumococcal conjugate vaccine 13 delivered as one primary and one booster dose (1 + 1) compared with two primary doses and a booster (2 + 1) in UK infants: a multicentre, parallel group randomised controlled trial. Lancet Infect Dis 2018;18:171–9.
[69] Temple B, Toan NT, Dai VTT, et al. Immunogenicity and reactogenicity of ten-valent versus 13-valent pneumococcal conjugate vaccines among infants in Ho Chi Minh City, Vietnam: a randomised controlled trial. Lancet Infect Dis 2019;19:497–509.
[70] ClinicalTrials.gov. Identifier: NCT02943902, reduced PCV dosing schedules in South African Infants. ClinicalTrials.gov; 2016. https://clinicaltrials.gov/ct2/show/NCT02943902 [published online Oct 25, accessed Feb 18, 2019].
[71] Leaning J. Ethics of research in refugee populations. The Lancet 2001;357:1432–3.
[72] Ojal J, Griffiths U, Hammitt LL, et al. TSustaining pneumococcal vaccination after transitioning from Gavi support: a modelling and cost-effectiveness study in Kenya. Lancet Glob Health 2019;7:644–54.
[73] Baguelin M, Flasche S, Camacho A, Demiris N, Miller E, Edmunds WJ. Assessing optimal target populations for influenza vaccination programmes: an evidence synthesis and modelling study. PLoS Med 2013;10:e1001527.
[74] Flasche S, Jit M, Rodríguez-Barraquer I, et al. The long-term safety, public health impact, and cost-effectiveness of routine vaccination with a recombinant, live-attenuated dengue vaccine (Dengvaxia): a model comparison study. PLOS Med 2016;13:e1002181.
[75] Le Polain De O, Waroux, Edmunds WJ, Takahashi K, et al. Predicting the impact of pneumococcal conjugate vaccine programme options in Vietnam. Hum Vaccines Immunother 2018;14:1939–47.
[76] Havumaki J, Meza R, Phares CR, Date K, Eisenberg MC. Modeling cholera transmission and vaccination in a refugee camp. Epidemiology 2019. https://doi.org/10.1101/514406.
[77] Finger F, Funk S, White K, Siddiqui MR, Edmunds WJ, Kucharski AJ. Real-time analysis of the diphtheria outbreak in forcibly displaced Myanmar nationals in Bangladesh. BMC Med 2019;17:58.
[78] Finger F, Bertuzzo E, Luquero FJ, et al. The potential impact of case-area targeted interventions in response to cholera outbreaks: a modeling study. PLOS Med 2018;15:e1002509.
[79] Kucharski AJ, Eggo RM, Watson CH, Camacho A, Funk S, Edmunds WJ. Effectiveness of ring vaccination as control strategy for Ebola virus disease. Emerg Infect Dis 2016;22:105–8.
[80] Cooper BS, White LJ, Siddiqui R. Reactive and pre-emptive vaccination strategies to control hepatitis E infection in emergency and refugee settings: a modelling study. PLoS Negl Trop Dis 2018;12:e0006807.
[81] Mossong J, Hens N, Jit M, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med 2008:5. https://doi.org/10.1371/journal.pmed.0050074.
[82] Le Polain O, de Waroux, Cohuet S, Ndazima D, et al. Characteristics of human encounters and social mixing patterns relevant to infectious diseases spread by close contact: a survey in Southwest Uganda. BMC Infect Dis 2017;18. https://doi.org/10.1101/121665.
[83] Althouse BM, Hammitt LL, Grant L, et al. Identifying transmission routes of *Streptococcus pneumoniae* and sources of acquisitions in high transmission communities. Epidemiol Infect 2017;145:2750–8.
[84] Weinberger DM, Pitzer VE, Regev-Yochay G, Givon-Lavi N, Dagan R. Association between the decline in pneumococcal disease in unimmunized adults and vaccine-derived protection against colonization in toddlers and preschool-aged children. Am J Epidemiol 2019;188:160–8.
[85] Lopez Bernal JA, Andrews N, Amirthalingam G. The use of quasi-experimental designs for vaccine evaluation. Clin Infect Dis 2019;68:1769–76.
[86] Silaba M, Ooko M, Bottomley C, et al. Effect of 10-valent pneumococcal conjugate vaccine on the incidence of radiologically-confirmed pneumonia and clinically-defined pneumonia in Kenyan children: an interrupted time-series analysis. Lancet Glob Health 2019;7:e337–46.
[87] Bruhn CAW, Hetterich S, Schuck-Paim C, et al. Estimating the population-level impact of vaccines using synthetic controls. Proc Natl Acad Sci U S A 2017;114:1524–9.
[88] UNHCR population statistics – Data – Demographics. http://popstats.unhcr.org/en/demographics [accessed July 29, 2019].

Chapter 2

Social contacts and other risk factors for respiratory infections among internally displaced people in Somaliland

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1604011	**Title**	Mr
First Name(s)	Kevin		
Surname/Family Name	van Zandvoort		
Thesis Title	Strategies for pneumococcal conjugate vaccine use in humanitarian crises		
Primary Supervisor	Stefan Flasche		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Epidemics		
When was the work published?	29/08/2022		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	Prior to the survey, I led the study design, conducted a scoping visit to the field site, prepared all electronic Open Data Kit survey forms, and worked with collaborators at the Murdoch Children's Research Institute, London School of Hygiene and Tropical Medicine, and Save the Children International to prepare and ship required study items (electronic tablets, anthropometric measurement tools, etc) to the field site. I spent a total of ten weeks in Somaliland. Initially this was to finalize agreements with collaborators at Save the Children and the Ministry of Health Development, training data collectors on the study, sampling process, requesting informed consent, electronic data collection using Open Data Kit, and to participate in a community engagement activity in Digaale. During the survey, I led the daily data collection in Digaale: supervising the five data collection teams during data collection, synchronizing data between electronic tablets and an external server, assessing and validating collected data in the evening, discussing any issues with the data collection teams the next morning, and regularly updating local collaborators on the progress of the study. Once data collection was completed, I cleaned all the data and conducted all statistical analyses. I then interpreted the results, produced all figures and tables, wrote the original manuscript draft, and edited and prepared the final manuscript for submission.

SECTION E

Student Signature	Kevin van Zandvoort
Date	03/10/2024

Supervisor Signature	Stefan Flasche
Date	03/10/2024

Epidemics 41 (2022) 100625

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

Social contacts and other risk factors for respiratory infections among internally displaced people in Somaliland

Kevin van Zandvoort [a,b,*], Mohamed Omer Bobe [c], Abdirahman Ibrahim Hassan [d], Mohamed Ismail Abdi [c], Mohammed Saed Ahmed [c], Saeed Mohamood Soleman [d], Mohamed Yusuf Warsame [d], Muna Awil Wais [c], Emma Diggle [e], Catherine R. McGowan [e,f], Catherine Satzke [g,h], Kim Mulholland [a,g], Mohamed Mohamoud Egeh [c], Mukhtar Muhumed Hassan [c], Mohamed Abdi Hergeeye [d], Rosalind M. Eggo [a,b], Francesco Checchi [a], Stefan Flasche [a,b]

[a] Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom
[b] Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom
[c] Save the Children International, Maansoor area, Jig-jiga yar, Hargeisa, Somaliland
[d] Republic of Somaliland Ministry of Health Development, 26 June District, Presidential Road, Hargeisa, Somaliland
[e] Save the Children UK, 1 St John's Lane, London EC1M 4AR, United Kingdom
[f] Department of Public Health, Environments and Society, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom
[g] Infection and Immunity, Murdoch Children's Research Institute, The University of Melbourne Department of Paediatrics at the Royal Children's Hospital, Flemington Road, Parkville, Victoria 3052, Australia
[h] Department of Microbiology and Immunology, The University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Victoria 3010, Australia

ARTICLE INFO

Keywords:
Contact data
Contact pattern
Acute Respiratory Infection
Internally displaced people
Humanitarian crises

ABSTRACT

Background: Populations affected by humanitarian crises experience high burdens of acute respiratory infections (ARI), potentially driven by risk factors for severe disease such as poor nutrition and underlying conditions, and risk factors that may increase transmission such as overcrowding and the possibility of high social mixing. However, little is known about social mixing patterns in these populations.

Methods: We conducted a cross-sectional social contact survey among internally displaced people (IDP) living in Digaale, a permanent IDP camp in Somaliland. We included questions on household demographics, shelter quality, crowding, travel frequency, health status, and recent diagnosis of pneumonia, and assessed anthropometric status in children. We present the prevalence of several risk factors relevant to transmission of respiratory infections, and calculated age-standardised social contact matrices to assess population mixing.

Results: We found crowded households with high proportions of recent self-reported pneumonia (46% in children). 20% of children younger than five are stunted, and crude death rates are high in all age groups. ARI risk factors were common. Participants reported around 10 direct contacts per day. Social contact patterns are assortative by age, and physical contact rates are very high (78%).

Conclusions: ARI risk factors are very common in this population, while the large degree of contacts that involve physical touch could further increase transmission. Such IDP settings potentially present a perfect storm of risk factors for ARIs and their transmission, and innovative approaches to address such risks are urgently needed.

1. Introduction

Prior to the COVID-19 pandemic, acute respiratory infections contributed to over 2 million deaths annually (Troeger et al., 2018). These pathogens are transmitted via direct or indirect (e.g. via fomites) contact with respiratory droplets. Unsurprisingly, patterns of transmission for respiratory pathogens correlate strongly with social contact patterns among different groups of people (Melegaro et al., 2011; le

* Corresponding author at: Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom.
E-mail address: Kevin.Van-Zandvoort@lshtm.ac.uk (K. van Zandvoort).

https://doi.org/10.1016/j.epidem.2022.100625
Received 31 January 2022; Received in revised form 16 August 2022; Accepted 25 August 2022
Available online 29 August 2022

Polain de Waroux et al., 2018). In addition, other factors such as malnutrition (Rytter et al., 2014) modulate host responses to infection thereby affecting both individual susceptibility to infection and the rate of pathogen shedding, i.e. infectiousness to others.

Accordingly, epidemiological transmission models commonly stratify populations into groups, usually by age, and assume different contact rates among these groups. To accurately parametrize such models, several surveys have collected empirical contact data by asking participants to report the frequency with which they come into direct contact with people from their and other groups (Mossong et al., 2008; le Polain de Waroux et al., 2017; Kiti et al., 2014; Melegaro et al., 2017; Johnstone-Robertson et al., 2011). The majority of these studies have been conducted in high-income countries, with limited information from lower- and middle-income countries (Hoang et al., 2019). To our knowledge, no contact surveys have captured the experience of populations living in refugee or internally-displaced persons' (IDP) camps, despite these settings experiencing a high burden of endemic and epidemic-prone infections, including respiratory diseases (Bellos et al., 2010; Lam et al., 2015).

Nearly 80 million people were forcibly displaced due to insecurity and war in 2019 (United Nations High Commissioner for Refugees, 2020). Many of these people live in overcrowded camps where residents are often unemployed or unable to work, not all children may have access to school, hygiene may be curtailed by inadequate water and sanitation services, and people may be required to gather in large groups during relief distributions or at queues at water points, etc. These factors likely affect social contact patterns, in turn impacting the spread of infections. Food insecurity and resultant acute malnutrition are also a major threat, particularly among IDPs, and increase both transmission and case-fatality of most respiratory infections (Rodríguez et al., 2011).

Epidemiological transmission models are an increasingly appreciated tool to quantify infection transmission and efficiently explore the impact of possible interventions. In the context of a project to study the effectiveness of pneumococcal conjugate (PCV) vaccination among IDPs, we aimed to parameterise a model of *Streptococcus pneumoniae* transmission (van Zandvoort et al., 2019). We therefore conducted a contact and risk factor survey in an IDP camp in Somaliland in 2019 to quantify factors commonly associated with transmission of respiratory infections.

2. Methods

2.1. Study population

We conducted a cross-sectional survey on social contacts and prevalence of risk factors in the Digaale IDP camp near Hargeisa, the capital of Somaliland. The camp was established in 2014, in response to a large influx of displaced people in Somaliland following an acute food insecurity crisis. It is situated 3 km from Hargeisa airport and 4 km from the city borders, has an area of 15 ha and is surrounded mainly by desert and shrubland. Digaale has a small primary healthcare centre, that operates under the Somaliland Ministry of Health Development, and a primary school. There are 894 shelters in Digaale that are all constructed from corrugated sheets. Each household has access to a private latrine and private water tank. Several photos of Digaale IDP camp are provided in section A of the Supplementary Material.

2.2. Study design and sample selection

The survey was nested within a larger study of pneumococcal carriage, and aimed to include a representative sample of 100 individuals in each of the following age groups: <1, 1, 2–5, 6–14, 15–29, 30–49, 50 + years old. Purposes and procedures of the study were presented to and discussed with the elders and other representatives from the community in advance. Data collection was conducted by the Somaliland Ministry of Health Development in collaboration with Save the Children International, who provide health and nutrition services in the camp.

A team of 12 enumerators were trained for a week and undertook data collection between October- November 2019. Data collection was done from 8AM until 4PM on all days of the week except Friday. A community leader notified the population living in the block that would be visited on specific days, and answered any questions regarding the presence of enumerators. Adults provided informed written consent, including on behalf of their children under the age of 18. Children aged >12 and <18 years were asked for assent.

As there was no sampling frame, we employed quota sampling to reach the desired sample size in each age group. We visited all shelters in Digaale and conducted a household survey in consenting households. A household was defined as individuals who live together and share a common source of food. The survey followed the Standardised Monitoring and Assessment of Relief and Transitions (SMART) guidance to collect the age and gender of those currently living in the household, and retrospectively assessed who had been born or died, and who migrated in or out of the household in the preceding six months (SMART, 2017). A localized event calendar was used to aid in the recollection of timing of events. In addition, the survey included questions to ascertain the presence of household level risk factors: total number of rooms, leakage or draft in the shelter, cooking fuel used, ventilation used in the cooking area, water source, and substance use in the household. At completion of the survey, household members were assigned sampling weights based on their age. These weights were then used to sample household members, who were invited to take part in the second stage of the survey. We asked those consenting to remember all people they would contact on the following day, and agreed a time for a return visit on the day thereafter.

During the return visit, we conducted a second survey that assessed individual travel patterns and direct social and physical contacts. The survey was developed as an extension of contact surveys conducted in non-IDP populations (Mossong et al., 2008; le Polain de Waroux et al., 2017). Parents or caregivers acted as proxies for young children (<10 years) who were unable to answer for themselves. Participants were asked to first list (nick)names or initials of all their contactees. A contactee was defined as any individual who was met in person during the 24 h before waking up on the day being surveyed, and with whom the participant had at least a short conversation in close proximity (direct contact). For each contactee, participants were asked to list several characteristics: estimated age (in years) and gender, relationship to the contactee, setting where the contact occurred, type of contact (physical touch or nonphysical), duration of the contact (chosen from a category: <15 m, 15 m - 1 h, 1–2 h, 2–4 h, >4 h), and typical frequency with which participant and contactee have contact. Physical touch was defined as any form of skin-to-skin contact. Participants were also asked to estimate the total number of indirect in-person contacts they made that did not fit the criteria for a direct contact, which are people who were met in close proximity but without having a conversation or skin-to-skin contact, and to list any health conditions diagnosed by a health professional. If pneumonia was listed as a health condition, we asked whether diagnosis occurred within the six months preceding the survey.

For all participants aged 6–59 months, we measured height (or length if less than 85 cm), weight, middle-upper arm circumference (MUAC) and presence of bilateral oedema on the dorsum of both feet as described in the SMART guidance. Anthropometric instruments and observations were standardised prior to data collection on a convenience sample of children seen at the health facility.

Shelters where no individual was present on the first visit were revisited on different days throughout the study period, up to a total of five times. To assess selection bias for shelters where no person was present on all visits, we asked neighbours of a random sample of 96 of these shelters whether their neighbouring shelter was still inhabited. Participants who were not available on the return visit were also revisited up to a total of five times. If participants were no longer available, or withdrew consent, they were replaced with another

individual of the same household within the same age group, where possible.

All data were collected on electronic tablets using Open Data Kit software (Hartung et al., 2010). Ethical approval for the study was granted by the Somaliland Ministry of Health Development, Directorate of Planning, Policy, and Strategic Information, and by the London School of Hygiene & Tropical Medicine. The funding sources had no role in the study design; collection, analysis, and interpretation of data; or in writing the report.

2.3. Data analysis

As the sample for the participants included in the second phase of our survey was not self-weighting, and we deliberately oversampled individuals in the youngest and oldest age groups, we used post-stratification weights to calculate population representative estimates of individual-level data (Lumley, 2011). These weights are the inverse of a participant's probability of selection in the sample, which was the inverse of the estimated proportion of the population included in the sample in their respective age- and gender stratum. We further applied an additional weight to correct for imbalances in the distribution of days of the week within the final survey sample when estimating contact rates. We censored sample weights below or above the 5th and 95th percentile of all sample weights to those percentiles. We explored inclusion of household size as an additional variable in calculating post-stratification weights, but this resulted in too sparse strata due to a small number of participants. Stratified random sampling and a larger sample size would allow doing so in future studies. We assessed the sensitivity of post-stratification weights on estimated contact rates in section D of the Supplemental Material.

As we sampled a large proportion of households (65%) and individuals (17%) living in Digaale, we used finite population corrections (FPC) to calculate standard errors. The total population size of Digaale used in the FPC, for all ages and within each age group, was estimated by multiplying the total number of household members reported in the survey with a correction factor. This factor was calculated as the survey-estimated proportion of inhabited shelters included in the survey, the latter being the upper bound of the survey-estimated proportion of non-abandoned shelters. More detailed information on the sampling design, finite population correction, and post-stratification is provided in Supplemental Material section B. The *survey* package in R was used to perform the weighting and to apply the FPC when estimating weighted means, proportions, and quantiles where applicable (Lumley, 2020).

We estimate crude birth (live births per 1000 people per year), death (number of deaths per 1000 people per year), and migration rates (number of people migrating in- or out of the population per 1000 people per year) from events occurring in the six months preceding the survey. To do so, we calculated person-time for all household members reporting to live in Digaale during the period, including those who had died or migrated out, by assigning six months (the full recall period) to any individual who lived continuously within the household during the recall period. As we did not record exact dates when deaths, births and migration events occurred, we imputed person-time for individuals experiencing these events by randomly sampling from uniform distributions $U(0,t)$ for all experienced events, where t was set to 6 months for the first event, or the result of the previous sampled value for all subsequent events. We repeated this process to generate 10,000 datasets. Weighted rates were estimated by fitting Poisson generalized linear models using the natural logarithm of person-time as an offset. We were not able to account for censoring after migration out of the household when estimating death rates.

Nutritional data were analysed by calculating age- and sex-standardised z-scores for a range of anthropometric indices based on the World Health Organization Growth Reference standards (World Health Organization, 2006), which were subsequently categorized to assess malnutrition status. The *zscorer* package in R was used for the anthropometric analyses (Myatt and Guevarra, 2019).

We constructed contact matrices to visualise age-stratified daily contact rates (average number of daily contacts with individuals in a population group) and per capita contact rates (probability of a contact occurring with a single member of a population group on an average day), which were adjusted for reciprocity in the total number of contacts using the method by Wallinga et al (Wallinga et al., 2006). Uncertainty in the contact matrix was quantified by taking 10,000 bootstrap samples of all participants in the survey. Contacts for contactors living in the same household were assumed to be independent. Methods are described in more detail in Supplemental Material section C.

All analyses were conducted in R 4.0 (Core Team, 2021). Analysis scripts, anonymized data, and questionnaire scripts are available on GitHub via https://github.com/kevinvzandvoort/espicc-somaliland-digaale-survey-2019. The anonymized contact data has also been uploaded to the respective repository on Zenodo (https://doi.org/10.5281/zenodo.5226281) to be accessible through the socialmixr package in R for epidemiological contact surveys (Funk, 2018).

3. Results

We visited all 894 shelters in Digaale IDP camp. On all occasions visited, no individuals were present in 405 shelters. We randomly sampled 96 of these empty shelters and asked neighbours whether the shelter was occupied. Using this sample, we estimate that 12% (6 19) of empty shelters were either a shop or combined with another shelter already included in the survey, resulting in a conservative estimate of 872 unique shelters. 50% (95%CI 39–60) of the empty shelters had been uninhabited for a long time, while 38% (28 49) of the empty shelters were occupied. Using the upper bound of these estimates, we conservatively assume that 715 of the 872 unique shelters in Digaale were inhabited.

Twenty-five households declined consent. We thus collected demographic information from 2049 individuals who were living in 464 households at the time of the survey (65% of inhabited unique shelters), with additional information collected regarding 166 individuals who migrated out of these households and 34 individuals who had died in the six months preceding the survey. In the contact survey we enrolled 509 participants from 426 households, who provided information regarding 4857 contacts. Of all participants included in the contact survey, we collected anthropometric estimates from 171 children aged 6–59 months. Despite their inclusion and consent in the first household visit, individuals from 22 households were lost to follow up for participation in the contact survey, with a further two individuals declining consent. Section B in the Supplemental Material explains the sampling of individuals in more detail.

3.1. Demographic characteristics

The median age among household study participants was 15 years; 25% were younger than 7 years old and 75% were younger than 34 years old (Fig. 1). There were notably fewer adult men than women. The male to female gender ratio among enrolled household members of all ages was 1:1.2, and 1:1.5 in adults. The respective gender ratio was 1:1.9 among the participants enrolled in the contact study and 1:1.9 among their reported contacts.

The crude birth and death rates were estimated as 32 and 33 per 1000 per year, respectively, in the six months preceding the survey. The crude under 5 years death rate was estimated at 57 per 1000 per year, and 106 per 1000 per year in those aged 50 years and older. Crude in- and out-migration rates were high at 139 and 161 per 1000 per year in the same period (Table 1). The majority of households (79%) settled in Digaale more than 3 years prior to the survey, while 6% settled in the year prior to the survey. The median household size was 4 individuals, ranging from 1 (83 households) to 12 (3 households) (Fig. 1A).

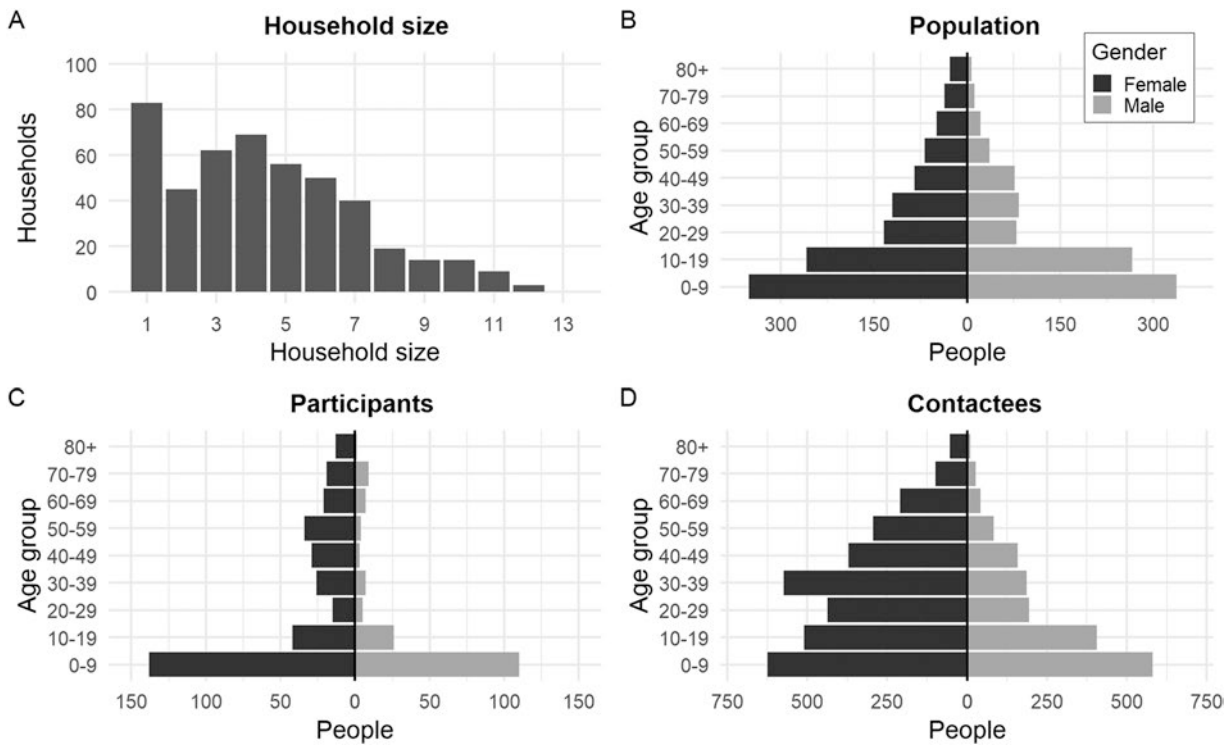


Fig. 1. Household and population distribution. A. Frequency distribution of household sizes. B. age and gender distribution of all household members in participating households, which was used as the assumed population distribution. C. Age and gender distribution of participants included in the contact survey. D. Age and gender distribution of contactees listed by participants.

3.2. Living conditions

The majority of households had only a single room (93%), with 6% having two rooms and 2% three or four rooms. Most households reported having draughts (73%) and leakage (74%) in their shelter. Both firewood (84%) and charcoal (60%) were commonly reported as cooking fuels. The majority of households usually cook outside (66%) or in a ventilated area inside (11%), with a minority (23%) of households reporting to cook in an unventilated area. Khat (36%) and tobacco (27%) were reported to be consumed by at least one household member in around one third of all households.

3.3. Travel patterns

We estimate that just under half of the population never travels outside the Digaale camp, while one fifth does so at least once per week. Those who do travel predominantly go to nearby locations, including the city of Hargeisa. Less than 10% reported travelling further than 10 km from the camp and less than 3% more than once a week (Table 2).

3.4. Anthropometric status

We estimate that 14% of all children under 5 years old in the population were underweight for their age, and 4% of all children severely underweight. Prevalence of global acute malnutrition (wasting, GAM) was 9% when assessed using weight-for-height Z score, with 2% severely malnourished. In contrast, GAM was 3% when assessed by MUAC, with no children severely malnourished and none diagnosed with oedema. 20% of children were stunted.

3.5. Social contact patterns

The average number of direct daily contacts was relatively homogeneous by age (Table 3), and was highest for those aged 2–5 years (10.7) and lowest for those aged 50 + years (8.8). A large proportion (>77%) of these contacts were physical. The proportion of physical contacts was relatively homogeneous across contact frequency, duration, relationship, and setting (Fig. 2). Nearly all direct contacts were made with previously known individuals (99% of all contacts), mostly household members (34% of all contacts), relatives not in the household (25%), and friends (30%). Most contactees were met daily or almost daily (88%), and most (42%) reported contacts lasted longer than four hours. Very few (8) contacts were reported to be made with people never met before. All age groups reported a high number of indirect casual contacts in addition to their direct contacts, with little variability across age-groups. Prior to adjusting the contacts for reciprocity, contact intensities reported by contactors aged 0–9, 10–19, and 60 + years were substantially higher than contact intensities with contactees of these age groups as reported by contactors of other ages (Supplemental Figure D2).

The vast majority (>82%) of direct contacts were made at a home or in another house, with few reported contacts at school or work, or in other settings. The average number of school or work contacts was higher when restricted to participants who report at least one school or work contact, but remains lower compared to contacts made at home or in another house in all age groups (Supplemental Table D3).

The contact matrices in Fig. 2 show who contacts whom. We observe the highest average daily number of contacts (2 E) within the same age group to be among children. The higher overall contact rates made with children reflects the relatively high proportion of the population that are children, while the per capita contact rates (2 F) give an estimate of assortativity after adjusting for the population size. Overall, contacts were mostly age-assortative, especially in children, with more intergenerational contacts in adults. Out of 4857 contacts, only 91 were reported to occur outside of Digaale, all of which occurred in Hargeisa. 95% uncertainty estimates around the contact rates are shown in Supplemental Figure D5. Note that the high per-capita contact rates in the oldest age group in Fig. 2 F results from a small population size in this age-group.

Table 1
Characteristics of participating households and prevalence of risk factors in Digaale IDP camp.

Variable	N[a]	Estimate[b]	
Demographic characteristics			
Median household size	464	4	2–6 (IQR)
Median age	2049	15	7–34 (IQR)
Crude in-migration rate[c]	144	139.3	127.1–152.8
Crude out-migration rate[c]	166	160.6	147.4–175
Crude birth rate[c]	33	31.9	26.3–38.7
Crude death rate (all ages)[c]	34	32.9	27.2–39.7
Crude death rate by age[c]			
< 5 years old	9	56.5	39.1–81.8
6–14 years old	0	0	0–0
15–29 years old	6	26.4	16.8–41.3
30–49 years old	4	22.1	12.7–38.3
50 +years old	14	105.9	78.8–142.2
Years since household settled in Digaale			
< 1 year	27	5.8%	4.6–7.1
1–2 years	39	8.4%	6.9–9.9
2–3 years	31	6.7%	5.3–8
> 3 years	367	79.1%	76.9–81.3
Quality of shelters			
Total number of rooms (mean)	505	1.1	1.1–1.1
Reported draught in shelter	338	72.8%	70.4–75.2
Reported leakage in shelter	341	73.5%	71.1–75.9
Indoor air pollution			
Cooking fuel used			
Charcoal	275	59.3%	56.6–61.9
Firewood	388	83.6%	81.6–85.6
Ventilation in cooking area			
Cook outside	308	66.4%	63.8–68.9
Ventilation absent	107	23.1%	20.8–25.3
Ventilation present	49	10.6%	8.9–12.2
Primary water source			
Tanker truck delivery	456	98.3%	97.6 – 99.0
Rainwater collection	7	1.5%	0.9–2.2
Substance use in household			
None	289	62.3%	59.7–64.9
Khat	168	36.2%	33.6–38.8
Smoke	126	27.2%	24.8–29.6
Snuff	29	6.2%	4.9–7.6
Alcohol	0	0%	0–0
Prevalence of malnutrition in U5			
Weight for age			
Not underweight (z > 2)	136	81.8%	77.4–86.3
Underweight (z ≤ 2)	27	14.1%	10.1–18.1
Severely underweight (z ≤ 3)	8	4.1%	1.9–6.3
Height for age			
Not stunted (z > 2)	124	71.7%	65.8–77.6
Stunted (z ≤ 2)	31	19.5%	14.1–24.9
Severely stunted (z ≤ 3)	16	8.8%	5.3–12.2
Weight for height			
Not wasted (z > 2)	150	89.2%	85.5–92.9
Wasted (z ≤ 2)	18	9.3%	5.8–12.8
Severely wasted (z ≤ 3)	3	1.5%	0.1–2.8
Middle-Upper Arm Circumference			
Not wasted (≥ 125 mm)	165	96.8%	94.5–99.1
Wasted (< 125 mm)	5	3.2%	0.9–5.5
Severely wasted (< 115 mm)	0	0%	0–0
Cumulative incidence of self-reported pneumonia[d]			
In the last six months			
< 2 years old	32	42.8%	34.5–51
2–5 years old	31	26%	19.2–32.8

(*continued on next page*)

Table 1 (*continued*)

Cumulative incidence of self-reported pneumonia[d]			
6–14 years old	11	12.4%	5.7–19.1
15–29 years old	4	6.6%	0.3–13
30–49 years old	7	10.6%	2.4–18.8
50 + years old	3	3%	0–6.1

[a] Total number of observations
[b] Central estimates are the weighted mean, median, or percentage. The uncertainty interval next to the central estimates is the 95% confidence interval for proportions and means, and interquartile range (IQR) for median estimates. A finite population correction was applied in calculating confidence intervals.
[c] Rates are per 1000 people per year
[d] Pneumonia in young children (<10 years) was reported by an adult or caregiver.

Supplemental Figure D6 shows contact patterns by setting. They are relatively age-assortative outside the household, while more age-disassortative within the household. Contacts were also gender-assortative (Supplemental Figure D7). Children had higher contact rates with adult women than adult men, and age-assortative mixing was higher in men than in women.

3.5.1. *Self-reported pneumonia illness*

There was a strong decreasing trend by age for self-reported cumulative incidence of pneumonia diagnosis (Table 1). 46% of children under 2 were reported to have been diagnosed with pneumonia within the last six months. We assessed the association between the exposure variables and self-reported pneumonia cases separately, results of which are provided in section E of the Supplemental Material.

4. Discussion

To our best knowledge this is the first study to collect data on social contacts in any IDP camp, and to describe risk factors relevant for the spread of infectious respiratory diseases in a Somaliland IDP camp. We found that the majority of households had been living in Digaale for over three years, while estimated crude migration rates were relatively high. There was a high female to male ratio in adults living in Digaale.

We estimated a low mean age and corresponding high crude death rates, especially in those younger than five years. The estimated crude death rates in the six months preceding the survey (per 1000 per year) were 33 for all ages and 57 for children younger than 5 years, and are considerably higher than the average crude death rates reported for Somalia including Somaliland between 2013 and 2018 (16, 0–59; and 24,0 to 91) (Warsame et al., 2020) There are challenges in comparing death rates across space and time, but our findings underscore the consistently higher mortality observed among IDPs, compared to other population groups (Heudtlass et al., 2016). Although we do not know the aetiology of these deaths, a high proportion of participants reported a historical pneumonia diagnosis, especially in younger children.

Several known risk factors for respiratory infections in children are prevalent (Sonego et al., 2015; Mulholland and Weber, 2016). First, approximately 14% of children in Digaale were underweight. Although no children were found to have severe acute malnutrition, food insecurity may vary substantially over time. Nearly a fifth of children had stunted growth, reflecting long-term undernutrition, which may indicate periods of inadequate access to food for their families. These data are similar to estimates in IDP populations in Somalia (Grijalva-Eternod et al., 2018).

Second, individuals live in relatively poor-quality shelters, while local minimum temperatures can drop to 5 °C. Firewood and charcoal are the only cooking fuels used in Digaale, and both can raise levels of indoor air pollution. However, only a small proportion of households cooks in an unventilated area indoors, which mitigates this risk. At least one household member smokes in one third of all households, which could further affect levels of indoor air pollution.

Third, levels of crowding within Digaale are substantial. While the average household size of four people is below the national average (Central Statistics Department, 2020), most households share only a single small room. Increased human contact facilitates the transmission of pathogens causing respiratory diseases. On average, individuals have

Table 2
Frequency of travel outside Digaale IDP camp.

Travel distance	Most days of the week		At least once per week		At least once per month		Less than once per month		Never	
< 5 km	18.1%	13.8–22.4	21.6%	17–26.1	13.4%	9.5–17.3	3.8%	1.9–5.7	43.1%	38.1–48.2
5–10 km	9.9%	6.1–13.7	12.5%	8.9–16.1	16.9%	13.3–20.6	5.8%	3.6–7.9	54.9%	49.5–60.2
> 10 km	1.3%	0–2.8	1.5%	0–3	3.6%	1–6.2	3.9%	1.4–6.3	89.7%	85.9–93.5

Estimates are the weighted proportion and corresponding 95% confidence interval. A finite population correction was applied in calculating confidence intervals.

Table 3
Mean number of reported daily contacts by age, contact type and contact setting.

Age group	Contact type						Contact setting (Direct)[a]					
	Total (Direct)[a]		Physical (Direct)[b]		Total (Indirect)[c]		Home or another house		School or work		Other	
< 2[d]	9.1	8.6–9.6	6.6	6.1–7.2	11.4	10.4–12.4	8.4	8–8.9	0.1	0–0.1	0.7	0.5–0.9
2–5[d]	10.6	10–11.1	8.6	8–9.2	11.7	10.8–12.6	9.9	9.4–10.5	0.2	0–0.4	0.6	0.4–0.8
6–14[d]	10.2	9.2–11.1	8.3	7.3–9.4	13.0	11.7–14.4	7.1	6.2–7.9	2.3	1.5–3.1	1.0	0.5–1.4
15–29	9.2	8.3–10.1	7.3	6.2–8.4	14.3	12.9–15.8	5.9	5.1–6.8	1.3	0.6–1.9	2.3	1.5–3.2
30–49	9.9	8.5–11.2	7.0	5.7–8.3	14.6	12.6–16.6	6.6	5.6–7.6	1.3	0.2–2.5	2.1	1–3.1
50 +	8.6	8–9.3	5.9	5.3–6.6	12.3	11–13.6	6.5	5.9–7	0.5	0.2–0.7	1.9	1.5–2.3

Estimates are the weighted mean and corresponding 95% confidence interval. A finite population correction was applied in calculating confidence intervals.
[a] Direct contacts were defined as in-person contacts with whom the participant had at least a short conversation.
[b] Physical contacts were direct contacts involving physical touch.
[c] Indirect contacts were additionally reported contacts that did not fit the definition of a direct contact.
[d] Contacts made by young children (<10 years) were proxy reported by adults or caregivers.

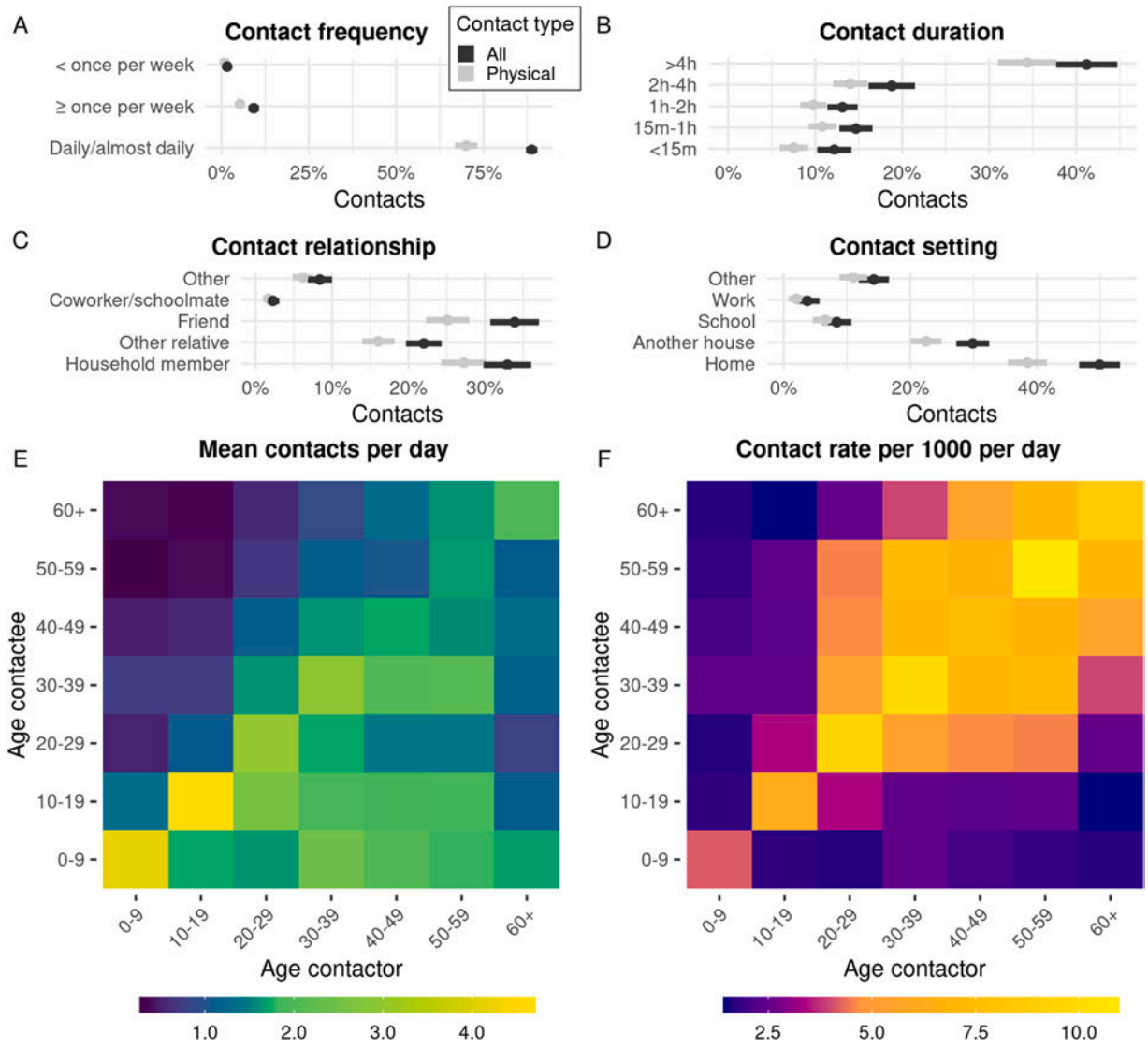


Fig. 2. Contact patterns and matrices. Panels show the reported frequency by which contactors meet contactees (A), the duration of contacts (B), the relationships between contactors and contactees (C), and the setting where contacts occurred (D). Direct contacts are stratified by contact type. Contact matrices show the weighted mean number of daily contacts made by contactors with contactees of certain age groups (E), and the age-specific weighted daily per-capita contact rates (F), reported per 1000 people (i.e. the rate at which any two individuals are assumed come into contact each day). Contacts made by young children (<10 years) were reported by proxy by adults or caregivers. Both matrices are adjusted for reciprocity of contacts.

10 direct contacts each day, which are age-assortative, and the majority of these contacts involve physical touch. The average number of reported direct contacts was lower in this setting when compared to contact surveys conducted in Kenya (Kiti et al., 2014) or a South African township (Johnstone-Robertson et al., 2011), though this likely reflect a difference in survey design, as we excluded casual contacts from our contact definition. In Digaale, individuals reported on average a further 13 indirect contacts per day, though these are less likely to result in respiratory transmission.

The average proportion of direct contacts that were physical was higher in Digaale (78%) compared to a crowded township in South Africa (27%) (Johnstone-Robertson et al., 2011), and a peri-urban township in Zimbabwe (57%) (Melegaro et al., 2017), but similar to a rural setting in Uganda (73%) (le Polain de Waroux et al., 2017). Compared to these settings, we find that contacts in Digaale are more likely to occur at the home, with very few made at school or work.

There are several limitations to our study. First, our estimates may be affected by selection bias, as the proportion of male participants of working age included in the survey was smaller than the proportion of men of this age in the population. We could only conduct data collection during daylight hours, and may therefore have missed individuals who work outside Digaale, as many leave the camp very early in the morning and only return late at night. Our contact matrices partially account for this by i) applying post-stratification weights to those individuals of this subpopulation that were included, and ii) adjusting the contact matrices to account for the reciprocity of contacts. However, contacts made between both contactors and contactees in this subpopulation, and with contactees not living in Digaale, are not observed and therefore not represented in the contact matrices. Practitioners should be aware of this when applying these estimates in their models, especially if they believe social contacts within this subpopulation to be of importance in the spread of a particular pathogen. Generally, contact matrices assume a closed population, which may be less appropriate for populations with large numbers of contacts outside the population. This may similarly have affected the reported travel patterns, with frequent travellers least likely to be sampled for study participation. Of the 35% inhabited shelters not included in the survey, many were (according to neighbours) single occupancy dwellings whose residents were absent during

working hours.

Second, parents or caregivers completed the survey on behalf of young children, which may have biased the estimates of non-household contacts of those children, and explain why e.g. reported direct contacts at school are very low. We found a discrepancy between contacts reported by contactors aged 0–9, 10–19, and 60 + years, and contacts reported with contactees aged 0–9, 10–19, and 60 + years, but were not able to assess the accuracy of proxy reporting further. The consistencies of these discrepancies suggest that these are not only due to sampling error. Similar discrepancies are present in other contact surveys (Mossong et al., 2008; le Polain de Waroux et al., 2017; Melegaro et al., 2017; Béraud et al., 2015; Leung et al., 2017; Horby et al., 2011; Grijalva et al., 2015; Zhang et al., 5, 2020) (Supplemental Figure D3), often in the same age groups, though the direction of discrepancies differs between studies.

Third, we designed the survey to collect retrospective data on social contacts, which could result in a lower reported number of contacts due to recall bias (Hoang et al., 2019). We asked participants to remember their contacts one day in advance, and used structured interviews with specific prompts in an attempt to minimize any recall issues. We assessed contact rates for potential underreporting by comparing the reported intra-household contact rates with the expected intra-household contact rates based on household demographics (Supplemental Figure D6). While the reported contact patterns were very similar to the expected contact patterns, their absolute values were lower, with the dominant eigenvalue of the matrix 43% lower than that of the expected matrix. This may reflect underreporting of household contacts, but could also reflect household members having less than one contact per day, on average. Mortality, birth, and migration rates were estimated retrospectively using SMART survey methodology. We aimed to minimize recall bias using local event calendars and structured questionnaires, but cannot rule it out.

Fourth, our sample size was chosen to estimate age-specific pneumococcal carriage prevalence, and therefore oversampled young children. While this was still sufficient to estimate contact patterns with adequate precision, a higher sample size in older individuals would have allowed us to estimate contact patterns in those age groups with more detail.

5. Conclusion

We find a high prevalence of risk factors for lower respiratory tract infections in an IDP camp in Somaliland. Crude death rates in the camp exceeded the already high rates in the host population of Somaliland. Compared to other settings, a large degree of contacts are of a physical nature, and the vast majority of contacts are made within homes. We find social mixing to be assortative by both age and gender, but there is low variability in total number of contacts by age. Malnutrition is prevalent, and indoor air pollution is likely high, while individuals live in crowded shelters of poor quality. This study illustrates that such IDP settings potentially present a perfect storm of risk factors for lower respiratory tract infections and their transmission, often combined with inadequate access to curative or preventive health care. Innovative approaches to address such risks are urgently needed.

Sources of Funding

This work was supported by Elrha's Research for Health in Humanitarian Crises (R2HC) Programme, which aims to improve health outcomes by strengthening the evidence base for public health interventions in humanitarian crises. The R2HC programme is funded by the UK Government (DFID), the Wellcome Trust, and the UK National Institute for Health Research (NIHR). SF acknowledges a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (grant: 208812/Z/17/Z). In addition, RME acknowledges an HDR UK Innovation Fellowship (grant: MR/S003975/1), MRC (grant: MC_PC 19065), and NIHR (grant: NIHR200908) for the Health Protection Research Unit in Modelling and Economics at LSHTM. MCRI was supported by the Victorian Government's Operational Infrastructure Support Program. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care.

CRediT authorship contribution statement

KvZ: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **MOB**: Investigation, Supervision, Methodology, Writing – review & editing. **AIH**: Investigation, Supervision, Writing – review & editing. **MIA:** Investigation, Conceptualization, Writing – review & editing. **MSA**: Investigation, Writing – review & editing. **SMM**: Investigation, Writing – review & editing. **MYW**: Investigation, Writing – review & editing. **MAW:** Investigation, Writing – review & editing. **ED**: Conceptualization, Resources, Writing – review & editing. **CRM**: Resources, Writing – review & editing. **CS:** Conceptualization, Resources, Writing – review & editing. **KM:** Conceptualization, Resources, Writing – review & editing. **MME:** Investigation, Writing – review & editing. **MMH**: Investigation, Supervision, Writing – review & editing. **MAH:** Investigation, Supervision, Writing – review & editing. **RME:** Writing – original draft, Writing – review & editing. **FC:** Funding acquisition, Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **SF:** Funding acquisition, Conceptualization, Methodology, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

KM and CS are investigators on a research-led study on PCV13 and adult pneumonia in Mongolia funded by Pfizer. CS and KM are investigators on a Merck Investigator Studies Program grant funded by MSD on pneumococcal serotype epidemiology in children with empyema. All other authors report no conflicts of interest.

Data Availability

Analysis scripts, anonymized data, and questionnaire scripts are available on GitHub via https://github.com/kevinvzandvoort/espicc_somaliland_digaale_contact_survey_2019.

Acknowledgements

The authors thank Hamse Shaban Ahmed, Ayaan Ismail Adan, Khadar Abiib Ahmed, Ayaan Mohamoud Ali, Hamda Awil Garaad, Ahmed Yusuf Mohamed, Suaad Abdi Osman, Amiin Abdi Ismail, Nimco Abdilahi Ismail, AbdiFatah Mohamed Ahmed, Mustafe Shugri Mohamed, and Nimco Mohamed Abdi from the Republic of Somaliland Ministry of Health Development for their data collection efforts. Electronic data solutions were provided by LSHTM Open Data Kit (odk.lshtm.ac.uk).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.epidem.2022.100625.

References

Bellos, A., Mulholland, K., O'Brien, K.L., Qazi, S.A., Gayer, M., Checchi, F., 2010. The burden of acute respiratory infections in crisis-affected populations: a systematic review. Confl. Health 4 (1), 3. https://doi.org/10.1186/1752-1505-4-3.

Béraud, G., Kazmercziak, S., Beutels, P., et al., 2015. The french connection: the first large population-based contact survey in france relevant for the spread of infectious diseases. PLOS ONE 10 (7), e0133203. https://doi.org/10.1371/journal.pone.0133203.

Central Statistics Department, Ministry of Planning and National Development, Somaliland Government. The Somaliland Health and Demographic Survey 2020. Published online 2020.

R. Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2021. ⟨https://www.R-project.org/⟩.

Funk S. Socialmixr: Social Mixing Matrices for Infectious Disease Modelling. Published January 15, 2018. Accessed June 27, 2020. ⟨https://datacompass.lshtm.ac.uk/646/⟩.

Grijalva, C.G., Goeyvaerts, N., Verastegui, H., et al., 2015. A household-based study of contact networks relevant for the spread of infectious diseases in the highlands of Peru. PLOS ONE 10 (3), e0118457. https://doi.org/10.1371/journal.pone.0118457.

Grijalva-Eternod, C.S., Jelle, M., Haghparast-Bidgoli, H., et al., 2018. A cash-based intervention and the risk of acute malnutrition in children aged 6–59 months living in internally displaced persons camps in Mogadishu, Somalia: a non-randomised cluster trial. PLOS Med. 15 (10), e1002684 https://doi.org/10.1371/journal.pmed.1002684.

Hartung, C., Lerer, A., Anokwa, Y., Tseng, C., Brunette, W., Borriello, G., 2010. Open data kit: tools to build information services for developing regions. Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development. ICTD '10. Association for Computing Machinery,, pp. 1–12. https://doi.org/10.1145/2369220.2369236.

Heudtlass, P., Speybroeck, N., Guha-Sapir, D., 2016. Excess mortality in refugees, internally displaced persons and resident populations in complex humanitarian emergencies (1998–2012) – insights from operational data. Confl. Health 10 (1), 15. https://doi.org/10.1186/s13031-016-0082-9.

Hoang, T., Coletti, P., Melegaro, A., et al., 2019. A systematic review of social contact surveys to inform transmission models of close-contact infections. Epidemiology 30 (5), 723. https://doi.org/10.1097/EDE.0000000000001047.

Horby, P., Thai, P.Q., Hens, N., et al., 2011. Social contact patterns in vietnam and implications for the control of infectious diseases. PLOS ONE 6 (2), e16965. https://doi.org/10.1371/journal.pone.0016965.

Johnstone-Robertson, S.P., Mark, D., Morrow, C., et al., 2011. Social mixing patterns within a south african township community: implications for respiratory disease transmission and control. Am. J. Epidemiol. 174 (11), 1246–1255. https://doi.org/10.1093/aje/kwr251.

Kiti, M.C., Kinyanjui, T.M., Koech, D.C., Munywoki, P.K., Medley, G.F., Nokes, D.J., 2014. Quantifying age-related rates of social contact using diaries in a rural coastal population of Kenya. In: Borrmann, S. (Ed.), PLoS ONE, 9, e104786. https://doi.org/10.1371/journal.pone.0104786.

Lam, E., McCarthy, A., Brennan, M., 2015. Vaccine-preventable diseases in humanitarian emergencies among refugee and internally-displaced populations. Hum. Vaccin. Immunother. 11 (11), 2627–2636. https://doi.org/10.1080/21645515.2015.1096457.

Leung, K., Jit, M., Lau, E.H.Y., Wu, J.T., 2017. Social contact patterns relevant to the spread of respiratory infectious diseases in Hong Kong. Sci. Rep. 7 (1), 7974. https://doi.org/10.1038/s41598-017-08241-1.

Lumley T. Survey: Analysis of Complex Survey Samples.; 2020.

Lumley T. Complex Surveys: A Guide to Analysis Using R. John Wiley & Sons; 2011.

Melegaro, A., Jit, M., Gay, N., Zagheni, E., Edmunds, W.J., 2011. What types of contacts are important for the spread of infections? using contact survey data to explore European mixing patterns. Epidemics 3 (3), 143–151. https://doi.org/10.1016/j.epidem.2011.04.001.

Melegaro, A., Del Fava, E., Poletti, P., et al., 2017. Social contact structures and time use patterns in the manicaland Province of Zimbabwe. In: Nishiura, H. (Ed.), PLoS ONE, 12, e0170459. https://doi.org/10.1371/journal.pone.0170459.

Mossong, J., Hens, N., Jit, M., et al., 2008. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLOS Med. 5 (3), e74 https://doi.org/10.1371/journal.pmed.0050074.

Mulholland K., Weber M. Pneumonia in Children: Epidemiology, Prevention and Treatment.; 2016.

Myatt M., Guevarra E. *Zscorer: Child Anthropometry z-Score Calculator*.; 2019. ⟨https://CRAN.R-project.org/package=zscorer⟩.

le Polain de Waroux, O., Cohuet, S., Ndazima, D., et al., 2017. Characteristics of human encounters and social mixing patterns relevant to infectious diseases spread by close contact: a survey in Southwest Uganda. BMC Infect. Dis. 18 (172) https://doi.org/10.1101/121665.

le Polain de Waroux, O., Flasche, S., Kucharski, A.J., et al., 2018. Identifying human encounters that shape the transmission of *Streptococcus pneumoniae* and other acute respiratory infections. Epidemics 25, 72–79. https://doi.org/10.1016/j.epidem.2018.05.008.

Rodríguez, L., Cervantes, E., Ortiz, R., 2011. Malnutrition and gastrointestinal and respiratory infections in children: a public health problem. Int J. Environ. Res Public Health 8 (4), 1174–1205. https://doi.org/10.3390/ijerph8041174.

Rytter, M.J.H., Kolte, L., Briend, A., Friis, H., Christensen, V.B., 2014. The immune system in children with malnutrition—a systematic review. PLOS ONE 9 (8), e105017. https://doi.org/10.1371/journal.pone.0105017.

SMART. Standardized Monitoring and Assessment for Relief and Transitions Manual 2.0. Published online 2017.

Sonego, M., Pellegrin, M.C., Becker, G., Lazzerini, M., 2015. Risk factors for mortality from acute lower respiratory infections (ALRI) in children under five years of age in low and middle-income countries: a systematic review and meta-analysis of observational studies. PLOS ONE 10 (1), e0116380. https://doi.org/10.1371/journal.pone.0116380.

Troeger, C., Blacker, B., Khalil, I.A., et al., 2018. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Infect. Dis. 18 (11), 1191–1210. https://doi.org/10.1016/S1473-3099(18)30310-4.

United Nations High Commissioner for Refugees. Figures at a Glance. Accessed October 8, 2020. ⟨https://www.unhcr.org/figures-at-a-glance.html⟩.

van Zandvoort, K., Checchi, F., Diggle, E., et al., 2019. Pneumococcal conjugate vaccine use during humanitarian crises. Vaccine 37 (45), 6787–6792. https://doi.org/10.1016/j.vaccine.2019.09.038.

Wallinga, J., Teunis, P., Kretzschmar, M., 2006. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. Am. J. Epidemiol. 164 (10), 936–944. https://doi.org/10.1093/aje/kwj317.

Warsame A., Frison S., Gimma A., Checchi F. Retrospective estimation of mortality in Somalia, 2014–2018: a statistical analysis. Published online 2020.

World Health Organization. WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age; methods and development. Published online 2006.

Zhang, J., Klepac, P., Read, J., et al., 2020. Social contact data for China mainland. Publ. Online. https://doi.org/10.5281/zenodo.3878754.

Supplemental Material for

Social contacts and other risk factors for respiratory infections among internally displaced people in Somaliland

Kevin van Zandvoort*, Mohamed Omer Bobe, Abdirahman Ibrahim Hassan, Mohamed Ismail Abdi, Mohammed Saed Ahmed, Mohamed Yusuf Warsame, Muna Awil Wais, Emma Diggle, Catherine Satzke, Kim Mulholland, Mohamed Mohamoud Egeh, Mukhtar Muhumed Hassan, Mohamed Abdi Hergeye, Rosalind M Eggo, Francesco Checchi, Stefan Flasche

* Corresponding author: Kevin van Zandvoort, Kevin.Van-Zandvoort@lshtm.ac.uk

Additional analysis scripts, anonymized data, and questionnaire scripts are available on GitHub via https://github.com/kevinvzandvoort/espicc_somaliland_digaale_contact_survey_2019. These can be used to recreate all Tables and Figures in both the main manuscript and this Supplemental Material.

TABLE OF CONTENTS

Section A. Photos of Digaale IDP camp

Several shelters seen from a distance near the border of Digaale.

Two shelters seen from the front. All shelters are made from corrugated sheets. Several shelters have not been inhabited for some time.

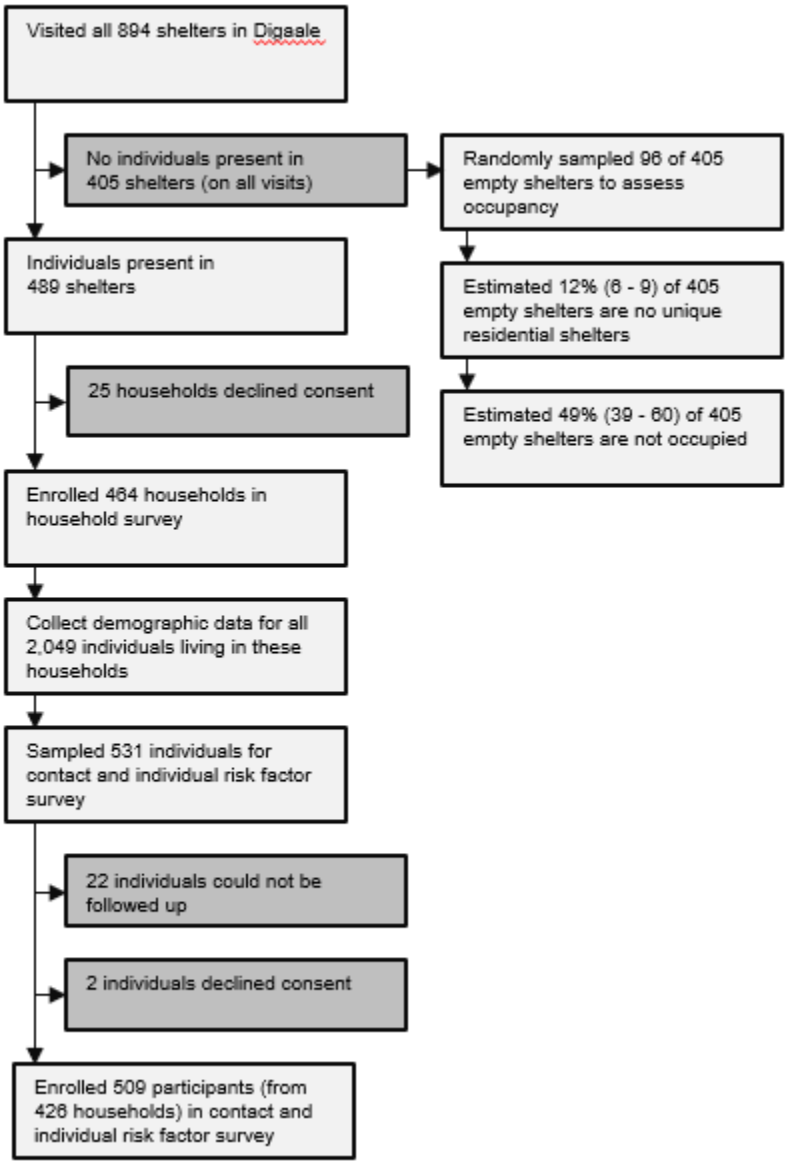
Two shelters in Digaale. There is a water tank on the left-hand side of each shelter. A traditional Somali hut has been constructed next to the right shelter to increase space.

The yellow concrete blocks are latrines. Each shelter has its own private latrine.

Digaale is surrounded by desert and shrubland.

There is a primary school in the center of Digaale.

Section B. Sampling design, assessment of missing households, finite population correction, and poststratification



SUPPLEMENTAL FIGURE B1. Flowchart showing all sampling steps used to collect the final sample for the cross-sectional survey in Digaale. Additional sampling steps used to collect nasopharyngeal swabs are not shown.

Sampling design

Supplemental Figure B1 shows a flowchart explaining the sampling steps households and participants in Digaale. We visited all 894 shelters. Shelters where no individuals were present

were revisited on different days throughout the study period, up to a total of five times. For 405 shelters, no individuals were present on any of the visit to the shelter. We took a random sample of 96 of these shelters to assess the occupancy of these shelters, results of which are provided in Supplemental Table B2 and described in the next section. In shelters where individuals were present, we asked an adult in households for consent to take part in the study. 25 households declined consent, and we enrolled 464 households in the household survey. We asked for the age and gender of all individuals currently living in the household, and thus collected information about all 2,049 individuals living in these households. All household members were assigned sampling weights based on their age. These weights were then used to sample household members, who were invited to take part in the second stage of the survey. We sampled a total of 531 individuals using this method (lower than the initially proposed 700 individuals, mainly because of the higher than expected number of vacant shelters, and lower number of individuals aged <2). Discrepancies between the target and realized sample size are provided in Supplemental Table B1. We asked those consenting to remember all people they would contact on the following day, and agreed a time for a return visit on the day thereafter. Despite the agreed return visit, we were not able to follow up 22 individuals, while 2 individuals declined consent on the later visit. 509 participants were enrolled in the final contact and risk factor survey.

SUPPLEMENTAL TABLE B1. Target and realised sample size in contact and individual risk factor survey

Shelters	Target sample size	Realised sample size
<2	200	75
2-5	100	119
6-14	100	83
15-29	100	59
30-49	100	65
50+	100	108

Assessing occupancy of empty shelters

Supplemental Table B2 shows the results of an assessment of a random sample of 96 shelters, sampled from a total of 405 shelters where no individual was present on multiple visits. We asked people living in neighbouring shelters whether they knew if the shelter was occupied. We could not retrieve any information for 23 shelters. Of the 73 shelters for which we did retrieve information, 36 were listed as having been vacant for a long time (empty). Twenty-eight were occupied, but residents either only returned back to Digaale at night, or were travelling to care for animals. Nine were no unique shelters, but were either a shop or the shelter was combined with a different shelter already included in the survey.

Using these data, we can estimate population totals and use the lower bound of these estimates to conservatively assume that at least 157 shelters are empty and at least 22 shelters are a shop or existing household. Thus, we estimate that out of all 894 shelters in Digaale, only 715 are inhabited with unique households (Supplemental Table B2).

SUPPLEMENTAL TABLE B2. Reported status of 73 shelters from a random sample of shelters where no individual was present on multiple visits

Status of shelters not present [a]	n [b]	% [c]	(95% CI)	N [d]	(95% CI)
Empty	36	49.3%	38.9 - 59.8	200	157 - 242
Occupied	28	38.4%	28.2 - 48.5	155	114 - 197
Shop or same household	9	12.3%	5.5 - 19.2	50	22 - 78
Total	73			405	

[a] Status reported by individuals living in neighbouring shelters.
[b] Number of recorded observations.
[c] Estimated proportion of shelters in Digaale with this status.
[d] Estimated total number of shelters in Digaale with this status.

Finite population correction

Of the 715 households that we assumed to be inhabited (Supplemental Table B3), 489 households were invited to participate in the study and 464 consented to do so. Without any additional information, we assumed that the demographics in the households that were included in the survey were similar to those of the inhabited households not included in the survey, and calculated a correction factor for the population size as 715/464, which was subsequently used to estimate the total population size and used for all finite population corrections in our analyses.

SUPPLEMENTAL TABLE B3. Total number of shelters and households in Digaale and included in the survey

Shelters	N	%
Total shelters in Digaale	894	
Assumed inhabited shelters in Digaale	715	80% [a]
Household present during survey	489	68% [b]
Households consented that were present	464	95% [c]

[a] Percentage of total shelters that is assumed to be inhabited.
[b] Percentage of total assumed inhabited shelters where a household was present and visited during the survey.
[c] Percentage of present households that consented to participate.

Poststratification

As participants were not selected using a stratified random sampling design, we post-stratified our results, calculating poststratification weights for stratum *s* as follows[1]:

$$w_s = \frac{N_s}{n_s}$$

where n_s is the total number of participants in stratum *s* in our sample, and N_s is the total number of estimated people living in stratum *s* in Digaale IDP camp. Data were stratified by gender (male or female) and the following age groups: <2, 2-5, 6-14, 15-29, 30-49, 50+ years of age. We censored poststratification weights below or above the 5th and 95th percentile of all poststratification weights to the values of poststratification weights at those percentiles, in order to reduce the weight of individual observations in strata with low numbers of participants. Poststratification and finite population corrections were calculated and implemented using the *Survey*[2] package in *R.*

Section C. Contact matrix weighting and calculations

We used post-stratification weights to adjust the contact data when calculating contact matrices, in order to improve the representativeness of our sample to the population living in Digaale IDP camp. Supplemental Material section B explains the calculation of post-stratification weights in more detail.

In order to make the contact matrices representative for a typical day, we recalculated post-stratification weights used in these analyses to account for contacts reported on week- (Sunday to Thursday) and weekend- (Friday and Saturday) days, in addition to post-stratifying for discrepancies between the age- and gender-distribution in the population and our sample. Contacts made on the five weekdays are weighted $\frac{5}{2}$ times higher than contacts made on the two weekend-days.

For all contactors *p* in stratum *s*, poststratification weights were recalculated as follows:

$$w_p = \frac{g_s}{\pi_p}$$

where:

$$\pi_p = \begin{cases} 5/2, & contactor\ p\ reported\ contacts\ for\ a\ weekday \\ 1, & contactor\ p\ reported\ contacts\ for\ a\ weekend-day \end{cases}$$

and:

$$g_s = \frac{N_s}{\mathrm{O}_s}$$

in which:

$$O_s = \sum_{p=1}^{n_s} \pi_p$$

We first calculate matrix **T**, with elements t_{ij} which sums the reported number of contacts made by people in age group *j* with those in age group *i,* multiplied by the poststratification weights of the contactors in age group *j*:

$$t_{ij} = \sum_{p=1}^{n_j} c_{pi} w_p$$

where *p* is the index of a contactor in age group *j*, n_j is the total number of contactors in the sample who are in age group *j*, c_{pi} is the total number of contacts reported by contactor *p* made with contactees in age group *i*, and w_p is the post-stratification weight assigned to contactor *p*.

We also calculate vector **x** with elements x_j, which sum all post-stratification weights of contactors in age group *j*:

$$x_j = \sum_{p=1}^{n_j} w_p$$

We then calculate matrix **A**, with elements a_{ij} which is the weighted average number of daily contacts made by contactors in age group *j,* with contactees in age group *i*:

$$a_{ij} = \frac{t_{ij}}{x_j}$$

In a closed population, we would expect reciprocity of contacts and equality of the total number of population-wide contacts (i.e. $a_{ij}N_j = a_{ji}N_i$ where N_i and N_j are the total number of individuals in the population in age groups *i* and *j*). I.e. if person A would report a contact with

person B, person B would also report a contact with person A. However, because of sampling errors or inconsistencies in the reporting of contacts, the estimated population-wide number of contacts may differ without further corrections. Therefore, we estimate matrix ***Z*** where elements z_{ij} adjust elements a_{ij} for reciprocity by averaging the total number of population-wide contacts, so that $z_{ij}N_j = z_{ji}N_i$:

$$z_{ij} = \frac{a_{ij}N_j + a_{ji}N_i}{2N_j}$$

All contact matrices that show the daily contact rates are calculated as matrix ***Z***, e.g. Figure 2E in the main manuscript.

We also calculate matrix ***F***, with elements f_{ij} which are the per-capita daily contact rates between a single individual in age group *j* and a single individual in age group *i:*

$$f_{ij} = \begin{cases} z_{ij}/N_i\,, & i \neq j \\ z_{ij}/(N_i - 1)\,, & i = j \end{cases}$$

Note that the denominator is subtracted by 1 when *i* and *j* are from the same age group (which are the diagonal values in the contact matrix), because individuals cannot contact themselves.

f_{ij} can be interpreted as the rate at which two individuals of these age groups come into contact with each other, assuming homogeneous mixing. Whereas $z_{ij} \neq z_{ji}$ (unless the population size N_j is the same for all age groups *j)*, $f_{ij} = f_{ji}$.

Section D. Additional contact pattern analyses

Assessing the choice of post-stratification weights on contact intensities

Contact intensities in our manuscript and this supplement are, unless stated otherwise, weighted by the calculated post-stratification weights on age and gender, and the day of the week (see Sections B and C).

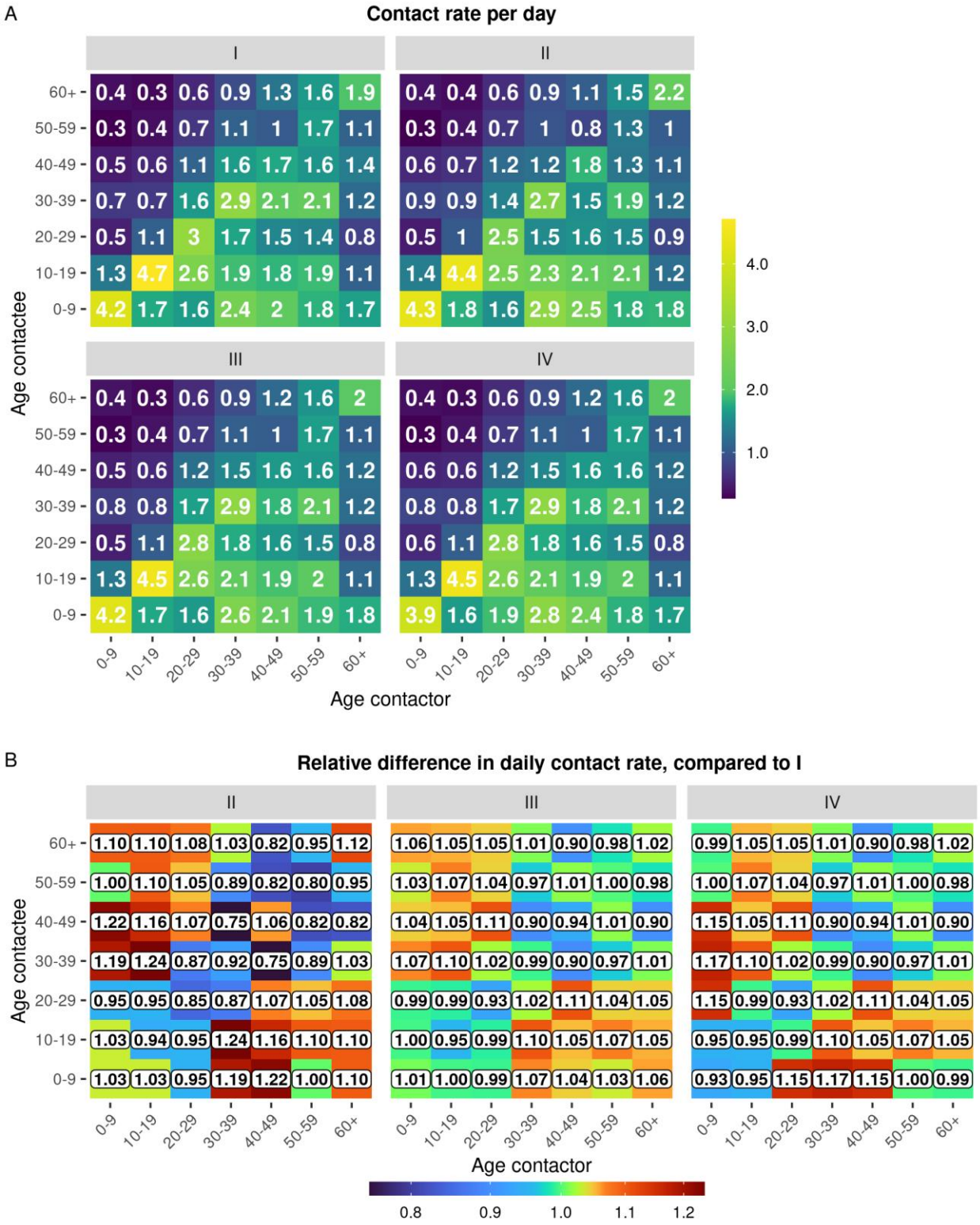
Here, we assess the sensitivity of the chosen post-stratification weights on our estimated contact intensities by comparing contact matrices weighted by the parameters in Supplemental Table D1.

SUPPLEMENTAL TABLE D1. Post-stratification weights used to calculate contact intensities.

Weights	Day of the week [a]	Age group [b]	Gender [c]	Household size [d]
I [e]	✓	✓	✓	×
II	✓	✓	×	✓
III	✓	✓	×	×
IV	✓	×	×	×

Variables used in calculating the weights used to construct contact intensities.
[a] *Weekday (Sunday to Thursday) or weekend-day (Friday and Saturday)*
[b] *Categorised as <2, 2-5, 6-14, 15-29, 30-49, 50+ years of age*
[c] *Female or male*
[d] *Categorised by quantiles: 1-2, 3-4, 5-6, and 7-12 household members*
[e] *Main weights used in the manuscript*

Panel A in Supplemental Figure D1 shows the contact matrix estimated for each type of post-stratification weights listen in Supplemental Table D1. Contact matrix I is the same as that shown in Figure 2E in the main manuscript. While there are some slight differences in estimated contact rates within the same contactor- and contactee age-group pair (cells), the overall patterns are qualitatively consistent with most contacts clustered in those aged <20 years and increased intergenerational mixing in older individuals. The dominant eigenvalues for each matrix are 10.100, 10.020, 10.127, and 10.151, respectively.



SUPPLEMENTAL FIGURE D1. A. Estimated daily contact rates where contactor data are post-stratified according to variables in Supplemental Table D1. B. The relative difference in daily contact rates for contactors and contactees in each age group *ij* comparing matrices II, III, and IV to matrix I.

Panel B directly compares each cell in matrices II, III, and IV to those in matrix I. They are calculated as $\mathbf{A} \oslash \mathbf{B}$, where $\oslash$ denotes a Hadamard division (element-wise division). Values below 1 indicate a lower average daily contact rate compared to matrix I, and values above 1 a higher daily contact rate. We observe that differences are relatively small, and primarily occur in the estimated contacts between children and adults. The largest differences are with contact matrix II, which does not weight for gender, but does weight for household size. Overall, contact patterns seem insensitive to the choice of post-stratification weights.

Total number of reported contacts by age

Supplemental Table D2 shows the total number of contacts reported by contactors with contactees of specific age groups. Estimates are similar to those that would be calculated in matrix **T** in Supplemental Material section C, but without post-stratification by age, sex, and day of the week. These estimates can be used to calculate completely unweighted contact matrices.

SUPPLEMENTAL TABLE D2. Total number of contacts reported by contactors in age groups used in contact matrices.

Contactors			**Contactees**						
Age group	**N** [a]	**n** [b]	**0-9**	**10-19**	**20-29**	**30-39**	**40-49**	**50-59**	**60+**
0-9	1006	248	993	385	255	338	203	121	148
10-19	766	68	79	312	91	73	47	34	32
20-29	313	20	8	33	52	37	24	13	7
30-39	297	33	32	39	51	95	52	36	26
40-49	237	32	28	50	51	44	52	33	32
50-59	152	38	13	34	57	77	62	57	50
60+	225	70	51	62	73	96	89	84	145

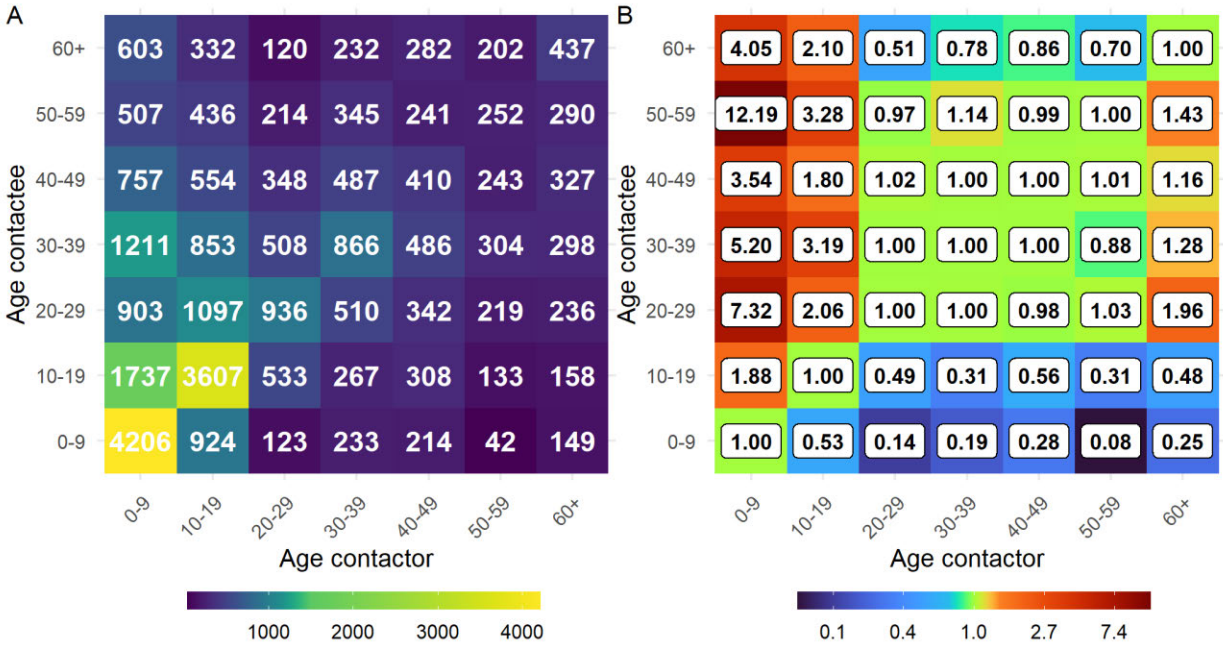
Estimates are the total number of contacts reported with contactees in the age-groups in columns by contactors in the age-groups in rows.

[a] *Total estimated number of people in the entire population in the contactor age-group*

[b] *Total number of contactors in the sample in the contactor age-group*

Assessing discrepancies and potential biases in reported contacts

We can compare the diagonals of the total number of contacts (or per-capita contact rates) to assess discrepancies and potential biases in the reported number of contacts in our sample. As explained in Supplemental Material section C, $a_{ij}N_j \neq a_{ji}N_i$ because of sampling errors and biases in estimating these rates. We calculated the total contact matrix with the total number of population-wide contacts in each age-group after post-stratifying the data, but without adjusting the matrix for reciprocity of contacts. It is equivalent to $\boldsymbol{A} \cdot \boldsymbol{Diag}(\mathbf{p})$ where $x_j = n_j$ (n_j is the total number of contactors in age group *j* in our sample, column 'n' in Supplemental Table D2), and where the elements p_j of vector **p** are the population size estimates for the entire population in age group j (column 'N' in Supplemental Table D2). In other words, all columns in matrix **A** are multiplied with the total population size estimates in vector **p**.



SUPPLEMENTAL FIGURE D2. A: Total number of expected population-level contacts between individuals in different age groups. B: The ratio between the total number of expected population-level contacts in panel A and its transpose. Estimates are not adjusted for reciprocity of contacts, but weighted for age, gender, and day of the week.

Supplemental Figure D2 shows this matrix (left-hand side), which compares the contacts reported *by* those of a certain age group (in columns), with the contacts reported *with* those of a certain age group (in rows), converted to total population-wide contacts. If all contacts in the population were observed, this matrix should be symmetrical. The matrix on the right-hand side shows the relative difference between the matrix and its transpose. It is calculated as $\mathbf{T} \oslash \mathbf{T}'$ where $\oslash$ denotes a Hadamard division (i.e. element-wise division) of the two-matrices. Values below 1 indicate that contactors of age group *j* reported more contacts with contactees of age group *i*, than contactors of age group *i* did with contactees of age group *j*.

We observe that the total number of contacts reported by children aged 0-9 and 10-19 years are substantially higher than those reported with children aged 0-9 years by other age groups. Contacts that were made by children 0-9 years were proxy reported by their parents or caregivers, which may have overestimated the true number of contacts made by their children. However, this does not explain the discrepancy seen with those aged 10-19 years.

We observe a similar (but less extreme) pattern in contacts with 60+ year olds, where contacts reported by this age group are higher (with the exception of contacts reported with children 0-9 and 10-19 years) than the contacts reported by contactors of this age group. In addition, it is noteworthy that we observe very similar total contact estimates between all other age-groups.

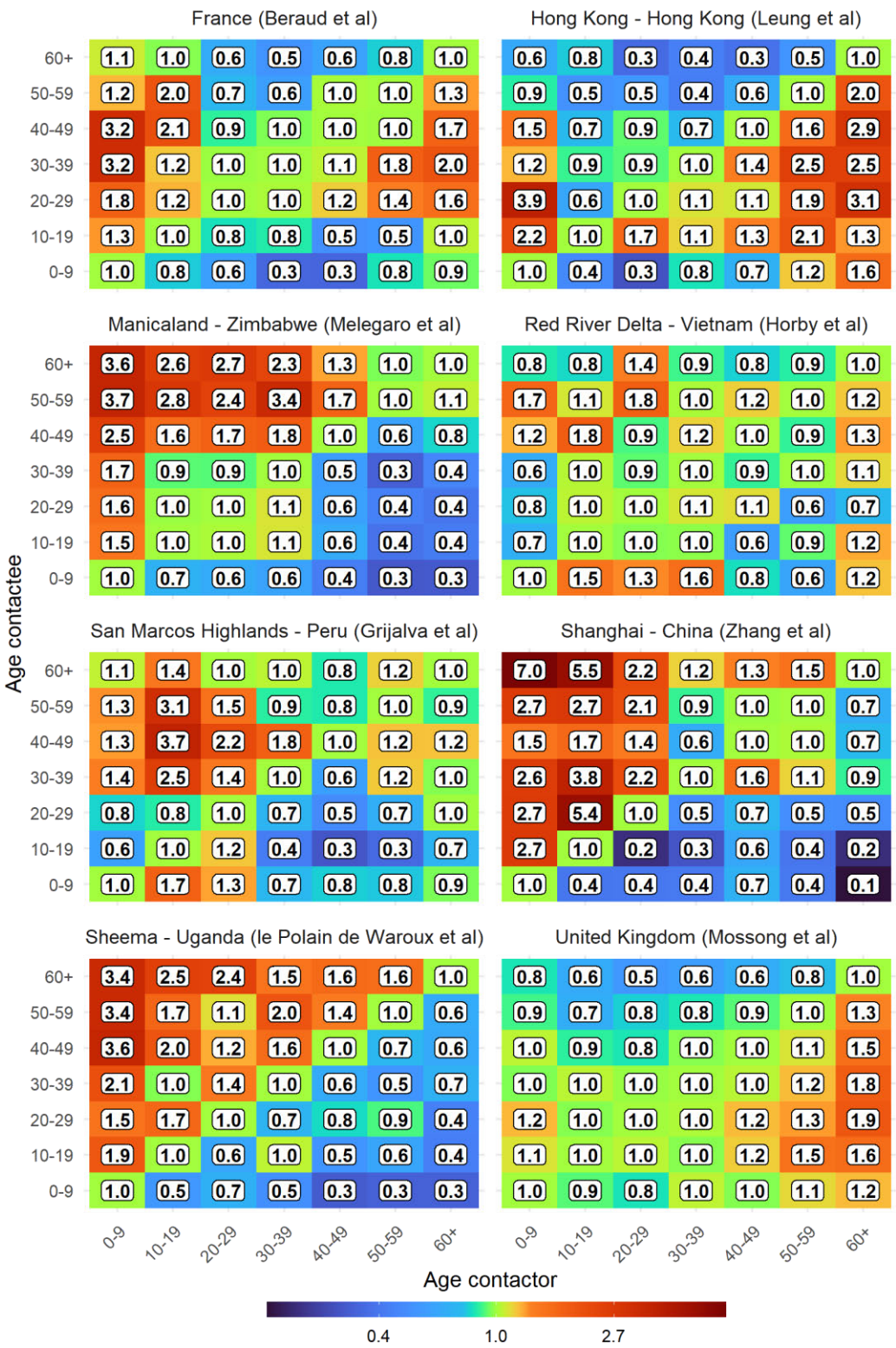
We replicated the analysis using the *socialmixr*[3] package with data from eight other surveys[4–11] where we could access individual level data. Contact matrices were weighted for day of the week where applicable. Where available, we combined the unadjusted contact rates with total population size estimates as provided in the study, and UN WPP[12] estimates for the respective country and year of the study otherwise. Matrices of the ratio of the total expected population-level contacts, without adjusting contact rates for reciprocity, are provided in Supplemental Figure D3. These can be compared to the matrix in Supplemental Figure D2B.

The differences in the reporting of contact intensities by age group are not unique to our study. The lowest discrepancies are found in the UK POLYMOD data. A large number of ratios are close to 1 in this dataset, with the exception of contacts reported by contactors aged 60+ years, who on average reported greater contact intensities with other age groups, than other age groups reported with this age group.

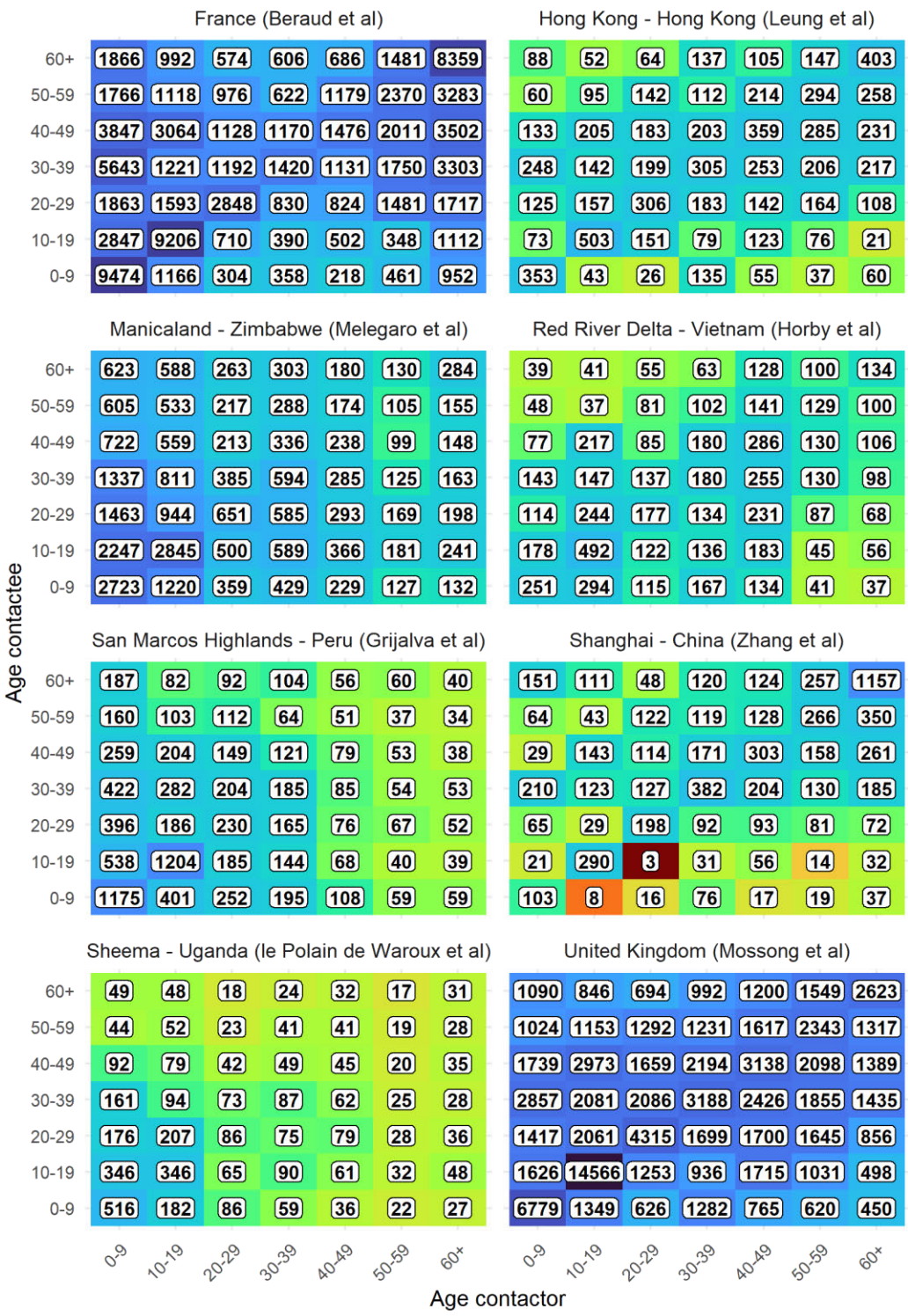
There appears to be a pattern where, as was also observed in our study, contact intensities made by children <10 years are commonly reported to be higher than contact intensities reported with this age group. This ratio above 1 is observed in all datasets, except for the UK POLYMOD and Red River Delta (Vietnam) datasets. Larger differences in contacts made by or with the 60+ year age group are also common. The direction of those ratios seems consistent within datasets, but diverge between datasets.

Notably, contact intensities tend to be consistently higher for contacts made with contactees who are older than the contactor in the Sheema (Uganda), San Marcos Highlands (Peru), Manicaland (Zimbabwe), and Shanghai (China) datasets. In the Hong Kong and to a lesser extent the UK POLYMOD datasets however, these tend to be consistently higher for contacts made with contactees who are younger than the contactor. No such pattern is apparent in the French, Red River Delta (Vietnam), or our Digaale dataset.

It is not clear what is driving these consistent discrepancies. If only driven by sampling error, the expected discrepancies would be randomly distributed between different age pairs. Most matrices are constructed from a sufficiently high number of reported contacts within each contactor and contactee age pair in the chosen age groups (Supplemental Figure D4). There were only 2 age-pairs where contact rates were based on fewer than 10 reported contacts (both in the Shanghai dataset).



SUPPLEMENTAL FIGURE D3. Ratio of the total number of expected age-specific contacts between individuals in different age groups without adjusting for reciprocity, for eight studies conducted in different settings.

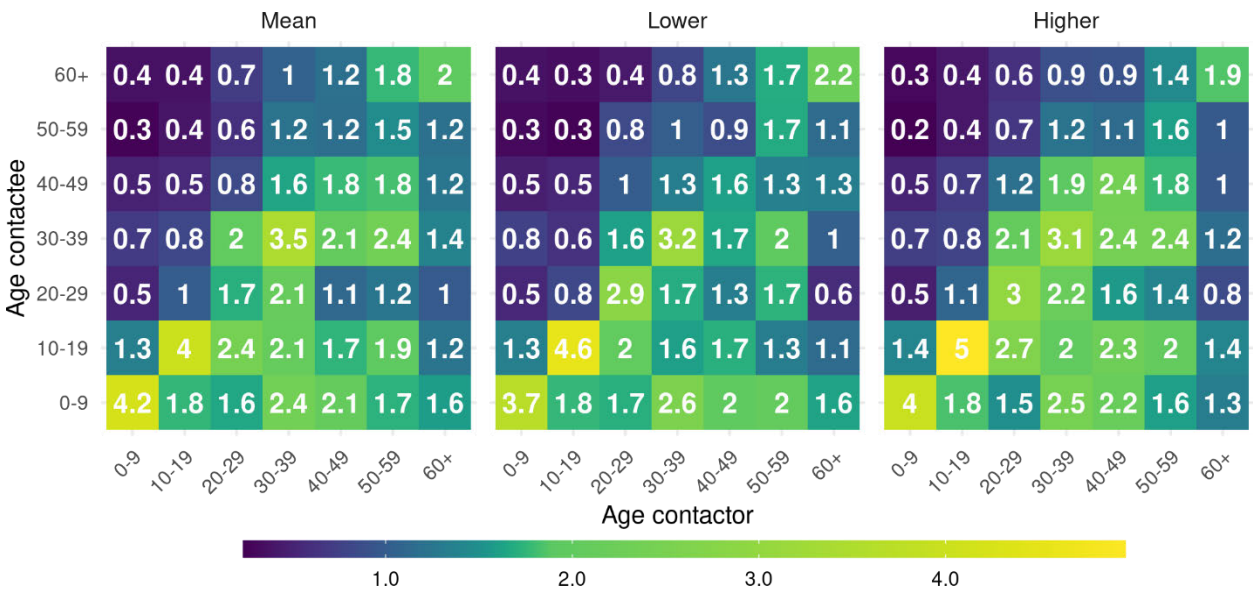


SUPPLEMENTAL FIGURE D4. Total number of observations used to calculate age-specific contact intensities between age pairs, for eight studies conducted in different settings.

Assessing uncertainty in the contact matrices

The basic reproduction number (R_0), is the average number of secondary infections arising from a typical single infection in a completely susceptible population, and can be estimated as the dominant eigenvalue of the next generation matrix (NGM). The exact form of the NGM is model dependent, but often a function of the age-specific number of daily contacts, the probability that a single contact leads to transmission, and the total duration of infectiousness. Therefore, the dominant eigenvalue of the contact matrix is proportional to R_0, and can be used as a useful summary measure to describe a contact matrix.[13,14]

To assess uncertainty in the contact matrix, we took 10,000 bootstrap samples where we resampled participants in the contact data with replacement. We recalculated sample weights in each bootstrap sample, and calculated the corresponding contact matrix. We then calculated the dominant eigenvalue for each bootstrapped contact matrix, and selected the contact matrices whose dominant eigenvalue (shown in brackets) lay closest to the mean (10.1), 2.5th percentile (9.6), and 97.5th (10.7) percentile of all eigenvalues. A procedure explaining the bootstrapping steps that resulted in the selected contact matrices is described in detail in pseudocode in Algorithm 1. These matrices are represented as the mean, lower, and higher contact matrix in Supplemental Figure D5. Note that the variation between these matrices is relatively low, and patterns are similar.



SUPPLEMENTAL FIGURE D5. Mean daily age-specific contact matrix and lower and upper 95% uncertainty estimates of age dependent social contact matrices adjusted for reciprocity of contacts in Digaale IDP camp, Somaliland.

Generating bootstrap contact matrix estimates

Algorithm 1 provides the pseudocode explaining how bootstrapped contact estimates were generated, where θ^0 is a set of all participant identifiers (PIDs) included in the sample, N is the length of set θ^0 (equivalent to the total number of participants), and M is the total number of bootstrap samples (set to 10,000 in our analysis).

First, we generate the m^{th} bootstrap set θ^m by randomly resampling all PIDs included in the survey with replacement. We then pass this set to the *getReciprocalContactMatrix* function, which is a function that i) generates a new bootstrap sample using the PIDs passed to the function, ii) recalculates poststratification weights based on the bootstrap sample as in Section B, and iii) calculates and returns the reciprocal matrix with daily contact rates Z_m following the equations in Section C. We then calculate E_m, the dominant eigenvalue of matrix Z_m.

Once we have calculated the dominant eigenvalues for all M bootstrap samples, we calculate the mean (μ), 2.5th (ρ_l) and 97.5th (ρ_h) percentile of all dominant eigenvalues E. We calculate

the absolute difference between each dominant eigenvalue E_m and μ, ρ_l, and ρ_h, and retrieve the index of the dominant eigenvalue that is closest to these values in a_M, b_M, and c_M respectively. We then return the reciprocal contact matrices Z_m at these indexes, which are presented in Supplemental Figure D5.

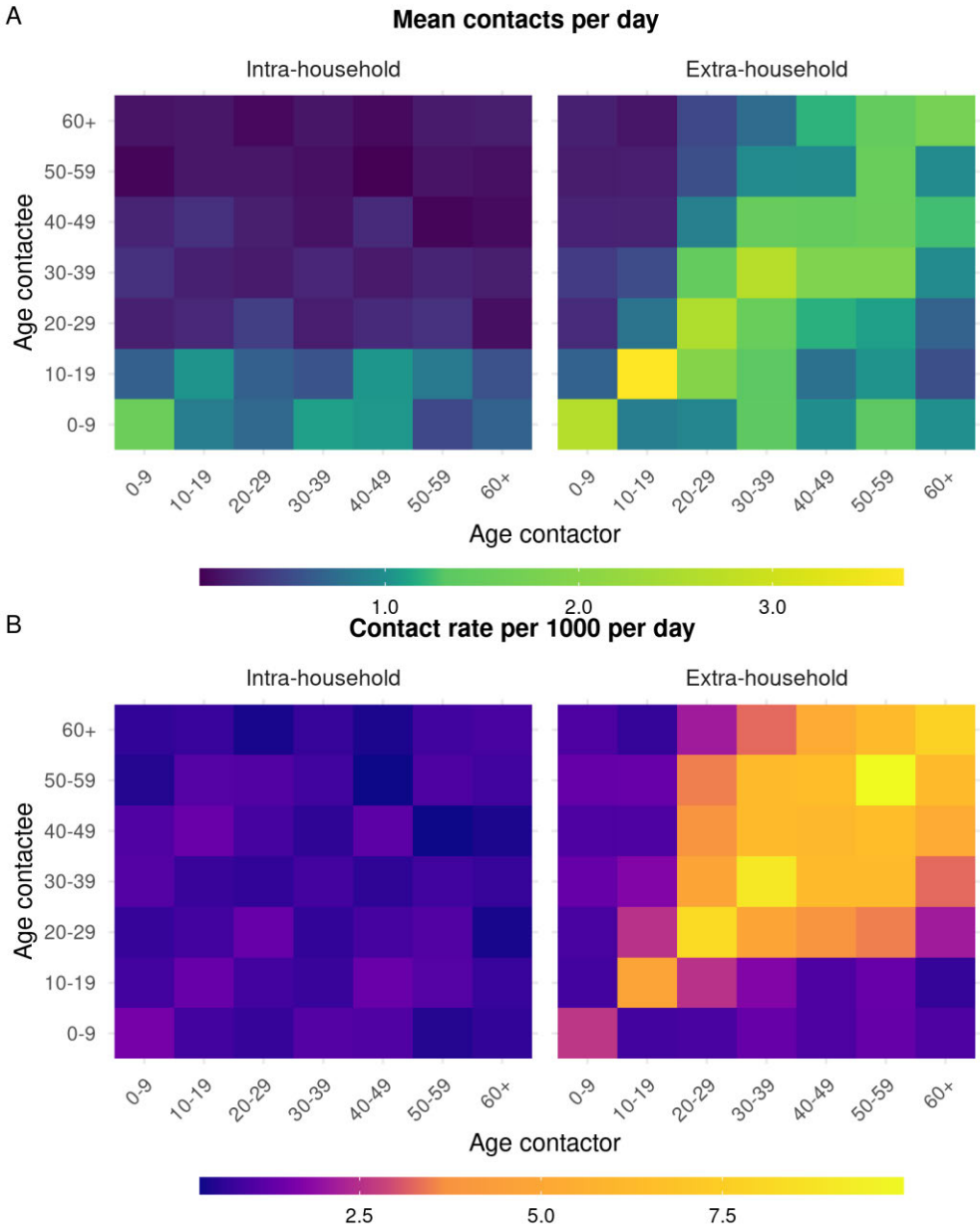
Algorithm 1 Generating bootstrap contact matrix estimates

Given θ^0, N, and M
for $m = 1$ to M **do**
  for $i = 1$ to N **do**
    $\alpha_i \sim \lfloor \mathrm{U}(0,1) \rfloor$
    $j_i = 1 + \alpha_i N$
    $\theta_i^m = \theta_{j_i}^0$
  end for
  $\boldsymbol{Z_m} = getReciprocalContactMatrix(\theta^m)$
  $E_m = largest\ eigenvalue\ of\ \boldsymbol{Z_m}$
end for
$\mu = \sum_{m=1}^{M} E_m / M$
$\rho_l = P_{2.5}(E)$
$\rho_h = P_{97.5}(E)$
$a_0 = b_0 = c_0 = 10^6$
for $m = 1$ to M **do**
  $\delta_m^\mu = |E_m - \mu|$
  $a_m = \begin{cases} \delta_m^\mu, & \delta_m^\mu < a_{m-1} \\ a_{m-1}, & \delta_m^\mu \geq a_{m-1} \end{cases}$
  $\delta_m^l = |E_m - \rho_l|$
  $b_m = \begin{cases} \delta_m^\mu, & \delta_m^\mu < b_{m-1} \\ b_{m-1}, & \delta_m^\mu \geq b_{m-1} \end{cases}$
  $\delta_m^h = |E_m - \rho_h|$
  $c_m = \begin{cases} \delta_m^\mu, & \delta_m^\mu < c_{m-1} \\ c_{m-1}, & \delta_m^\mu \geq c_{m-1} \end{cases}$
end for
return $\boldsymbol{Z_{a_M}, Z_{b_M}, Z_{c_M}}$

Comparing intra- and extra-household contacts

Supplemental Figure D6 shows i) the contact matrix and ii) per-capita contact matrix for contacts made with (intra-household) and without (extra-household) household members. We

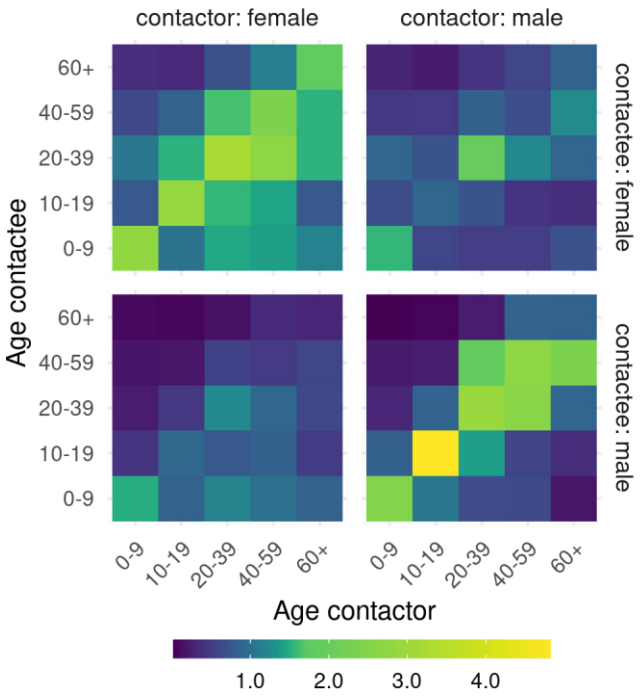
observe that extra-household contacts are very assortative, but intra-household contacts are mostly homogeneous across ages.



SUPPLEMENTAL FIGURE D6. Weighted mean number of contacts per day between contactors and contactees of certain ages (A), and age-specific weighted daily contact rates (reported per 1000 people) for contacts with household members (intra-household) and non-household members (extra-household) in Digaale IDP camp, Somaliland. Both matrices are adjusted for reciprocity of contacts.

Comparing contacts by age and gender

Supplemental Figure D5 shows daily contact rates by age and gender. Contact rates are assortative by both age and gender, with less mixing between the two genders compared to mixing within each gender. Female adults make more contacts with children than to male adults.



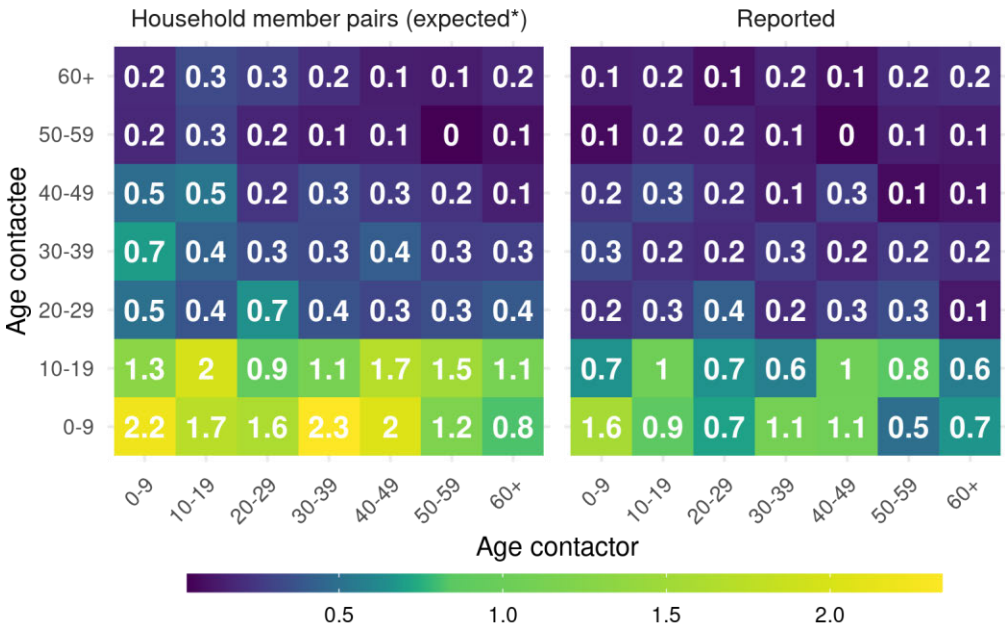
SUPPLEMENTAL FIGURE D7. Daily contact rates [a] by age and gender

[a] Matrices are adjusted for reciprocity, meaning that e.g. the total number of contacts of females aged *j* with males aged *i* are the same as the total number of contacts of males aged *i* with females aged *j*.

Comparing observed to expected number of household contacts

Supplemental Figure D8 shows the results of an analysis where we assumed that all contactors in the survey make one contact per day with all of their household members, and compare these with the number of household-contacts actually reported by them. While overall patterns are very similar, contact rates are lower in the empirical data compared to the data that assumes a single contact of household members per day. The dominant eigenvalue of the matrix based on the empirical data (2.9) is 43% lower than that of the matrix with expected intra-household contacts under this assumption (5.2). The lower observed number of contacts may

indicate that household members do not contact each other every day, for instance because some members may travel outside of the household for several day for work.

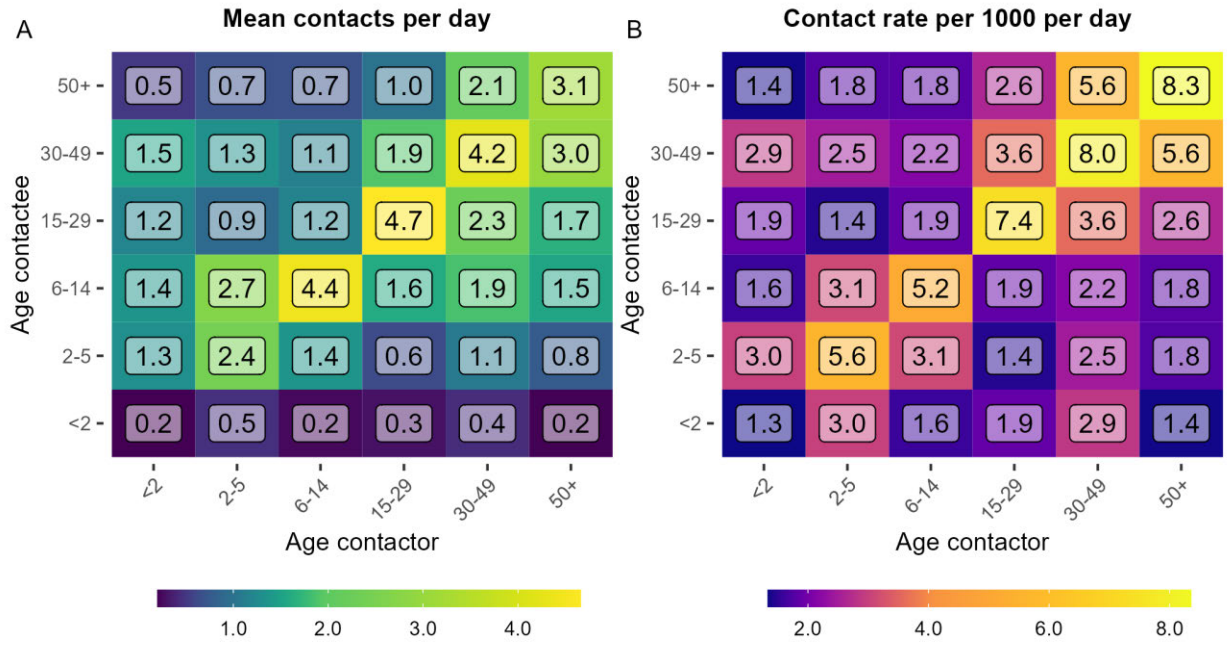


SUPPLEMENTAL FIGURE D8. Expected[a] and observed contact matrices for contacts made with household members

[a] Expected values indicate the contact matrix of a hypothetical scenario where we assume that all contactors make exactly one contact per day with all of their household members.

Contacts in youngest age group

Supplemental Figure D9 shows the contact matrices by age groups for which quota sampling was conducted. Contact rates for contactors aged <2 years were substantially lower than those for age groups. Due to the different sized age groups, caution must be taken when comparing the mean number of contacts between age groups. In contrast to other age groups, children <2 years were least likely to contact other children in the same age groups and more likely to contact older children and adults, particularly those aged 2-5 and 30-49 years old.



SUPPLEMENTAL FIGURE D9. Contact matrices with alternative age groups. Matrices show the weighted mean number of daily contacts (A) and per-capita daily contact rates, reported by 1000 people (B). Both matrices are adjusted for reciprocity of contacts.

Contacts at school and work for attending contactors

Supplemental Table D3 shows the estimated average daily contact rate at school or work by age, excluding individuals who reported no work or school contacts. Estimates are substantially higher in all age group compared to estimates where contactors with 0 contacts are not excluded (Table 3 in main manuscript).

SUPPLEMENTAL TABLE D3. Average daily number of (direct) contacts reported at school or work settings, excluding individuals who reported no school and work contacts

Age group	Total (Direct) contacts at school or work
<2	1.0 1 - 1
2-5	3.6 2.7 - 4.5
6-14	5.4 4.3 - 6.6
15-29	3.8 2.4 - 5.1
30-49	5.8 2.4 - 9.2
50+	2.1 1.3 - 2.9

Estimates are the weighted mean and 95% confidence interval.

Section E. Self-reported pneumonia incidence analyses

We asked participants whether they were ever diagnosed with pneumonia. For those who were, we asked whether this diagnosis occurred in the last six months. Results are shown in Supplemental Table E1, stratified by age group.

SUPPLEMENTAL TABLE E1. Cumulative incidence of self-reported pneumonia by age

Variable	N [a]	Estimate [b]	
Cumulative incidence of self-reported pneumonia [d]			
Ever			
<2 years old	39	52.1%	43.8 - 60.4
2-5 years old	62	52%	44.3 - 59.8
6-14 years old	25	28.8%	19.5 - 38.2
15-29 years old	18	32.5%	20.2 - 44.7
30-49 years old	15	24.6%	12.7 - 36.5
50+ years old	10	11.6%	5.8 - 17.5
In the last six months			
<2 years old	32	42.8%	34.5 - 51
2-5 years old	31	26%	19.2 - 32.8
6-14 years old	11	12.4%	5.7 - 19.1
15-29 years old	4	6.6%	0.3 - 13
30-49 years old	7	10.6%	2.4 - 18.8
50+ years old	3	3%	0 - 6.1

[a] Total number of observations

[b] Central estimates are the weighted mean, median, or percentage. The uncertainty interval next to the central estimates is the 95% confidence interval for proportions and means, and interquartile range (IQR) for median estimates. A finite population correction was applied in calculating confidence intervals.

[c] Rates are per 1000 people per year

[d] Pneumonia in young children (<10 years) was reported by an adult or caregiver.

We assessed the association between several household- and individual level risk factors with self-reported cumulative pneumonia incidence in the six months preceding the survey (Supplemental Table E2), or ever (Supplemental Table E3) through logistic regression.

The cumulative incidence patterns show a strong non-linear trend by age (Supplemental Table E1). We decided to adjust all risk factors for age and gender a priori, to understand whether there was an additional association with any of the listed risk factors. We fitted multiple multivariable models, each with i) age and gender as a-priori defined confounders and ii) one of the listed exposure variables. Our study and analysis is not able to infer a causal relationship between any of the risk factors and self-reported pneumonia. We did not correct for a finite population in the regression models.

A likelihood ratio test found good evidence to include age as a categorical variable ($p = 0.012$) for models regressing risk factors against cumulative pneumonia incidence in the last six months (Supplemental Table E2), but not for models regressing against cumulative pneumonia incidence ever (Supplemental Table E3). Age group was included as a categorical variable for both outcomes. In models where data was restricted to children younger than five years of age (those that include anthropometric estimates), age in months was used as a continuous variable instead.

After adjusting for age and gender alone, we did not find evidence of any increase in the odds of cumulative incidence of self-reported pneumonia diagnosis for the majority of risk factors listed in Table 1 (main manuscript) nor the total number of contacts. There was some evidence for an increased odds of self-reported pneumonia diagnosis in both the six months preceding the survey (OR 2.12; 95%CI 1 – 4.7) or ever (OR 2.6; 95%CI 1.1 - 6.0) for those living in households where snuff was consumed. Our study was not designed nor powered to assess the relationship between risk factors and self-reported pneumonia. Note that self-reported

pneumonia may be an unreliable outcome measure, and it was unclear what case definitions were used when making the diagnoses.

SUPPLEMENTAL TABLE E2. Association between risk factors and self-reported pneumonia diagnosis in the six months preceding the survey

Variable	OR [a]	95% CI	p-value
Demographic characteristics			
Household size	1.03	0.94 - 1.12	0.567
Household members <5y	0.98	0.76 - 1.25	0.856
Household members <2y	0.65	0.4 - 1.04	0.079
Years since household settled in Digaale			
>3 years	*ref*		
2-3 years	0.41	0.17 - 0.9	0.034
1-2 years	0.57	0.28 - 1.12	0.111
<1 year	0.71	0.31 - 1.58	0.413
Quality of shelter			
Total number of rooms	0.81	0.44 - 1.4	0.467
Reported draught in shelter			
no	*ref*		
yes	0.87	0.56 - 1.35	0.526
Reported leakage in shelter			
no	*ref*		
yes	0.98	0.62 - 1.56	0.934
Indoor air pollution			
Use charcoal for cooking			
no	*ref*		
yes	0.86	0.56 - 1.31	0.479
Use firewood for cooking			

SUPPLEMENTAL TABLE E2. Association between risk factors and self-reported pneumonia diagnosis in the six months preceding the survey

Variable	OR [a]	95% CI	p-value
no	*ref*		
yes	1.04	0.62 - 1.75	0.886
Use ventilation when cooking			
no	*ref*		
yes	0.86	0.42 - 1.82	0.693
cook outside	1.04	0.54 - 2.04	0.910
Substance use in household			
Household member who uses khat			
no	*ref*		
yes	0.83	0.54 - 1.26	0.388
Household member who smokes			
no	*ref*		
yes	0.61	0.38 - 0.96	0.035
Household member who uses snuff			
no	*ref*		
yes	2.12	0.97 - 4.69	0.060
Contact behaviour			
Total number of direct contacts	1.01	0.96 - 1.07	0.629
Total number of physical contacts	1.03	0.97 - 1.08	0.312
Total number of contacts at home	1.02	0.96 - 1.08	0.613
Total number of contacts at school	0.98	0.86 - 1.11	0.791
Total number of contacts at work	1.01	0.8 - 1.21	0.950
Total number of contacts at other settings	1.01	0.91 - 1.12	0.826
Malnutrition in U5			
Weight by age			

SUPPLEMENTAL TABLE E2. Association between risk factors and self-reported pneumonia diagnosis in the six months preceding the survey

Variable	OR [a]	95% CI	p-value
Not underweight (z > -2)	*ref*		
Underweight (z ≤ -2)	1.66	0.65 - 4.55	0.302
Underweight (z ≤ -3)	0.65	0.12 - 3.12	0.591
Height by age			
Not stunted (z > -2)	*ref*		
Stunted (z ≤ -2)	1.23	0.54 - 2.9	0.626
Severely stunted (z ≤ -3)	1.26	0.4 - 4.13	0.693
Weight by height			
Not wasted (z > -2)	*ref*		
Wasted (z ≤ -2)	0.79	0.26 - 2.44	0.675
Severely wasted (z ≤ -3)	0.00		
Middle-Upper Arm Circumference			
Not wasted (≥ 125mm)	*ref*		
Wasted (< -125mm)	1.50	0.23 - 12.09	0.668

[a] All odds ratios are adjusted for age and gender.

SUPPLEMENTAL TABLE E3. Association between risk factors and self-reported pneumonia diagnosis (ever)

Variable	OR [a]	95% CI	p-value
Demographic characteristics			
Household size	1.06	0.95 - 1.18	0.274
Household members <5y	1.05	0.78 - 1.4	0.758
Household members <2y	0.81	0.46 - 1.4	0.469
Years since household settled in Digaale			
>3 years	*ref*		
2-3 years	0.63	0.22 - 1.59	0.361
1-2 years	0.71	0.3 - 1.57	0.419
<1 year	0.82	0.3 - 2.02	0.674
Quality of shelter			
Total number of rooms	1.03	0.49 - 1.91	0.929
Reported draught in shelter			
no	*ref*		
yes	0.81	0.48 - 1.38	0.421
Reported leakage in shelter			
no	*ref*		
yes	1.35	0.76 - 2.5	0.322
Indoor air pollution			
Use charcoal for cooking			
no	*ref*		
yes	0.93	0.56 - 1.58	0.798
Use firewood for cooking			
no	*ref*		
yes	1.05	0.56 - 2.04	0.889
Use ventilation when cooking			
no	*ref*		

SUPPLEMENTAL TABLE E3. Association between risk factors and self-reported pneumonia diagnosis (ever)

Variable	OR [a]	95% CI	p-value
yes	2.77	0.86 - 12.43	0.121
cook outside	3.63	1.23 - 15.6	0.039
Substance use in household			
Household member who uses khat			
no	*ref*		
yes	1.07	0.64 - 1.79	0.781
Household member who smokes			
no	*ref*		
yes	0.99	0.57 - 1.7	0.978
Household member who uses snuff			
no	*ref*		
yes	2.64	1.12 - 6.01	0.023
Contact behaviour			
Total number of direct contacts	0.99	0.92 - 1.07	0.834
Total number of physical contacts	1.03	0.96 - 1.1	0.458
Total number of contacts at home	1.00	0.92 - 1.07	0.920
Total number of contacts at school	0.82	0.59 - 1.02	0.138
Total number of contacts at work	0.62	0.17 - 1.1	0.308
Total number of contacts at other settings	1.07	0.93 - 1.21	0.310
Malnutrition in U5			
Weight by age			
Not underweight (z > -2)	*ref*		
Underweight (z ≤ -2)	1.12	0.41 - 2.91	0.817
Underweight (z ≤ -3)	0.69	0.09 - 3.44	0.666
Height by age			
Not stunted (z > -2)	*ref*		

SUPPLEMENTAL TABLE E3. Association between risk factors and self-reported pneumonia diagnosis (ever)

Variable	OR [a]	95% CI	p-value
Stunted (z ≤ -2)	0.86	0.34 - 2.08	0.746
Severely stunted (z ≤ -3)	1.20	0.36 - 3.8	0.753
Weight by height			
Not wasted (z > -2)	*ref*		
Wasted (z ≤ -2)	0.46	0.1 - 1.61	0.263
Severely wasted (z ≤ -3)	0.0		
Middle-Upper Arm Circumference			
Not wasted (≥ 125mm)	*ref*		
Wasted (< -125mm)	1.82	0.28 - 14.63	0.531

[a] All odds ratios are adjusted for age and gender.

References used in Supplemental Material

1. Lumley T. *Complex Surveys: A Guide to Analysis Using R*. John Wiley & Sons; 2011.

2. Lumley T. *Survey: Analysis of Complex Survey Samples*.; 2020.

3. Funk S. Socialmixr: Social Mixing Matrices for Infectious Disease Modelling. Published January 15, 2018. Accessed June 27, 2020. https://datacompass.lshtm.ac.uk/646/

4. Melegaro A, Del Fava E, Poletti P, et al. Social Contact Structures and Time Use Patterns in the Manicaland Province of Zimbabwe. Nishiura H, ed. *PLoS ONE*. 2017;12(1):e0170459. doi:10.1371/journal.pone.0170459

5. le Polain de Waroux O, Cohuet S, Ndazima D, et al. Characteristics of human encounters and social mixing patterns relevant to infectious diseases spread by close contact: a survey in Southwest Uganda. *BMC Infect Dis*. 2018;18(1):172. doi:10.1186/s12879-018-3073-1

6. Mossong J, Hens N, Jit M, et al. Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLOS Medicine*. 2008;5(3):e74. doi:10.1371/journal.pmed.0050074

7. Béraud G, Kazmercziak S, Beutels P, et al. The French Connection: The First Large Population-Based Contact Survey in France Relevant for the Spread of Infectious Diseases. *PLOS ONE*. 2015;10(7):e0133203. doi:10.1371/journal.pone.0133203

8. Leung K, Jit M, Lau EHY, Wu JT. Social contact patterns relevant to the spread of respiratory infectious diseases in Hong Kong. *Sci Rep*. 2017;7(1):7974. doi:10.1038/s41598-017-08241-1

9. Horby P, Thai PQ, Hens N, et al. Social Contact Patterns in Vietnam and Implications for the Control of Infectious Diseases. *PLOS ONE*. 2011;6(2):e16965. doi:10.1371/journal.pone.0016965

10. Grijalva CG, Goeyvaerts N, Verastegui H, et al. A Household-Based Study of Contact Networks Relevant for the Spread of Infectious Diseases in the Highlands of Peru. *PLOS ONE*. 2015;10(3):e0118457. doi:10.1371/journal.pone.0118457

11. Zhang J, Klepac P, Read J, et al. Social contact data for China mainland. Published online June 5, 2020. doi:10.5281/zenodo.3878754

12. United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects - Population Division - United Nations. Published 2019. Accessed April 19, 2020. https://population.un.org/wpp/Download/Standard/Population/

13. Jarvis CI, van Zandvoort K, Gimma A, et al. Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK | medRxiv. Accessed April 19, 2020. https://www.medrxiv.org/content/10.1101/2020.03.31.20049023v1

14. Diekmann O, Heesterbeek JAP, Roberts MG. The construction of next-generation matrices for compartmental epidemic models. *J R Soc Interface*. 2010;7(47):873-885. doi:10.1098/rsif.2009.0386

Chapter 3

Pre-vaccination carriage prevalence of *Streptococcus pneumoniae* serotypes among internally displaced people in Somaliland

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1604011	**Title**	Mr
First Name(s)	Kevin		
Surname/Family Name	van Zandvoort		
Thesis Title	Strategies for pneumococcal conjugate vaccine use in humanitarian crises		
Primary Supervisor	Stefan Flasche		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	BMC Pneumonia
Please list the paper's authors in the intended authorship order:	Kevin van Zandvoort, Abdirahman Ibrahim Hassan, Mohamed Bobe, Casey Pell, Mohamed Saed Ahmed, Belinda Ortika, Saed Ibrahim, Mohamed Ismail Abdi, Mustapha A Karim, Rosalin M Eggo, Sulieman Yusuf, Saeed Mohamood Soleman, Rachael Cummings, Catherine

	McGowan, Kim Mulholland, Mohamed Abdi Hergeye, Catherine Satzke, Francesco Checchi, Stefan Flasche
Stage of publication	**In press**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	Prior to the survey, I led the study design, conducted a scoping visit to the field site, prepared all electronic Open Data Kit survey forms, and worked with collaborators at the Murdoch Children's Research Institute, London School of Hygiene and Tropical Medicine, and Save the Children International to prepare and ship required study items (nasopharyngeal swabs, vials with STGG, electronic tablets, etc) to the field site. I spent a total of ten weeks in Somaliland. Initially this was to finalize agreements with collaborators at Save the Children and the Ministry of Health Development, training data collectors on the study, sampling process, requesting informed consent, electronic data collection using Open Data Kit and nasopharyngeal swab collection, and to participate in a community engagement activity in Digaale. During the survey, I led the daily data collection in Digaale: supervising the five data collection teams during data collection, transporting collected swabs to the cold-chain facility, synchronizing data between electronic tablets and an external server, assessing and validating collected data in the evening, discussing any issues with the data collection teams the next morning, and regularly updating local collaborators on the progress of the study. I also organized the sample shipment communicating with several shippers before, during, and after data collection as we encountered several challenges with shipment of samples. Once data collection was completed, I cleaned all the data and conducted all statistical analyses. This took several stages, due to delays in sample shipment and redirecting my work to the Covid-19 pandemic. I then interpreted the results, produced all figures and tables, wrote the original manuscript draft, and edited and prepared the final manuscript for submission.

SECTION E

Student Signature	Kevin van Zandvoort
Date	03/10/2024

Supervisor Signature	Stefan Flasche
Date	03/10/2024

Pre-vaccination carriage prevalence of *Streptococcus pneumoniae* serotypes among internally displaced people in Somaliland: a cross-sectional study

Kevin van Zandvoort[1]*, Abdirahman Ibrahim Hassan[2], Mohamed Omer Bobe[3], Casey L. Pell[4], Mohamed Saed Ahmed[3], Belinda D. Ortika[4], Saed Ibrahim[3], Mohamed Ismail Abdi[3], Mustapha A Karim[2], Rosalind M Eggo[1], Saleban Yousuf Ali[2], Jason Hinds[5, 6], Saeed Mohamood Soleman[2], Rachael Cummings[7], Catherine R. McGowan[1,7], E Kim Mulholland[1,4], Mohamed Abdi Hergeye[2], Catherine Satzke[4,8,9], Francesco Checchi[1], Stefan Flasche[1, 10]

* *Corresponding author: Kevin.Van-Zandvoort@lshtm.ac.uk*

1. Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine. United Kingdom.
2. Republic of Somaliland Ministry of Health Development. Somaliland.
3. Save the Children International Somaliland. Somaliland.
4. Infection, Immunity and Global Health, Murdoch Children's Research Institute, Melbourne, Australia.
5. Institute for Infection and Immunity, St. George's, University of London, United Kingdom
6. BUGS Bioscience, London Bioscience Innovation Centre, London, United Kingdom
7. Save the Children UK. United Kingdom.
8. Department of Microbiology and Immunology at the Peter Doherty Institute for Infection and Immunity, The University of Melbourne, Melbourne, Australia
9. Department of Paediatrics, The University of Melbourne, Melbourne, Australia
10. Centre for Global Health, Charité - Universitätsmedizin Berlin, Germany

Abstract

Background

Populations affected by humanitarian crises likely experience high burdens of pneumococcal disease. *Streptococcus pneumoniae* carriage estimates are essential to understand pneumococcal transmission dynamics and the potential impact of pneumococcal conjugate vaccines (PCV). Over 100 million people are forcibly displaced worldwide, yet here we present only the second pneumococcal carriage estimates for a displaced population.

Methods

In October 2019, we conducted a cross-sectional survey among internally displaced people (IDP) living in Digaale, a permanent IDP camp in Somaliland where PCV has not been implemented. We collected nasopharyngeal swab samples from 453 residents which were assessed for presence of pneumococci and serotyped using DNA microarray.

Results

We found that pneumococcal carriage prevalence was 36% (95%CI 31 – 40) in all ages, and 70% (95%CI 64 – 76) in children under 5. The three most common serotypes were vaccine serotypes 6B, 19F, and 23. We estimated that the serotypes included in the 10-valent PNEUMOSIL vaccine were carried by 41% (95%CI 33 – 49) of all pneumococcal carriers and extrapolated that they caused 52% (95%CI 35 – 70) of invasive pneumococcal disease. We found some evidence that pneumococcal carriage was associated with recent respiratory symptoms, the total number of physical contacts made, and with malnutrition in children under 5. Through linking with a nested contact survey we projected that pneumococcal exposure of children under 2 was predominantly due to contact with children aged 2-5 (39%; 95%CI 31 – 48) and 6-14 (25%; 95%CI 17 – 34).

Conclusions

These findings suggest considerable potential for direct and indirect protection against pneumococcal disease in Digaale through PCV use in children and potentially adolescents.

Keywords

Streptococcus pneumoniae; displaced population; humanitarian crisis; pneumococcal conjugate vaccine; Somaliland

Introduction

The United Nations High Commissioner for Refugees estimates that over 100 million people were forcibly displaced worldwide in 2022, of whom over half are internally displaced and more than 40% are children [1]. These people typically live in overcrowded settings, have poor access to hygiene and healthcare, and thus experience high morbidity and mortality from respiratory diseases [2–4] including invasive disease from *Streptococcus pneumoniae* (pneumococcus) [5]. There is a lack of health research in these populations in general [6], and for pneumococcal disease specifically [2, 4]. Pneumococcal prevalence estimates are only available for one other displaced population globally, living in Mae La refugee camp, Thailand [7].

Pneumococcal Conjugate Vaccines (PCV) are highly efficacious and have been used routinely for protecting children against pneumococcal colonization and disease in most countries worldwide. Five PCVs are currently used for childhood immunization: Prevenar 20 (20 valent, targeting 20 of the over 100 serotypes), Vaxneuvance (15 valent), Prevenar 13 (13 valent), Synflorix and PNEUMOSIL (both 10 valent) [8–10]. However, despite the high prevalence of risk factors for severe disease and intense transmission, they are rarely offered to populations that have become displaced as a result of food insecurity, conflict, natural disasters, or other emergencies [2].

Pneumococcal colonization is common and is a precursor to disease, thus providing an opportunity to assess the likely risk for disease, even in small populations, without costly disease surveillance programmes that would be difficult to establish in displaced populations [11]. Such information is crucial to inform the design of effective pneumococcal immunization strategies in displaced populations where routine immunization is rarely possible [2].

We conducted a cross-sectional survey to estimate nasopharyngeal carriage prevalence and related risk factors in Digaale, a camp for internally displaced people (IDPs) in Somaliland [2]. PCVs had not yet been introduced in Somaliland at the time of writing.

Methods

Study population and sampling method

We conducted a cross-sectional survey in October-November 2019. The Digaale IDP camp was established in 2014 and, at the time of our study, housed an estimated population of 3,000 people largely displaced due to drought and food insecurity during 2013 and 2014 [12]. The camp is located about 4km south-east of Hargeisa, the capital of Somaliland, and consists of corrugated-steel shelters. The population is served by a school and primary health centre.

We visited all 894 shelters in Digaale and invited households to participate in our survey. We first administered a structured household survey among consenting households to establish their composition and collect household-level information on shelter conditions, pneumococcal risk factors, and retrospective mortality and demographic changes. We then selected and invited individual household members for participation in a survey on carriage, contact and individual-level risk factors. We aimed to sample 100 individuals in each of the following age groups: <1, 1, 2-5, 6-14, 15-29, 30-49, and ≥50 years (y) old to detect age-specific pneumococcal prevalence within 10% precision. We purposively oversampled young children who have the highest incidence of pneumococcal carriage and disease. Quota sampling by age was used to select individual household members.

We returned to individual participants two days after the household survey and study enrolment to conduct the contact and risk factor survey, in which we asked about individual-level risk factors including social contacts, and measured anthropometry for children aged 6-59 months. Further details about the design and sampling method, as well as detailed social contact and household-level findings, are described in Van Zandvoort et al [12]. In the final week of data collection, one to four weeks after participation in the contact survey, we followed up participants to collect a nasopharyngeal swab sample. To compensate for loss-

to-follow up, additional participants were sampled from household members of participating households.

All swabs were collected at a community hall in the centre of the camp. During swab collection, we asked participants whether they had experienced any respiratory symptoms (cough, sore throat, sneezing, wheezing, headache, or fever) or used any antibiotics in the two weeks prior. Responses for children under 10y were provided by an adult parent or caregiver. Participants with a contraindication for a nasopharyngeal swab such as facial trauma were excluded.

Sample collection, storage, and shipment

Trained nurses collected nasopharyngeal swabs from each participant using flexible paediatric- (Ultra Minitip) or adult-size (Flexible Minitip) flocked swabs (FLOQSwabs; Copan Diagnostics, USA), following WHO recommendations [13, 14]. Paediatric-sized swabs were used for children under 15y. The nasopharyngeal swabs were stored in screw-capped tubes containing 1 ml of skim milk-tryptone-glucose-glycerol (STGG) medium and kept on wet ice in cool boxes immediately after collection. Samples were transferred to a -20°C freezer at the Ministry of Health Development national cold chain facility within eight hours of collection [15]. Within two weeks after collection, all samples were transferred to an ultra-low temperature (ULT) freezer at the culture laboratory of the Somaliland National Tuberculosis Hospital. Due to technical issues with the ULT freezer, samples were temporarily stored at -20°C for 15 days in March 2020.

We used a prequalified shipping solution with phase change materials (PCMs) that provided passive cooling to keep contents below -15°C for up to 96h as per manufacturer specifications (Schaumaplast, DE) [16]. We found that it maintained temperatures below -15°C up to 160h in an empty trial shipment from London to Hargeisa when conditioned at ULT (Supplemental Material Section B). Samples were first shipped to Nairobi, Kenya,

where they were repackaged and placed on dry ice for further transport to the Murdoch Children's Research Institute (MCRI) in Melbourne, Australia. We successfully completed a pilot shipment of 81 samples in May 2021, and shipped all remaining samples in December 2021. Transit delays during the second shipment extended the period during which only passive cooling was provided by PCMs to 11 days. There was no temperature monitoring during this second shipment, but we project that temperatures may have increased to >0°C for up to 2.5 days based on temperature measurements from the pilot shipment. We explored any difference in carriage estimates between the two shipments in a sensitivity analysis.

Microbiological analysis

Upon arrival at the MCRI laboratory, STGG swab samples were immediately stored at ULT until testing. Briefly, samples were thawed, vortexed, and DNA extracted from 100 µl using the QIAcube HT machine (Qiagen) following a protocol described previously [17]. Concurrently, each sample was cultured overnight on selective agar, and growth harvested if alpha-haemolytic colonies were present [18]. Real-time quantitative PCR, targeting the *lytA* gene (*lytA* qPCR) [19], was conducted as previously described, except for using 5 µl template DNA and the AriaMX PCR system (Agilent) [18]. For samples with presumptive pneumococci (*lytA* qPCR cycle threshold <40 and alpha-haemolytic growth), DNA was extracted from harvested colonies using the QIAcube HT machine (Qiagen) and the resultant DNA serotyped using microarray (Senti-SP version 1.5, BUGS Bioscience) [18]. Pneumococcal density was calculated using a genomic DNA standard curve [18]. Serotype-specific density was calculated by multiplying the pneumococcal density (as determined by *lytA* qPCR) by the relative abundance of the serotype (as determined by microarray). Carriage density estimates were log10-transformed and reported as $\log_{10}$ genome equivalents/ml (GE/ml). Non-encapsulated serotypes (NESp) were categorized by their genetic lineage[20].

Statistical analysis

Serotypes were grouped as NESp, vaccine serotypes (VT) for each of PNEUMOSIL, Synflorix, Prevenar 13, Vaxneuvence, and Prevenar 20 vaccines, or non-vaccine serotypes (NVT). Unless stated otherwise, we used PNEUMOSIL targeted serotypes to define VT in our base case analysis, because PNEUMOSIL was specifically developed to provide a lower cost alternative that targets the main serotypes causing pneumococcal disease in low and middle income countries [21]. Analyses that defined VT using the other vaccines are presented in Section C in the Supplemental Material.

We estimated the population and age-specific prevalence of pneumococcal serotypes with 95% binomial confidence intervals. To account for multiple serotype carriage, serotypes were weighted by their relative abundance within a sample to calculate the serotype distribution including co-carried serotypes, so that the weights of all serotypes in a sample summed to one. Unweighted distributions of all and dominant-only serotypes were assessed in a sensitivity analysis. We define serotypes ranked with the highest relative abundance in a single sample as dominant serotypes, and used logistic regression to assess the odds that serotypes were dominant.

As we sampled a large proportion of households (65%) and individuals (17%) living in Digaale, we used finite population corrections (FPC) to calculate standard errors used in population-level estimates. To correct for imbalance resulting from quota sampling and thus improve the representativeness, we applied poststratification weights based on age group and gender [12, 22]. We assessed the sensitivity of our population level estimates to poststratification in additional analyses. We used the *Survey* package in *R* to perform poststratification weighting and to apply the FPC when estimating weighted means, proportions, and quantiles where applicable [23].

To project the proportion of current IPD cases caused by serotypes covered by PCVs, we combined the dataset with age- and serotype-specific estimates of invasiveness by Løchen et al [24], and computed confidence intervals by bootstrapping our dataset and invasiveness estimates. More details are provided in Supplemental Material Section D.

To estimate the likely contribution of different age groups to pneumococcal exposure and transmission, we followed a method developed by Qian et al [25] and calculated the proportion of colonisations attributable to contact with different age groups by linking the estimates of carriage prevalence with previously reported contact patterns [12] in those individuals and the likelihood of colonisation among those contacts. As the contact and carriage datasets featured different statistical error processes, we computed confidence intervals for this analysis through bootstrapping from each parameter's uncertainty distribution.

We used logistic regression to test the univariate association of likely risk factors with the odds of pneumococcal carriage and used linear regression to test their association with mean logged overall carriage density in pneumococcal carriers. Age and gender were included as *a priori* defined confounding variables in both analyses, but no other multivariate analyses were conducted due to data sparsity. We also used logistic regression to test for any association between the odds of multiple carriage and age, and to assess whether the odds of carrying NESp serotypes differed by age. FPC and poststratification weights were not applied in regression analyses.

All analyses were conducted in *R4.2.2*.

Results

Participant sampling and enrolment

We enrolled 464 (65%) households and collected demographic data from 2,049 individuals living in these households. A contact and individual-level risk-factor survey was conducted among 509 participants. We collected a nasopharyngeal swab from 365 of these participants. An additional 88 nasopharyngeal swabs were collected from other individuals from consenting households. In total, 453 swabs were collected (Fig 1).

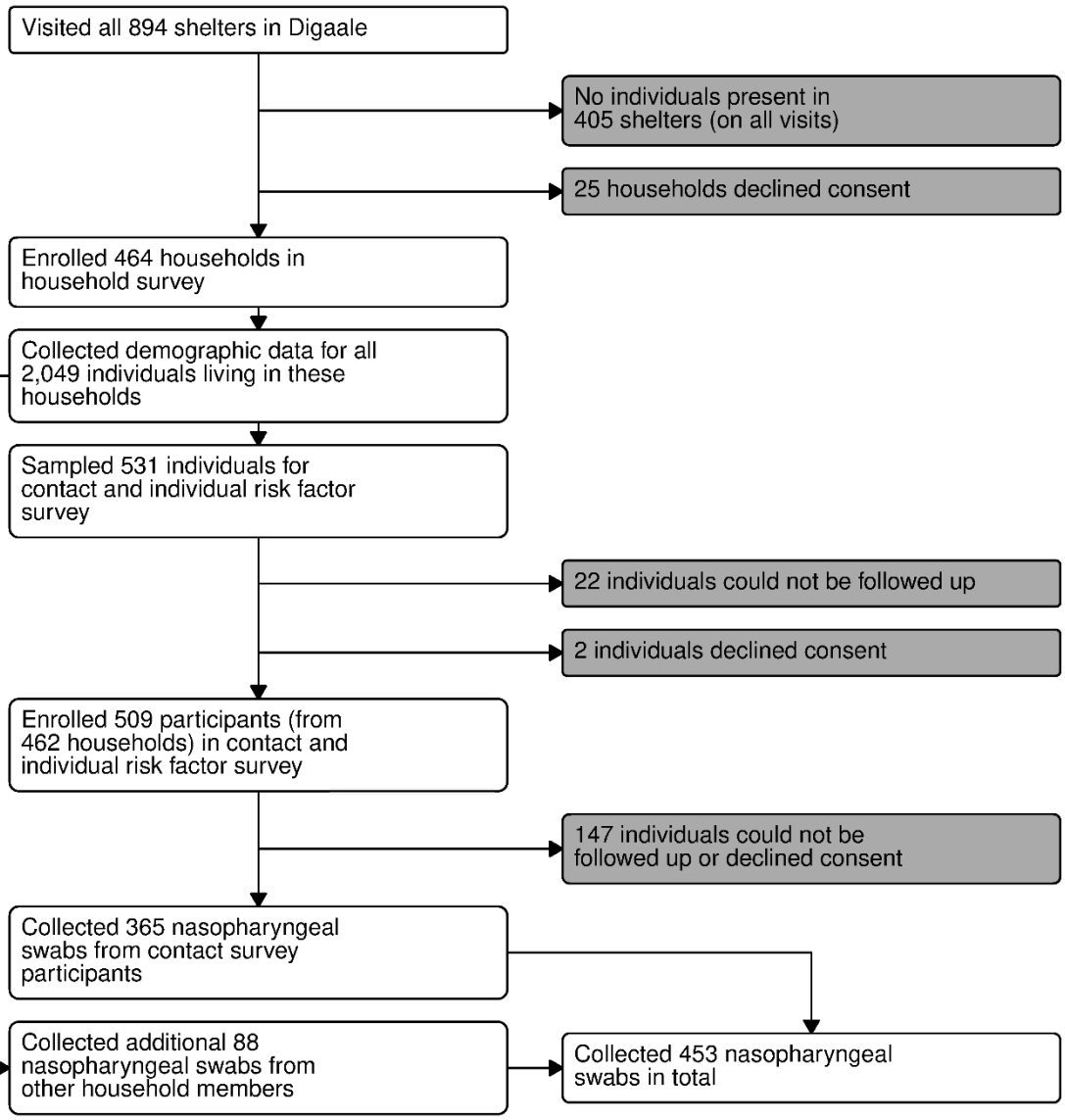


Fig 1. Flowchart and sampling procedure.

All households that were present were invited to participate in the survey. A structured household

survey was conducted in consenting households. Individual household members were selected using quota sampling and invited for a contact and individual-level risk factor survey. Participants were followed up after one to four weeks to administer a nasopharyngeal swab. Additional household members were sampled from consenting households using quota sampling to compensate for loss to follow-up.

Sample characteristics

Two (0.4%) nasopharyngeal swab samples were excluded due to insufficient sample volume for laboratory testing. An additional two (0.4%) samples were *lytA* positive, indicating pneumococcal colonization, but with no growth prohibiting microarray serotyping; these were included only in non-serotype specific carriage prevalence analyses. Due to data entry errors, 53 swabs (12%) could not be fully linked to records in contact or household datasets. This included 16 swabs that could not be linked to their household information, and 6 samples without information on age and gender (Supplemental Material Section A). Data from these swabs are excluded in analyses that require linking to those data, but included otherwise.

67% of the study participants from whom swabs were collected were female (Table 1). Median age was 13y and median household size was 5 people. Pneumococci were detected in 39% (175/445) of samples, with at least one PNEUMOSIL VT present in 49% (85/175) of positive samples. We estimated the overall carriage prevalence in Digaale as 36% (95%CI 31 – 40), with 41% (95%CI 33 – 49) of carriers carrying at least one VT. Estimated prevalence was 70% (95%CI 64 – 76) in children under 5y, with VTs carried by 61% (95%CI 53 – 69) of carriers. Large proportions of participants reported respiratory symptoms (64%) and antibiotic use (36%) in the two weeks leading up to the sampling.

Table 1. Sample characteristics, carriage prevalence, and invasive disease likely caused by vaccine serotypes.

Variable	Obs[a]	Sample value	Population est[b]		Population est (<5y)[b]	
Sample characteristics (unweighted)						
Median age (y)	447	13 (IQR 3 – 40)	.		.	
Percentage female	298/447	66.7%	.		.	
Median household size (people)	433	5 (IQR 3 – 6)	.		.	
Median household members <5y (people)	433	0 (IQR 0 – 2)	.		.	
Potential risk factors						
Respiratory symptoms	261/410	63.7%	61.1%	55.9 – 66.2	73.5%	67.0 – 79.9
Antibiotic use	148/409	36.2%	31.6%	27.0 – 36.3	45.9%	38.6 – 53.2
Direct contacts	364	9.6	9.6	9.1 – 10.0	9.9	9.4 – 10.3
Carriage prevalence						
Pneumococcal carriage	175/445[c]	39.3%	35.8%	31.1 – 40.4	70.0%	63.7 – 76.4
Non-encapsulated carriage	19/175	10.9%	13.2%	6.9 – 19.5	6.6%	2.8 – 10.5
Proportion of carriers with VT						
PNEUMOSIL[d]	85/175	48.6%	40.8%	32.9 – 48.7	60.9%	52.7 – 69.1
Synflorix[e]	84/175	48.0%	40.0%	32.1 – 47.8	59.7%	51.4 – 68.0
Prevenar 13[f]	98/175	56.0%	51.0%	43.1 – 58.9	62.9%	54.8 – 71.0
Vaxneuvance[g]	98/175	56.0%	51.0%	43.1 – 58.9	62.9%	54.8 – 71.0
Prevenar 20[h]	110/175	62.9%	58.0%	50.1 – 65.9	68.3%	60.4 – 76.3
Projected proportion of IPD covered by PCVs[i]						
PNEUMOSIL[d]	.	52.9%	51.9%	34.8 – 69.1	67.8%	50.9 – 81.6
Synflorix[e]	.	60.8%	60.1%	41.8 – 75.0	74.1%	57.1 – 85.6
Prevenar 13[f]	.	72.4%	74.9%	60.0 – 85.3	77.9%	61.0 – 88.6
Vaxneuvance[g]	.	73.0%	75.5%	60.8 – 85.8	78.9%	61.9 – 89.0
Prevenar 20[h]	.	80.9%	81.6%	66.9 – 89.9	88.0%	77.0 – 93.7

a. Total number of observations in the dataset. For proportions, both numerators and denominators are shown.
b. Mean value and corresponding 95%CI. Post-stratification weights and finite population corrections have been used to calculate population estimates.
c. Six observations excluded due to missing age or gender data prohibiting post-stratification weighting.
d. Serotypes 1, 5, 6A, 6B, 7F, 9V, 14, 19A, 19F, and 23F.
e. Serotypes 1, 4, 5, 6B, 7F, 9V, 14, 18C, 19F, and 23F.
f. Serotypes 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, and 23F.
g. Serotypes 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, 22F, 23F, and 33F.
h. Serotypes 1, 3, 4, 5, 6A, 6B, 7F, 8, 9V, 10A, 11A, 12F, 14, 15B, 18C, 19A, 19F, 22F, 23F, and 33F.
i. Estimated from serotype-specific invasiveness estimates. See Section D in the Supplemental Material for more details.

Serotype distribution

The three most prevalent pneumococcal serotypes overall (6B, 19F, and 23F) were VTs included in all PCVs, followed by three serotypes (6C, 11A, 16F) of which 11A is only

included in Prevenar 20, and all others are NVTs (Fig 2). A similar serotype distribution was observed in children under 5y, while 15B/C, 6B, and 6C were the most prevalent serotypes in people over 5y (Fig C2 in the Supplemental Material). There was little difference in the serotype distribution when restricting analysis to dominant serotypes alone, or without weighting for relative abundance in multiple serotype carriers (Fig C1 in the Supplemental Material). We extrapolate that 52% (95%CI 35 – 70) of all IPD cases and 68% (95%CI 51 – 82) of IPD cases in <5y were caused by VT serotypes, with slightly increased proportions for higher valency vaccines (Table 1). Out of all 248 identified serotypes in all samples, 20 were identified as NESp. These were categorized as NT2 (8), NT3b (6), NT4a (4), and NT4b (2) genetic lineages.

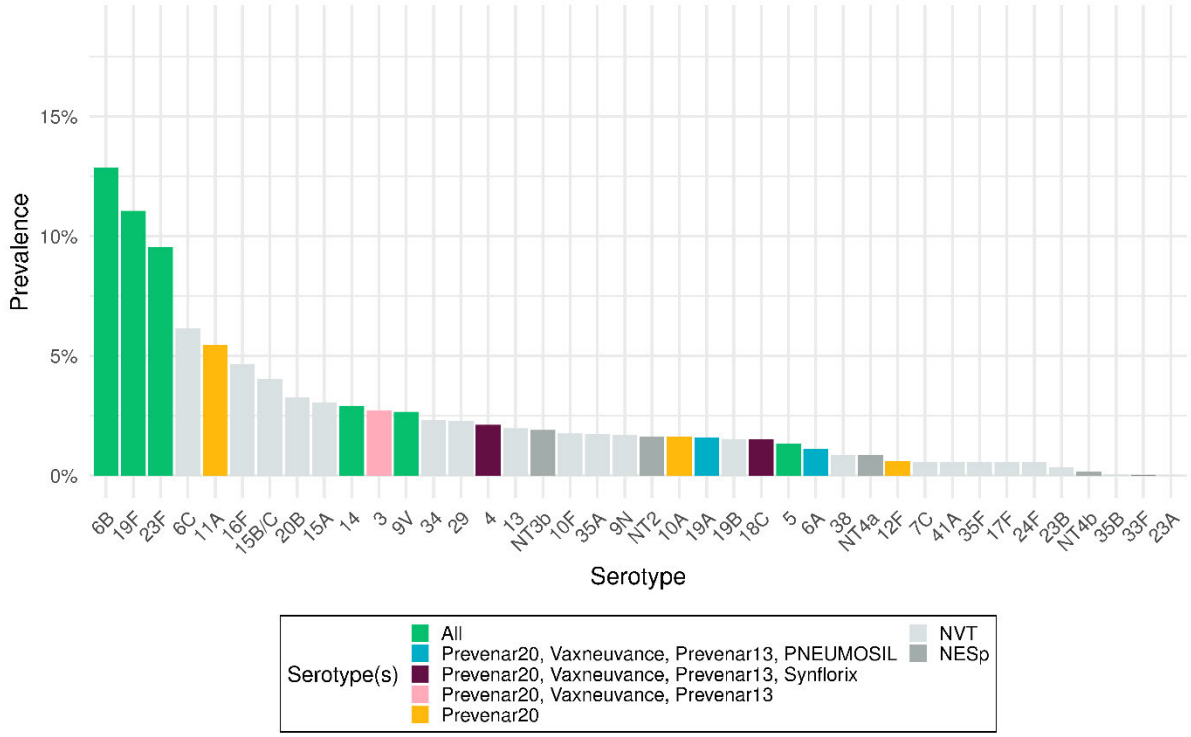


Fig 2. Pneumococcal serotype distribution.
Bars show the carriage prevalence of pneumococcal serotypes identified in Digaale IDP camp, weighted by their relative abundance. Coloured bars show the serotypes included in the five PCVs, dark grey bars non-encapsulated pneumococci, and light grey bars other serotypes not included in the PCVs. Error bars show 95% confidence intervals for each estimate.

Multiple serotype carriage

Co-colonization with more than one serotype was detected in 30% (52/176) of samples with pneumococci, corresponding to a population-level prevalence of 39% (95%CI 35 – 44). There were 36 samples in which two serotypes were detected, 13 samples with three serotypes, and three samples with more than three serotypes. The odds that the serotype was dominant was 2.0 (95%CI 1.1 – 3.7) times higher for VTs than for NVTs in all carriers, and 2.8 (95%CI 1.0 – 8.0) times higher among carriers colonized with both VT and NVT (Table C2 in the Supplemental Material).

Sensitivity analyses

We found no evidence of a difference in pneumococcal carriage, density of pneumococcal serotypes, number of carried serotypes, or serotype distribution between our two sample shipments (Supplemental Material Section B). Post stratifying our estimates did not substantially affect pneumococcal carriage prevalence estimates (Table C5 in the Supplemental Material).

Antimicrobial resistance

Microarray assays detected the presence of select antimicrobial resistance genes in 30% (95%CI 21 – 41) of samples (Table C2 in the Supplemental Material). In those samples, the most common detected genes typically associated with antimicrobial resistance were *tetM* (28% [95%CI 19 – 39]) and *ermB* (9% [95%CI 4 – 16]). We restricted this analysis to samples in which no other species and only a single pneumococcal serotype were detected

Carriage prevalence and serotype distribution by age

Overall carriage prevalence was 79% (95%CI 70 – 88) and 66% (95%CI 57 – 75) in children under 2 and 2-5y, respectively. Carriage prevalence was 41% (95%CI 31 – 51) in children

aged 6-14y, 25% (95%CI 14 – 36) in people aged 15-29y, 17% (95%CI 9 – 26) in adults aged 30-49y and 10% (95%CI 4 – 15) in adults aged ≥50y (Fig 3). Co-colonization rates decreased by age alongside reductions in overall prevalence, although this reduction was not statistically significant. The proportion of VT among all carriers was similar across age groups and robust to the definition of VT for different PCV products (Fig C3 in the Supplemental Material).

We found weak evidence that the odds of carrying a NESp serotype among carriers was 4.8 (95%CI 1.01 - 22.34) times higher in adults 30-49y compared to children under 2y, but no evidence for a difference in NESp carriage other age groups.

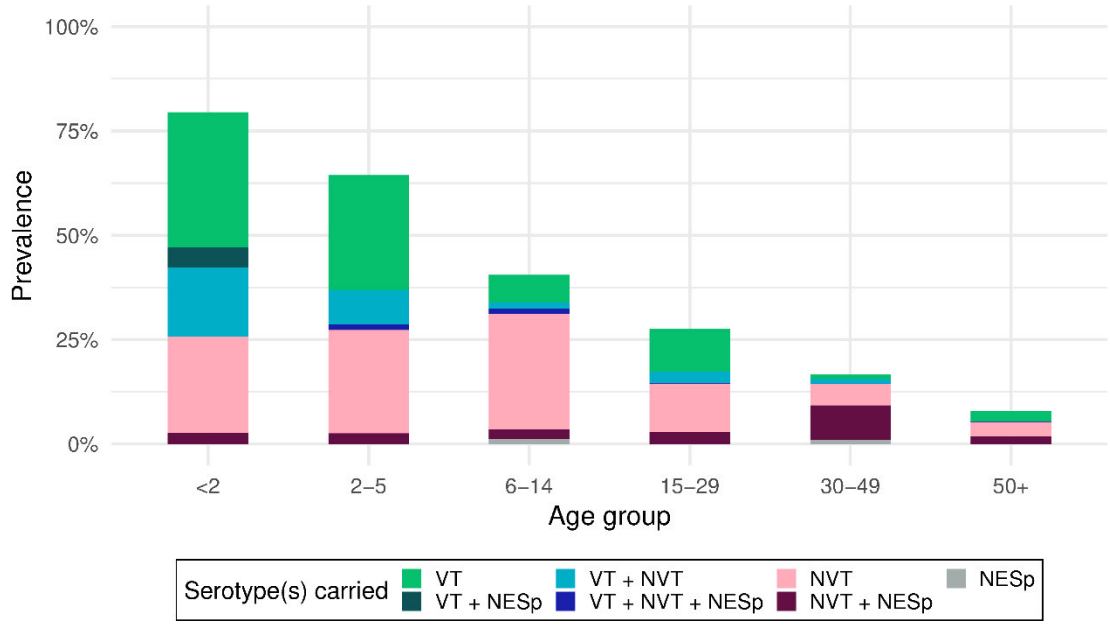


Fig 3. Prevalence and serotype distribution by age.
Bars show the estimated prevalence of pneumococcal serotypes by age group, weighted for age and gender. Error bars show 95% confidence intervals around overall pneumococcal carriage prevalence. Colours show the prevalence of serotypes that are carried; VT: only vaccine type(s), NVT: only non-vaccine type(s); NESp: only non-encapsulated type(s); VT + NVT: both vaccine- and non-vaccine type(s). Multiple carriage with non-encapsulated type(s) is shown as a darker shading.

Contribution of different age groups to pneumococcal exposure

We projected that a large proportion (39% [95%CI 31 – 48]) of pneumococcal exposure of children <2y may be attributed to contact with 2-5y children, followed by school age children aged 6-14y (25% [95%CI 17 – 34]) (Fig 4 and Table C6 in the Supplemental Material). A

similar contribution was made by carriers of these age groups to exposure of children aged 2-5y (45% [95%CI 37 – 52]; and 30% [95%CI 23 – 38]). Most of the exposure of school age children, however, was found to be attributable to other school age children (52% [95%CI 42 – 60]), followed by 2-5y olds (26% [95%CI 20 – 32]). While carriage prevalence was high in children <2y, this age group was found to contribute relatively little to onward transmission to any age group.

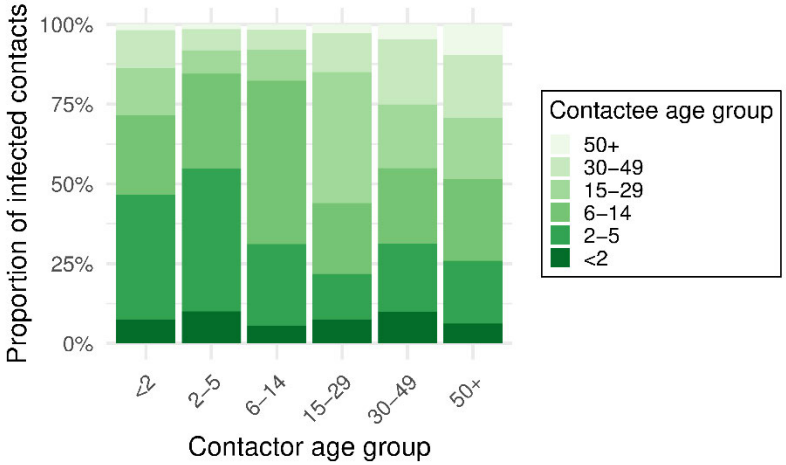


Fig 4. The contribution of different age groups towards the age-specific exposure to pneumococcus.
Bars show the average proportion of contacts made by a contactor of age *i* (x-axis), with pneumococci-carrying contactees of different age groups. Shades of green stratify into age group of the contactee, i.e. the person potentially transmitting to the contactor.

Association of risk factors with pneumococcal carriage and carriage density

We found no evidence that the number of overall household members increased the odds of pneumococcal carriage, but weak evidence that living with one additional household member <5y of age increased the odds of carriage by 1.3 (95%CI 1.0 – 1.8) (Table 2). The odds of carriage were 2.0 (95%CI 1.2 – 3.3) times higher in people with recent respiratory symptoms. Having a cough (2.0 [95%CI 1.2 – 3.3]) had the strongest association, followed by having a sore throat (1.7 [95%CI 1.0 – 3.0]). There was some evidence that the odds of carriage increased by 1.1 (95%CI 1.0 – 1.2) for every additional physical contact reported.

277 We found good evidence for a reduction in the odds of carriage for improved scores of
278 weight-for-age (0.6 [95%CI 0.4 – 0.9]) and height-for-age (0.6 [95%CI 0.4 – 0.8]) among
279 children 6-59 months old, but no evidence for an association with weight-for-height or
280 middle-upper arm circumference. Notably, we did not find any evidence of an association
281 with self-reported antibiotic use.

Table 2. Association between risk factors and pneumococcal carriage.

Variable	OR[a]	95%CI	p-value	N[b]
Demographic characteristics				
Household size	1.05	0.95 – 1.17	0.336	431
Household members <5y	1.32	0.99 – 1.76	0.054	431
Shelter quality				
House leakage	1.15	0.69 – 1.95	0.591	431
House draught	0.68	0.41 – 1.13	0.140	431
Indoor air pollution				
Use firewood as fuel	1.03	0.55 – 1.92	0.938	434
Use charcoal as fuel	0.75	0.46 – 1.19	0.223	434
Ventilation				431
yes	0.49	0.20 – 1.16		
cook outside	0.56	0.25 – 1.24	0.348	
Current health[c]				
Antibiotic use	1.28	0.78 – 2.10	0.324	407
Respiratory symptoms	1.99	1.20 – 3.32	0.008	408
Cough	2.00	1.23 – 3.25	0.005	408
Sore throat	1.69	0.95 – 3.00	0.072	408
Headache	0.74	0.41 – 1.32	0.309	408
Fever	1.17	0.70 – 1.94	0.545	408
Diarrhoea	1.62	0.73 – 3.73	0.244	408
Morbidities[d]				
Pneumonia 6m[e]	1.28	0.68 – 2.41	0.451	363
Sickle Cell	1.25	0.38 – 3.77	0.697	363
Asthma	0.96	0.11 – 5.73	0.968	363
Diabetes	1.09	0.05 – 8.57	0.939	363
Individual substance use				
Tobacco	0.50	0.03 – 2.92	0.524	363
Khat	0.63	0.03 – 3.93	0.681	363
Household substance use[f]				
Smoking	1.42	0.84 – 2.39	0.186	434
Snuff	1.64	0.66 – 4.17	0.287	434
Khat	1.31	0.81 – 2.12	0.265	434

Table 2. Association between risk factors and pneumococcal carriage.

Variable	OR[a]	95%CI	p-value	N[b]
Contact behaviour				
Total number of direct contacts	1.04	0.97 – 1.12	0.278	362
Total number of physical contacts	1.08	1.01 – 1.15	0.034	362
Malnutrition in <5y olds				
Weight-for-age z-score	0.59	0.37 – 0.90	0.018	112
Weight-for-height z-score	1.16	0.79 – 1.70	0.440	112
Height-for-age z-score	0.61	0.43 – 0.82	0.003	112
MUAC[g] (in cm)	0.80	0.53 – 1.21	0.229	112

a. Estimates are adjusted for age and gender.
b. Total number of records used in regression.
c. Self-reported antibiotic use and symptoms in 2 weeks preceding the survey.
d. Self-reported diagnosed morbidities.
e. Pneumonia diagnosis in the 6m preceding the survey.
f. Substance use by at least one household member.
g. Middle-Upper-Arm-Circumference

We also tested the association between these risk factors and the density of pneumococcal carriage (Table C3 in the Supplemental Material) and found weak evidence that living with one additional household member <5y was associated with a small 0.2 (95%CI -0.0 – 0.4) increase in mean log10 GE/ml, while associations with respiratory symptoms were either non-significant or negative: the mean log10 density was 0.40 (95%CI -0.8 – 0.0) lower in participants reporting a sore throat in the two weeks preceding the survey. Again, we did not find any significant association with self-reported antibiotic use. There was very weak evidence for an increase in children's mean log10 density with a one-unit increase in weight-for-height z-score (0.2 [95%CI -0.0 – 0.4]).

Discussion

This is the first study to have estimated pneumococcal serotype prevalence in Somaliland and in an IDP camp. We find high carriage prevalence of 36% in all age groups, and 70% in children under 5y. Between 40 and 58% of pneumococcal carriers carried serotypes included in PCVs, depending on the PCV product, and the three most prevalent serotypes were covered by all PCVs. The majority of exposure to pneumococcal carriers in children

younger than 15y may have been attributable to carriers aged 2-5y and 6-14y, with little exposure from carriers aged younger than two years of age. We found that pneumococcal carriage was associated with the number of household members aged <5y, a recent cough, the total number of physical contacts in all age groups, and with stunting in children aged <5y. We estimate that all PCVs cover a substantial proportion of serotypes likely causing IPD in this population.

While we did not find local evidence of a significant association for all, many risk factors previously found to be associated with pneumococcal carriage are present in this population [26]. Residents in Digaale live in overcrowded conditions and likely experience high levels of indoor air pollution. On average, one in five children are malnourished, and residents report a high frequency of direct contacts involving physical touch [12, 27, 28]. While carriage prevalence was high, it is similar to that observed in non-displaced populations in other high-transmission settings in east Africa, and not as high as prevalence observed in rural Gambia where high carriage prevalence is sustained into older adulthood [29–32]. Despite a high disease burden, displaced populations are understudied, and we are only aware of one other published carriage survey conducted in Mae La, a long-term camp for displaced people in Thailand, where carriage prevalence was estimated at a similar 80% in children <2y [7].

The most prevalent serotypes in Digaale (6B, 19F, and 23F) have often been observed to dominate carriage in other PCV-naïve populations, although the relatively high prevalence of 6C and low prevalence of 6A and 19A in our study is unusual [29–31, 33]. Around 50% of serotypes we detected were VTs, and the prevalence of observed serotypes included in the 10-valent Synflorix and PNEUMOSIL PCVs were similar. The proportion of VTs increased with valency of the vaccine. However, due to our relatively small sample size, serotype-specific confidence intervals are very wide. We estimate that any of the five PCVs are likely to cover the serotypes causing the majority of the pneumococcal disease burden. While serotype replacement would mitigate the overall PCV impact, substantial reductions in

pneumococcal disease have been observed where PCVs have been introduced with sufficient coverage [34, 35]. While we did not collect data on the pneumococcal disease burden in Digaale, 43% of children under two years of age were reported to have been diagnosed with pneumonia in the six months preceding the survey [12], and pneumococci were one of the leading causes of childhood pneumonia globally in the pre-PCV era [36].

The combined contact and prevalence estimates showed that pneumococcal transmission in the <2y in Digaale was mostly driven by children aged 2-5y (39%) and 6-14y (25%), with little contribution to transmission from children younger than 2y old who have fewer social contacts. This could be important when designing vaccine strategies, especially those that partially rely on controlling pneumococcal transmission by indirect effects or need to prolong campaign effects in settings where continued vaccination through routine EPI schedule is not possible, as this requires extending the age range of the eligible population [2]. While Digaale is an established camp that is safe and easy to access, this is not the case for many other displaced populations. In conflict settings, it is often not feasible to introduce routine immunization, and policy makers may choose alternative strategies that aim to immunize the subgroups that drive transmission, thereby indirectly protecting other subgroups at highest risk of severe disease.

Participants reported high rates of antibiotic usage in the two weeks preceding sample collection. This may be associated with the high proportion of participants with respiratory symptoms in the same period. However, we cannot rule out reporting bias. We found no association of antibiotic use with reduced carriage contrary to findings in other settings [30, 37, 38]. Although in this study we do not have estimates of phenotypic pneumococcal resistance, microarray testing identified genes typically associated with pneumococcal resistance in a third of all samples, which may be consistent with high antibiotic pressure. The *tetM* gene, known to encode tetracycline resistance, was identified in 28% of pneumococci, mirroring its high prevalence in other studies [39, 40]. The *ermB* gene was found in 9% of pneumococci, and is associated with macrolide resistance [41]. Future

353 improved understanding of antimicrobial resistance in pneumococci would be useful to better
354 understand the impact of more clinically-relevant antibiotics for standard care as well as the
355 potential impact of mass-drug administration campaigns, a potential alternative intervention
356 to reduce the pneumococcal disease burden proposed for crisis settings [42].

357 We assessed the relationship between a number of known risk factors with the odds of
358 pneumococcal carriage and the mean pneumococcal carriage density. We found
359 relationships in the expected direction for some risk-factors, such as an increased odds of
360 carriage for participants with a higher number of household members under 5y of age, those
361 with more direct contacts, and those with recent respiratory symptoms. Asymptomatic
362 pneumococcal carriage has previously been found to be associated with rhinitis, and may be
363 affected by other respiratory infections [43]. It should be noted that confidence intervals were
364 very wide due to a relatively low sample size and low variability within many risk factors.
365 Moreover, we only adjusted our estimates for age and gender, and they are likely affected by
366 residual confounding, while the large number of significance tests means that some spurious
367 associations may have been estimated.

368 There are several limitations to our study. The study population was substantially smaller
369 than expected as many shelters were uninhabited at the time of the study. Thus, we only
370 reached 65% of our target sample size of 700 participants, particularly in young children,
371 where we only reached 24% and 43% of our target sample size of 100 each in children aged
372 <1 and 1y [12]. We therefore pooled the <1 and 1y age groups in a single <2y age group,
373 which allowed us to estimate age-specific prevalence with sufficient precision, but a larger
374 sample would have resulted in more detailed estimates. We could only conduct data
375 collection during daylight hours and may have missed older individuals who work outside
376 Digaale, as many leave the camp very early in the morning and return late at night. This is
377 likely reflected in the gender distribution of the recruited sample, but unlikely to have affected
378 our carriage estimates, as prevalence is low in these older age groups and unlikely to differ
379 substantially from those who were present in Digaale. We post-stratified our estimates to

adjust for any imbalances in our sample and did not detect differences. Although, pneumococcal carriage is generally consistent across seasons [44], we only conducted a single cross-sectional survey and do not know how estimates may change throughout the year. Carriage prevalence was similar to that in general populations in East Africa, but no other prevalence estimates exist for Somaliland, and we cannot infer how results may differ from the general Somaliland population. Our study was not powered to detect relationships between carriage and risk factors, which may explain why we did not find statistically significant effects in most univariate analyses, and a prospective cohort design would be more suitable to infer causality. Finally, many risk factors were self-reported, and their accuracy may be affected by reporting bias.

Ideally, pneumococci are stored at ULT to maintain long-time viability, but we experienced several challenges related to sample storage and shipment. Sample shipment was substantially delayed, partly due to the COVID-19 pandemic, and after several months of storage at ULT, swabs had to be temporarily transferred to a -20°C freezer to allow for ULT freezer repairs. We were not able to transport samples at ULT as local airlines did not accept shipments of dry ice. However, we have shown separately that effect on pneumococcal viability is limited if stored at -20°C for up to three weeks [15], which was maintained in our study in the periods that ULT storage was not feasible. We further monitored sample viability by incorporation of *lytA* qPCR, a molecular screening assay that is not expected to be affected by culture viability, and did not observe a large number of non-culturable samples that were *lytA* positive. Despite transit delays during the second shipment of most of our samples during which temperatures may have exceeded 20°C for up to 2.5 days, we found no difference between carriage and VT prevalence estimates between these samples and those transported during the first shipment. Hence, we believe any effect on sample viability was limited and did not greatly affected our results, supported by the ability to culture at high prevalence and with detection of serotypes carried at low abundance.

406 We collected nasopharyngeal swabs assessed using *lytA* qPCR and microarray consistent
407 with the WHO guideline [14]. We did not examine oropharyngeal carriage which may have
408 underestimated prevalence in adults [45]. Future studies could consider collecting
409 oropharyngeal swab or saliva samples in addition to nasopharyngeal swabs for increased
410 sensitivity in adults.

411

412 ## Conclusion

413 We found high pneumococcal carriage prevalence in a PCV-naïve population living in an
414 IDP camp in Somaliland, consistent with carriage rates in non-displaced populations in other
415 high transmission settings. About half of all circulating pneumococci were included in
416 currently available PCVs. We estimate that at least half of all resulting IPD cases in this
417 population were caused by serotypes included in PCVs, indicating the potential for
418 substantial vaccine effects. Transmission was primarily driven by children 2-5 years and 6-
419 14 years old, partially exceeding the proposed age eligibility for PCV campaigns that aim to
420 temporarily reduce transmission in crisis-affected populations. These findings advance our
421 understanding of pneumococcal carriage in crisis-affected populations and provide important
422 evidence for the design of future vaccination strategies.

423

424 ## List of abbreviations

425 FPC: finite population correction

426 GE/ml: genome equivalents/ml

427 IDP: internally displaced people

428 IPD: invasive pneumococcal disease

429 MCRI: Murdoch Children's Research Institute

NESp: non-encapsulated serotype

NVT: non-vaccine serotype

PCM: phase-change material

PCV: pneumococcal conjugate vaccine

STGG: skim milk-tryptone-glucose-glycerol

ULT: ultra-low temperature

VT: vaccine serotype

Declarations

Ethics approval and consent to participate

Ethical approval for the study was granted by the Research Ethical Committee of the London School of Hygiene and Tropical Medicine (16577) and the Republic of Somaliland Ministry of Health Development (2/13075/2019).

Consent for publication

Not applicable.

Availability of data and materials

The anonymised aggregated data generated and analysed during the current study, including all analysis scripts, are available in the Zenodo repository https://doi.org/10.5281/zenodo.11493116

Competing interests

EKM and CS are investigators on a clinical research collaboration with Pfizer on PCV vaccination in Mongolia and are investigators on a Merck Investigator Studies Program grant funded by MSD on pneumococcal serotype epidemiology in children. All other authors report no competing interests.

Funding

This work was supported by Elrha's Research for Health in Humanitarian Crises (R2HC) Programme, which aims to improve health outcomes by strengthening the evidence base for public health interventions in humanitarian crises. The R2HC programme is funded by the UK Government (DFID), the Wellcome Trust, and the UK National Institute for Health Research (NIHR). SF acknowledges a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (grant: 208812/Z/17/Z). In addition, RME acknowledges an NIHR (grant: NIHR200908) for the Health Protection Research Unit in Modelling and Economics at LSHTM. MCRI was supported by the Victorian Government's Operational Infrastructure Support Program. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care. The funding sources had no role in the study design; collection, analysis, and interpretation of data; or in writing the report.

Author's contributions

Conceptualization and funding acquisition: FC, SF, EKM, and CS; methodology: FC, SF, CS, and KvZ; data collection implementation: KvZ, AIH, MOB, MSA, and SY; supervision: KvZ, AIH, MOB, SMS, RC, and MH; statistical analysis: KvZ; microbiological analysis: CLP, JH,

CS, and BDO; writing-original draft: KvZ; writing-review and editing: KvZ, CLP, RME, JH, RC, CRM, EKM, FC, and SF. All authors read and approved the final manuscript.

Acknowledgements

We sincerely thank the Digaale community who participated in this study, as well as Hamse Shaban Ahmed, Ayaan Ismail Adan, Khadar Abiib Ahmed, Ayaan Mohamoud Ali, Hamda Awil Garaad, Ahmed Yusuf Mohamed, Suaad Abdi Osman, Amiin Abdi Ismail, Nimco Abdilahi Ismail, AbdiFatah Mohamed Ahmed, Mustafe Shugri Mohamed, and Nimco Mohamed Abdi from the Republic of Somaliland Ministry of Health Development for their data collection efforts. We also thank the Somaliland National TB Reference Laboratory in Hargeisa for the storage of samples. Electronic data solutions were provided by LSHTM Open Data Kit (odk.lshtm.ac.uk).

Reference list

1. United Nations High Commissioner for Refugees. UNHCR - Refugee Statistics. https://www.unhcr.org/refugee-statistics/. Accessed 2 Mar 2023.

2. van Zandvoort K, Checchi F, Diggle E, Eggo RM, Gadroen K, Mulholland K, et al. Pneumococcal conjugate vaccine use during humanitarian crises. Vaccine. 2019;37:6787–92.

3. Heudtlass P, Speybroeck N, Guha-Sapir D. Excess mortality in refugees, internally displaced persons and resident populations in complex humanitarian emergencies (1998–2012) – insights from operational data. Conflict and Health. 2016;10:15.

4. Bellos A, Mulholland K, O'Brien KL, Qazi SA, Gayer M, Checchi F. The burden of acute respiratory infections in crisis-affected populations: a systematic review. Conflict and Health. 2010;4:3.

5. O'Brien KL, Baggett HC, Brooks WA, Feikin DR, Hammitt LL, Higdon MM, et al. Causes of severe pneumonia requiring hospital admission in children without HIV infection from Africa and Asia: the PERCH multi-country case-control study. The Lancet. 2019;394:757–79.

6. Kohrt BA, Mistry AS, Anand N, Beecroft B, Nuwayhid I. Health research in humanitarian crises: an urgent global imperative. BMJ Global Health. 2019;4:e001870.

7. Turner P, Turner C, Jankhot A, Helen N, Lee SJ, Day NP, et al. A Longitudinal Study of *Streptococcus pneumoniae* Carriage in a Cohort of Infants and Their Mothers on the Thailand-Myanmar Border. PLOS ONE. 2012;7:e38271.

8. Prequalified vaccines. WHO - Prequalification of Medical Products (IVDs, Medicines, Vaccines and Immunization Devices, Vector Control). 2019. https://extranet.who.int/pqweb/vaccines/prequalified-vaccines. Accessed 2 Mar 2023.

9. Pfizer. U.S. FDA Approves PREVNAR 20™, Pfizer's Pneumococcal 20-valent Conjugate Vaccine for Adults Ages 18 Years or Older | Pfizer. 2021. https://www.pfizer.com/news/press-release/press-release-detail/us-fda-approves-prevnar-20tm-pfizers-pneumococcal-20-valent. Accessed 14 Sep 2021.

10. Merck. Merck Announces U.S. FDA Approval of VAXNEUVANCE™ (Pneumococcal 15-valent Conjugate Vaccine) for the Prevention of Invasive Pneumococcal Disease in Adults 18 Years and Older Caused by 15 Serotypes. Merck.com. 2021. https://www.merck.com/news/merck-announces-u-s-fda-approval-of-vaxneuvance-pneumococcal-15-valent-conjugate-vaccine-for-the-prevention-of-invasive-pneumococcal-disease-in-adults-18-years-and-older-caused-by-15-serot/. Accessed 14 Sep 2021.

11. Bogaert D, De Groot R, Hermans PWM. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. Lancet Infect Dis. 2004;4:144–54.

12. van Zandvoort K, Bobe MO, Hassan AI, Abdi MI, Ahmed MS, Soleman SM, et al. Social contacts and other risk factors for respiratory infections among internally displaced people in Somaliland. Epidemics. 2022;41:100625.

13. FLOQSwabs | COPAN Diagnostics Inc. https://www.copanusa.com/sample-collection-transport-processing/floqswabs/. Accessed 2 Mar 2023.

14. Satzke C, Turner P, Virolainen-Julkunen A, Adrian PV, Antonio M, Hare KM, et al. Standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*: Updated recommendations from the World Health Organization Pneumococcal Carriage Working Group. Vaccine. 2013;32:165–79.

15. Pell CL, Ortika BD, Van Zandvoort K, Wee-Hee A, Mulholland EK, Checchi F, et al. Effect of transport conditions on recovery of *Streptococcus pneumoniae*. 2022.

16. THERMOCON CLASSIC 15.5 - THERMOCON. https://thermocon-coldchain.com/en/produkt/thermocon-classic-15-5/. Accessed 2 Mar 2023.

17. Dunne EM, Murad C, Sudigdoadi S, Fadlyana E, Tarigan R, Indriyani SAK, et al. Carriage of *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Moraxella catarrhalis*, and *Staphylococcus aureus* in Indonesian children: A cross-sectional study. PLOS ONE. 2018;13:e0195098.

18. Satzke C, Dunne EM, Choummanivong M, Ortika BD, Neal EFG, Pell CL, et al. Pneumococcal carriage in vaccine-eligible children and unvaccinated infants in Lao PDR two years following the introduction of the 13-valent pneumococcal conjugate vaccine. Vaccine. 2019;37:296–305.

19. Carvalho M da GS, Tondella ML, McCaustland K, Weidlich L, McGee L, Mayer LW, et al. Evaluation and Improvement of Real-Time PCR Assays Targeting *lytA*, *ply*, and *psaA* Genes for Detection of Pneumococcal DNA. Journal of Clinical Microbiology. 2007;45:2460–6.

20. Salter SJ, Hinds J, Gould KA, Lambertsen L, Hanage WP, Antonio M, et al. Variation at the capsule locus, cps, of mistyped and non-typable *Streptococcus pneumoniae* isolates. Microbiology (Reading). 2012;158 Pt 6:1560–9.

21. Phase 3 Study of 10-valent Pneumococcal Conjugate Vaccine (PNEUMOSIL) in Healthy Infants - No Study Results Posted - ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/results/NCT03197376. Accessed 24 Jun 2019.

22. Lumley T. Complex Surveys: A Guide to Analysis Using R. John Wiley & Sons; 2011.

23. Lumley T. survey: analysis of complex survey samples. 2020.

24. Løchen A, Truscott JE, Croucher NJ. Analysing pneumococcal invasiveness using Bayesian models of pathogen progression rates. PLoS Comput Biol. 2022;18:e1009389.

25. Qian G, Toizumi M, Clifford S, Le LT, Papastylianou T, Satzke C, et al. Association of pneumococcal carriage in infants with the risk of carriage among their contacts in Nha Trang, Vietnam: A nested cross-sectional survey. PLoS Med. 2022;19:e1004016.

26. Neal EFG, Chan J, Nguyen CD, Russell FM. Factors associated with pneumococcal nasopharyngeal carriage: A systematic review. PLOS Glob Public Health. 2022;2:e0000327.

27. le Polain de Waroux O, Flasche S, Kucharski AJ, Langendorf C, Ndazima D, Mwanga-Amumpaire J, et al. Identifying human encounters that shape the transmission of *Streptococcus pneumoniae* and other acute respiratory infections. Epidemics. 2018;25:72–9.

28. Neal EFG, Flasche S, Nguyen CD, Ratu FT, Dunne EM, Koyamaibole L, et al. Associations between ethnicity, social contact, and pneumococcal carriage three years post-PCV10 in Fiji. Vaccine. 2020;38:202–11.

29. Hill PC, Akisanya A, Sankareh K, Cheung YB, Saaka M, Lahai G, et al. Nasopharyngeal Carriage of *Streptococcus pneumoniae* in Gambian Villagers. Clin Infect Dis. 2006;43:673–9.

30. Nackers F, Cohuet S, le Polain de Waroux O, Langendorf C, Nyehangane D, Ndazima D, et al. Carriage prevalence and serotype distribution of *Streptococcus pneumoniae* prior to 10-valent pneumococcal vaccine introduction: A population-based cross-sectional study in South Western Uganda, 2014. Vaccine. 2017;35:5271–7.

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

574 32. Heinsbroek E, Tafatatha T, Phiri A, Swarthout TD, Alaerts M, Crampin AC, et al. Pneumococcal
575 carriage in households in Karonga District, Malawi, before and after introduction of 13-valent
576 pneumococcal conjugate vaccination. Vaccine. 2018;36:7369–76.

577 33. Clifford S, Knoll MD, O'Brien KL, Pollington TM, Moodley R, Prieto-Merino D, et al. Global
578 landscape of *Streptococcus pneumoniae* serotypes colonising healthy individuals worldwide before
579 vaccine introduction; a systematic review and meta-analysis. 2023;:2023.03.09.23287027.

580 34. Cohen R, Cohen JF, Chalumeau M, Levy C. Impact of pneumococcal conjugate vaccines for
581 children in high- and non–high-income countries. Expert Review of Vaccines. 2017;16:625–40.

582 35. Flasche S, Hoek AJV, Sheasby E, Waight P, Andrews N, Sheppard C, et al. Effect of
583 Pneumococcal Conjugate Vaccination on Serotype-Specific Carriage and Invasive Disease in
584 England: A Cross-Sectional Study. PLOS Medicine. 2011;8:e1001017.

585 36. Troeger C, Blacker B, Khalil IA, Rao PC, Cao J, Zimsen SRM, et al. Estimates of the global,
586 regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195
587 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The
588 Lancet Infectious Diseases. 2018;18:1191–210.

589 37. Murad C, Dunne EM, Sudigdoadi S, Fadlyana E, Tarigan R, Pell CL, et al. Pneumococcal
590 carriage, density, and co-colonization dynamics: A longitudinal study in Indonesian infants. Int J Infect
591 Dis. 2019;86:73–81.

592 38. Abdullahi O, Karani A, Tigoi CC, Mugo D, Kungu S, Wanjiru E, et al. The prevalence and risk
593 factors for pneumococcal colonization of the nasopharynx among children in Kilifi District, Kenya.
594 PLoS One. 2012;7:e30787.

595 39. Lupien A, Gingras H, Bergeron MG, Leprohon P, Ouellette M. Multiple mutations and increased
596 RNA expression in tetracycline-resistant *Streptococcus pneumoniae* as determined by genome-wide
597 DNA and mRNA sequencing. Journal of Antimicrobial Chemotherapy. 2015;70:1946–59.

598 40. Ousmane S, Diallo BA, Ouedraogo R. Genetic Determinants of Tetracycline Resistance in Clinical
599 *Streptococcus pneumoniae* Serotype 1 Isolates from Niger. Antibiotics. 2018;7:19.

600 41. Okitsu N, Kaieda S, Yano H, Nakano R, Hosaka Y, Okamoto R, et al. Characterization of *ermB*
601 Gene Transposition by Tn1545 and Tn917 in Macrolide-Resistant *Streptococcus pneumoniae*
602 Isolates. J Clin Microbiol. 2005;43:168–73.

603 42. Rolfe RJ, Shaikh H, Tillekeratne LG. Mass drug administration of antibacterials: weighing the
604 evidence regarding benefits and risks. Infectious Diseases of Poverty. 2022;11:77.

605 43. Rodrigues F, Foster D, Nicoli E, Trotter C, Vipond B, Muir P, et al. Relationships between rhinitis
606 symptoms, respiratory viral infections and nasopharyngeal colonization with *Streptococcus*
607 *pneumoniae*, *Haemophilus influenzae* and *Staphylococcus aureus* in children attending daycare.
608 Pediatr Infect Dis J. 2013;32:227–32.

609 44. Bojang A, Jafali J, Egere UE, Hill PC, Antonio M, Jeffries D, et al. Seasonality of Pneumococcal
610 Nasopharyngeal Carriage in Rural Gambia Determined within the Context of a Cluster Randomized
611 Pneumococcal Vaccine Trial. PLOS ONE. 2015;10:e0129649.

612 45. Miellet WR, Almeida ST, Trzciński K, Sá-Leão R. *Streptococcus pneumoniae* carriage studies in
613 adults: Importance, challenges, and key issues to consider when using quantitative PCR-based
614 approaches. Front Microbiol. 2023;14:1122276.

615

Additional files

Supplemental Material (pdf). Includes more details on data collection and matching datasets, temperature-controlled sample shipment, additional microbiological and statistical analyses, and invasive pneumococcal disease projections.

Supplemental Material for

Pre-vaccination carriage prevalence of *Streptococcus pneumoniae* serotypes among internally displaced people in Somaliland

Kevin van Zandvoort*, Abdirahman Ibrahim Hassan, Mohamed Omer Bobe, Casey L. Pell, Mohamed Saed Ahmed, Belinda D. Ortika, Saed Ibrahim, Mohamed Ismail Abdi, Mustapha A. Karim, Rosalind M. Eggo, Saleban Yousuf Ali, Jason Hinds, Saeed Mohamood Soleman, Rachael Cummings, Catherine R McGowan, E Kim Mulholland, Mohamed Abdi Hergeye, Catherine Satzke, Francesco Checchi, Stefan Flasche

* Corresponding author: Kevin van Zandvoort, Kevin.Van-Zandvoort@lshtm.ac.uk

Analysis scripts and anonymized aggregated data are available on GitHub: https://github.com/kevinvzandvoort/espicc-somaliland-digaale-survey-2019-carriage. These can be used to recreate all Tables and Figures in both the main manuscript and this Supplemental Material.

TABLE OF CONTENTS

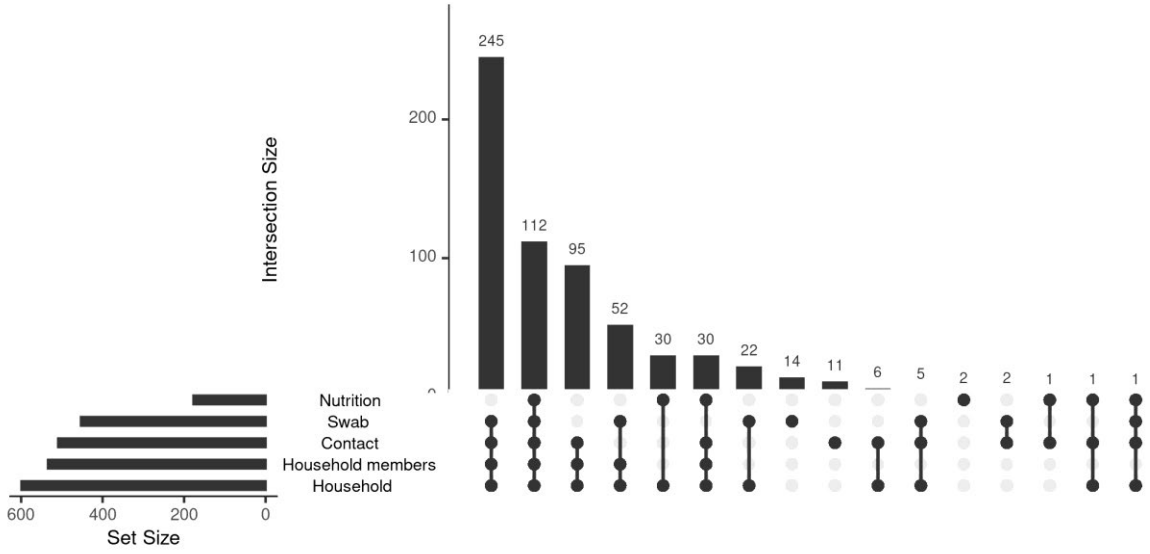
Section A. Data collection and matching datasets

Data was collected in multiple stages and using multiple Open Data Kit (ODK) forms. All shelters in Digaale IDP camp were visited. In shelters where at least one adult was present who consented to participation, a household survey was conducted using a household survey form which collected data on the household level (stored in the *household* dataset) and individual demographic data for all household members living in the household (stored in the *household members* dataset). Quota sampling was used to invite individual household members to participate in the contact survey, and data collectors returned to the household to conduct the contact survey two days after the household survey. For participants in the contact survey aged 6m to 59m, a nutritional assessment was conducted immediately after the contact survey. These data were collected on separate contact survey and nutrition forms, and stored in the *contact* and *nutrition* datasets, respectively. Finally, data collectors returned to all contact survey participants in the final two weeks of the survey, and asked them to have a nasopharyngeal swab taken. Swab-related data were collected using a swab form, and subsequently stored in the *swab* dataset. Additional household members who were not initially selected for participation in the contact survey were sampled for participation in nasopharyngeal swab selection (and a nutritional assessment, for age-eligible children).

Unique household and individual identifiers were automatically assigned in the household form, and manually re-entered in subsequent forms. Due to human error, some identifiers were incorrectly re-entered in subsequent forms, resulting in unmatched data between datasets for some records.

Supplemental Figure A1 shows an UpSet plot (1) with the number of matched records between multiple datasets. The majority (400; 88%) of records for collected swabs could be matched between all datasets for which forms were completed: i) 243 records with household, contact, and swab data for people age-ineligible for the nutritional assessment, ii)

112 records with household, contact, nutrition, and swab data for people age-eligible for the nutritional assessment, and iii) 45 records with household and swab data for people age-ineligible for the nutritional assessment who were not invited for the contact survey.



Supplemental Figure A1. Matching records between datasets.
UpSet plot showing the number of matching records between the different datasets, matches show the number of records that can be linked to the same person in the household member, contact, nutrition, and swab datasets, and to the same household in the household dataset. Data collected for households and household members that did not participate in an individual survey is not shown. Nutrition data is only collected for participants aged 6 to 59 months.

There were 53 records for collected swabs that could not be completely matched between all datasets for which forms were completed: i) there were two records for people age-eligible for the nutritional assessment, for who records could be linked to the household, contact, and swab, dataset, but not the nutrition dataset; ii) there were seven records for collected swabs that could be linked to household and swab data for people age-eligible for the nutritional assessment, that could not be linked to records in the nutrition dataset; iii) there were 22 records for collected swabs that could be linked to records for an individual household, but not to records for an individual in any of the other datasets, including household members; iv) 14 records for swabs could not be linked to any of the other datasets; v) 5 could be linked to an individual record in the contact dataset and to records for an individual household, but not to records for individual household members, and two of

these were age-eligible for the nutritional assessment; vi) there were two records that could only be linked to records in the contact dataset; and vii) there was one record that could be linked to all but the household members datasets.

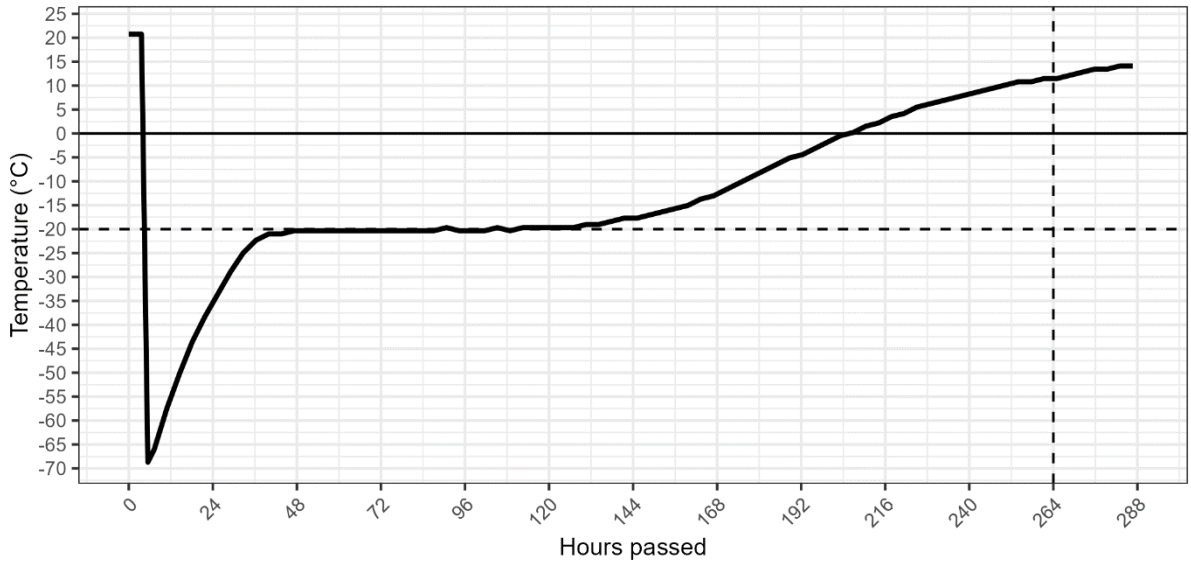
The inability to match certain records prevented the inclusion of data from some swabs in certain analyses. resulting in missing data for some covariates. Future surveys should aim to minimize challenges with matching identifiers.

Section B. Sample shipment

Passive temperature-controlled shipment

Due to logistical challenges, we were not able to ship collected samples on ultra-low-temperatures (ULT), which is the gold standard. Instead, samples were shipped using Thermocon Classic 15, a prequalified shipping solution utilizing phase change materials (PCMs), that provided passive cooling to keep temperatures at below -15°C for up to 96h as per manufacturer instructions (Schaumaplast, DE).

Manufacturer instructions state to precondition PCMs at -20°C prior to shipment. Instead, we tested preconditioning of PCMs at -70°C in a ULT freezer and measured the temperature of the shipping solution during an empty test-shipment of the shipping solution from London to Hargeisa in November 2020 using a Testo 184 T4 (Testo, DE) temperature logger (Supplemental Figure B1 and Supplemental Table B1). Temperatures were maintained at below -15°C for up to 160h during this test shipment. Temperatures rose above 0°C after 204h. We did not measure the ambient temperatures during this shipment.



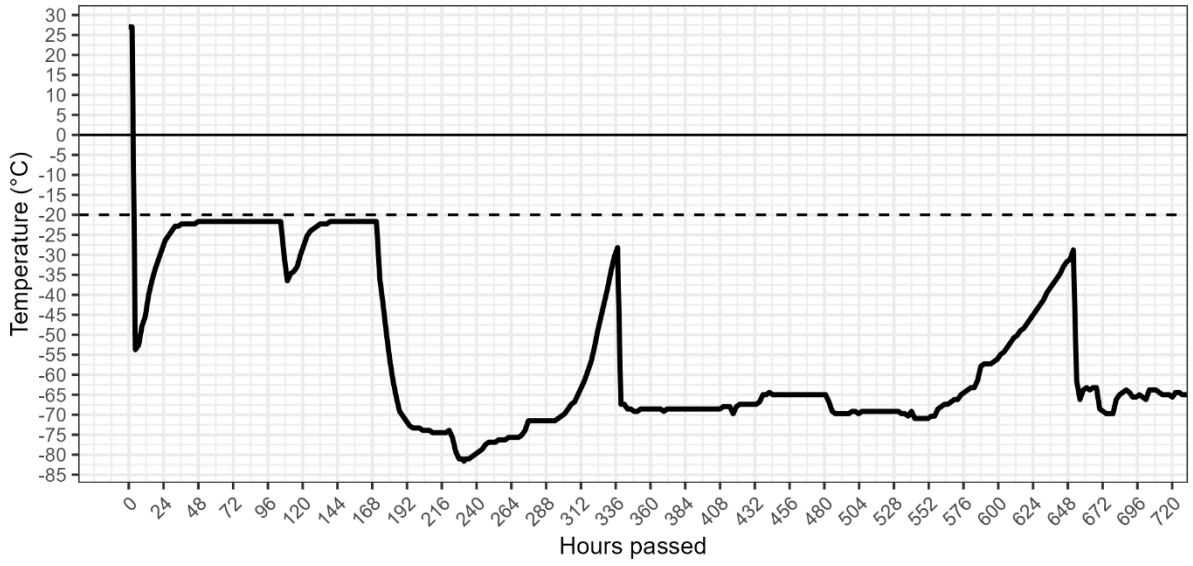
Supplemental Figure B1. Temperature of test shipment over time.
Temperature (in degrees Celsius) inside a Thermocon Classic 15 over time, during a shipment from London to Hargeisa. The dotted vertical line shows the temperature after 264 hours (11 days).

Supplemental Table B1. Time until temperature exceedance.

Temperature exceeded	Hours passed	Days passed
-20 °C	91	3.6
-15 °C	163	6.8
-10 °C	178	7.4
-5 °C	192	8.0
0 °C	207	8.6
5 °C	225	9.4
10 °C	250	10.4
15 °C	301	12.5

Temperature during sample shipments

We initially tested our shipment route using the PCMs with a pilot shipment of 81 samples in May 2021. Samples arrived in Nairobi, Kenya within 4 days, where they were placed in a -20°C freezer and shipped on dry ice after a further 3 days to the Murdoch Children's Research Institute (MCRI) in Melbourne, Australia for storage and analysis. Samples remained below -20°C for the entire duration of the pilot shipment (Supplemental Figure B2).



Supplemental Figure B2. Temperature of pilot shipment over time.
Temperature (in degrees Celsius) of a pilot shipment from Hargeisa, Somaliland to Melbourne, Australia transiting through Nairobi, Kenya.

We subsequently shipped all remaining samples in December 2021. This shipment got delayed in transit to Nairobi, and samples were only temperature controlled through the

passive cooling from PCMs for a total of 264h (11 days). No temperature data was available for this shipment. However, if temperatures during this shipment were similar as during the test shipment, this may indicate that the temperature of the swabs may have been >0°C for up to 2.5 days and may have risen to 12.5°C (Supplemental Figure B1). This extended transit, and potential rise in temperature, may have affected the viability of the samples. Pell et al. previously found that pneumococcal isolates at low density remained detectable when stored on flocked swabs in skim milk-tryptone-glucose glycerol medium (STGG) at -20°C for up to three weeks, though at reduced levels compared to the gold standard of ULT storage (2). They also found that isolates remained viable when stored in STGG at 4°C for up to 2 days, but that viability rapidly declined after >2 days.

Assessing carriage between shipments

To understand whether the extended period at which the 372 swabs transported during the second shipment were likely exposed to higher temperatures which may have affected the viability of these samples, we compared the pneumococcal carriage rates estimated from these swabs to those of the 81 samples shipped during the first pilot shipment. Sample-specific poststratification weights on age and gender were used to calculate population level prevalence estimates.

Overall pneumococcal carriage prevalence was similar in estimates from the two shipments (Supplemental Table B2), at 42.0% (95%CI 31.3 – 52.6) compared to 38.9% (34.3 – 43.6) in all ages, and 67.9% (50.8 – 84.9) compared to 72.6% (65.8 – 79.4) in children <5y. There are some difference in the proportion of VT-covered serotype carriers, at 55.9% (39.3 – 72.4) compared to 50.0% (42.3 – 57.7) for PNEUMOSIL-covered serotype carriers in all age-groups, and 36.8% (15.4 – 58.3) compared to 61.1% (52.4 – 69.8) in children <5y. Due to the small sample sizes, uncertainties around those estimates are wide, and sampling error cannot be ruled out to explain those differences.

We assessed whether there was any statistical evidence of a difference in pneumococcal carriage between the samples shipped in the two samples. There was no evidence of a difference in the odds of pneumococcal carriage for samples shipped during the pilot shipment compared to those shipped in the second shipment (OR 1.14; 95%CI 0.69 – 1.85; p=0.61), or of a difference in the odds of carrying PNEUMOSIL VTs in those who carried pneumococci (OR: 0.70; 0.09 – 14.44; p=0.76), as assessed using a logistic regression model. There also was no evidence of a difference in the mean log10 density in samples with pneumococci (-0.25; -0.64 – 0.15; p=0.22), as assessed using a linear regression model. Finally, there also was no evidence of a relative difference in the total number of serotypes identified (0.86; 0.61 – 1.19; p=0.371), in those carrying at least one serotype, as assessed using a Poisson regression model.

Supplemental Table B2. Carriage prevalence in pilot and second shipment.

Variable	Obs[a]	Sample value	Population est[b]		Population est (<5y)[b]	
Pilot shipment						
Overall carriage	34/81	42.0%	42.0%	31.3 - 52.6	67.9%	50.8 - 84.9
PNEUMOSIL-covered serotype carriers	15/34	44.1%	55.9%	39.3 - 72.4	36.8%	15.4 - 58.3
Synflorix-covered serotype carriers	16/34	47.1%	52.9%	36.3 - 69.6	36.8%	15.4 - 58.3
Prevenar 13-covered serotype carriers	19/34	55.9%	44.1%	27.6 - 60.7	31.6%	10.9 - 52.2
Vaxneuvance-covered serotype carriers	19/34	55.9%	44.1%	27.6 - 60.7	31.6%	10.9 - 52.2
Prevenar 20-covered serotype carriers	22/34	64.7%	35.3%	19.4 - 51.2	21.1%	2.9 - 39.2
Second shipment						
Pneumococcal carriage	144/370	38.9%	38.9%	34.3 - 43.6	72.6%	65.8 - 79.4
PNEUMOSIL-covered serotype carriers	72/144	50.0%	50.0%	42.3 - 57.7	61.1%	52.4 - 69.8
Synflorix-covered serotype carriers	70/144	48.6%	48.6%	41 - 56.3	60.0%	51.3 - 68.7
Prevenar 13-covered serotype carriers	81/144	56.2%	56.2%	48.7 - 63.8	63.3%	54.7 - 71.9
Vaxneuvance-covered serotype carriers	81/144	56.2%	56.2%	48.7 - 63.8	63.3%	54.7 - 71.9
Prevenar 20-covered serotype carriers	90/144	62.5%	62.5%	55.1 - 69.9	68.9%	60.6 - 77.1

Section C. Additional microbiological and statistical analyses

Multiple serotype carriage, density, and resistant genes

The median number of serotypes in samples in which pneumococci were detected by microarray was 1 (IQR 1 – 2). In 30% (95%CI 23 – 37) of these samples, more than one serotype was detected (Supplemental Table C1). The median density was 6.5 (IQR 5.4 – 7.0) $\log_{10}$ GE/ml for all samples. In 30% (21 – 41) of pneumococci, at least one resistance gene was detected. The most common resistant gene was *tetM*, present in 28% (19 – 39) of pneumococci. Samples in which more than one pneumococcal serotype, or in which other species were detected, were excluded from estimates of the proportion with resistant genes.

Supplemental Table C1. Other microbiological results.

Variable	Obs	Sample value	
Multiple serotype carriage			
Median number of serotypes[b]	176	1	1 - 2 (IQR)
Multiple serotype carriage	52/176	30%	23 - 37
Pneumococcal density (log10 GE/ml; median)[a]			
All samples	176	6.45	5.38 - 7.00 (IQR)
Samples with only 1 serotype	124	6.43	5.32 - 6.93 (IQR)
Samples with > 1 serotype	52	6.58	5.67 - 7.06 (IQR)
Per serotype	248	6.07	5.02 - 6.84 (IQR)
Other species detected by microarray			
All	63/191	33%	.
Resistant genes[c]			
Any	28/92	30%	21 - 41
Gene-specific			
aphA3	1/92	1%	0 - 6
cat	1/92	1%	0 - 6
ermB	8/92	9%	4 - 16
ermC	0/92	0%	0 - 4
mefA	2/92	2%	0 - 8
tetK	1/92	1%	0 - 6
tetL	0/92	0%	0 - 4
tetM	26/92	28%	19 - 39
tetO	0/92	0%	0 - 4
sat4	1/92	1%	0 - 6

a. Median and IQR of log10 transformed density estimates.
b. In those who carry pneumococci
c. Restricted to samples with only one pneumococci, and no other species detected - this has excluded all samples with NEPs

The odds that an individual serotype was the dominant serotype among all carried serotypes in a sample was 2.0 (1.1 – 3.7) times higher for VTs than for NVTs. The odds were similar when restricting the dataset to samples in which both VTs and NVTs were detected, these odds increased to 2.8 (1.0 – 8.0).

Supplemental Table C2. Association between serotype and dominant carriage.

Variable	OR	95% CI	p-value	N
Serotype is dominant (all participants)				
NVT	1.00			248
VT	2.02	1.10 - 3.73	0.020	248
Serotype is dominant (in those carrying both VT and NVT)				
NVT	1.00			65
VT	2.79	1.01 - 8.03	0.051	65

Association between risk factors and pneumococcal density

We used linear regression to assess the association between observed risk factors and the mean log10 density of pneumococcal serotypes (Supplemental Table C3). We found very weak evidence that living with one additional household member aged <5y was associated with a 0.18 (95%CI -0.01 – 0.37) increase in mean log10 density, and for an association between having a sore throat in the 2 weeks preceding the survey and a 0.40 reduction (0.00 – 0.79) in mean log10 density. There was some very weak evidence for a reduction in mean log10 density for pneumococci carried by children with improved weight-for-height z-scores, but no evidence for an association with any other potential risk factor.

Supplemental Table C3. Association between risk factors and pneumococcal density.

Variable	Mean difference[a,b]	95% CI	p-value	N[c]
Demographic characteristics				
Household size	-0.02	-0.10 - 0.05	0.533	172
Household members <5y	0.18	-0.01 - 0.37	0.064	172
Shelter quality				
House leakage				
no	*ref*			
yes	-0.19	-0.56 - 0.18	0.307	172
House draft				
no	*ref*			
yes	0.27	-0.07 - 0.61	0.116	172

Supplemental Table C3. Association between risk factors and pneumococcal density.

Variable	Mean difference[a,b]	95% CI	p-value	N[c]
Indoor air pollution				
Fuel firewood				
no	*ref*			
yes	0.02	-0.41 - 0.45	0.934	172
Fuel charcoal				
no	*ref*			
yes	-0.06	-0.39 - 0.27	0.713	172
Ventilation				
no	*ref*			
yes	0.23	-0.32 - 0.78		172
cook outside	0.08	-0.41 - 0.57	0.628	172
Current health[d]				
Antibiotic use				
no	*ref*			
yes	0.09	-0.23 - 0.42	0.585	163
Respiratory symptoms				
no	*ref*			
yes	-0.10	-0.46 - 0.26	0.587	164
Cough				
no	*ref*			
yes	0.05	-0.27 - 0.36	0.778	164
Sore throat				
no	*ref*			
yes	-0.40	-0.79 - 0.00	0.050	164
Headache				
no	*ref*			
yes	-0.17	-0.63 - 0.28	0.460	164
Fever				
no	*ref*			
yes	-0.17	-0.50 - 0.17	0.340	164
Diarrhoea				
no	*ref*			
yes	-0.06	-0.50 - 0.38	0.793	164
Morbidities[e]				
Pneumonia 6m[f]				
no	*ref*			
yes	0.12	-0.26 - 0.51	0.528	152
Sickle Cell				
no	*ref*			
yes	0.10	-0.76 - 0.96	0.826	152
Asthma				
no	*ref*			
yes	1.49	-0.02 - 3.01	0.056	152
Diabetes				
no	*ref*			
yes	1.53	-0.61 - 3.67	0.163	152

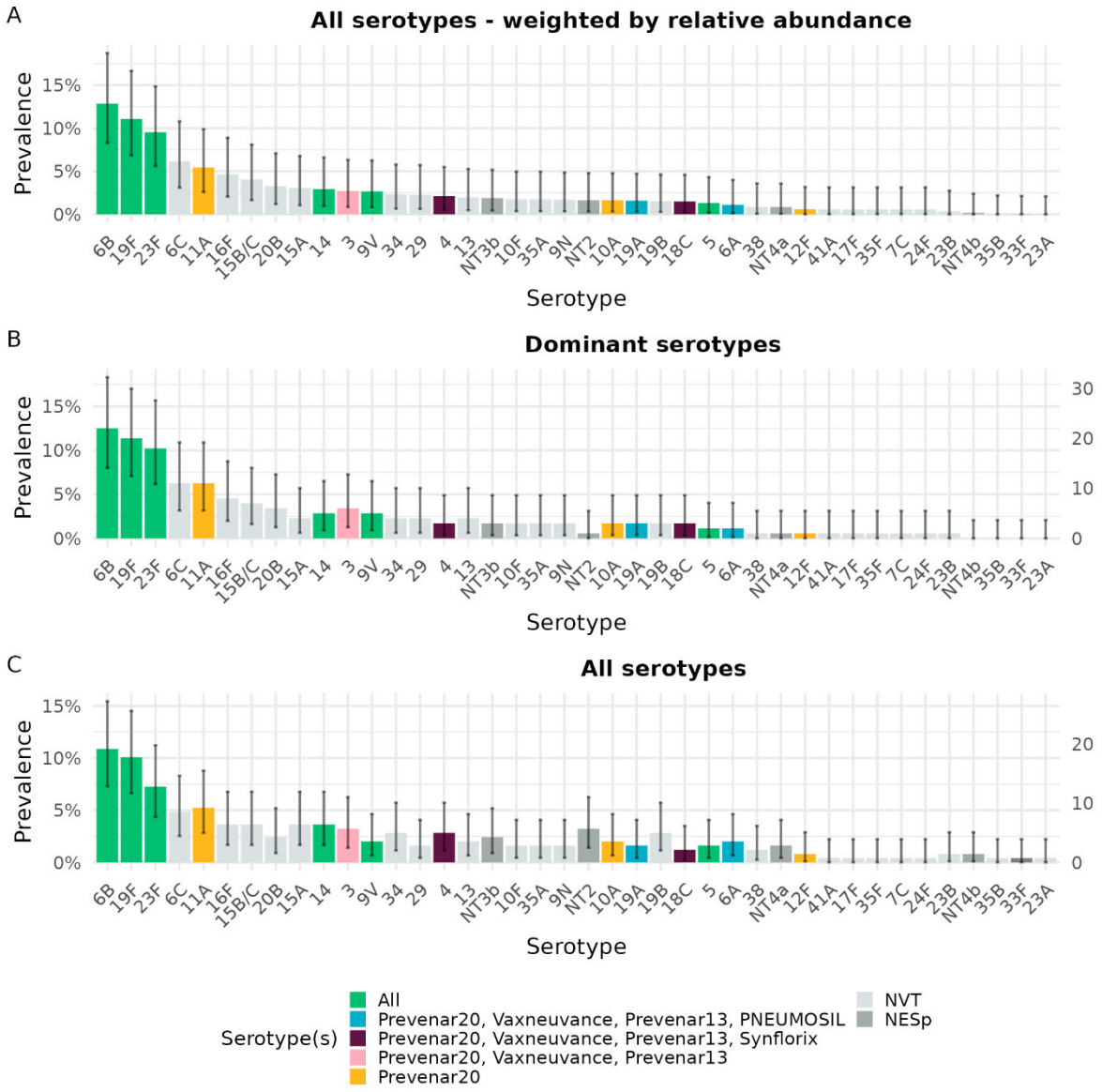
Supplemental Table C3. Association between risk factors and pneumococcal density.

Variable	Mean difference[a,b]	95% CI	p-value	N[c]
Individual substance use				
Tobacco				
no	*ref*			
yes	-0.20	-2.37 - 1.98	0.861	152
Khat				
no	*ref*			
yes	-0.20	-2.37 - 1.98	0.861	152
Household substance use[g]				
Household smoke				
no	*ref*			
yes	-0.09	-0.43 - 0.25	0.602	172
Household snuff				
no	*ref*			
yes	0.02	-0.54 - 0.58	0.949	172
Household khat				
no	*ref*			
yes	-0.26	-0.58 - 0.06	0.107	172
Contact behaviour				
Total number of direct contacts	-0.03	-0.08 - 0.02	0.303	151
Total number of physical contacts	-0.01	-0.06 - 0.04	0.602	151
Malnutrition in <5y				
Weight-for-age z-score	0.10	-0.12 - 0.33	0.358	81
Weight-for-height z-score	0.16	-0.02 - 0.35	0.088	81
Height-for-age z-score	-0.05	-0.23 - 0.12	0.556	81
MUAC[h]-for-age z-score	0.01	-0.23 - 0.25	0.933	81
MUAC [h] (in cm)	0.02	-0.18 - 0.23	0.823	81

a. Density estimates were log10 transformed prior to linear regression.
b. Estimates are adjusted for age and gender.
c. Total number of records used in regression.
d. Self-reported antibiotic use and symptoms in 2 weeks preceding the survey.
e. Self-reported diagnosed morbidities.
f. Pneumonia diagnosis in the 6m preceding the survey.
g. Substance use by at least one household member.
h. Middle Upper Arm Circumference

Impact of weighting on serotype distribution

In our main analysis, we weighted the pneumococcal serotype distribution by their relative abundance. We assessed the sensitivity to this approach by comparing it to the serotype distribution of only dominant serotypes, and all serotypes (unweighted). There were no major differences between the three distributions (Supplemental Figure C1), with serotypes 6B, 19F, and 23F as the most common serotypes in all three.

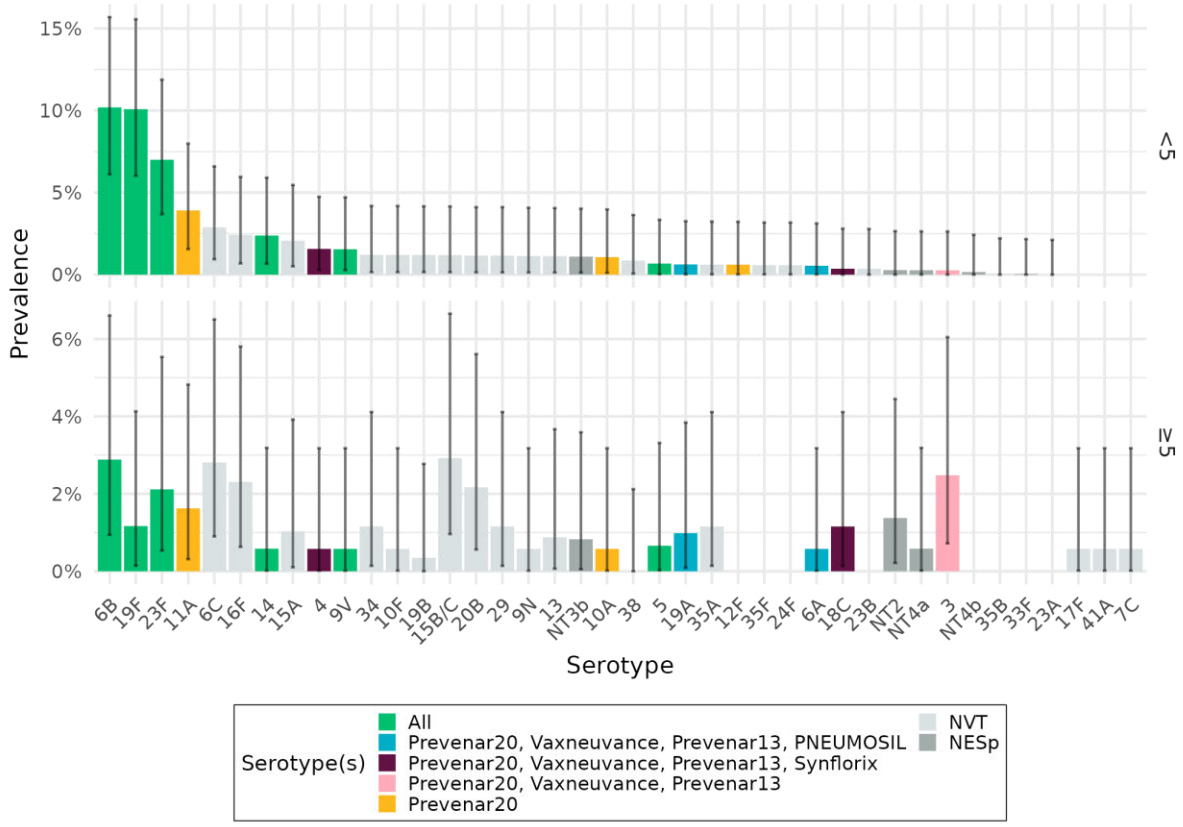


Supplemental Figure C1. Pneumococcal serotype distribution.
Bars show the proportion of serotypes among all identified pneumococci in Digaale IDP camp. A: serotype distribution of all serotypes, weighted by their relative abundance of carriage; B: serotype distribution of dominant serotypes only; C: unweighted serotype distribution of all serotypes. Coloured bars show the serotypes included in the three licensed PCVs, dark grey bars non-encapsulated pneumococci, and light grey bars other serotypes not included in the PCVs. Error bars show 95% confidence intervals for each estimate.

Serotype distribution by age

We also stratified the serotype distribution by age, in those carried by children <5y, and by people aged ≥5y (Supplemental Figure C2). The serotype distribution differed substantially between age groups, although confidence intervals around the point estimates are wide. In children <5y, the five most common serotypes were 6B (10%; 95%CI 6 – 16), 19F (10%; 6 –

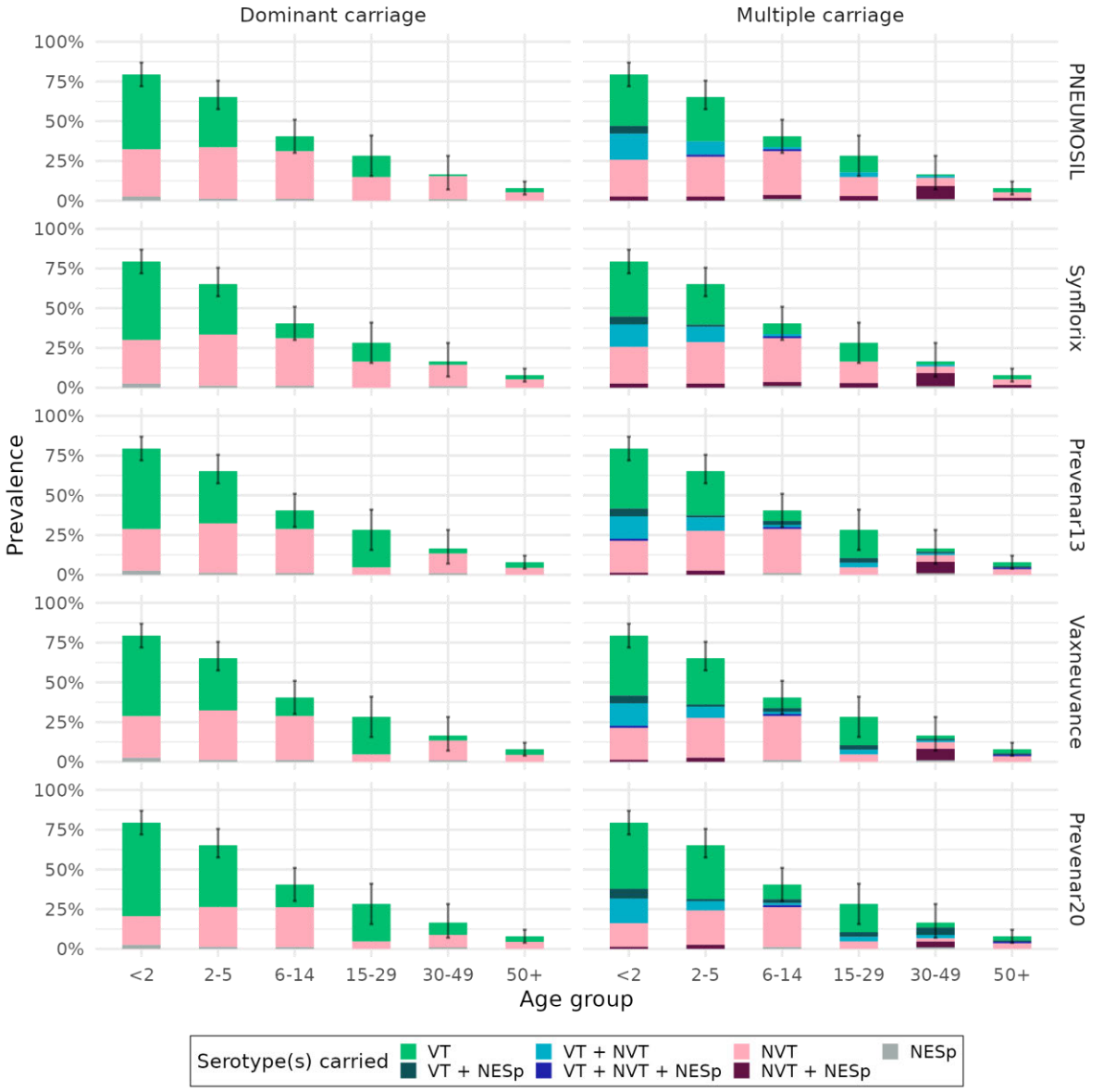
16), 23F (7%; 4 – 12), 11A (4%; 2 – 8), and 6C (3%; 1 – 7). In contrast, the relative proportion of these serotypes substantially decreased in people ≥5y, with a much more uniform distribution. In these older people, the five most common serotypes were 15B/C (3%; 1 – 7), 6B (3% 1 – 7), 6C (3%, 1 – 7), 3 (2%, 1 – 6), and 16F (2% 1 – 6).



Supplemental Figure C2. Pneumococcal serotype distribution by age.
Bars show the proportion of serotypes among all identified pneumococci in Digaale IDP camp, weighted by their relative abundance, in participants aged <5y (A), and ≥5y (B). Coloured bars show the serotypes included in the three licensed PCVs, dark grey bars non-encapsulated pneumococci, and light grey bars other serotypes not included in the PCVs. Error bars show 95% confidence intervals for each estimate.

Carriage prevalence by age

Supplemental Figure C3 shows overall pneumococcal prevalence, and prevalence of VTs, NVTs, and NESp by age, for different PCV formulations. We also show the distributions for dominant serotype carriage only.



Supplemental Figure C3. Prevalence and serotype distribution by age.
Facet columns show the serotype distribution of dominant serotypes only, and of all serotypes in multiple serotype carriers. Facet rows use different definitions for vaccine types: PNEUMOSIL, Synflorix, Prevenar 13, Vaxneuvance, Prevenar 20. Bars show the estimated prevalence of pneumococcal serotypes by age group, weighted for age and gender. Error bars show 95% confidence intervals around overall pneumococcal carriage prevalence. Colours show the prevalence of serotypes that are carried; VT: only vaccine type(s), NVT: only non-vaccine type(s); NESp: only non-encapsulated type(s); and where applicable VT + NVT: both vaccine- and non-vaccine type(s). Multiple carriage with non-encapsulated type(s) is shown as a darker shading of the grouping without NESp.

Assessing the choice of post-stratification weights on prevalence estimates

As participants were not selected using a stratified random sampling design, we post-stratified our results to calculate population-level estimates. Data were stratified by gender

(male or female) and the following age groups: <2, 2-5, 6-14, 15-29, 30-49, 50+ years of age, and poststratification was implemented using the *Survey* package in R (3). Prevalence estimates in our main manuscript are weighted by the calculated poststratification weights on age and gender, unless otherwise stated.

We assessed the sensitivity to the choice of variables used to weight the data by comparing prevalence estimates weighted by the variables listed in Supplemental Table C4.

Supplemental Table C4. Post-stratification weights used to calculate population-level estimates.

Weights	Age group [a]	Gender [b]	Household size [c]
I [d]	✓	✓	×
II	✓	×	✓
III	✓	×	×
IV	×	×	×

Variables used in calculating the weights used to construct contact intensities

a. Categorised as <2, 2-5, 6-14, 15-29, 30-49, 50+ years of age
b. Female or male
c. Categorised by quantiles: 1-2, 3-4, 5-6, and 7-12 household members
d. Main weights used in the analysis

There was little difference in the estimated pneumococcal prevalence by different post-stratification weights. The unweighted sample estimate of 39.5% (35.3 – 43.6) resulted in a estimated prevalence after applying post-stratification weights ranging between 35.0% – 35.8%, while the unweighted sample estimate of 71.7% in children <5y did not substantially change, with weighted estimates ranging between 70.0% – 70.4%.

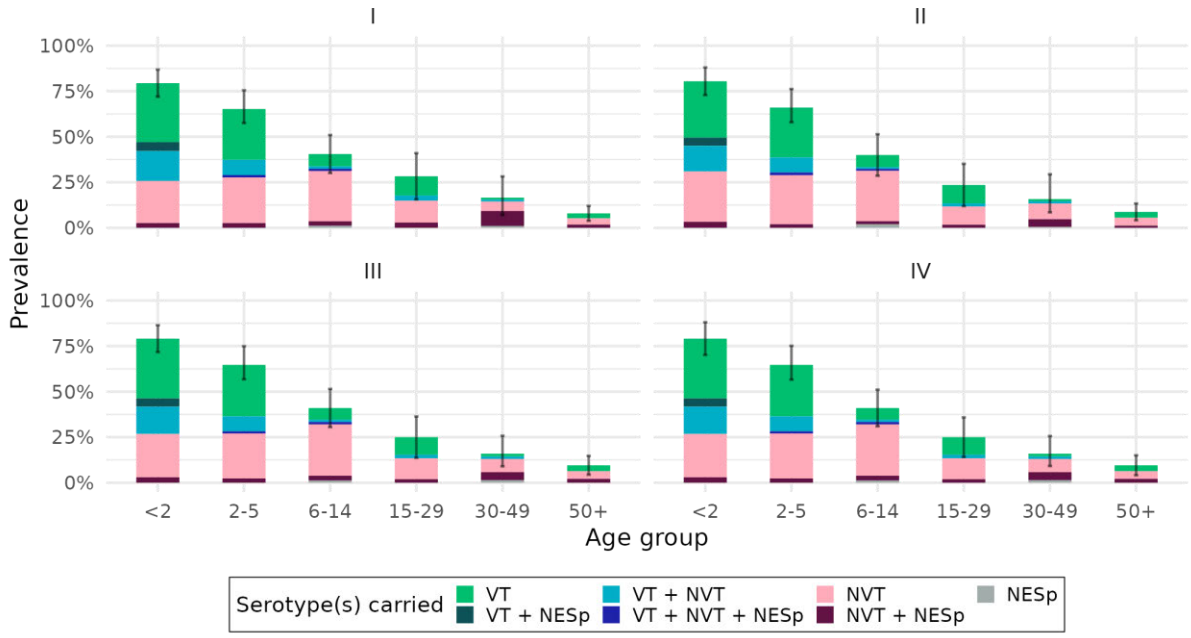
Supplemental Table C5. Pneumococcal prevalence estimates by different post-stratification weights.

Weights	Obs[a]	Sample value	Population est		Population est (<5y)	
I	175/445	39.3%	35.8%	31.1 - 40.4	70.0%	63.7 - 76.4
II	174/443	39.3%	35.4%	30.8 - 40.1	70.4%	64.0 - 76.8
III	175/445	39.3%	35.0%	30.6 - 39.4	70.0%	63.5 - 76.5
IV	178/451	39.5%	39.5%[b]	35.3 - 43.6	71.7% [b]	65.8 - 77.6

a. Number of observations included in sample. Observations with missing values for variables used in weighting are excluded from the dataset.
b. Population estimate is the same as the sample estimate, as estimates are unweighted.

There were also no substantial differences in the serotype distribution by age using different weights. The most prominent difference was an increase in the proportion of NVTs and NTs

co-carried in people aged 30-49 when weighting on age and gender, but overall prevalence in this age groups is small.

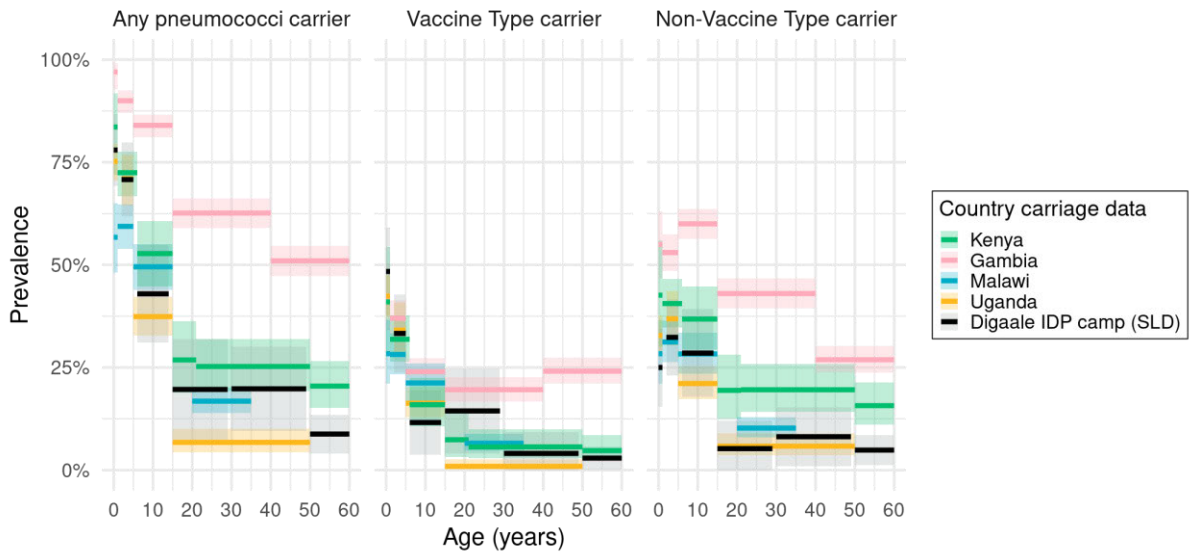


Supplemental Figure C4. Prevalence and serotype distribution by age using different weights. Facets show the serotype distribution weighted by I) Age and gender, II) Age and household size, III) Age only, and IV) unweighted. Bars show the estimated prevalence of pneumococcal serotypes by age group, weighted for age and gender. Error bars show 95% confidence intervals around overall pneumococcal carriage prevalence. Colours show the prevalence of serotypes that are carried; VT: only vaccine type(s), NVT: only non-vaccine type(s); NT: only non-encapsulated type(s); and where applicable VT + NVT: both vaccine- and non-vaccine type(s). Multiple carriage with non-encapsulated type(s) is shown as a darker shading.

Carriage prevalence compared to other settings

We compared overall carriage prevalence and VT and NVT carriage prevalence by age with that observed in Kilifi County in Kenya, Brikama Local Government Area in the Gambia, Karonga District in Malawi, and Sheema North sub-district in Uganda (4–7). Overall carriage prevalence by age is very similar to carriage prevalence in Kenya, Malawi, and Uganda, but lower than prevalence reported in Gambia, for all age groups. A similar pattern is observed for VTs, though NVT prevalence is more similar to Malawi and Uganda than for Kenya, where it was higher. Note microbiological techniques and VT definitions (as PNEUMOSIL carriage in our Digaale dataset, as Prevenar-13 carriage in the Malawi dataset, as PCV7

carriage in the Gambian dataset, and as Synflorix carriage in the Kenyan and Ugandan datasets) differed between studies, which limits comparability.



Supplemental Figure C5. Prevalence by age compared to different settings.
Pneumococcal prevalence by age in Digaale IDP camp (black), compared to prevalence observed in Kenya (green), the Gambia (pink), Malawi (blue), and Uganda (yellow). Prevenar 13 serotypes are used to define VTs. Thick lines show the age-specific carriage prevalence, and shaded areas their associated 95% binomial confidence interval. Post-stratification weights on age and gender are applied to the Digaale estimates. Studies used different age categorisations: these are shown as the width of each estimate. Facet columns show overall pneumococcal carriage prevalence, prevalence of vaccine types only, and prevalence of non-vaccine types.

Contribution of different age groups towards the age-specific exposure to pneumococcus

We estimated the contribution of different age groups towards the age-specific exposure to pneumococcus by combining the estimated contact matrices (8) with prevalence estimates by age (Figure 4 in main manuscript). We assessed the uncertainty around these estimates by taking 1,000 bootstrap samples of the dataset to calculate carriage prevalence and contact matrices, and calculated bootstrapped confidence interval as the 2.5% and 97.5% quantiles of estimated values over all datasets (Supplemental Table C6).

Supplemental Table C6. The contribution of different age groups towards the age-specific exposure to pneumococcus.

Contactee age group	Contactor age group					
	<2	*2-5*	*6-14*	*15-29*	*30-49*	*50+*
<2	7.5% (4.1 - 12.4)	10.1% (7.8 - 12.9)	5.5% (3.8 - 7.6)	7.6% (5.3 - 10.7)	9.9% (7.7 - 12.7)	6.3% (4.3 - 8.8)
2-5	39.2% (31.3 - 47.9)	44.5% (37.2 - 51.9)	25.6% (20.0 - 32.3)	14.5% (10.4 - 19.9)	21.8% (16.5 - 27.5)	19.6% (14.6 - 26.0)
6-14	24.9% (17.4 - 33.6)	30.1% (22.6 - 37.6)	51.6% (41.5 - 60.4)	22.4% (15.5 - 31.6)	23.9% (17.2 - 32)	26.0% (17.9 - 36.2)
15-29	14.0% (7.7 - 21.6)	7.0% (3.8 - 10.8)	9.3% (5.2 - 14.4)	39.7% (26.1 - 52)	19.1% (11.9 - 27.7)	18.0% (11.0 - 26.9)
30-49	11.7% (5.7 - 18.7)	6.6% (3.1 - 10.8)	6.2% (2.9 - 10.1)	12.1% (6.0 - 20.0)	20.0% (10.2 - 30.2)	19.7% (9.9 - 29.5)
50+	1.8% (0.7 - 3.3)	1.4% (0.6 - 2.6)	1.6% (0.6 - 3.1)	2.7% (1.2 - 4.9)	4.7% (2.0 - 8.0)	9.4% (4.0 - 15.6)

Values denote mean and bootstrapped 95% confidence intervals over 1,000 bootstrap samples of the estimated proportion of all contacts made by a contactor of age *j* (columns), that are with contactees that are carrying pneumococci of age *i* (rows).

Section D. Invasive pneumococcal disease projections

We estimated the likely proportion of invasive pneumococcal disease in the population in Digaale covered by different PCV products by applying serotype specific estimates of invasiveness to our observed carriage estimates.

Age- and serotype-specific invasiveness

Løchen et al (9) estimated the progression rate from carriage to invasive disease as the number of cases per carrier per year in a meta-analysis of several *Streptococcus pneumoniae* datasets using their *progressionEstimation* RStan package. They estimated progression rates separately for children and adults.

We took their reported median and 95% credible intervals values from invasiveness estimates for serotypes in children and adults, and fitted lognormal distributions to each set of values in order to recover their posterior distributions.

For each serotype *s*, in each dataset *a*, we assume that the invasiveness values can be described by a lognormal distribution with mean $\mu_{s,a}$ and standard deviation $\sigma_{s,a}$ (on the log-scale).

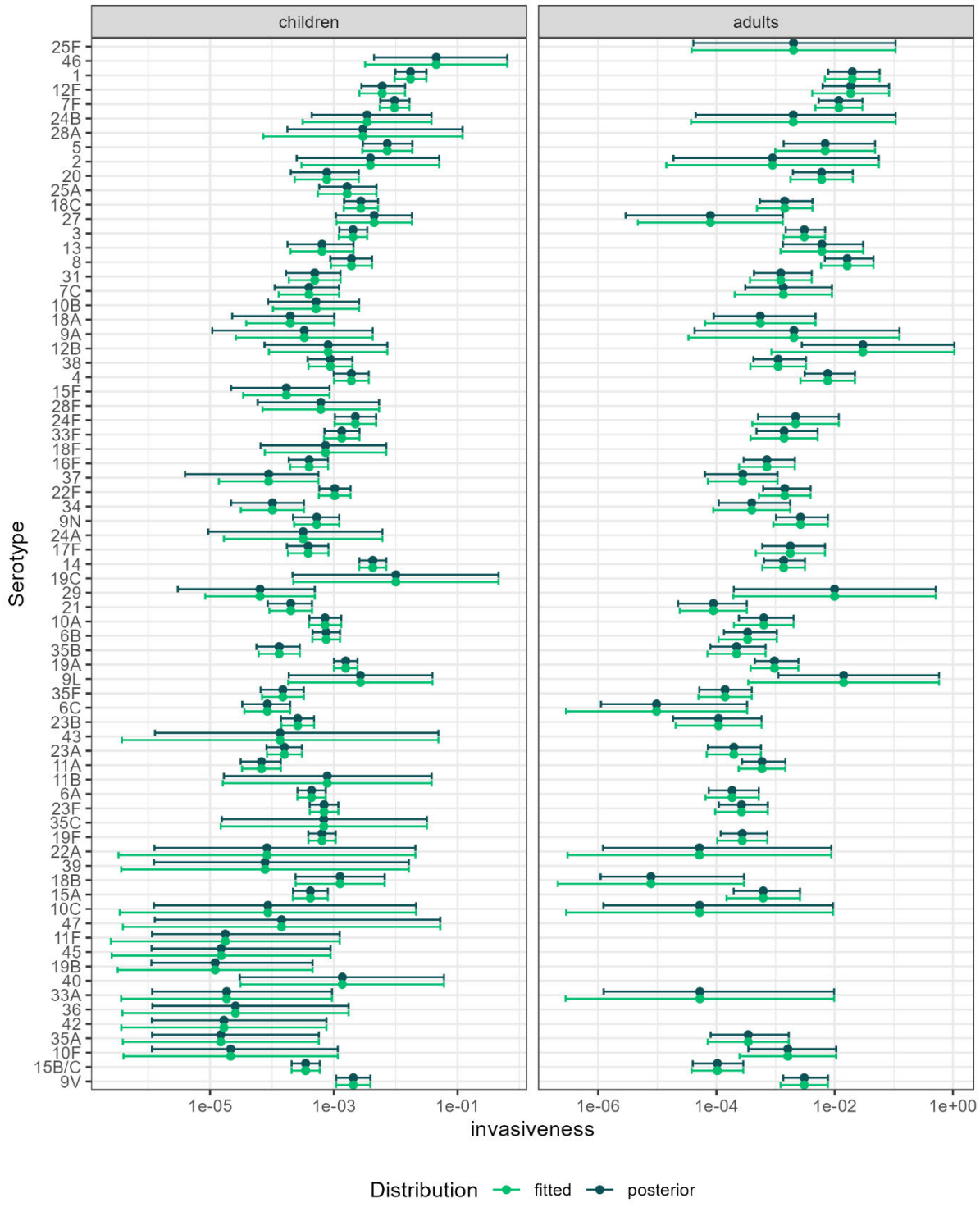
$$v_{s,a} \sim \text{lognormal}(\mu_{s,a}, \sigma_{s,a})$$

As the median value of any lognormal distribution is equivalent to e^{μ}, we parameterized the log-mean for $v_{s,a}$ as $\mu_{s,a} = \log y_{s,a,0.5}$, where $y_{s,a,0.5}$ is the reported median estimate for serotype *s* in dataset *a.* We then optimized the value for $\sigma_{s,a}$ using the *optimize* function in base *R* to minimize $g_{(\sigma_{s,a})}$: the log-least squares of the difference between the reported boundaries of the 95% credible interval for $v_{s,a}$, $y_{s,a,0.025}$ and $y_{s,a,0.975}$, and corresponding quantile distribution for the lognormal distribution evaluated at 0.025 and 0.975. We took the log of the values to minimize the large difference in scale between the lower and upper boundary value.

$$g(\sigma_{s,a}) = \log\left(\left(y_{s,a,0.025} - Q_{s,a}(0.025)\right)^2 + \left(y_{s,a,0.975} - Q_{s,a}(0.975)\right)^2\right)$$

Where $Q_{s,a}(p)$ is the quantile distribution of the lognormal distribution used to describe $v_{s,a}$ evaluated at *p.*

Supplemental Figure D1 shows the fitted lognormal distributions against the reported values estimated by Løchen et al. The 95% quantile values of the fitted distributions were in broad agreement with the reported 95% credible intervals, with only some minor discrepancies at primarily the lower tails of some distributions.



Supplemental Figure D1. Age and serotype specific invasiveness values. Dark-green values show the median and 95% credible intervals reported by Løchen et al (9). Light-green values show the median and 95% quantile values from the refitted lognormal distributions.

Estimating invasiveness in Digaale

We applied the invasiveness estimates to our carriage estimates by generating 10,000 bootstrap samples from our dataset where we resampled participants with replacement. In each bootstrap sample, we sampled age- and serotype-specific invasiveness values from the fitted invasiveness distributions, and applied these values to carriers of that serotype in the bootstrapped dataset. Poststratification weights were recalculated in each bootstrap dataset to ensure representativity of population-level estimates.

We assumed the same invasiveness value for any carried pneumococci of the same serotype, regardless of their abundance or the number of other serotypes carried by an individual. We applied invasiveness values sampled from invasiveness distributions for children to serotypes carried by individuals aged <18y, and values from distributions for adults to all other serotypes. We ignored any invasive disease caused by non-encapsulated serotypes, by assuming an invasiveness value of 0.

The Løchen et al dataset provided invasiveness values for 33/34 encapsulated serotypes identified in children, and for 17/20 encapsulated serotypes identified in adults. When available, we replaced any missing values by the average invasiveness of serotypes in the same age and serogroup. Values for serotype 20B in both children and adults were replaced with age-specific estimated values for serotype 20. Values for serotype 19B in adults were replaced by the mean values for serotypes 19A and 19F. Values for serotype 41A in adults were replaced by the median of all adult invasiveness values, as no estimate was available for any serotype in serogroup 41.

The total number of IPD cases expected within one year in bootstrap sample *i* excluding those in PCV product *p* was calculated as:

$$d_{i,p} = \sum_{x=1}^{N} \left(w_{i,x} \sum_{s=1}^{S} v_{i,s,a} I_x(s) \left(1 - V_p(s)\right) \right)$$

Where *N* is the total number of individuals in the dataset, $w_{i,x}$ is the post-stratification weight calculated for individual *x* in bootstrap sample *i*, *S* is the total number of unique serotypes identified in Digaale, $v_{i,s,a}$is the invasiveness value sampled for serotype s in bootstrap sample i, for individuals of age *a*, *a* is the index of the age-group for individual x (<18y or ≥18y), $I_x(s)$ is an indicator function that returns 1 if individual x carries serotype s, and 0 otherwise, and $V_p(s)$ is an indicator function that returns 0 if serotype *s* is not included in PCV product *p*, and 1 if it is included. V_0 denotes no vaccination, and returns 0 for all serotypes *s*. for serotype s and individual x, which returns 1 if individual estimate for serotype.

For each vaccine PNEUMOSIL, Synflorix, Prevenar 13, Vaxneuvance, and Prevenar 20, we calculated the proportion of the total invasiveness i) unweighted in the sampled dataset and ii) at the population level by applying post-stratification weights in all ages, those <5y, and those in age groups <2y, 2-5y, 6-14y, 15-29y, 30-49y, and 50+y.

The proportion of total invasiveness caused by serotypes covered by vaccine product *p* in bootstrap sample *i* was then calculated as $1 - \frac{d_{i,p}}{d_{i,0}}$, where $d_{i,p}$ are the total number of IPD cases excluding those caused by serotypes included in vaccine product *p*, and $d_{i,0}$ are all IPD cases including those caused by serotypes in vaccine product *p*.

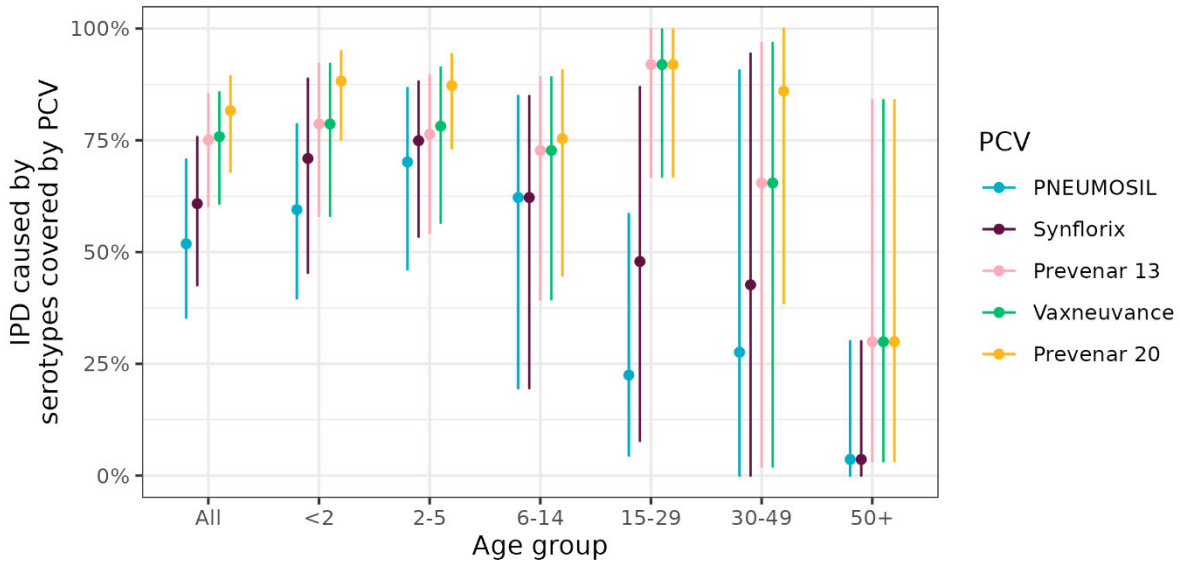
We reported the median and 95% quantile values of that proportion across all bootstrap samples.

Results

We estimate that between 52% (95%CI 35 – 70) of all IPD cases were caused by serotypes included in the PNEUMOSIL vaccine. The estimated proportion for serotypes covered by Synflorix serotypes was slightly higher (60%; 42 – 75), with higher proportions for serotypes included in higher valency vaccines (Prevenar 13: 75%, 60 – 85; Vaxneuvance: 76%, 61 –

86; and Prevenar 20: 82%, 67 – 90) (Supplemental Figure D2 and Supplemental Table D1). While the proportion tended to be the lowest for serotypes included in the PNEUMOSIL vaccine and the highest for serotypes included in the Prevenar 20 vaccine in all age groups, this difference was especially apparent in estimated IPD in people aged 15y and older.

These results indicate that a substantial proportion of current IPD is likely caused by serotypes included in any of the PCVs. We caution against overinterpretation of these results, as no data is available to validate these estimates, and uncertainty around the estimates is wide. These estimates are not estimating the potential impact of PCVs, which depend on a multitude of factors including their vaccine coverage, serotype-specific vaccine efficacy, immunological cross-reactivity with uncovered serotypes, and indirect effects including serotype replacement of VTs with NVTs.



Supplemental Figure D2. Estimated proportion of IPD cases caused by serotypes covered by PCVs. Points shows the median and lines the 95% uncertainty interval estimated from 10,000 bootstrapped samples, for different PCVs, by age group.

Supplemental Table D1. Proportion of current IPD covered by PCVs

Age group	PNEUMOSIL[a]	Synflorix[b]	Prevenar 13[c]	Vaxneuvance[d]	Prevenar 20[e]
All	51.9% (34.8 - 69.1)	60.1% (41.8 - 75)	74.9% (60.0 - 85.3)	75.5% (60.8 - 85.8)	81.6% (66.9 - 89.9)
<2	59.4% (39.8 - 78.1)	69.9% (47.8 - 88.5)	78.4% (60.0 - 91.6)	78.4% (60.0 - 91.6)	88.5% (76.1 - 95.3)
2-5	70.2% (46.3 - 86.2)	74.8% (52.4 - 88.1)	76.4% (53.4 - 89.5)	78.2% (56.6 - 91.1)	87.6% (73.2 - 94.7)
6-14	60.9% (19.8 - 83.9)	60.9% (19.8 - 83.9)	72.7% (39.5 - 89.4)	72.7% (39.5 - 89.4)	74.9% (41.2 - 90.4)
15-29	22.0% (4.7 - 63.2)	45.3% (6.1 - 86.6)	92.1% (63.0 - 99.4)	92.1% (63 - 99.4)	92.1% (63.0 - 99.4)
30-49	30.2% (0.0 - 92.1)	43.9% (0.0 - 94.3)	66.1% (3.5 - 97.6)	66.1% (3.5 - 97.6)	86.0% (39.3 - 99.9)
50+	3.7% (0.0 - 31.6)	3.7% (0.0 - 31.6)	29.1% (2.9 - 82.5)	29.1% (2.9 - 82.5)	29.1% (2.9 - 82.5)

Median estimate and 95% quantile values of proportion of current IPD cases that are caused by serotypes covered by PCVs, across 10,000 bootstrap samples of participant datasets and age-and serotype specific invasiveness estimates.

a. Serotypes 1, 5, 6A, 6B, 7F, 9V, 14, 19A, 19F, and 23F.
b. Serotypes 1, 4, 5, 6B, 7F, 9V, 14, 18C, 19F, and 23F.
c. Serotypes 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, and 23F.
d. Serotypes 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, 22F, 23F, and 33F.
e. Serotypes 1, 3, 4, 5, 6A, 6B, 7F, 8, 9V, 10A, 11A, 12F, 14, 15B, 18C, 19A, 19F, 22F, 23F, and 33F.

References used in Supplemental Material

1. Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: Visualization of Intersecting Sets. IEEE Transactions on Visualization and Computer Graphics. 2014 Dec;20(12):1983–92.

2. Pell CL, Ortika BD, Van Zandvoort K, Wee-Hee A, Mulholland EK, Checchi F, et al. Effect of transport conditions on recovery of *Streptococcus pneumoniae*. Abstract presented at: The 12th International Symposium on Pneumococci and Pneumococcal Diseases (ISPPD); 2022 Jun; Toronto, Canada.

3. Lumley T. survey: analysis of complex survey samples. 2020.

4. Hammitt LL, Akech DO, Morpeth SC, Karani A, Kihuha N, Nyongesa S, et al. Population effect of 10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of *Streptococcus pneumoniae* and non-typeable *Haemophilus influenzae* in Kilifi, Kenya: findings from cross-sectional carriage studies. The Lancet Global Health. 2014 Jul 1;2(7):e397–405.

5. Nackers F, Cohuet S, le Polain de Waroux O, Langendorf C, Nyehangane D, Ndazima D, et al. Carriage prevalence and serotype distribution of *Streptococcus pneumoniae* prior to 10-valent pneumococcal vaccine introduction: A population-based cross-sectional study in South Western Uganda, 2014. Vaccine. 2017 18;35(39):5271–7.

6. Hill PC, Akisanya A, Sankareh K, Cheung YB, Saaka M, Lahai G, et al. Nasopharyngeal Carriage of *Streptococcus pneumoniae* in Gambian Villagers. Clin Infect Dis. 2006 Sep 15;43(6):673–9.

7. Heinsbroek E, Tafatatha T, Phiri A, Swarthout TD, Alaerts M, Crampin AC, et al. Pneumococcal carriage in households in Karonga District, Malawi, before and after introduction of 13-valent pneumococcal conjugate vaccination. Vaccine. 2018 Nov 19;36(48):7369–76.

8. Van Zandvoort K, Bobe MO, Hassan AI, Abdi MI, Ahmed MS, Soleman SM, et al. Social contacts and other risk factors for respiratory infections among internally displaced people in Somaliland. Epidemics. 2022 Dec 1;41:100625.

9. Løchen A, Truscott JE, Croucher NJ. Analysing pneumococcal invasiveness using Bayesian models of pathogen progression rates. PLoS Comput Biol. 2022 Feb;18(2):e1009389.

Chapter 4

Effectiveness of pneumococcal vaccination campaigns in humanitarian settings: a modelling study

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1604011	**Title**	Mr
First Name(s)	Kevin		
Surname/Family Name	van Zandvoort		
Thesis Title	Strategies for pneumococcal conjugate vaccine use in humanitarian crises		
Primary Supervisor	Stefan Flasche		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	The Lancet Global Health
Please list the paper's authors in the intended authorship order:	Kevin van Zandvoort, Mohamed Bobe, Abdirahman Ibrahim Hassan, Rachael Cummings, Abdihamid Warsame, Casey Pell, Mohamed Ismail Abdi, Mohamed Abdi Hergeeye, Catherine McGowan, E Kim Mulholland, Rosalind M Eggo, Catherine Satzke, Francesco Checchi, Stefan Flasche

Stage of publication	**Not yet submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I led the design of the modelling study and developed a generalizable pneumococcal transmission model in R and C++ that can flexibly be extended to any number of age-groups, vaccinated strata, populations, and even model structures. The model has since been used for six different PCV studies, including the current study in Somaliland. As described in the previous chapters, I led the data collection and data cleaning of key parameters to model pneumococcal transmission in the Digaale IDP camp, and sourced the literature to parameterize the model with relevant data for other required parameters. I then fitted the model, conducted all analyses, interpreted the results, produced all figures and tables, wrote the original manuscript draft, and edited and prepared the final manuscript for submission.

SECTION E

Student Signature	Kevin van Zandvoort
Date	03/10/2024

Supervisor Signature	Stefan Flasche
Date	03/10/2024

Effectiveness of pneumococcal vaccination campaigns in humanitarian settings: a modelling study

Kevin van Zandvoort [1], Mohamed O Bobe [2], Abdirahman I Hassan [3], Rachael Cummings [4], Abdihamid Warsame [1], Casey L Pell [5, 6], Mohamed Ismail Abdi [2], Mohamed Abdi Hergeye [3], Catherine R McGowan [1], E Kim Mulholland [1, 5, 6], Rosalind M Eggo [1], Catherine Satzke [5, 6], Francesco Checchi [1], Stefan Flasche [1, 7]

1. Department for Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, United Kingdom
2. Health and Nutrition, Save the Children International, Somaliland
3. Republic of Somaliland Ministry of Health Development, Somaliland
4. Save the Children, United Kingdom
5. Infection, Immunity and Global Health, Murdoch Children's Research Institute, The University of Melbourne Department of Paediatrics at the Royal Children's Hospital, Australia
6. Department of Microbiology and Immunology, The University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Australia
7. Centre for Global Health, Charité - Universitätsmedizin Berlin, Germany

Summary

Background

A large and increasing number of people are forcibly displaced worldwide because of war or conflict, food insecurity, natural disasters, and other crises. *Streptococcus pneumoniae* (the pneumococcus) likely causes a substantial health burden in these populations, but pneumococcal vaccines (PCVs) are rarely part of the humanitarian response. We evaluated the potential impact of logistically feasible PCV campaigns in such settings.

Methods

We conducted a cross-sectional pneumococcal carriage, malnutrition and social contacts survey in a camp for displaced people near Hargeisa, Somaliland, to parameterise an age-structured dynamic transmission model accounting for population movements. We projected the effect of PCV mass-vaccination campaigns using one or two doses and vaccinating children <1, 2, 5, 10 or 15 years (y) of age in the Somaliland camp and topologically different crisis settings.

Findings

Our model projects that a single-dose PCV campaign with high vaccine coverage in <5y old children can partially control transmission of vaccine-serotypes for up to three years, preventing 28% (24 – 32) of severe pneumococcal disease. A campaign vaccinating only <1y or <2y olds failed to induce similarly substantial indirect effects. In settings with more migration, or more interaction with an unvaccinated host population, expanded age eligibility was needed to achieve comparable protection. A campaign targeting children <5y was the most efficient use of PCV with 217 (163 - 287) vaccines needed to prevent one case of severe pneumococcal disease within the first year. If such a strategy was to be implemented in all affected populations about 40 million doses would be needed globally in the next 5 years.

Interpretation

Single-dose PCV mass-vaccination campaigns offer feasible and pragmatic protection if routine vaccination is not possible.

Funding

Elrha's Research for Health in Humanitarian Crises (R2HC) Programme, which aims to improve health outcomes by strengthening the evidence base for public health interventions in humanitarian crises.

Research in context

Evidence before this study

Acute respiratory infections are a leading cause of morbidity and mortality in crisis-affected populations, but pneumococcal conjugate vaccines are rarely used in humanitarian responses. Routine vaccination is often not feasible in these settings due to a lack of access or security, and there is little guidance on alternative delivery options for these vaccines. A literature search on Apr 23, 2024, using the terms ("pneumococcal conjugate vaccine*" OR "PCV") AND ("humanitarian" AND ("cris*" OR "emergenc*")) returned 7 results on Embase and 6 results on PubMed. PCV mass vaccination strategies have previously been shown to be cost-effective, but no study has directly compared the effect and efficiency of different age eligibility and dosing regimens in crisis-affected populations.

Added value of this study

We previously collected key primary data that would allow us to model the effect of different PCV campaigns on the transmission of vaccine-type pneumococci. To our knowledge, this is the first analysis that has assessed the combined direct and indirect effect of PCV campaigns in crisis-settings. We show the importance of achieving high levels of indirect protection for a substantial and sustained impact, which can be realized by vaccinating the main transmitters in addition to the children at highest risk of severe disease. This can likely be done using only a single-dose strategy, which greatly improves the feasibility of a campaign.

Implications of all the available evidence.

Single-dose mass-vaccination PCV campaigns that achieve high coverage in children up to 4 years of age or older can provide substantial health benefits in crisis-affected populations. Temporarily disrupting transmission of vaccine-type pneumococci is crucial to protect unvaccinated children born or migrated after the vaccination campaign. Integration into multi-antigen vaccination strategies could reduce logistical overheads and costs.

Introduction

Over 100 million people were forcibly displaced worldwide in 2022 by war or conflict, food insecurity, natural disasters, and other crises. Of these people, over half were internally displaced, and more than 40% were children[1]. Displaced populations often live in overcrowded settings where infectious disease outbreaks are common, with poor access to healthcare, and increased levels of morbidity and mortality. Acute respiratory infections are a leading cause of mortality in children under five, and *Streptococcus pneumoniae* (the pneumococcus) likely causes a sizeable proportion of this burden[2,3]. However, pneumococcal conjugate vaccines (PCVs) are rarely used during humanitarian responses[4].

The WHO recommends that PCVs be used in crisis settings for children <1 year of age, and that they be considered in children <5 years of age[5,6]. Routine vaccination schedules tend to cover infants (children aged <12 months), while children <5 years of age have been included in campaigns coinciding with PCV introduction into the Expanded Programme on Immunization (EPI). However, crisis settings may require other vaccination strategies, as limited access to services, disruption of the cold or supply chain , lack of sufficient trained personnel and/or reduced safety of health care workers can make routine vaccination challenging to implement[4].

Mass vaccination campaigns (MVC) are used in many humanitarian responses to rapidly provide direct and indirect protection against pathogens including measles, polio, and cholera.[7] MVCs often involve giving a single-dose of vaccine across many age groups as a pragmatic approach to maximize vaccination coverage[8]. This approach immunizes those most at risk of severe disease, as well as potentially interrupting transmission and boosting herd immunity. High levels of herd protection would i) mitigate the suboptimal protection induced by a single PCV dose in infants, and ii) have the added benefit of providing indirect protection to unvaccinated cohorts of newborns or newly migrated individuals who entered the population after an MVC. This is especially relevant in crisis-affected settings with high birth rates and migration[4,9]

To assess the impact of PCV vaccination strategies that may feasibly be used in crisis settings, we previously collected pneumococcal carriage[10] and risk factor data[11] in a camp for internally displaced people (IDP) in Somaliland informing a pneumococcal transmission model. We here report our modelled findings of the effectiveness and efficiency of potential PCV vaccination strategies in a range of crisis typologies constructed from publicly available information, including the Somaliland IDP camp.

Methods

Transmission Model

We developed a compartmental model to simulate pneumococcal transmission within and between a displaced and host population. We grouped pneumococcal serotypes as vaccine-type (VT) or non-vaccine-type (NVT), depending on their inclusion in the Pneumosil vaccine (serotypes 1, 5, 6A, 6B, 7F, 9V, 14, 19A, 19F, and 23F), which is a low-cost PCV targeting serotypes commonly circulating in low- and middle-income countries. Populations were stratified by age, malnutrition status and vaccination status (Figure 1). In our model, susceptible individuals (S) can become infected with VT (V) or NVT (N) following effective contact with a VT or NVT carrier. Individuals carrying VT or NVT can acquire a superinfection (B) through additional infection with serotype(s) in the other serotype group. Acquisition of superinfection occurs at a lower rate as serotypes compete. As serotypes are grouped, compartments V, N, and B represent individuals carrying at least one VT serotype, individuals carrying at least one NVT serotype, and individuals carrying at least one VT and NVT serotype, respectively. More details including model equations are provided in supplemental material section A.

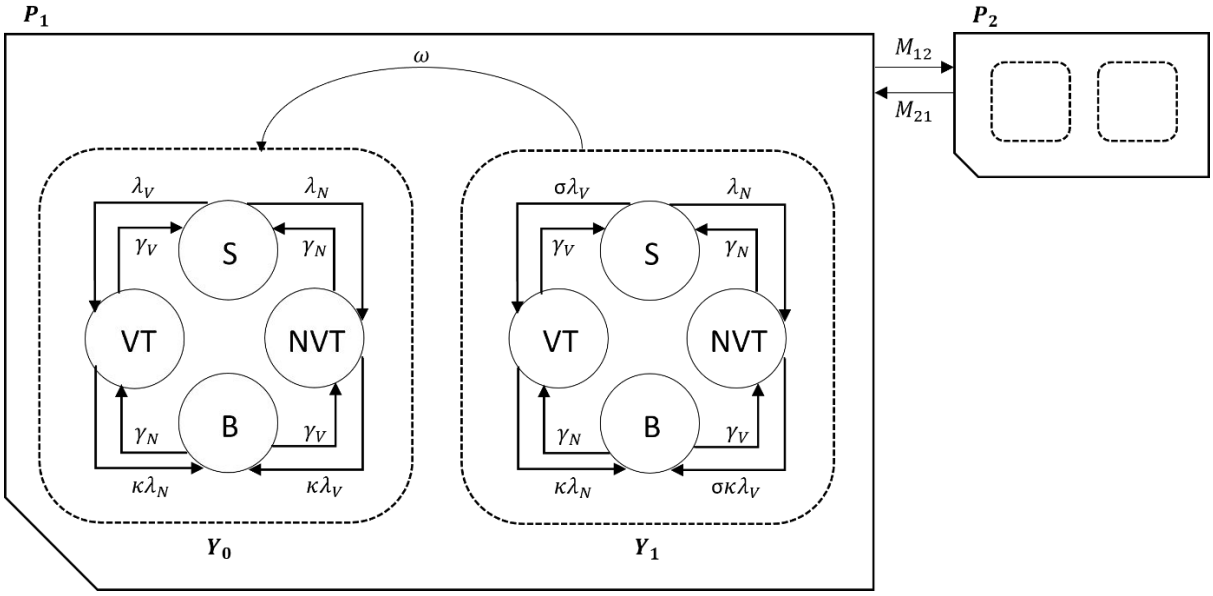


Figure 1. Structure of the compartmental pneumococcal transmission model. Pneumococcal serotypes are grouped as those included in the vaccine (V), or not included in the vaccine (N). At any time *t*, susceptible individuals (S) acquire serotypes in either group at rates $\lambda_V(t)$ and $\lambda_N(t)$, and become infectious carriers themselves. Carriers of serotypes in one group remain partially susceptible to infection with serotypes in the other group, and may develop a superinfection (B) at rates $\kappa\lambda_V(t)$ and $\kappa\lambda_N(t)$, where κ models competition between serotypes. Carriers are assumed to clear their infection of all serotypes in either group at rates γ_V for VTs and γ_N for NVTs. Individuals may be without vaccine derived protection (Y_0), or with complete vaccine derived protection (Y_1). Those with vaccine protection acquire VT at a reduced rate, σ, and lose their protection at rate ω. Different populations P_1 and P_2 represent a displaced and host population. Individuals of population *p* migrate

to population q at rate, M_{qp}. Age groups, malnourished strata, ageing, and transmission between vaccine protection strata, populations, and age groups are not shown.

We implemented mixing between age groups by an age-dependent daily contact rate (Table 1), informed by setting-specific contact matrices using data from the Digaale IDP camp and other displaced populations (described later in methods). Contact between populations is governed by a travel matrix, which incorporates a constant parameter for the proportion of extra-household contacts made by individuals between populations.

Table 1. Overview of model parameters.

Parameter	Description	Value	Source
Epidemiological parameters			
$\gamma_{s,a}$	Duration of VT carriage (days) for different serotypes (s: VT or NVT) and age groups. See supplemental material section A.	VT: a < 21m: 125; 22≤ a <40m: 67; 40≤ a <59m: 28; a ≥60m: 22. NVT: a < 21m: 61; 22≤ a <40m: 50; 40≤ a <59m: 34; a ≥60m: 24.	[12]
κ	Competition parameter, which determines the degree by which acquisition of heterologous carriage is reduced in carriers of only VT or NVT compared to susceptible individuals	Fitted, single estimate	N/A
β_a	Probability of transmission per contact with an infectious individual carrying NVT for different age groups (a)	Fitted, three values for contactors aged a <5y, 5≤ a <15y, and a ≥15y.	N/A
$\beta_{VT/NVT,a}$	Rate ratio for probability of transmission with contacts carrying VT, compared to NVT in age group a.	Fitted, two values for a <5y and a ≥5y	N/A
β_m	Rate ratio for probability of transmission in malnourished compared to non-malnourished individuals. Explored in sensitivity analysis (supplemental material section D). See supplemental material section A.	1.2 (values 1.0 to 2.0 explored in a sensitivity analysis for the base model)	Assumption based on [13], [14], and [15]
$\tau_{s,a}$	Case Carrier Ratio: Proportion of carriage episodes that result in severe pneumococcal disease, for a given serotype (s) and age group (a). Estimates were adjusted for β_m. See supplemental material section A.	We applied the age- and serotype specific posterior estimates from Kilifi, Kenya to calculate IPD cases. These were multiplied by 5.98 to add non-bacteraemic pneumonia cases.	[16–19]
ψ	Rate ratio of CCR ($\tau_{s,a}$) for malnourished compared to non-malnourished individuals.	2.0	[20–22]
Vaccine parameters			
$1-\sigma_a$	Vaccine efficacy against VT carriage acquisition. Explored in sensitivity analysis (supplemental material section D).	50% in <1y who received two doses, and in all ≥1y; 25% in <1y who received a single dose (values 0 to 60% were explored for the base model in a sensitivity analysis)	Assumption based on [23]
$1-\phi_a$	Total vaccine efficacy (combined vaccine efficacy against carriage acquisition and severe pneumococcal disease). See supplemental material section A.	80% in <1y who received two doses, and in all ≥1y; 56% in <1y who received a single dose	[24]
$\frac{1}{w}$	Mean duration of vaccine protection (y). Explored in sensitivity analysis (supplemental material section A).	6 years, assumed to follow a negative exponential distribution (values 2 to 9 years were explored in a sensitivity analysis for the base model)	[25]
ρ_a	Vaccine coverage in age group *a*. Explored in sensitivity analysis (supplemental material	85% if age group included in PCV campaign, 0% otherwise (values 45% to	Assumption

	section D). We assume that all populations are unvaccinated prior to the MVC.	95% were explored in a sensitivity analysis for the base model)	
Demographic parameters			
a	Modelled age group	25 age groups: 12 monthly age groups for <1y, five annual age groups for 1-5y, and 8 broader age groups for 5-6y, 6-10y, 10-15y, 15-20y, 20-30y, 30-40y, 40-50y, 50+y	N/A
N_a^p	Population size of age group *a* in population *p*.	Setting specific, see table 2	
$c_{a,b}^{p,l}$	Average number of daily contacts between contactors aged *b* in population *p*, and contactees aged *a* of location *l* (intra- or extra-household).	Setting specific, see table 2	
T_{qp}	Proportion of extra-household contacts made by contactors in the population *p* with contactees in population *q*. See supplemental section A. Adjusted in sensitivity analysis (supplemental material section D).	Setting specific, see table 2 (values 0, 0.03, 0.05, 0.1, 0.2, and 0.4 were explored for the base model in a sensitivity analysis)	
M_{qp}	Daily migration rate from population *p* to population *q*. See supplemental section A. Adjusted in sensitivity analysis (supplemental material section D).	Setting specific, see table 2 (values for an average of 2, 4, 6, 8, and 10 years until migration, as well as no migration, were explored for the base model in a sensitivity analysis)	

We modelled the proportion of people within each population and age group in the different model compartments, assumed a constant population size and age distribution, and ignored deaths directly related to pneumococcal transmission as they are likely to be numerically negligible. Equal and opposite rates of in- and out-migration (additional to travel-related contact) occur between populations.

Ageing between age groups is implemented using a constant rate, which implicitly assumes a memoryless exponentially distributed rate of ageing. To reduce extreme values in this exponential distribution where a substantial proportion of children may age too slow and too fast, we used twelve monthly age groups for infants so that the duration an infant would remain below one year of age would approximate a gamma distribution. This minimizes the proportion of vaccinated and unvaccinated infants that would be allocated to an incorrect age group, while still allowing continuous ageing without resorting to the use of partial differential equations.[26]

Epidemiological assumptions

We used average age-specific duration of carriage estimates for VTs and NVTs, estimated from a longitudinal study in Kilifi, Kenya[12]. Malnutrition is an important risk factor for pneumococcal disease that is often highly prevalent in crisis-affected populations, and may affect the age-distribution of cases.[4] We stratified populations by acute malnutrition status (not

malnourished, and moderate or severe acute malnutrition) to account for increased risk of pneumococcal transmission and disease for malnourished individuals. We assumed that malnourished individuals had a 20% increase in acquisition rates compared to non-malnourished individuals and a 100% increase in invasive pneumococcal disease (IPD)[14,21]. We calculated the incidence of IPD by applying VT- and NVT-specific case-carrier ratios (CCR) estimates from Kilifi, Kenya[27], which were applied to the modelled incidence of carriage acquisitions. We followed Ojal et al and assumed that for every IPD case, there would be 4.98 cases of non-bacteraemic pneumococcal pneumonia as has been estimated in Kenya[17–19] (Table 1 and supplemental material section A). The total number of severe pneumococcal disease cases was then calculated as the sum of IPD and non-bacteraemic pneumococcal pneumonia cases.

Vaccine protection

We assumed that two doses in infants and a single dose in those aged 1 year and older were sufficient to elicit a complete immune response corresponding to a vaccine efficacy (VE) against carriage acquisition of VTs of 50%, a total VE against pneumococcal disease of 80%, and an average duration of vaccine protection of 6 years.[23–25] For infants who only received a single dose of PCV, we assumed that these values were halved.[23] VE was implemented as leaky vaccine protection (Table 1). We assumed no difference in VE and duration of protection for malnourished vaccinees.[28]

Base scenario: Digaale IDP camp, Somaliland

To parameterize our transmission model, we previously conducted a cross-sectional survey in Digaale, an IDP camp in Somaliland, in 2019.[4,10,11] Digaale is a permanent camp established in 2013 following droughts and famines, and is located near the city of Hargeisa, the capital of Somaliland. It has an estimated 3000 people living across 700 shelters. We collected data on social contacts (a proxy for pneumococcal transmission pathways[29]), pneumococcal carriage prevalence, and demographic estimates including mortality, migration, and malnutrition rates (Table 2). The survey, contact data, and carriage data, are described in more detail elsewhere[10,11]. Ethical approval was granted by the Somaliland Ministry of Health Development, Directorate of Planning, Policy, and Strategic Information, and by the London School of Hygiene & Tropical Medicine. We used data from the 2020 Somaliland Demographic and Health Survey, UN World Population Prospects, and synthetic contact matrices to inform

demographic estimates for the host population.[30–32] As we collected detailed data from Digaale, we used it as our base scenario in all analyses.

Alternative crisis typology scenarios

Digaale is one specific setting for which we have good data, but that may differ from other settings. While every crisis is unique, crisis-affected populations may be categorized in different typologies: i) mass-displacement in an IDP or refugee camp, typically with high rates of overcrowding and migration, and ii) entrapment in an existing community, which may be rural or urban.[33] To account for a wider diversity of distinct crisis-affected settings, we conducted a sensitivity analysis in which we modelled one example of a population in each typology: i) an acute-phase IDP camp based on the Bentiu protection of civilian (PoC) site in South Sudan in 2015, which is much bigger than Digaale, ii) an acute-phase urban setting with mixed IDP and host communities based on the city of Maiduguri in North-East Nigeria in 2016, and iii) a protracted-phase rural setting based on Bambari town.

Bentiu PoC is a large IDP camp near the town of Bentiu in South Sudan and was established following the start of the 2013-2020 South Sudanese Civil War. When conflict escalated in 2015, the population grew rapidly to 150,000 resulting in overcrowding, poor hygiene, malnutrition.[34]

Maiduguri is the capital of Borno state, Nigeria. Armed conflict in the northeastern states of Nigeria resulted in mass displacement. In 2016, 800,000 displaced individuals were housed with family members, in schools, or in incomplete housing projects throughout the city with a host population of 1 million people. This resulted in high malnutrition rates, high levels of population movement, and high levels of mixing between the displaced and host populations.[35]

Bambari is the capital of the rural Ouaka prefecture in the Central African Republic. It has suffered decades of conflict and instability. While there was no mass displacement and little acute malnutrition in 2019, the protracted crisis has resulted in poor healthcare access with a need to maintain immunity over years with only limited EPI services.[36]

For each population, we identified and extracted data on population demographics, malnutrition, and migration (Table 2).[37,38] Our sensitivity analysis then ensures realistic values for these different variables that may occur in these typologies, allowing us to compare the effectiveness of vaccination strategies under different conditions. As in the base model, we modelled a distinct host and displaced population in all settings with the exception of Bambari. In the absence of other pneumococcal prevalence estimates, we assumed prevalence rates in the displaced population to be the same as estimated in Digaale.

Table 2. Scenario specific model parameters.

Parameter		Digaale 2019	Bentiu PoC 2015	Maiduguri 2016	Bambari 2019
$\sum_{a=1}^{25} N_a^p$	Total population size in population *p* (combined for malnourished and non-malnourished)	Displaced: 3K [11]; host: 1.2M [30]	Displaced: 150K [39]; host:70K [40]	Displaced: 800K [41]; host: 1M [41]	375K [37] (no discrepancy made between host and displaced)
h	Average household size in displaced population	4.5 [11]	7.8 [42]	5.03 [37]	N/A
$c_{a,b}^{p,h}$	Average number of daily contacts between contactors aged b in population p, and contactees aged a and in location l (within or outside of the household).	Displaced: estimated in Digaale.[11] Host: synthetic matrix for Ethiopia.[32]	Displaced: adjusted from Digaale.[11] Host: synthetic matrix for South Sudan.[32]	Displaced: adjusted from Digaale.[11] Host: synthetic matrix for Nigeria.[32]	Synthetic matrix for the Central African Republic. [32]
T_{qp}	Proportion of extra-household contacts made by contactors in the population p with contactees in population q.	2.7% [11]	Assumed to be the same as in Digaale.	55.6% (Assumed homogenous mixing).	N/A
M_{qp}	Daily migration rate from population p (displaced) to population q (host).	3.5/10,000/year [11]	12.6/10,000/year[43]	13.7/10,000/year[44]	N/A
χ_p	Proportion of population that is acutely malnourished.	Displaced: 1.5% [11]; host: 20% [30]	Displaced: 17.9% [43]; host: 17.6% [43]	Displaced: 19.2% [45]; host: 19.2% [45]	4.2% [46]

We informed contact rates in the displaced population by extrapolating the contact rates from Digaale, adjusting for the population distribution and household sizes (supplemental material section A). Synthetic contact matrices were used to model contact rates within the host populations.

Model fitting

We fitted five parameters for the age- and serotypes specific probability that contact between a susceptible and carrying individual resulted in transmission, and one competition parameter to scale these probabilities for acquisition of heterologous carriage.

We used an MCMC algorithm to fit our model to age-specific carriage estimates from Digaale in 2019, comparting the model-predicted steady state prevalence with the observed age-dependent prevalence in the S, VT, NVT, and B compartments. We used the resulting samples from the joint-posterior distribution for our model simulations. The model was refitted for each crisis typology scenario to ensure similar carriage prevalence with the different model parameters (supplemental material section D).

Vaccination strategies

We modelled a range of mass-vaccination campaign (MVC) strategies using different age targeting: offering PCVs to children aged 6 wks to 11 mths, and extending this to children aged up to 1, 4, 9 and 14 y. We assumed a single-dose regimen in all age groups as our base case and assumed that every MVC reaches 85% coverage in the eligible age groups.

We compared strategies by their modelled impact on VT carriage prevalence and on severe pneumococcal disease cases. We quantified the efficiency of different strategies by calculating the average number needed to vaccinate (NNV) to prevent one case of severe pneumococcal disease, given the cumulative cases predicted in the no-vaccination scenario. We also assessed the incremental NNV between vaccination strategies, i.e. the number of additional vaccine doses needed to prevent one additional case of severe pneumococcal disease, where each strategy is compared to the strategy with the next smaller campaign.

Sensitivity analyses

In addition to extending the model to alternative crisis typology scenarios, we assessed the sensitivity of our results against different parameter values for i) the number of vaccine doses, ii) the vaccination coverage, iii) duration of vaccine protection, iv) VE against carriage, v) migration rate, vi) mixing with the host population, and vii) the effect of malnutrition on transmission. We also assessed the structural uncertainty of our compartmental model to an implementation that is structurally neutral (supplemental section D).

Global demand forecast

To understand the stockpile requirements for PCV use in humanitarian settings, we estimated the demand for the annual humanitarian PCV doses needed in low- and lower middle-income countries over a five-year horizon, if mass vaccination PCV campaigns would commonly be used. We assumed that an initial MVC would be conducted in all currently displaced and newly displaced populations, and that impact would then be maintained by vaccinating all unvaccinated infants born after the MVC annually (supplemental material section E).

Role of the funding source

The funders were not involved in the study design; collection, analysis, and interpretation of data; writing of the paper; and the decision to submit it for publication. All authors had full access to data in the study, and final responsibility for the decision to submit for publication.

Code availability

All analyses were implemented in R and C++. The model code can be found at https://github.com/kevinvzandvoort/espicc-metavax-pcv-humanitarian-crises.

Results

The model reproduced the age-specific pre-vaccination prevalence and distribution of VT and NVT serotypes as observed in Digaale (supplemental material section B).

Impact on VT prevalence

In our base case modelling the Digaale IDP camp, all strategies reduced VT prevalence compared to the no-vaccination baseline. Widening the age eligibility of the PCV campaign increased both the peak and duration of impact on VT carriage prevalence and thus severe pneumococcal disease (Figure 2). For MVCs including children <1, <2, <5, <10, or <15 y, peak impact on infant VT prevalence was reached on day 93 (81 – 106), 163 (120 – 440), 293 (211 – 418), 327 (243 – 438), and 336 (252 – 446) following the MVC, respectively. Infant VT carriage prevalence reduced from 60% (50 – 68) in the absence of vaccination to 56% (47 – 64), 54% (45 – 63), 48% (37 – 58), 42% (31 – 54), and 41% (29 – 52) at peak impact, respectively. Without subsequent vaccination, VT carriage prevalence returned to its pre-MVC baseline after about one year for the <1y and <2y strategies, while the effect was still evident after 3 years for strategies using wider age targeting.

Impact on disease

The cumulative impact of the PCV campaigns on disease cases increased substantially with widening of the age eligibility for vaccination (Table 3). Campaigns in infants and toddlers only had a small impact on disease, and only in the first year. The <5y campaign did prevent 5 (4 - 7) severe pneumococcal disease cases per 10,000 people in all ages, and 23 (14 - 39) cases per 10,000 infants in the first year, while the <10 and <15y campaigns extended the prevented number of cases to 6 (5 - 9) and 7 (5 - 9) per 10,000 people in all ages, and 29 (18 - 49) and 31 (19 - 52) cases per 10,000 infants in the same period.

Restricted to infant cases, there was only a small benefit of the <2y over the <1y campaign. Strategies vaccinating children up to 5y of age or greater were able to achieve sufficient indirect protection to substantially protect unvaccinated infants born after the MVC in the years after the MVC (Figure 1 and Table 3).

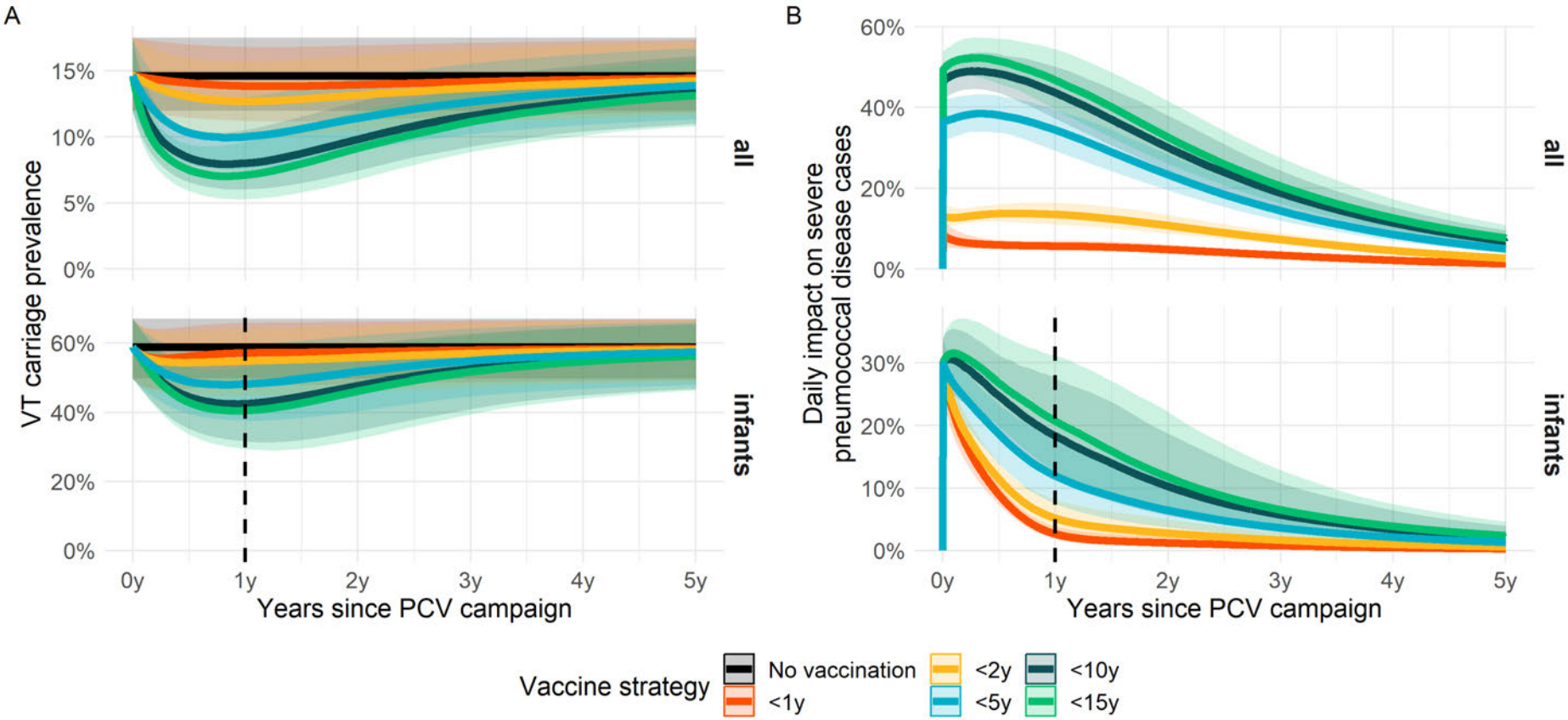


Figure 2. Impact of PCV mass vaccination campaigns on prevalence and incidence. A: VT carriage prevalence (VT + B) in all age groups and infants only. B: reduction in daily severe pneumococcal disease cases compared to no vaccination, in all age groups and infants only. In all plots, thick lines show the median estimates and shaded areas corresponding 95% credible intervals from 500 model posteriors. A dotted vertical line is plotted for infants 1 year after the PCV campaign. Infants to the right of this line have been born after the PCV campaign and are thus all unvaccinated. The reductions in these birth cohorts result from indirect vaccine protection.

Efficiency of PCV doses

The <5y campaign had the lowest NNV to prevent one case of severe pneumococcal disease in the first 6 months (425; 316 - 570) and 1-year (217; 163 - 287), indicating a more efficient use of PCV than any other campaign, while the <2y campaign had the lowest NNV over a 2-year (120; 92 – 154) and 3-year (89; 68 – 115) period (Table 3). As no further PCVs were administered, the NNV decreased with time for all vaccine strategies considered as prevented cases accrue. The NNV of the <15y campaign was highest over all periods, reducing to 206 (154 - 274) over a 3-year period.

Table 3. Cumulative disease impact and efficiency of PCV campaigns, over different periods

Age	Vaccine strategy	Period			
		0-6m	*0-1y*	*0-2y*	*0-3y*
Severe pneumococcal disease cases prevented per 10,000 people per year[a]					
All ages	<1y	1 (1 - 2, 7%)	1 (1 - 1, 6%)	1 (1 - 1, 6%)	1 (1 - 1, 5%)
	<2y	2 (2 - 3, 13%)	2 (2 - 3, 13%)	2 (2 - 3, 13%)	2 (2 - 3, 12%)
	<5y	6 (5 - 9, 38%)	6 (5 - 8, 37%)	6 (4 - 7, 33%)	5 (4 - 6, 28%)
	<10y	8 (6 - 11, 48%)	8 (6 - 10, 47%)	7 (6 - 9, 42%)	6 (5 - 8, 36%)
	<15y	9 (7 - 11, 51%)	9 (6 - 11, 50%)	8 (6 - 10, 45%)	7 (5 - 9, 39%)
Infants	<1y	24 (16 - 39, 16%)	16 (10 - 25, 11%)	9 (6 - 15, 6%)	7 (4 - 11, 4%)
	<2y	26 (17 - 43, 18%)	19 (12 - 31, 13%)	12 (8 - 21, 8%)	9 (6 - 16, 6%)
	<5y	34 (22 - 56, 23%)	28 (18 - 47, 19%)	21 (12 - 36, 14%)	16 (9 - 28, 11%)
	<10y	41 (26 - 67, 28%)	36 (22 - 61, 25%)	29 (17 - 50, 19%)	23 (13 - 42, 15%)
	<15y	43 (27 - 71, 30%)	39 (24 - 66, 27%)	32 (18 - 56, 21%)	26 (14 - 47, 17%)
NNV: Number of PCV doses needed to prevent one severe pneumococcal disease case[b]					
All ages	<1y	443 (308 - 612)	235 (172 - 309)	126 (94 - 161)	93 (70 - 118)
	<2y	471 (351 - 607)	230 (175 - 294)	120 (92 - 154)	89 (68 - 115)
	<5y	425 (316 - 570)	217 (163 - 287)	122 (92 - 159)	95 (72 - 124)
	<10y	688 (524 - 911)	353 (270 - 459)	198 (150 - 257)	154 (116 - 201)
	<15y	933 (714 - 1,200)	477 (363 - 629)	267 (200 - 350)	206 (154 - 274)
Infants	<1y	631 (390 - 965)	480 (296 - 736)	414 (254 - 650)	383 (231 - 609)
	<2y	1,200 (745 - 1,900)	849 (519 - 1,300)	658 (387 - 1,100)	581 (337 - 948)
	<5y	2,400 (1,500 - 3,800)	1,500 (879 - 2,300)	1,000 (578 - 1,700)	860 (488 - 1,500)
	<10y	4,200 (2,600 - 6,600)	2,400 (1,400 - 3,900)	1,500 (869 - 2,600)	1,200 (691 - 2,200)
	<15y	5,700 (3,500 - 9,200)	3,200 (1,900 - 5,200)	2,000 (1,100 - 3,500)	1,600 (882 - 2,900)

a. Estimates show the median absolute cumulative impact per 10,000 people per year, in each respective period. The 95% uncertainty intervals and median relative impact are provided in brackets. Estimates are taken from 500 model runs.

b. Estimates show the total number of PCV doses needed to prevent one case of severe pneumococcal disease. The 95% uncertainty intervals are provided in brackets. Estimates are taken from 500 model runs. Values greater than 1000 are rounded to the nearest hundred.

The additional doses administered to children 1y in an <2y campaign, and those administered to children 2-4y in an <5y campaign were no less effective in preventing additional severe pneumococcal disease cases compared to the <1y strategy. While <10y and <15y campaigns result in larger impacts, the additional doses administered to children 5-9y and 10-14y prevent 73% (63 - 83) and 89% (82 - 94) fewer disease cases per dose administered (Figure 3).

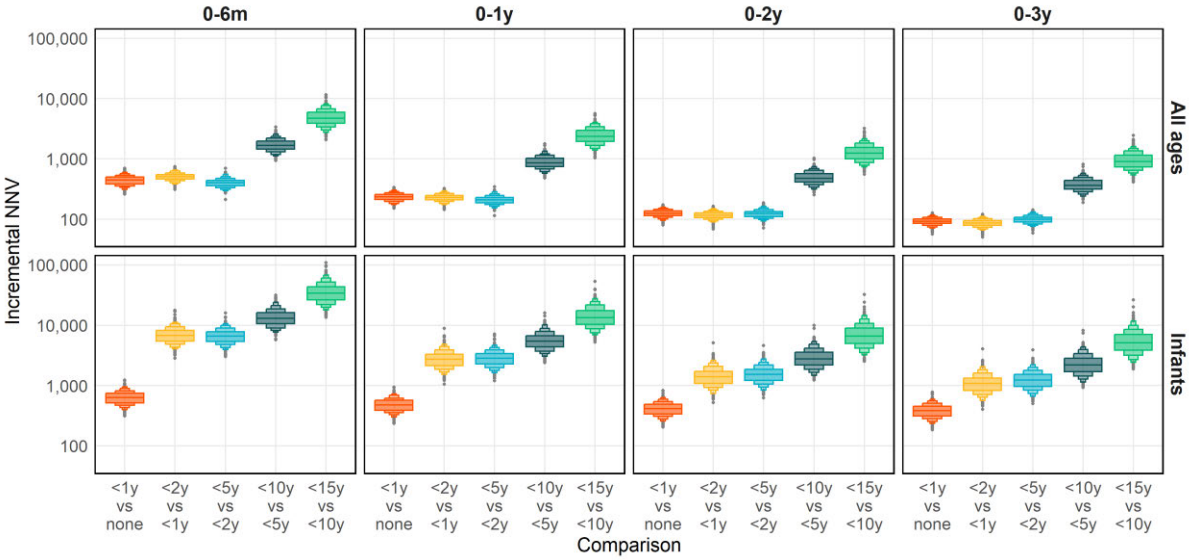


Figure 3. Incremental NNV of vaccine strategies. Estimates show the additional number of PCV doses needed to prevent one additional severe pneumococcal disease case, compared to a strategy with narrower age targeting. Boxen plots show the distributions of the estimated incremental NNVs from 500 model iterations.

The <1y campaign had the lowest NNV to prevent one case of infant disease in all periods considered, with 383 (231 – 609) over a 3-year period. The NNV gradually increased with wider age targeting, to 1,600 (882 – 2,900) for an <15y campaign over 3 years. While the rank for the NNV to prevent one infant disease case remained the same, the relative difference between strategies reduced as time progressed, due to the substantially higher impact of strategies with wider age targeting at later durations.

Dosing used in PCV campaigns

Campaigns where infants received two doses were more effective in reducing infant severe pneumococcal disease cases (Figure 4), especially within the first year, but had a minimal effect on cases in other age groups. With wider age targeting, the relative difference with single-dose campaign reduced. A single-dose <5y campaign prevented an equivalent number of infant cases as an <2y campaign where infants received two doses, but prevented a

substantially larger proportion of cases in all age groups and for a longer period with a lower NNV.

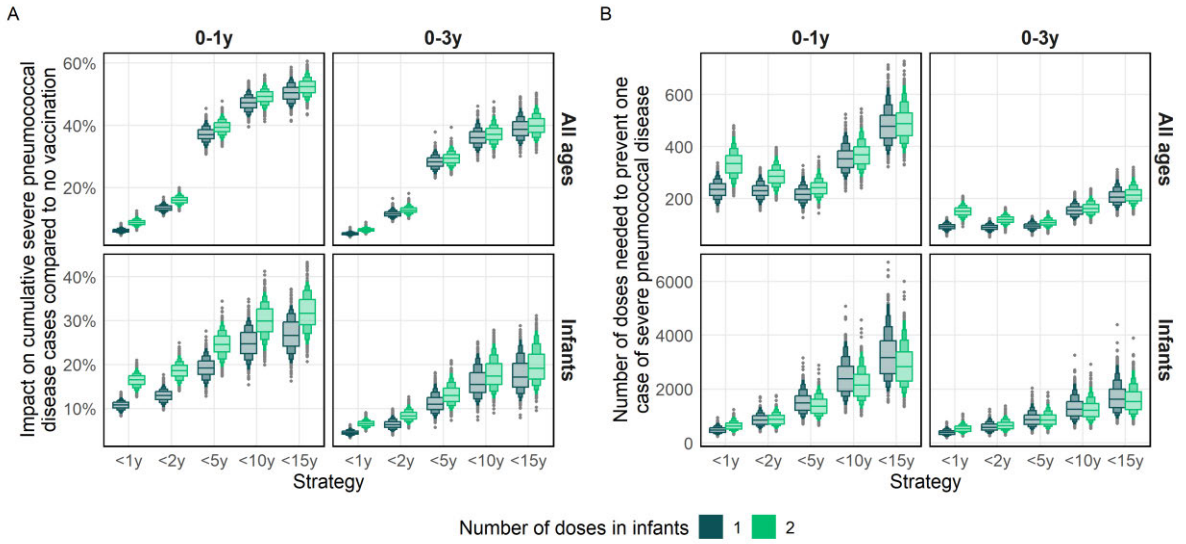


Figure 4. Comparing single-dose strategies to strategies where infants received two doses of PCV. A: impact on cumulative severe pneumococcal disease cases over a 1y and 3y period, in all age groups and infants only. B: NNV of vaccine strategies over a 1y and 3y period, in all age groups and infants only. Boxen plots show the distributions of the estimated values from 500 model iterations.

Sensitivity analyses

The modelled impact on severe pneumococcal disease cases in Digaale over 3 years was sensitive to the degree of interaction with the host population, migration rate, vaccine coverage, VE against transmission, and duration of vaccine protection, but not to the rate ratio for the probability of transmission in malnourished (supplemental material section D). The sensitivity was greater for impact on infant cases than that in all age groups. In scenarios with decreased vaccination coverage, increased migration rates, or increased mixing between the displaced and host population, strategies with broader age targeting were needed to achieve substantial impacts, at the expense of higher NNV compared to the base scenario. Increased vaccination coverage resulted in increased impact but not increased NNV for all vaccination strategies, indicating little loss of efficiency for the additional doses provided to prevent the additional cases.

While an alternative structurally neutral model projected overall much greater impact of all PCV campaigns, with much longer duration, it did not qualitatively alter the relative rank of

each MCV strategy in terms of impact or efficiency compared to the base model (supplemental material section D).

Other crisis typologies

In all three settings, patterns were consistent with those estimated for Digaale (Figure 5). Bambari was the only setting where an <2y campaign was projected to have a substantial impact on infant VT carriage after 1y. The overall impact was most pronounced in Bambari and least in Maiduguri, highlighting the stark differences in their migration rates and interaction with an unvaccinated host population. Detailed estimates are provided in the supplemental material section C.

Global demand forecast

In 2022, an estimated 112 million people were forcibly displaced in low and lower middle-income countries, which has on average increased by 16 million people annually within the past five years.[1] If age appropriate single-dose <5y or <10y PCV campaigns were considered for all these populations, we estimate the global need of PCV to be 40 to 64 million doses over the next five years (supplemental material section E).

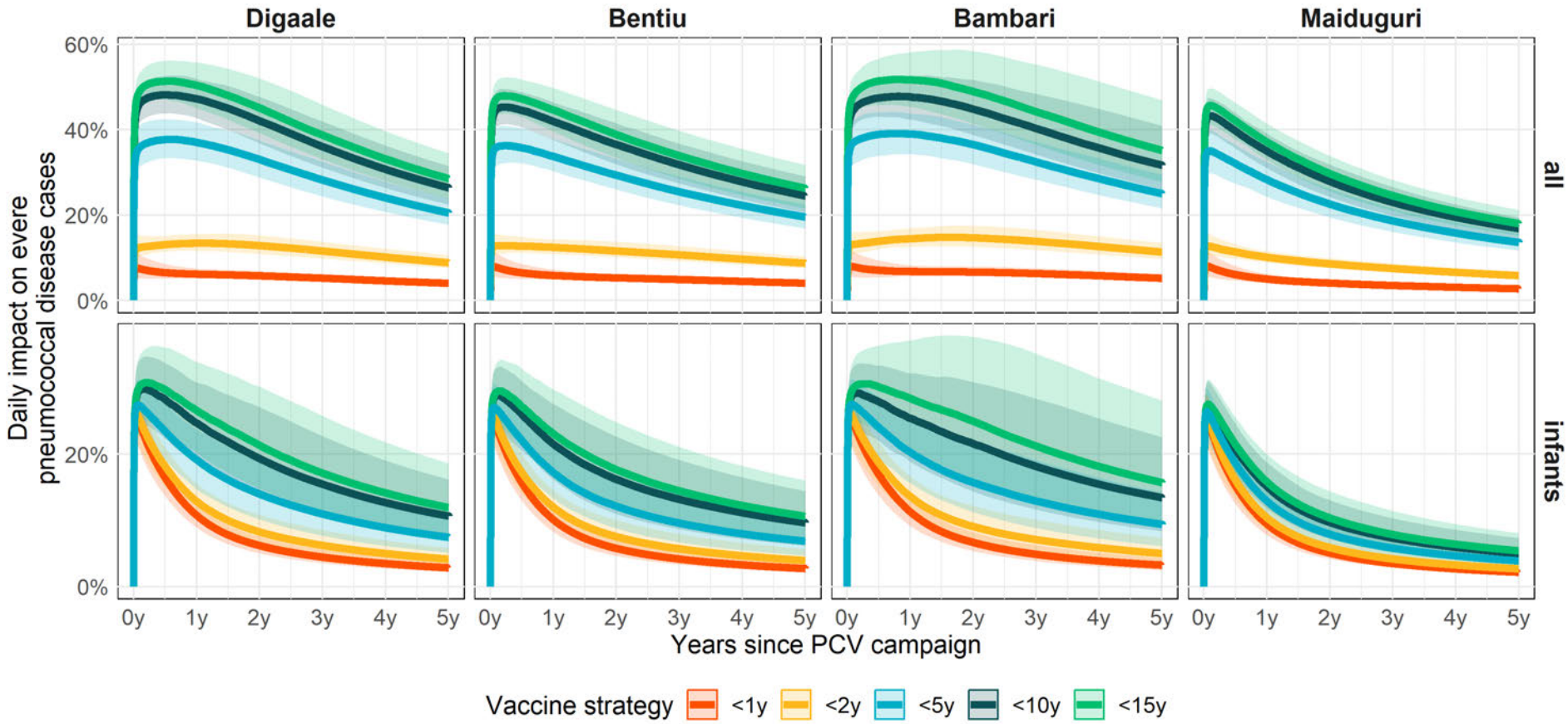


Figure 5. Cumulative impact of PCV campaigns in four settings. Plots show the cumulative impact over time in infants and all age groups for campaigns in Digaale, Bentiu, Bambari, and Maiduguri. In all plots, thick lines show the median estimates and shaded areas corresponding 95% credible intervals from 500 model iterations.

Discussion

We find that PCV campaigns could be an effective intervention to partially control pneumococcal transmission and prevent pneumococcal disease in crisis-affected populations where routine immunization is not feasible, and to sustain these reductions for up to two or three years. Campaigns that include children up to 5y or older are likely to achieve substantially better impact, in part because of substantial indirect protection that extends the impact to unvaccinated future birth cohorts. An <5y campaign was the most efficient use of PCV in the first year among the strategies considered, with the additional doses administered to children 2-5y being more effective in preventing severe pneumococcal disease than doses that would only be administered to infants and toddlers in strategies with narrower age targeting.

The large indirect effects estimated for strategies with extended age targeting can be explained by the high prevalence of pneumococci and contact interactions in these older age groups, as children aged 2-5y and 6-14y are the main drivers of transmission in Digaale.[11] Our results were consistent in all settings considered, although the overall impact was lower and NNV higher in settings with higher rates of mixing with an unvaccinated host population, or with high migration rates. In such settings, wider age targeting could partially mitigate the reduced impact of a campaign, as suggested by the WHO Framework for Decision-Making on Vaccination in Humanitarian Emergencies[47].

In our model, administering two doses in infancy resulted in a higher impact compared to the equivalent single-dose strategies. However, single-dose strategies where the MCV was extended up to the age of 5 or older resulted in greater and longer sustained overall impact on pneumococcal disease cases compared to two-dose strategies with narrower age targeting, while being a more efficient use of PCV per dose. It may be challenging and resource intensive to follow-up children for a second dose six weeks after their primary dose. Our results suggest that allocating resources towards vaccinating a larger group of children may prove to be more feasible, impactful, and cost-effective.

Gargano et al used static models to assess the cost-effectiveness of PCV10 immunization campaigns during humanitarian crises targeting children <1y in Somalia and <2y in a camp in South Sudan. They found PCV campaigns to be highly cost-effective using WHO cost effectiveness thresholds.[48,49] While results are not directly comparable, our results suggest that single-dose campaigns with extended age targeting up to 5y would be even more efficient and thus likely cost effective. Integration with a multi-antigen campaign, e.g. with measles, would further improve the cost-effectiveness of an MCV.

Increased vaccination coverage resulted in substantially greater impact without substantial loss off efficiency for the additional doses in <5y, <10y, and even <15y campaigns. While the feasibility of high PCV vaccine coverage may depend per setting, previous PCV campaigns in humanitarian settings have shown that high vaccine coverage for a single dose of >89% is feasible.[48,50]

In the absence of subsequent immunization activities or other changes, VT transmission will return to pre-vaccination baseline levels. If routine immunization cannot be established within 3-5 years after an MVC, alternative maintenance strategies need to be considered.

There are several limitations to our analysis. First, we defined pneumococcal serotypes as VT and NVT. This ignores differences between individual serotypes but is a common method that has been used to effectively understand PCV impact in other settings.[17,27,51,52] We assumed a similar VE as has been estimated for other PCV products, but there is few evidence on the VE on transmission of reduced-dose schedules, and none for Pneumosil. Overall, there is little data on pneumococcal transmission or disease burden for populations affected by humanitarian crises. While we collected key data from the Digaale IDP camp in Somaliland, we had to assume similar pneumococcal carriage prevalence in other settings and extrapolate the collected contact estimates. Digaale is a relatively small camp with a population not affected by an acute crisis, and we do not know how these estimates may differ in other populations. As there is no data on the pneumococcal disease burden in Digaale or other humanitarian settings, we had to extrapolate CCRs estimated in Kenya - a setting with similar

prevalence and age distribution of pneumococcal carriage as Digaale.[27] While we adjusted these estimates for differences in the prevalence of acute malnutrition, we do not know whether the presence of other risk factors may alter the invasiveness of VT and NVT serotypes in more fragile settings.

Overall, single dose PCV campaigns that achieve high coverage in children up to at least five years can effectively interrupt VT transmission and prevent pneumococcal disease in humanitarian settings for up to two or three years, while providing a feasible and pragmatic alternative if implementation of routine vaccination is not possible. If PCV campaigns are to be used regularly, 40 to 64 million doses may be needed to protect crisis-affected populations over the next five years.

Acknowledgements

We like to thank all study participants who dedicated their time to the cross-sectional survey and the staff conducting the survey. We also like to thank MSF for their participation in the inception of this work, and MCRI Translational Microbiology Group for advice and laboratory testing of samples.

Funding

The study was funded by Elrha's Research for Health in Humanitarian Crises (R2HC) Programme, which aims to improve health outcomes by strengthening the evidence base for public health interventions in humanitarian crises. SF was funded by a Sir Henry Dale Fellowship awarded jointly by the Royal Society and the Wellcome Trust (Grant number 208812/Z/17/Z)

Reference list

1 United Nations High Commissioner for Refugees. UNHCR - Refugee Statistics. https://www.unhcr.org/refugee-statistics/ (accessed March 2, 2023).

2 Troeger C, Blacker B, Khalil IA, *et al.* Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect Dis* 2018; **18**: 1191–210.

3 Bellos A, Mulholland K, O'Brien KL, Qazi SA, Gayer M, Checchi F. The burden of acute respiratory infections in crisis-affected populations: a systematic review. *Confl Health* 2010; **4**: 3.

4 van Zandvoort K, Checchi F, Diggle E, *et al.* Pneumococcal conjugate vaccine use during humanitarian crises. *Vaccine* 2019; **37**: 6787–92.

5 World Health Organization. Pneumococcal conjugate vaccines in infants and children under 5 years of age: WHO position paper – February 2019. *Wkly Epidemiol Rec* 2019; **94**: 85–104.

6 World Health Organization. Vaccination in Acute Humanitarian Emergencies: a framework for decision making. 2017. http://apps.who.int/iris/bitstream/handle/10665/255575/WHO-IVB-17.03-eng.pdf;jsessionid=2280CAAC607A17F83ABB56A52E9DA925?sequence=1 (accessed Nov 22, 2018).

7 Bellos A. Vaccination in Humanitarian Emergencies: Literature review and case studies. 2012. http://www.who.int/immunization/sage/meetings/2012/april/2_SAGE_WGVHE_SG1__Lit_Review_CaseStudies.pdf (accessed Nov 21, 2018).

8 Walldorf JA, Date KA, Sreenivasan N, Harris JB, Hyde TB. Lessons Learned from Emergency Response Vaccination Efforts for Cholera, Typhoid, Yellow Fever, and Ebola. *Emerg Infect Dis* 2017; **23**: S210–6.

9 Russell FM, Balloch A, Tang MLK, *et al.* Immunogenicity following one, two, or three doses of the 7-valent pneumococcal conjugate vaccine. *Vaccine* 2009; **27**: 5685–91.

10 Zandvoort K van, Hassan AI, Bobe M, *et al.* Pre-vaccination carriage prevalence of *Streptococcus pneumoniae* serotypes among internally displaced people in Somaliland. 2024; medRxiv: 2024.02.09.24302568.

11 van Zandvoort K, Bobe MO, Hassan AI, *et al.* Social contacts and other risk factors for respiratory infections among internally displaced people in Somaliland. *Epidemics* 2022; **41**: 100625.

12 Lipsitch M, Abdullahi O, D'Amour A, *et al.* Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal serotypes in Kenya with a Markov transition model. *Epidemiol Camb Mass* 2012; **23**: 510–9.

13 Verhagen LM, Gómez-Castellano K, Snelders E, *et al.* Respiratory infections in Eñepa Amerindians are related to malnutrition and *Streptococcus pneumoniae* carriage. *J Infect* 2013; **67**: 273–81.

14 Verhagen LM, Hermsen M, Rivera-Olivero IA, *et al.* Nasopharyngeal carriage of respiratory pathogens in Warao Amerindians: significant relationship with stunting. *Trop Med Int Health* 2017; **22**: 407–14.

15 Gebre T, Tadesse M, Aragaw D, *et al.* Nasopharyngeal Carriage and Antimicrobial Susceptibility Patterns of *Streptococcus pneumoniae* among Children under Five in Southwest Ethiopia. *Children* 2017; **4**: 27.

16 Flasche S, Ojal J, Le Polain de Waroux O, *et al.* Assessing the efficiency of catch-up campaigns for the introduction of pneumococcal conjugate vaccine: a modelling study based on data from PCV10 introduction in Kilifi, Kenya. *BMC Med* 2017; **15**: 113.

17 Ojal J, Griffiths U, Hammitt LL, *et al.* Sustaining pneumococcal vaccination after transitioning from Gavi support: a modelling and cost-effectiveness study in Kenya. *Lancet Glob Health* 2019; **7**: e644–54.

18 Hammitt LL, Etyang AO, Morpeth SC, *et al.* Effect of ten-valent pneumococcal conjugate vaccine on invasive pneumococcal disease and nasopharyngeal carriage in Kenya: a longitudinal surveillance study. *The Lancet* 2019; **393**: 2146–54.

19 Silaba M, Ooko M, Bottomley C, *et al.* Effect of 10-valent pneumococcal conjugate vaccine on the incidence of radiologically-confirmed pneumonia and clinically-defined pneumonia in Kenyan children: an interrupted time-series analysis. *Lancet Glob Health* 2019; **7**: e337–46.

20 Awori JO, PERCH Study Group. Association between malnutrition and pneumococcal pneumonia in the PERCH study. 2018.

21 von Mollendorf C, von Gottberg A, Tempia S, *et al.* Increased Risk for and Mortality From Invasive Pneumococcal Disease in HIV-Exposed but Uninfected Infants Aged <1 Year in South Africa, 2009–2013. *Clin Infect Dis* 2015; **60**: 1346–56.

22 von Mollendorf C, Cohen C, de Gouveia L, *et al.* Risk Factors for Invasive Pneumococcal Disease Among Children Less Than 5 Years of Age in a High HIV Prevalence Setting, South Africa, 2010 to 2012. *Pediatr Infect Dis J* 2015; **34**: 27.

23 Temple B, Tran HP, Dai VTT, *et al.* Efficacy against pneumococcal carriage and the immunogenicity of reduced-dose (0 + 1 and 1 + 1) PCV10 and PCV13 schedules in Ho Chi Minh City, Viet Nam: a parallel, single-blind, randomised controlled trial. *Lancet Infect Dis* 2023; **23**: 933–44.

24 Berman-Rosa M, O'Donnell S, Barker M, Quach C. Efficacy and Effectiveness of the PCV-10 and PCV-13 Vaccines Against Invasive Pneumococcal Disease. *Pediatrics* 2020; **145**: e20190377.

25 Le Polain De Waroux O, Flasche S, Prieto-Merino D, Goldblatt D, Edmunds WJ. The Efficacy and Duration of Protection of Pneumococcal Conjugate Vaccines Against Nasopharyngeal Carriage: A Meta-regression Model. *Pediatr Infect Dis J* 2015; **34**: 858–64.

26 Manfredi P, Williams JR. Realistic population dynamics in epidemiological models: the impact of population decline on the dynamics of childhood infectious diseases: Measles in Italy as an example. *Math Biosci* 2004; **192**: 153–75.

27 Flasche S, Ojal J, Le Polain de Waroux O, *et al.* Assessing the efficiency of catch-up campaigns for the introduction of pneumococcal conjugate vaccine: a modelling study based on data from PCV10 introduction in Kilifi, Kenya. *BMC Med* 2017; **15**. DOI:10.1186/s12916-017-0882-9.

28 Cohen C, Mollendorf C von, Gouveia L de, *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**: e359–69.

29 Qian G, Toizumi M, Clifford S, *et al.* Association of pneumococcal carriage in infants with the risk of carriage among their contacts in Nha Trang, Vietnam: A nested cross-sectional survey. *PLoS Med* 2022; **19**: e1004016.

30 Central Statistics Department, Ministry of Planning and National Development, Somaliland Government, Ministry of Planning and National Development, Somaliland Government. The Somaliland Health and Demographic Survey 2020. 2020.

31 United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects - Population Division - United Nations. 2019. https://population.un.org/wpp/Download/Standard/Population/ (accessed April 19, 2020).

32 Prem K, Zandvoort K van, Klepac P, *et al.* Projecting contact matrices in 177 geographical regions: An update and comparison with empirical data for the COVID-19 era. *PLOS Comput Biol* 2021; **17**: e1009098.

33 Checchi F, Warsame A, Treacy-Wong V, Polonsky J, Ommeren M van, Prudhon C. Public health information in crisis-affected populations: a review of methods and their use for advocacy and action. *The Lancet* 2017; **390**: 2297–313.

34 UN OCHA. South Sudan | OCHA. 2023; published online Dec 11. https://www.unocha.org/south-sudan (accessed Dec 18, 2023).

35 UN OCHA. Nigeria | OCHA. 2023; published online Nov 21. https://www.unocha.org/nigeria (accessed Dec 18, 2023).

36 Center for Preventative Action. Conflict in the Central African Republic. Glob. Confl. Tracker. https://cfr.org/global-conflict-tracker/conflict/violence-central-african-republic (accessed Dec 18, 2023).

37 IOM. Displacement Tracking Matrix. https://dtm.iom.int/ (accessed Dec 18, 2023).

38 UN OCHA. Situation Reports. https://reports.unocha.org/ (accessed Dec 18, 2023).

39 South Sudan: Cumulative Registration Data, January 2014 - August 2015 - South Sudan | ReliefWeb. 2015; published online Aug 31. https://reliefweb.int/report/south-sudan/south-sudan-cumulative-registration-data-january-2014-august-2015 (accessed April 16, 2024).

40 Checchi F, Testa A, Warsame A, Bs LQ, Burns R. Estimates of crisis-attributable mortality in South Sudan, December 2013-April 2018. 2018.

41 UN OCHA. Northeast Nigeria: Humanitarian emergency - Situation Report No. 1. 2016; published online Nov 28. https://www.unocha.org/publications/report/nigeria/northeast-nigeria-humanitarian-emergency-situation-report-no-1-28-november-2016 (accessed April 16, 2024).

42 South Sudan National Bureau of Statistics (NBS). National Baseline Household Survey 2009 Report. 2012.

43 Concern. Bentiu PoC - Nutrition Anthropometry & Retrospective Mortality Survey Report. 2015. https://info.undp.org/docs/pdc/Documents/SSD/20160124-180405_CWW%20-%20Bentiu%20PoC%20-%20Smart%20Survey%20-%20August%202015.pdf (accessed April 16, 2024).

44 Action Against Hunger. Report of Small-Scale SMART Survey in MMC, Jere LGAs, Borno. 2016.

45 Nigeria National Bureau of Statistics. Nutrition and food security surveillance: north east Nigeria - Emergency survey. 2019. https://fscluster.org/sites/default/files/documents/nfss_round_8_final_report_november_2019.pdf (accessed April 16, 2024).

46 Unicef. Multiple Indicator Cluster Surveys. https://mics.unicef.org/surveys (accessed April 16, 2024).

47 Rull M, Masson S, Peyraud N, *et al.* The new WHO decision-making framework on vaccine use in acute humanitarian emergencies: MSF experience in Minkaman, South Sudan. *Confl Health* 2018; **12**: 11.

48 Gargano LM, Hajjeh R, Cookson ST. Pneumonia prevention: Cost-effectiveness analyses of two vaccines among refugee children aged under two years, *Haemophilus influenzae type b*-containing and pneumococcal conjugate vaccines, during a humanitarian emergency, Yida camp, South Sudan. *Vaccine* 2017; **35**: 435–42.

49 Gargano LM, Hajjeh R, Cookson ST. Pneumonia Prevention during a Humanitarian Emergency: Cost-effectiveness of *Haemophilus Influenzae* Type B Conjugate Vaccine and Pneumococcal Conjugate Vaccine in Somalia. *Prehospital Disaster Med* 2015; **30**: 402–11.

50 Peyraud N, Quéré M, Duc G, *et al.* A post-conflict vaccination campaign, Central African Republic. *Bull World Health Organ* 2018; **96**: 540–7.

51 Le Polain De Waroux O, Edmunds WJ, Takahashi K, *et al.* Predicting the impact of pneumococcal conjugate vaccine programme options in Vietnam. *Hum Vaccines Immunother* 2018; **14**: 1939–47.

52 Choi YH, Jit M, Flasche S, Gay N, Miller E. Mathematical Modelling Long-Term Effects of Replacing Prevnar7 with Prevnar13 on Invasive Pneumococcal Diseases in England and Wales. *PLOS ONE* 2012; **7**: e39927.

Supplemental Material for

Effectiveness of pneumococcal vaccination campaigns in humanitarian settings: a modelling study

Kevin van Zandvoort, Mohamed O Bobe, Abdirahman I Hassan, Rachael Cummings, Abdihamid Warsame, Casey L Pell, Mohamed Ismail Abdi, Mohamed Abdi Hergeye, Catherine R McGowan, E Kim Mulholland, Rosalind M Eggo, Catherine Satzke, Francesco Checchi, Stefan Flasche

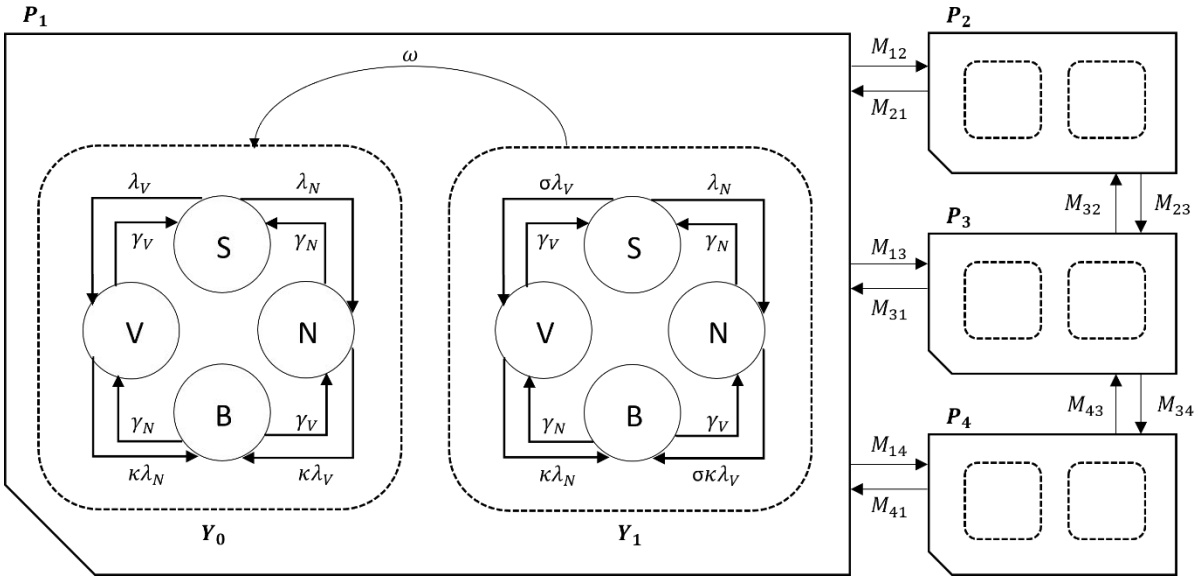
* Corresponding author: Kevin van Zandvoort, Kevin.Van-Zandvoort@lshtm.ac.uk

Analysis scripts and model code are available on GitHub: https://github.com/kevinvzandvoort/espicc-metavax-pcv-humanitarian-crises. These can be used to conduct all simulations and analyses in both the main manuscript and this Supplemental Material.

Table of Contents

Section A. Transmission model

Model structure and equations



Supplemental Figure A1. Compartmental model structure of the pneumococcal transmission model. Pneumococcal serotypes defined as those included in the PCV vaccine (vaccine-type, VT), or not included in the vaccine (non-vaccine type, NVT). At any time *t*, Susceptible individuals (S) acquire serotypes in either group at rates $\lambda_V(t)$ and $\lambda_N(t)$, and immediately become infectious carriers themselves. Carriers of VTs or NVTs remain partially susceptible to infection with serotypes in the other group, and may develop a superinfection (B) at rates $\kappa\lambda_V(t)$ and $\kappa\lambda_N(t)$, where κ represents competition between serotypes. Carriers are assumed to clear their infection of all serotypes in either group at rates γ_V for VTs and γ_N for NVTs. Individuals may be without vaccine derived protection (Y_0), or with full vaccine derived protection (Y_1). Those with vaccine protection acquire VT at a reduced rate *σ*, and lose their level of protection at rate ω. Different populations P_1, P_2, and P_3, P_4 represent malnourished and non-malnourished strata of a displaced and host population. Individuals of population *p* migrate to population *q* at rate M_{qp}. Age groups, ageing, and transmission between vaccine protection strata, populations, and age groups are not shown. Note, Figure 1 in the main manuscript does not show the stratification of each population by malnutrition for improved clarity, which is why the number of populations in this figure differs.

Modelled populations

We constructed two model populations for every population in a scenario, to stratify these populations by the proportion with acute malnutrition. In our mathematical notation, we will refer to these as the model population (p) and real population (r). For example, in our base scenario in Digaale IDP camp, we created four model populations ($N_{\text{pop}} = 4$) to represent two real populations ($N_{\text{real}} = 2$): a non-malnourished (p=1) and malnourished (p=2) population for the displaced population living in Digaale IDP camp (r=1), and a non-

malnourished (p=3) and malnourished (p=4) population for the host population in the city of Hargeisa (r=2). All populations (p and r) have the same number of age groups (N_{age}).

Supplemental Table A1. Modelled populations in the base model.

Real population index (r)	Model population index (p)	Description
1	1	Digaale – non malnourished
1	2	Digaale – malnourished
2	3	Hargeisa – non malnourished
2	4	Hargeisa - malnourished

We calculated the population size for age group *a* in model population p (P_a^p) by combining the population size for age group *a* in real population *r* (R_a^r) with the proportion of malnourished individuals in real population *r* (χ_r)*:*

$$P_a^{2(r-1)} = R_a^r(1-\chi_r) \quad 1$$

$$P_a^{2r} = R_a^r\chi_r \quad 2$$

Every model population is further stratified in two vaccinated strata (Y; $N_{\text{vac}} = 2$), where y=0 indicates the unvaccinated stratum, and $y = 1$ the vaccinated stratum, and in four epidemiological compartments: S (Susceptible), V (VT carrier), N (NVT carrier) and B (carrier of both VT and NVT). The values for epidemiological compartment $X_a^{p,y}$, where $X \in \{S, V, N, B\}$ (i.e. $S_a^{p,y}$, $V_a^{p,y}$, $N_a^{p,y}$, $B_a^{p,y}$), then hold the proportion of people in age group *a* in population *p*, that are in compartment X in vaccinated stratum *y*. We specifically model proportions, where all values for model compartments within a single age group in a model population sum to 1, i.e.

$$\sum_{y=0}^{N_{\text{vac}}} \left(S_a^{p,y} + V_a^{p,y} + N_a^{p,y} + B_a^{p,y}\right) = 1 \quad 3$$

Supplemental Table A2. Parameters used in the model equations.

Parameter	Description
N_{real}	Number of 'real populations' (e.g., a displaced or host population)
N_{pop}	Number of model populations
N_{age}	Number of age groups in model
r	Index of real population
p, q	Index of model population

a,b	Index of age group
R_a^r	Population size in age group a in real population r
P_a^p	Population size in age group a in model population p
χ_r	Proportion acutely malnourished in real population r.
N_{vac}	Number of vaccinated strata in model
y, z	Index of vaccinated stratum
t	Time
$X_a^{p,y}(t)$	Value of epidemiological compartment at time t; proportion of age group a in population p that is in epidemiological compartment X in vaccinated stratum y, where $X \in \{S, V, N, B\}$
$S_a^{p,y}(t)$	Epidemiological compartment for susceptibles
$V_a^{p,y}(t)$	Epidemiological compartment for VT carriers
$N_a^{p,y}(t)$	Epidemiological compartment for NVT carriers
$B_a^{p,y}(t)$	Epidemiological compartment for carriers of both VT and NVT
$\Delta D_{X_a^{p,y}}(t)$	Change in the demography of compartment $X_a^{p,y}$ at time t
α_a	Daily rate at which individuals age to the next age stratum.
l_a	Life expectance for age group a, i.e. the number of years an individual is expected to remain in age group a.
M_{qp}	Migration matrix: daily per capita rate at which individuals migrate out of model population p and into model population q
J_a^{qp}	Daily per capita rate at which an individual in age group a in population p migrates into population q
μ_a^p	Daily rate at which individuals in age group a migrate out of model population p
$\omega_a^{p,y}$	Daily rate at which vaccine protection wanes for individuals of age a in vaccinated stratum y in population p
$\Delta E_{X_a^{p,y}}(t)$	Change in the epidemiology of compartment $X_a^{p,y}$ at time t
$\sigma_a^{p,y}$	Relative risk (1 – vaccine efficacy) of acquiring VT serotypes compared to unvaccinated individuals for individuals of age a in vaccinated stratum y in population p
$\lambda_{V_a}^p(t)$	Force of infection of VT on people in age group a in population p, at time t
$\lambda_{N_a}^p(t)$	Force of infection of NVT on people in age group a in population p, at time t
γ_{V_a}	Daily rate at which a VT carrier in age group a clears VT pneumococci
γ_{N_a}	Daily rate at which an NVT carrier in age group a clears NVT pneumococci
κ	Competition parameter
$\beta_a^{p,V}$	Combined fitted values that adjust the force of infection of VT on people in age group a in population p
$\beta_a^{p,N}$	Combined fitted values that adjust the force of infection of NVT on people in age group a in population p
β_a	Probability of transmission per contact with an infectious individual carrying NVT for different age groups (a)
$\beta_{VT/NVT,a}$	Rate ratio for probability of transmission with contacts carrying VT, compared to NVT in age group a.
β_m	Rate ratio for probability of transmission in malnourished compared to non-malnourished individuals.
c_{ba}^p	Average number of daily contacts made by a contactor in age group a in population p with contactees in age group b
o^p	Proportion of daily contacts made by a contactor in population p that are made outside of the household
T_{qp}	Travel matrix: proportion of extra-household contacts made by a contactor in population p with contactees in population q

We separated our model equations in those that model the changes in demography between different strata, which is the same for each compartment X, and those that model the epidemiological transitions between the model compartments, which is the same for each stratum $X_a^{p,y}$.

Demographic transitions

We first model the change in the demography for each compartment $X_a^{p,y}$ as $\Delta D_{X_a^{p,y}}(t)$:

$$\begin{aligned}\Delta D_{X_a^{p,y}}(t) = &-\left(\alpha_a + \mu_a^p + \omega_a^{p,y} - \alpha_a\left(\mu_a^p + \omega_a^{p,y}\right) - \mu_a^p\omega_a^{p,y} - \alpha_a\mu_a^p\omega_a^{p,y}\right)X_a^{p,y}(t) \\ &+ \alpha_a\left(1-\mu_a^p\right)\left(1-\omega_{a-1}^{p,y}\right)X_{a-1}^{p,y}(t) \\ &+ \delta_{y,0}\left(1-\mu_a^p\right)\sum_{z=1}^{N_{\text{vac}}}(1-\alpha_a)\omega_a^{p,z}X_a^{p,z}(t) + \alpha_a\omega_a^{p,z}X_{a-1}^{p,z}(t) \\ &+ \sum_{q=1}^{N_{\text{pop}}} J_a^{qp}\left((1-\alpha_a)\left(1-\omega_a^{q,y}\right)X_a^{q,y}(t) + \alpha_a\left(1-\omega_{a-1}^{q,y}\right)X_{a-1}^{q,y}(t)\right) \\ &+ \delta_{y,0}\sum_{q=1}^{N_{\text{pop}}}\sum_{z=1}^{N_{\text{vac}}} J_a^{qp}\left((1-\alpha_a)\omega_a^{q,z}X_a^{q,z}(t) + \alpha_a\omega_a^{q,z}X_{a-1}^{q,z}(t)\right)\end{aligned} \quad 4$$

- α_a is the daily rate at which people age out of age group a. We apply a negative exponential rate to incorporate ageing, calculated as

$$\alpha_a = \frac{1}{365.25 l_a} \quad 5$$

- where l_a is the life expectancy or size for age group *a* in years, i.e. the number of years an individual should remain in age group *a*. For instance, if an individual in age group *a* is 2-3mo: $l_a = \frac{1}{12}$, while if they are 5-9yo: $l_a = 4$.
- μ_a^p is the daily rate at which people of age a migrate out of population *p*. It is calculated as

$$\mu_a^p = \sum_{q=1}^{N_{\text{pop}}} J_a^{qp} \quad 6$$

- where J_a^{qp} is the daily per capita migration rate at which individual in age group a migrate from population *p* to population *q*.
- $\omega_a^{q,z}$ is the rate at which vaccinated individuals lose vaccine protection and return to compartments in vaccinated stratum $y = 0$.
- $\delta_{y,0}$ denotes the Kronecker delta function which evaluates to 1 when $y = 0$, and evaluates to 0 for all other values of *y*.

The different elements of equation 4 can then be read as

- All those who leave compartment $X_a^{p,y}$ due to ageing, migration, or loss of vaccine protection, accounting for competing rates to keep population sizes consistent.
- Those who age into compartment $X_a^{p,y}$ from $X_{a-1}^{p,y}$. Note that newborns are a special case: when $X_a^{p,y} = S_0^{p,y}$, this term changes to $\alpha_a(1 - \mu_a^p)(1 - \omega_{a-1}^{p,y})$, while it changes to 0 when $X_a^{p,y} \neq S_0^{p,y}$(i.e. all newborns are born susceptible).
- Those from model population p who move back into compartment $X_a^{p,y}$due to loss of vaccine protection (only when $y = 0$).
- Those from other model populations q who migrate into compartment $X_a^{p,y}$.
- And those from other model populations q who lose their vaccine protection and simultaneously migrate into compartment $X_a^{p,y}$ (only when $y = 0$).

Note that the α_a rates applied to the $X_{a-1}^{p,y}$ compartments and the J_a^{qp} rates applied to the $X_a^{q,y}$ compartments are not typos; they adjust for any differences between l_a and l_{a-1} or between $P_a^p$and P_a^q, and thereby implicitly reweigh the values for compartment $X_a^{p,y}$.

Epidemiological transitions

With the demographic changes incorporated, we can then model the change in the epidemiology for each compartment $X_a^{p,y}$ as $\Delta E_{X_a^{p,y}}(t)$. To do so, we want to use the updated values accounting for the changes in demography:

$$\overline{X}_a^{p,y}(t) = X_a^{p,y}(t) + \Delta D_{X_a^{p,y}}(t) \quad 7$$

We can then calculate the epidemiological transitions. For ease of reading, we have omitted the time notations in the following set of equations. We have also given the notations for the age group, population, and vaccinated stratum a lighter colour:

$$\Delta E_{S_a^{p,y}}(t) = -\left(\sigma_a^{p,y}\lambda_{V\,a}^{\;p} + \lambda_{N\,a}^{\;p}\right)\overline{S}_a^{p,y} + \gamma_{V\,a}\overline{V}_a^{p,y} + \gamma_{N\,a}\overline{N}_a^{p,y}$$

$$\Delta E_{V_a^{p,y}}(t) = \sigma_a^{p,y}\lambda_{V\,a}^{\;p}\overline{S}_a^{p,y} - \left(\kappa\lambda_{N\,a}^{\;p} + \gamma_{V\,a}\right)\overline{V}_a^{p,y} + \gamma_{N\,a}\overline{B}_a^{p,y}$$

$$\Delta E_{N_a^{p,y}}(t) = \lambda_{N\,a}^{\;p}\overline{S}_a^{p,y} - \left(\sigma_a^{p,y}\kappa\lambda_{V\,a}^{\;p} + \gamma_{N\,a}\right)\overline{N}_a^{p,y} + \gamma_{V\,a}\overline{B}_a^{p,y}$$

$$\Delta E_{B_a^{p,y}}(t) = \kappa\lambda_{N\,a}^{\;p}\overline{V}_a^{p,y} + \sigma_a^{p,y}\kappa\lambda_{V\,a}^{\;p}\overline{N}_a^{p,y} - \left(\gamma_{V\,a} + \gamma_{N\,a}\right)\overline{B}_a^{p,y}$$

$$\lambda_{V\,a}^{\;p} = \beta_a^{p,V}\sum_{b=1}^{N_{\text{age}}}\sum_{q=1}^{N_{\text{pop}}}\left(c_{ba}^{p}\left(\delta_{q,p}(1-o^p) + o^p T_{qp}\right)\sum_{y=0}^{N_{\text{vac}}}\left(\overline{V}_b^{q,y} + \overline{B}_b^{q,y}\right)\right)$$

$$\lambda_{N\,a}^{\;p} = \beta_a^{p,N}\sum_{b=1}^{N_{\text{age}}}\sum_{q=1}^{N_{\text{pop}}}\left(c_{ba}^{p}\left(\delta_{q,p}(1-o^p) + o^p T_{qp}\right)\sum_{y=0}^{N_{\text{vac}}}\left(\overline{N}_b^{q,y} + \overline{B}_b^{q,y}\right)\right) \quad 8$$

- $\sigma_a^{p,y}$is the relative rate (1 – vaccine efficacy against carriage) that scales the force of infection for individuals in age group a, vaccine stratum y, and population p.
- $\gamma_{V\,a}$ and $\gamma_{N\,a}$are the rates at which VT and NVT carriers of age a clear their infection of that serotype.
- κ is the competition parameter and determines the degree by which acquisition of heterologous carriage is reduced in carriers of only VT or NVT compared to susceptible individuals.

- $\beta_a^{p,V}$ and $\beta_a^{p,N}$ are calculated from fitted parameters. They are the probabilities of transmission per contact by an individual in age group a with infectious individuals carrying VT or NVT respectively. They are calculated as:

$$\beta_a^{p,N} = \begin{cases} \beta_a \beta_m, & p \bmod 2 = 0 \\ \beta_a, & p \bmod 2 \neq 0 \end{cases} \quad 9$$

  - where β_a is the fitted probability of transmission per contact by an individual in age group a with infectious individual carrying NVT, and $\beta_m$is the rate ratio for this probability of transmission for malnourished compared to non-malnourished individuals;

  and

$$\beta_a^{p,V} = \beta_a^{p,N} \beta_{VT/NVT,a} \quad 10$$

  - where $\beta_{VT/NVT,a}$ is the fitted rate ratio for the probability of transmission with contacts carrying VT, compared to NVT, in age group *a*.

  See supplemental section B for more details about fitted model parameters.

- c_{ba}^{p} are the average number of daily contacts made by an individual in age group *a* in population *p*, with contactees in age group *b*. Within age groups, contact rates are assumed to be homogenous irrespective of the individual's vaccine protection or malnourished stratum.
- o^p is the proportion of contacts made by contactors in population *p* that take place outside of the household.
- $\delta_{q,p}$ denotes the Kronecker delta function which evaluates to 1 when $q = p$, and evaluates to 0 for all other values of *q*.
- T_{qp} is the proportion of extra-household contacts that are made by a contactor in population *p* with contactees in population *q*.

We then combine the demographic and epidemiological transitions, and calculate the change for each model compartment at time *t* as:

$$\frac{dX_a^{p,y}}{dt} = \Delta D_{X_a^{p,y}}(t) + \Delta E_{X_a^{p,y}}(t) \quad 11$$

Model parameter values

Extrapolating carriage to disease

For each age group and population, we first calculated the incidence of all VT and NVT acquisitions at time *t* as:

$$\text{acquisitions}_{\text{VT}_a}^{p,y}(t) = P_a^p\left(\sigma_a^{p,y}{\lambda_V}_a^p(t)\overline{S}_a^{p,y}(t) + \sigma_a^{p,y}\kappa{\lambda_V}_a^p(t)\overline{N}_a^{p,y}(t)\right)$$

$$\text{acquisitions}_{\text{NVT}_a}^{p,y}(t) = P_a^p\left({\lambda_N}_a^p(t)\overline{S}_a^{p,y}(t) + \kappa{\lambda_N}_a^p(t)\overline{V}_a^{p,y}(t)\right) \quad 12$$

We then applied age- and serotype specific case-carrier ratios (CCR) sampled from a posterior distribution estimated in Kilifi, Kenya by Flasche et al.[1].

We assumed that CCRs in malnourished individuals were increased by rate-ratio φ compared to non-malnourished individuals. To also account for a relative increase in the acquisition of pneumococci in malnourished individuals, β_m, we calculated φ so that the overall combined risk of IPD (increased risk in acquisition and disease) in malnourished individuals was increased by rate-ratio τ compared to non-malnourished individuals. as:

$$\varphi = \frac{\tau}{\beta_m} \quad 13$$

We then calculated the CCR for serotype *s* and age group *a* in non-malnourished individuals ($\text{CCR}_a^{s,0}$) as:

$$\text{CCR}_a^{s,0} = \frac{\text{CCR}_a^s}{\chi_0(\varphi - 1) + 1} \quad 14$$

Where CCR_a^s is a sampled CCR value (combined for malnourished and non-malnourished individuals) for serotype s (VT or NVT) and age group *a*, and χ_0 is the proportion of the population in Kilifi ($r = 0$) that was malnourished. In the baseline scenario, we assume that $\tau = 2.0$[2–4] and $\beta_m = 1.2$[5–7]. Furthermore, we assume that $\chi_0 = 0.182$[8].

The CCR for malnourished individuals was then calculated as:

$$\text{CCR}_a^{s,1} = \varphi \text{CCR}_a^{s,0} \quad (15)$$

These adjusted CCRs was then applied to each respective population, effectively scaling the estimated CCRs from Kilifi to settings with different rates of malnutrition.

Vaccine efficacy against disease

PCVs protect against both transmission and disease. Like we did for the rate-ratio φ for malnourished individuals, we calculate the vaccine efficacy (VE) against disease ($1 - \theta_a^{p,y}$) used in our model, given the VE against carriage ($1 - \sigma_a^{p,y}$) and the combined VE against carriage and disease ($1 - \phi_a^{p,y}$), which is often reported in studies of VE:

$$1 - \theta_a^{p,y} = 1 - \frac{\phi_a^{p,y}}{\sigma_a^{p,y}} \quad (16)$$

IPD cases

Total IPD cases at time *t* for age group *a* in vaccine stratum *y* in population *p* were then calculated as:

$$\text{cases}_{\text{VT}}{}_a^p(t) = \begin{cases} \text{acquisitions}_{\text{VT}}{}_a^{p,y}(t)\text{CCR}_a^{V,1}\theta_a^{p,y}, & p \bmod 2 = 0 \\ \text{acquisitions}_{\text{VT}}{}_a^{p,y}(t)\text{CCR}_a^{V,0}\theta_a^{p,y}, & p \bmod 2 \neq 0 \end{cases}$$

$$\text{cases}_{\text{NVT}}{}_a^p(t) = \begin{cases} \text{acquisitions}_{\text{NVT}}{}_a^{p,y}(t)\text{CCR}_a^{N,1}, & p \bmod 2 = 0 \\ \text{acquisitions}_{\text{NVT}}{}_a^{p,y}(t)\text{CCR}_a^{N,0}, & p \bmod 2 \neq 0 \end{cases} \quad (17)$$

In Kenya, it has been estimated that 4.98 cases of non-bacteraemic pneumonia are prevented for every case of IPD prevented.[9,10] We followed Ojal et al and calculated the total number of severe pneumococcal disease cases (the total number of IPD and non-bacteraemic pneumonia cases) as (1 + 4.98) times the estimated number of IPD cases.[11]

Duration of carriage

Lipsitch et al. estimated the age- and serotype-specific duration of carriage in children 0-22mo, 22-41mo, and 41-59mo from longitudinal data collected in Kenyan children.[12] Age-specific estimates for the 14 most common serotypes were provided in Figure 3 of their paper. We extracted these data and fitted negative binomial distributions to each set of values to recover their distributions.

For each serotype *s*, in each age *a*, we assume that the duration of carriage can be described by a negative binomial distribution with mean $\mu_{s,a}$ and dispersion parameter $k_{s,a}$.

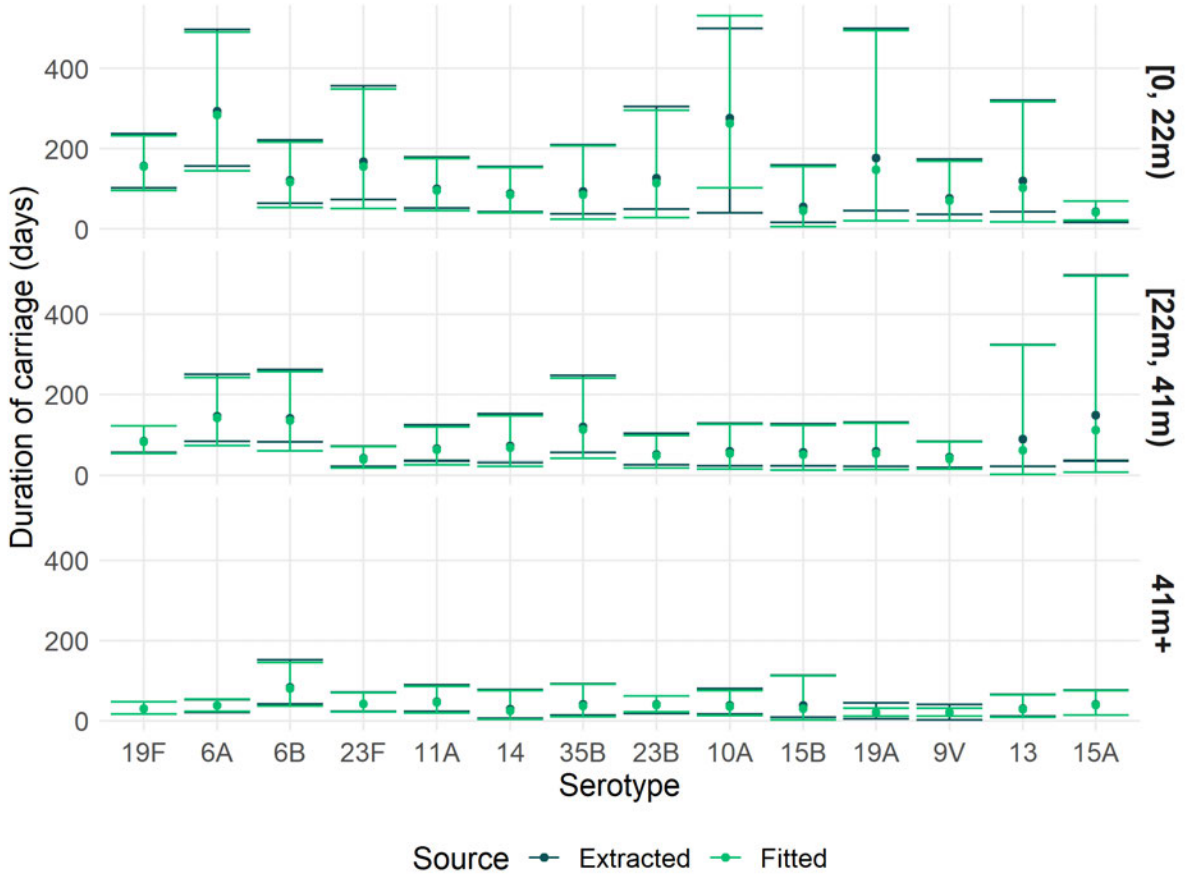
$$d_{s,a} \sim \text{nbinom}\left(\mu_{s,a}, k_{s,a}\right) \quad 18$$

We parameterized the mean value for $d_{s,a}$ as $\mu_{s,a} = y_{s,a,0.5}$, where $y_{s,a,0.5}$ is the extracted median estimate for serotype *s* in dataset *a.* We then optimized the value for $k_{s,a}$ using the *optimize* function in base *R* to minimize $g_{(d_{s,a})}$: the log-least squares of the difference between the extracted boundaries of the 95% confidence interval for $d_{s,a}$: $y_{s,a,0.025}$ and $y_{s,a,0.975}$, and corresponding quantile distribution for the negative binomial distribution evaluated at 0.025 and 0.975. We took the log of the values to minimize the large difference in scale between the lower and upper boundary value.

$$g\left(\sigma_{s,a}\right) = \log\left(\left(y_{s,a,0.025} - Q_{s,a}(0.025)\right)^2 + \left(y_{s,a,0.975} - Q_{s,a}(0.975)\right)^2\right) \quad 19$$

Where $Q_{s,a}(p)$ is the quantile distribution of the negative binomial distribution used to describe $d_{s,a}$ evaluated at *p.*

Supplemental Figure A2 shows the fitted negative binomial distributions against the extracted values estimated by Lipsitch et al. The 95% quantile values of the fitted distributions were in broad agreement with the reported 95% confidence intervals, with only some minor discrepancies at primarily the lower tails of some distributions.

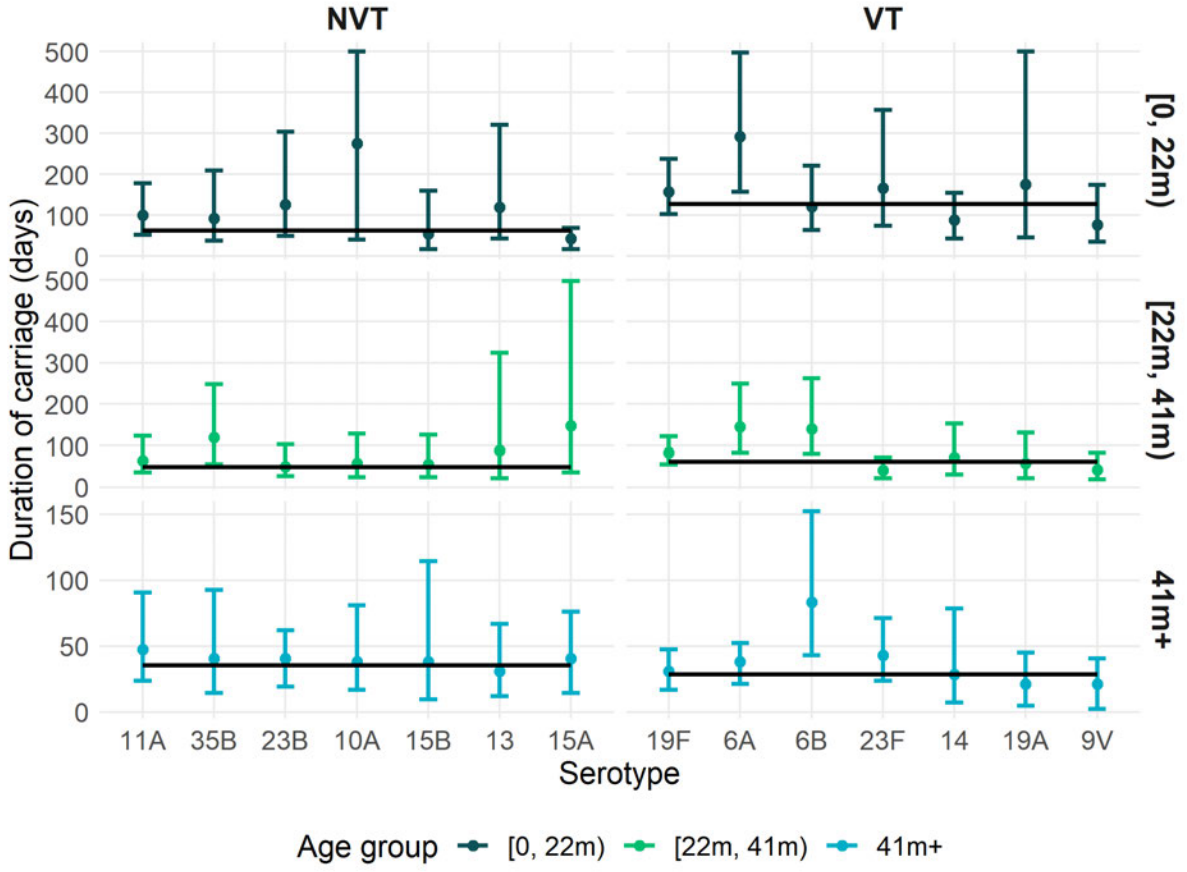


Supplemental Figure A2. Fitted distributions of age- and serotype-specific duration of carriage.

We then estimated 6 parameters independently to pool the average duration of carriage for VT and NVT in the three age groups, accounting for the uncertainty in the underlying distributions. We did so by finding the value of $\overline{\mu}_{\overline{s},a}$ that maximized the summed negative log-density of all fitted distributions $d_{s,a}$ maximized by the *optim* function, where the serotype was a VT or NVT (based on their inclusion in the Pneumosil vaccine).

The best fitting values for $\overline{\mu}_{\overline{s},a}$, for each age group and serotype group are described in Supplemental Table A3 and plotted against the serotype-specific distributions in Supplemental Figure A3.

For the duration of carriage in older people, we followed Flasche et al[1] and assumed that these were 36% lower for people aged 60m+ compared to those aged 24-59m, as estimated by Melegaro et al[13].



Supplemental Figure A3. Pool estimates of the duration of carriage for VT and NVT by age.

Supplemental Table A3. Estimated duration of carriage (days) for VT and NVT serotypes by age.

Age group	VT	NVT
[0, 22m)	125	61
[22m, 41m)	67	50
[41m, 59m)	28	34
≥ 60m	22	24

Transmission between populations

We used a travel matrix to inform extra-household contacts and thus transmission between populations. For the displaced populations (r=1, p≤2), we assumed the same values for all age groups, informed by the value in the travel matrix (T_{qp}). These values were themselves calculated as the proportion of extra-household contacts made by an individual in model population *p* with individuals in real population *r*, multiplied by the proportion of real population *s* that is in model population *q*. The latter implicitly assumes that mixing between malnourished and non-malnourished individuals is at random.

$$\mathrm{T}_a^{qp} = \begin{cases} \dfrac{\sum_{b=1}^{N_{\text{age}}} P_b^q}{\sum_{b=1}^{N_{\text{age}}} R_b^{\lceil q/2 \rceil}} T_{qp}, & p \leq 2 \\ \dfrac{\sum_{b=1}^{N_{\text{age}}} P_b^p}{\sum_{b=1}^{N_{\text{age}}} R_b^2} T_{pq} \dfrac{o^q \sum_{b=1}^{N_{\text{age}}} P_b^q c_{ab}^q}{o^p P_a^p \sum_{b=1}^{N_{\text{age}}} c_{ba}^p}, & p > 2 \end{cases} \quad 20$$

Migration between populations

For the displaced populations (r=1, p≤2) we assume the same migration rate for all age groups, informed by the value in the migration matrix (M_{qp}). For the host populations (r=2, p>2), the migration rate was calculated by the values in the displaced populations, adjusted for any difference in the age distributions and population sizes to ensure a fixed population size in all R_a^r:

$$J_a^{qp} = \begin{cases} M_{qp}, & p \leq 2 \\ M_{pq} \dfrac{P_a^q}{P_a^p}, & p > 2 \end{cases} \quad 21$$

Individuals in the malnourished stratum in the displaced population only migrate to the malnourished stratum in the host population, and vice versa.

Extrapolating contact estimates to other settings

As we only had observed contact estimates for Digaale, we adjust the observed contact matrices to other settings as follows:

Matrix $\boldsymbol{C_{pr}}$ is a contact matrix in population *p* and relationship *r*, where value $c_{pr,ij}$ is the average number of daily contacts between a contactor in population *p* aged *i*, and contactees aged *j* with relationship *r*.

We assume that $c_{pr,ij}$ can be split in three elements:

$$c_{pr,ij} = h_{p,j} q_{pr,ij} d_{pr,i} \qquad 22$$

Here, value $h_{p,j}$ is the expected proportion of all contacts made by a contactor in population *p* aged *i*, that are with contactees aged *j*, if mixing would be completely homogeneous and at random. It is the same for all contactor age groups *i*, and equal to the proportion of people in population *p* who are of age *j*, i.e. $h_{p,j} = \frac{n_{p,j}}{\sum_{x=1}^{a} n_{p,x}}$ where $n_{p,j}$ is the total number of people in population *p* who are of age *j*, and *a* is the total number of age groups.

Value $q_{pr,ij}$ is the relative assortativity of contacts between contactors in population *p* aged *i*, and contactees aged *j* with relationship *r*. It is calculated as the relative difference between the real and expected (assuming homogeneous mixing) proportion of contacts between contactors in population *p* aged *i*, and contactees aged *j*, i.e. $q_{pr,ij} = \frac{\left(\frac{c_{pr,ij}}{\sum_{x=1}^{a} c_{pr,ix}}\right)}{h_{p,j}}$. Values between 0 and 1 would represent age-dissasortative mixing, a value of 1 would represent perfect homogeneous mixing, and values greater than 1 would represent age-assortative mixing.

Finally, value $d_{pr,i}$ represents the total average number of daily contacts made by a contactor in population *p* aged *i*, made with all contactees aged *j* and relationship *r*, i.e. $d_{pr,i} = \sum_{x=1}^{a} c_{pr,ix}$

To extrapolate the estimated contact rates from Digaale (p=0) to populations in other settings, without any additional information, we assumed that $q_{pr,ij}$and $d_{0r,i}$ remain the same as in Digaale, but update $h_{p,j}$ to the respective values for population *p*, and calculate the new contact rates as

$$c_{pr,ij} = h_{p,j} q_{0r,ij} d_{0r,i} \qquad 23$$

Note that we can rewrite this expression as:

$$\begin{aligned} c_{pr,ij} &= h_{p,j} q_{0r,ij} d_{0r,i} \\ &= h_{p,j} \frac{\left(\frac{c_{0r,ij}}{\sum_{x=1}^{a} c_{0r,ix}}\right)}{h_{0,j}} \sum_{x=1}^{a} c_{0r,ix} \\ &= c_{0r,ij} \frac{h_{p,j}}{h_{0,j}} \end{aligned} \qquad 24$$

It is worth noting that, as matrix $\boldsymbol{C_{0r}}$ was adjusted for reciprocity of contacts so that the total number of contacts made between contactors/contactees and contactees/contactors aged *i* and *j* with relationship *r* are equal (i.e. $\boldsymbol{c_{0r,ij} n_{0,i} = c_{0r,ji} n_{0,j}}$), the same applies for the adjusted contact rates in population *p*:

$$\begin{aligned} c_{pr,ij} n_{p,i} &= c_{pr,ji} n_{p,j} \\ c_{0r,ij} \frac{h_{p,j}}{h_{0,j}} n_{p,i} &= c_{0r,ji} \frac{h_{p,i}}{h_{0,i}} n_{p,j} \\ c_{0r,ij} \frac{n_{p,j} \sum_{x=1}^{a} n_{0,x}}{n_{0,j} \sum_{x=1}^{a} n_{p,x}} n_{p,i} &= c_{0r,ji} \frac{n_{p,i} \sum_{x=1}^{a} n_{0,x}}{n_{0,i} \sum_{x=1}^{a} n_{p,x}} n_{p,j} \\ \frac{c_{0r,ij}}{n_{0,j}} &= \frac{c_{0r,ji}}{n_{0,i}} \\ c_{0r,ij} n_{0,i} &= c_{0r,ji} n_{0,j} \end{aligned} \qquad 25$$

Therefore, no further adjustments for reciprocity of contacts need to be made.

We adjust the intra- ($\boldsymbol{C_{p0}}$) and extra-household ($\boldsymbol{C_{p1}}$) contact matrix separately. We assume that the total number of extra-household contacts is the same as observed in Digaale ($d_{p1,i} = d_{01,i}$):

$$c_{p1,ij} = c_{01,ij}\frac{h_{p,j}}{h_{0,j}} \quad 26$$

We assume that the total number of intra-household contacts is proportional to the average household size in each setting ($d_{p0,i} = d_{00,i}\frac{s_p}{s_0}$), and calculate intra-household contacts as:

$$c_{p0,ij} = c_{00,ij}\frac{h_{p,j}}{h_{0,j}}\frac{s_p}{s_0} \quad 27$$

$$d_{p0,i} = d_{00,i}\frac{s_p}{s_0} \quad 28$$

Where s_p is the average household size in population *p*. Note that this implicitly assumes that the contact rate between two individual household members in population *p*, is the same as observed in Digaale:

$$d_{p0,i} = \frac{d_{00,i}}{s_0}s_p \quad 29$$

Section B. Model fit

Supplemental Table B1 describes the six fitted model parameters. We used the Differential Evolution Markov Chain with snooker updater (DEzs) algorithm implemented in the *BayesianTools* package in *R* to find parameter values that maximized the multinomial log-likelihood comparing observed age-dependent prevalence estimates in the S, VT, NVT, and B compartments with their model-predicted counterparts at steady state.[14]

We ran four independent DEzs chains in parallel, each with 36,000 model iterations distributed over 3 dependent chains, where the first 6,000 iterations were treated as burn-in and discarded. The MCMC acceptance rate was 15.1% and final joint effective sample size 3,054.

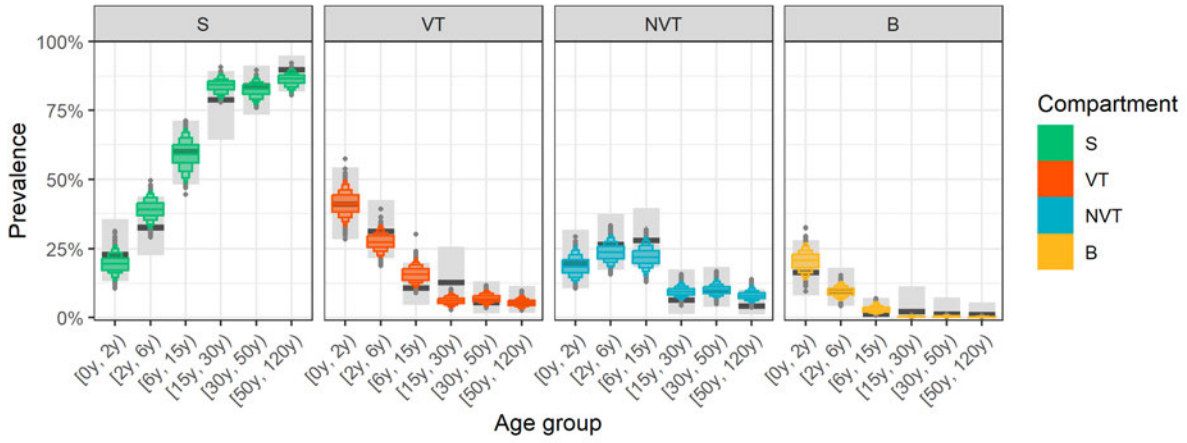
Supplemental Table B1. Fitted model parameters. Table shows the fitted model parameter, their assumed prior distribution, and the Gelman-Rubin convergence diagnostic.

Parameter	Description	Prior distribution	Notes	$\widehat{R}$[1] (upper CI)
beta_1	Proportion of effective contacts for children aged <5y	$\sim beta(0.1, 10)$	Vague prior	1.005 (1.008)
beta_2	Proportion of effective contacts for children aged 5-14y	$\sim beta(0.1, 10)$	Vague prior	1.004 (1.009)
beta_3	Proportion of effective contacts for people aged ≥15y	$\sim beta(0.1, 10)$	Vague prior	1.008 (1.016)
beta_VT_NVT_u5	Relative transmissibility of VTs compared to NVTs for children <5y	$\sim lognormal(1, 1)$	Vague prior	1.003 (1.006)
beta_VT_NVT_o5	Relative transmissibility of VTs compared to NVTs for people ≥5y	$\sim lognormal(1, 1)$	Vague prior	1.005 (1.009)
competition	Relative reduction in susceptibility to acquiring additional VT/NVT, when carrying NVT/VT	$\sim beta(1, 1)$	Vague prior	1.007 (1.013)

1. Gelman-Rubin convergence diagnostics, and the upper value of their 95% confidence interval. The multivariate $\widehat{R}$ was 1.010.

The posterior model fit to the observed data is shown for 500 posterior samples in Supplemental Figure B1. Our model was able to fit the observed data well with modelled prevalence close to the point estimates of observed prevalence for nearly all data points. The 95% credible intervals of modelled estimates did not include the observed point estimates of the proportion susceptible, carrying VTs, or carrying NVTs for those aged 15-

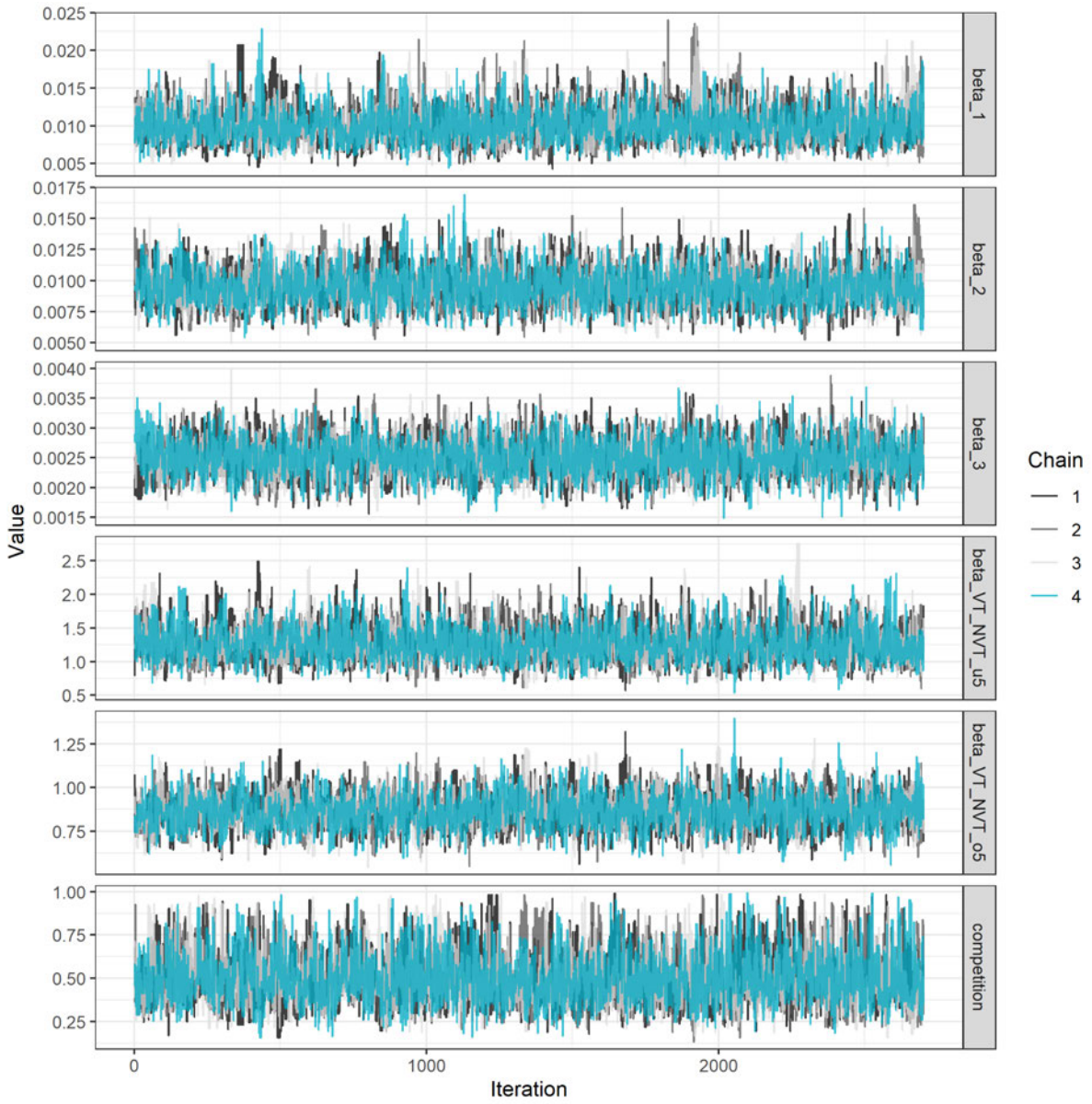
29y and 6-14y, but remained within the 95% confidence interval around all observed estimates.



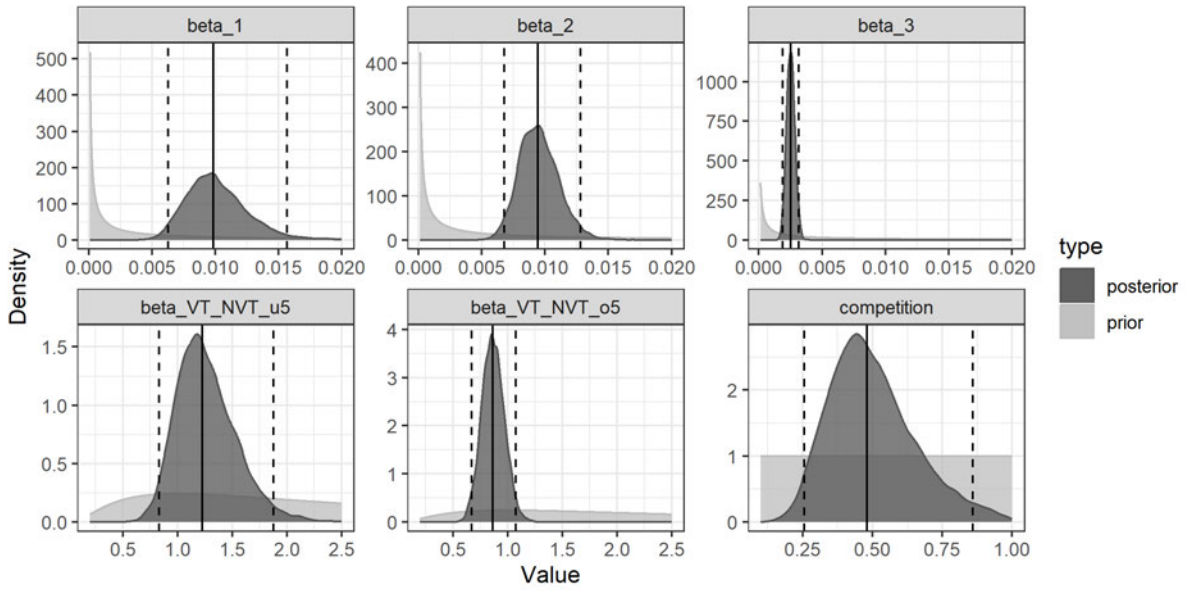
Supplemental Figure B1. Model fit compared to observed data. Prevalence by age is shown for S (susceptibles), VT (vaccine-type carriers), NVT (non-vaccine-type carriers) and B (both VT and NVT carriers). Observed estimates are shown as black bars with error bars indicating their 95% confidence intervals. Coloured boxen plots show the distribution of modelled prevalence from 1000 model runs sampling from the joint-posterior distribution.

Thinned trace plots of the four independent chains are shown in Supplemental Figure B2. We achieved good mixing in all chains, that all seemed to have converged. Gelman-Rubin convergence diagnostics ($\widehat{R}$) was <1.02 for all fitted parameters (Supplemental Table B1). Posterior and prior marginal density distributions are provided in Supplemental Figure B3. The posterior distribution for the competition parameter is relatively wide which may indicate a reduced identifiability for this parameter value, likely due to the limited amount of data for the prevalence of multiple carriage (B), which was low for all ages ≥6y.

Supplemental Table B2 shows the cross-correlation between the fitted parameters. Values for beta_1 and beta_VT_NVT_u5, that both affect transmission in children <5y, are relatively highly correlated (-0.52). The same is observed for values between the two beta_VT_NVT parameters (-0.83). There is also a relatively strong negative correlation between the competition parameter and beta_1 (-0.62), but not with the other beta parameters, which is probably due to the fact that as for overall prevalence, prevalence of superinfection (B) is highest in the <5y age group.



Supplemental Figure B2. Trace plots of 4 independent chains of fitted model parameters. Posterior samples are thinned by 10. Posteriors of the dependent DEMC chains within each independent chain are combined.



Supplemental Figure B3. Posterior and prior marginal density distributions of fitted model parameters. Light- and dark-grey areas show the prior- and posterior density distribution of each parameter. Solid vertical lines indicate the median of the posterior distribution, and dashed vertical lines its 95% credible interval.

Supplemental Table B2. Cross-correlations of posterior estimates.

	beta_1	beta_2	beta_3	beta_VT_NVT_u5	beta_VT_NVT_o5	competition
beta_1	1	-0.32	-0.12	-0.52	0.38	-0.62
beta_2	-0.32	1	-0.35	0.29	-0.28	-0.20
beta_3	-0.12	-0.35	1	0.27	-0.23	-0.09
beta_VT_NVT_u5	-0.52	0.29	0.27	1	-0.83	0.14
beta_VT_NVT_o5	0.38	-0.28	-0.23	-0.83	1	-0.07
competition	-0.62	-0.20	-0.09	0.14	-0.07	1

Section C. Impact in different settings

Supplemental Table C1 shows the total impact over a 2y and 3y period on all and infant severe pneumococcal disease cases in the different settings. Supplemental Table C2 shows the NNV to prevent one severe pneumococcal disease case in these settings.

Supplemental Table C1. Cumulative impact of PCV campaigns on severe pneumococcal disease cases in four different settings. Estimates show the median absolute cumulative impact per 10,000 people per year, in each respective period. The 95% uncertainty intervals and median relative impact are provided in brackets. Estimates are taken from 500 model runs.

Age	Vaccine strategy	*Severe pneumococcal disease cases prevented per 10,000 people per year*			
		Digaale 2019	*Bentiu 2015*	*Maiduguri 2016*	*Bambari 2019*
0-1y period					
All ages	U1: full	1 (1 - 1, 6%)	1 (1 - 1, 6%)	1 (1 - 1, 5%)	1 (1 - 1, 7%)
	U2: full	2 (2 - 3, 13%)	2 (1 - 2, 12%)	2 (1 - 2, 10%)	2 (1 - 3, 14%)
	U5: full	6 (5 - 8, 37%)	5 (4 - 7, 34%)	5 (3 - 6, 28%)	5 (4 - 7, 39%)
	U10: full	8 (6 - 10, 47%)	6 (5 - 9, 42%)	6 (4 - 7, 35%)	6 (5 - 8, 48%)
	U15: full	9 (6 - 11, 50%)	7 (5 - 9, 45%)	6 (5 - 8, 37%)	7 (5 - 9, 52%)
Infants	U1: full	16 (10 - 25, 11%)	14 (9 - 22, 10%)	13 (8 - 21, 9%)	14 (9 - 23, 11%)
	U2: full	19 (12 - 31, 13%)	16 (10 - 26, 12%)	14 (9 - 23, 10%)	17 (10 - 28, 14%)
	U5: full	28 (18 - 47, 19%)	23 (14 - 39, 17%)	18 (11 - 29, 13%)	25 (15 - 42, 20%)
	U10: full	36 (22 - 61, 25%)	29 (18 - 49, 22%)	20 (13 - 33, 15%)	32 (19 - 53, 25%)
	U15: full	39 (24 - 66, 27%)	31 (19 - 52, 23%)	22 (13 - 35, 16%)	35 (21 - 61, 28%)
0-3y period					
All ages	U1: full	1 (1 - 1, 5%)	1 (1 - 1, 5%)	1 (0 - 1, 4%)	1 (1 - 1, 6%)
	U2: full	2 (2 - 3, 12%)	2 (1 - 2, 11%)	1 (1 - 2, 8%)	2 (1 - 2, 14%)
	U5: full	5 (4 - 6, 28%)	4 (3 - 5, 25%)	3 (2 - 4, 19%)	4 (3 - 6, 32%)
	U10: full	6 (5 - 8, 36%)	5 (4 - 6, 32%)	4 (3 - 5, 23%)	5 (4 - 7, 40%)
	U15: full	7 (5 - 9, 39%)	5 (4 - 7, 34%)	4 (3 - 5, 25%)	6 (4 - 8, 44%)
Infants	U1: full	7 (4 - 11, 4%)	6 (4 - 9, 4%)	5 (3 - 8, 4%)	6 (4 - 10, 5%)
	U2: full	9 (6 - 16, 6%)	8 (5 - 14, 6%)	6 (4 - 9, 4%)	9 (5 - 15, 7%)
	U5: full	16 (9 - 28, 11%)	13 (8 - 24, 10%)	8 (5 - 13, 6%)	16 (9 - 29, 13%)
	U10: full	23 (13 - 42, 15%)	18 (10 - 32, 13%)	10 (6 - 17, 7%)	23 (12 - 41, 18%)
	U15: full	26 (14 - 47, 17%)	19 (11 - 34, 14%)	11 (6 - 18, 8%)	27 (14 - 51, 21%)

Supplemental Table C2. Number needed to vaccinate to prevent a single severe pneumococcal disease case in four different settings. Estimates show the total number of PCV doses needed to prevent one case of severe pneumococcal disease. The 95% uncertainty intervals are provided in brackets. Estimates are taken from 500 model runs. Values greater than 1000 are rounded to the nearest hundred.

Age	Vaccine strategy	*NNV: Number of PCV doses needed to prevent one severe pneumococcal disease case*			
		Digaale 2019	*Bentiu 2015*	*Maiduguri 2016*	*Bambari 2019*
	0-1y period				
All ages	U1: full	235 (172 - 309)	284 (206 - 372)	340 (240 - 454)	269 (195 - 354)
	U2: full	230 (175 - 294)	283 (213 - 365)	352 (260 - 456)	263 (196 - 337)
	U5: full	217 (163 - 287)	259 (192 - 349)	311 (232 - 416)	246 (185 - 325)
	U10: full	353 (270 - 459)	399 (297 - 529)	475 (360 - 628)	397 (300 - 531)
	U15: full	477 (363 - 629)	531 (395 - 710)	622 (474 - 820)	536 (403 - 718)
Infants	U1: full	480 (296 - 736)	549 (336 - 869)	591 (360 - 927)	534 (328 - 855)
	U2: full	849 (519 - 1,300)	974 (600 - 1,600)	1,100 (681 - 1,800)	929 (567 - 1,500)
	U5: full	1,500 (879 - 2,300)	1,700 (1,000 - 2,900)	2,200 (1,300 - 3,600)	1,600 (946 - 2,700)
	U10: full	2,400 (1,400 - 3,900)	2,600 (1,600 - 4,300)	3,600 (2,200 - 5,800)	2,500 (1,500 - 4,200)
	U15: full	3,200 (1,900 - 5,200)	3,500 (2,100 - 5,700)	4,800 (2,900 - 7,700)	3,300 (1,900 - 5,500)
	0-3y period				
All ages	U1: full	93 (70 - 118)	113 (86 - 145)	161 (119 - 206)	95 (73 - 122)
	U2: full	89 (68 - 115)	109 (83 - 143)	159 (120 - 206)	92 (70 - 118)
	U5: full	95 (72 - 124)	115 (87 - 155)	157 (118 - 211)	99 (75 - 130)
	U10: full	154 (116 - 201)	177 (133 - 236)	241 (183 - 314)	156 (115 - 207)
	U15: full	206 (154 - 274)	233 (175 - 313)	314 (238 - 411)	208 (150 - 281)
Infants	U1: full	383 (231 - 609)	441 (267 - 710)	530 (321 - 838)	408 (248 - 671)
	U2: full	581 (337 - 948)	682 (389 - 1,100)	915 (562 - 1,500)	590 (346 - 997)
	U5: full	860 (488 - 1,500)	1,000 (561 - 1,700)	1,600 (999 - 2,700)	821 (467 - 1,500)
	U10: full	1,200 (691 - 2,200)	1,400 (790 - 2,500)	2,400 (1,500 - 4,200)	1,200 (634 - 2,100)
	U15: full	1,600 (882 - 2,900)	1,900 (1,000 - 3,200)	3,100 (1,900 - 5,400)	1,500 (756 - 2,800)

Section D. Sensitivity analyses

Parametric sensitivity analyses

We assessed the sensitivity of model outcomes on our assumptions for several key parameters: i) the degree of mixing with the host community, ii) migration rate, iii) relative increased susceptibility of malnourished children, iv) vaccine coverage, v) duration of vaccine protection, and vi) vaccine efficacy against transmission.

As i, ii, and iii affect the pre-vaccination transmission dynamics, we first refitted our model for each new parameter value, running two independent DEzs chains for each model fit, fitting the same parameters using the same priors as provided in Supplemental Table B1. Each independent DEzs chain was ran for 36,000 model iterations distributed over 3 dependent chains, with the first 6,000 iterations were treated as burn-in and discarded. Gelman-Rubin convergence diagnostics ($\widehat{\boldsymbol{R}}$) for the fifteen model fits are shown in Supplemental Table D1. $\widehat{\boldsymbol{R}}$ values were well < 1.1 or all models.

Supplemental Table D1. Gelman-Rubin convergence diagnostics of refitted models.
Table shows for each refitted model (rows) the overall and parameter-specific Gelman-Rubin convergence diagnostics (columns).

Model	Mult. PSRF[a]	β 1	β 2	β 3	β VT/NVT <5y	β VT/NVT ≥5y	Comp
contact_host							
0	1.011	1.01 (1.025)	1.002 (1.005)	1.004 (1.009)	1.007 (1.015)	1.007 (1.017)	1.006 (1.013)
1.7[b]	.	.	.	.	.	.	.
5	1.012	1.011 (1.023)	1.002 (1.005)	1.004 (1.01)	1.004 (1.01)	1.004 (1.009)	1.005 (1.011)
10	1.017	1.003 (1.006)	1.006 (1.012)	1.008 (1.02)	1.003 (1.006)	1.009 (1.02)	1.005 (1.012)
20	1.014	1.007 (1.014)	1.002 (1.004)	1.003 (1.007)	1.003 (1.005)	1.005 (1.013)	1.009 (1.02)
40	1.014	1.007 (1.017)	1.009 (1.021)	1.005 (1.013)	1.007 (1.015)	1.003 (1.005)	1.006 (1.014)
malnutrition_transmission							
1	1.015	1.007 (1.017)	1.003 (1.007)	1.009 (1.018)	1.004 (1.009)	1.006 (1.014)	1.009 (1.022)
1.2[b]	.	.	.	.	.	.	.
1.4	1.007	1.006 (1.012)	1.004 (1.009)	1.004 (1.01)	1.005 (1.013)	1.005 (1.012)	1.002 (1.003)
1.6	1.009	1.011 (1.023)	1.002 (1.004)	1.003 (1.006)	1.006 (1.013)	1.004 (1.01)	1.006 (1.013)

1.8	1.018	1.008 (1.019)	1.004 (1.009)	1.003 (1.007)	1.012 (1.023)	1.008 (1.018)	1.008 (1.018)
2	1.012	1.012 (1.026)	1.007 (1.018)	1.003 (1.007)	1.004 (1.008)	1.005 (1.012)	1.008 (1.02)
Migration							
0	1.015	1.005 (1.012)	1.002 (1.004)	1.006 (1.016)	1.008 (1.018)	1.008 (1.019)	1.003 (1.006)
2	1.012	1.007 (1.014)	1.003 (1.007)	1.002 (1.006)	1.006 (1.014)	1.004 (1.009)	1.001 (1.004)
4	1.01	1.008 (1.017)	1.006 (1.014)	1.008 (1.02)	1.01 (1.021)	1.008 (1.019)	1.006 (1.014)
6	1.014	1.006 (1.014)	1.004 (1.009)	1.009 (1.021)	1.002 (1.005)	1.005 (1.009)	1.009 (1.019)
7.7[b]	.	.	.	.	.	.	.
10	1.01	1.01 (1.02)	1.007 (1.013)	1.006 (1.014)	1.009 (1.02)	1.007 (1.017)	1.003 (1.007)

Gelman-Rubin convergence diagnostics, and the upper value of their 95% confidence interval, for the six fitted parameters: beta_1, beta_2, beta_3, beta_VT_NVT_u5, beta_VT_NVT_o5, competition.

a. The upper value of the multivariate $\widehat{\boldsymbol{R}}$ is not shown.
b. The main model as not refitted

For each fitted model (including the baseline model), we simulated 250 post-vaccination runs sampling from their joint posterior and compared the univariate impact of the changed parameter value on six summary statistics: the maximum impact on infant VT prevalence, number of days since PCV campaign at which peak impact on infant VT prevalence is reached, the cumulative 3y impact on severe pneumococcal disease cases in infants and all ages, and their NNV, for each vaccine strategy <1y, <2y, <5y, <10y, and <15y, shown in Supplemental Figures D1 and D2.

Mixing with host population

Increasing the proportion of extra-household contacts made with the host community reduced the overall impact of vaccination on all outcomes. In general, it makes the impact of the different strategies more similar, reducing the additional benefit of wider age targeting with increased mixing with the host population, presumably as indirect protection is limited. The most drastic change is observed on the maximum impact of prevalence, and the number of days at which impact on prevalence peaks. Values for the <5y and <15y strategies shift from a peak impact of 17% (11 – 22) after 306 (191 – 441) days and 34% (22

– 47) after 366 (242 – 472) days in the absence of mixing with the host population, to 10% (8 – 11) after 149 (127 – 177) days and 15% (12 – 19) after 181 (152 – 1206) days when 40% of extra-household contacts are made with members of the host population. The reduced impact is also observed on the reduced impact on cumulative severe pneumococcal disease cases: the impact of <5y and <15y campaigns on all cases during the 3y following the campaign reduces from preventing 28% (25 – 32) and 42% (35 – 51) in the absence of mixing with the host population, to 22% (19 – 25) and 32% (28 – 35). While extended age targeting remains beneficial with increased mixing with the host population, the impact on infant cases is smaller. This difference reflects the fact that i) a substantial disease burden remains in older age groups, and that ii) carriers in older ages contribute to the transmission to those older age groups, in contrast to infants. While campaigns become less efficient when mixing with the host population increases, the change in efficiency is minimal.

Migration rate

Migration removes vaccinated individuals from the population and introduces new unvaccinated individuals. Therefore, increasing the migration rate reduces the overall impact of the different vaccination strategies. It reduces the relative difference between the strategies, though the overall impact of wider age targeting remains substantially higher for all migration values considered.

In the absence of migration, peak impact on infant VT carriage prevalence occurs 426 (304 – 544) days following the PCV campaign with a strategy targeting children <15y, and 368 (233 – 514) for an <5y campaign. These values reduce to 358 (233 – 483) and 296 (184 – 452) days when individuals remain displaced for, on average, 10 years, while they reduce even further to 238 (179 – 296) and 200 (147 – 261) days if individuals would move in, on average, 2 years. Migration also affects the maximum impact on infant VT prevalence. An

<15y campaign could reduce VT prevalence by 36% (25 – 49) in the absence of migration but would only reduce it by 24% (16 – 32) if migration occurs after 2 years (on average).

In the absence of migration, an <5y campaign would be able to prevent up to 32% (28 – 36) of all severe pneumococcal disease cases over a 3-year period. If individuals would migrate after, on average, 2 years, an <15y campaign would be required to achieve the same effect (28%, 23 – 32). In the absence of migration, the same <15y campaign would prevent 48% (41 – 53) of all cases over the 3-year period.

Migration especially affects the longevity of the impact of PCV campaigns with wider age targeting. Over a 2-year period, the relative difference between the scenarios with the highest considered migration rate and the absence of migration was 32% for an <15y campaign (a 36% compared to a 53% reduction). However, the cumulative impact of an <15y campaign over 5-years would be 45% (38 – 52) in the absence of migration, but only 21% (18 – 23) with the highest migration rate, a relative difference of 53%.

Increased acquisition malnourished

Increasing the relative acquisition rate of malnourished compared to non-malnourished individuals did not substantially affect any of the modelled outcomes. There may be a very small reduction in impact as the relative acquisition rate increases, most obvious in the cumulative 5y impact in infants, but it is a negligible effect.

This is not very surprising, as we refitted our models to the same carriage prevalence estimates in all models. We assume that carriage prevalence is an average of the total population (malnourished and non-malnourished) and refitting the model with the same proportion of the population that is malnourished, but an increased acquisition rate, has adjusted the fitted beta parameters so that the modelled overall prevalence remains close to the observed estimates. We would expect an increase in overall carriage prevalence if the

acquisition rate of a segment of the population would increase, but we don't have reliable data to parameterize such models accurately.

Vaccine coverage

Vaccine coverage affects both direct and indirect vaccine protection. Lower vaccine coverage would result in reduced vaccine impact, although the duration until peak impact of infant VT prevalence would be achieved is not substantially affected, especially for vaccine strategies with narrower age targeting. A vaccine coverage of only 45% would result in a peak impact prevalence after 252 (188 – 371) days and 287 (216 – 385) for an <5 and <15y campaign, whilst this would occur after 309 (224 – 423) and 360 (267 – 453) days if vaccine coverage would be 95%. In contrast, it would substantially affect the maximum impact on infant VT prevalence that would be achieved, from 20% (15 – 28) and 37% (25 – 48) for <5y and <15y campaigns with 95% coverage, to only 9% (6% - 12%) and 16% (11 – 22) for the same campaigns with 45% coverage.

Again, scenarios that result in lower indirect effects, in this case lower vaccine coverage, would reduce the impact of all strategies, but would especially diminish the additional benefit of increased age targeting. An <5y campaign with 95% coverage would prevent 31% (28 – 36) of all severe pneumococcal disease cases over a 3y period, but only 15% (13 – 18) at 45% coverage. In contrast, an <15y campaign at 95% coverage would prevent 43% (37 – 50) of all severe pneumococcal disease cases, but only 22% (19 – 26) at 45% coverage.

Similarly, impact on severe pneumococcal disease cases would only be 6% (4 – 8) and 9% (6 - 13) for an <5y and <15y campaign at 45% coverage, down from 13% (10 – 18) and 21% (14 – 29) at 95% coverage. This result it is not surprising but highlights the need for high vaccine coverage to maximize impact.

Duration of vaccine protection

The duration of vaccine protection primarily affects the longevity of the PCV campaign. While we have good estimates for the duration of vaccine protection of PCV, these are based on 3 or 4 dose schedules in infants, given in routine schedules. It is possible that the average duration of vaccine protection is lower following a reduced dose schedule, suboptimal vaccine storage, or in vaccinees with certain health conditions such as malnutrition. Our results do not suggest that the impact of a PCV campaign is very sensitive for this assumption, except for very low values for the duration of vaccine protection, that are probably unrealistic. As for migration rates, the assumed duration of vaccine protection primarily affects the longevity of a PCV campaign, e.g. the impact over a 3y period or longer.

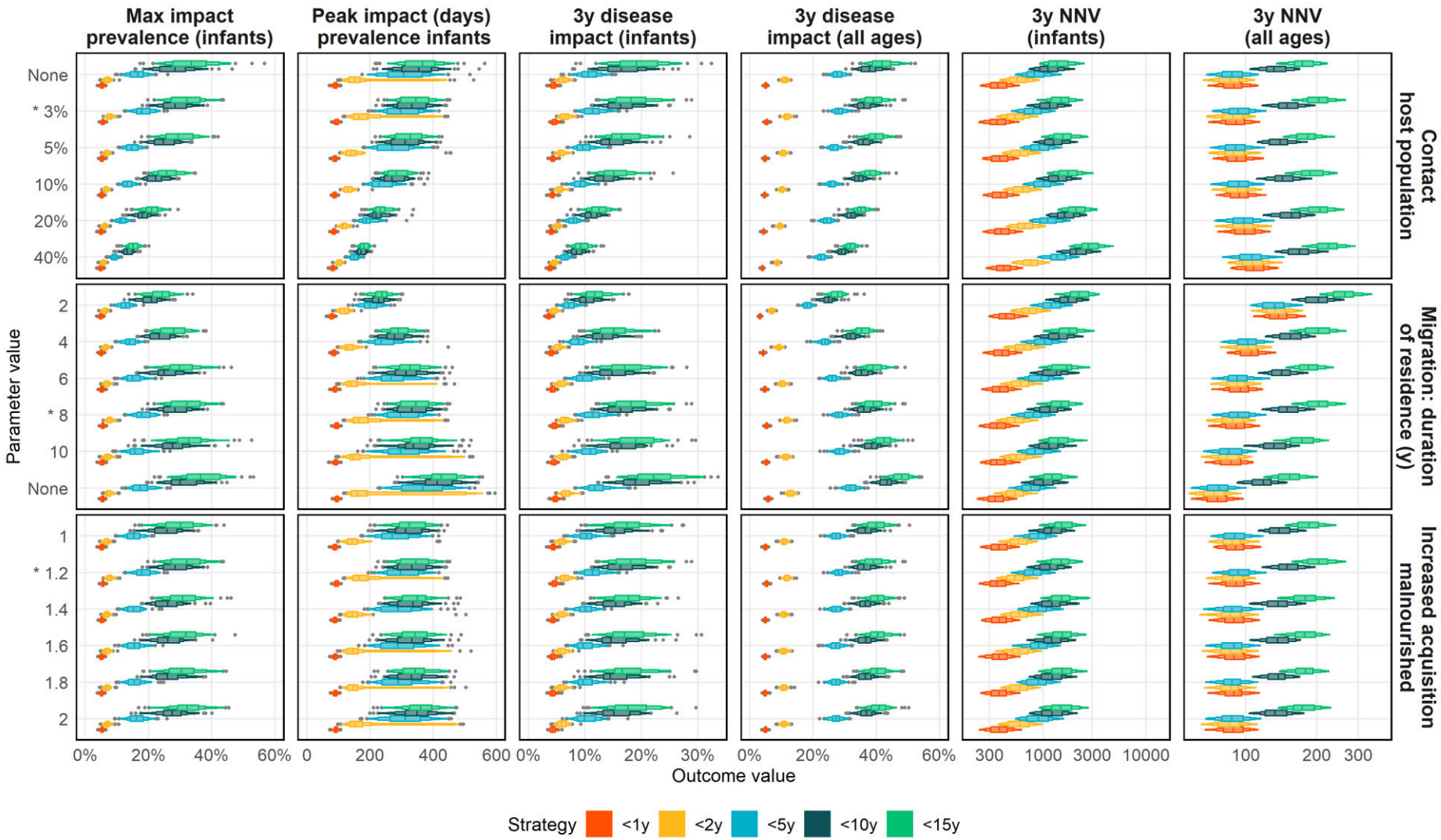
Vaccine efficacy against transmission

Finally, we assessed the sensitivity of our results against the assumed VE against transmission. Again, we have good estimates for the VE of PCVs, but it may be lower following a reduced dose schedule -especially in very young children, suboptimal storage or delivery of vaccine, or in people with certain underlying conditions. We also assessed a scenario where we assumed VE_c to be 0, not because we think this is realistic, but to highlight the importance of the indirect effect of the PCV campaign.

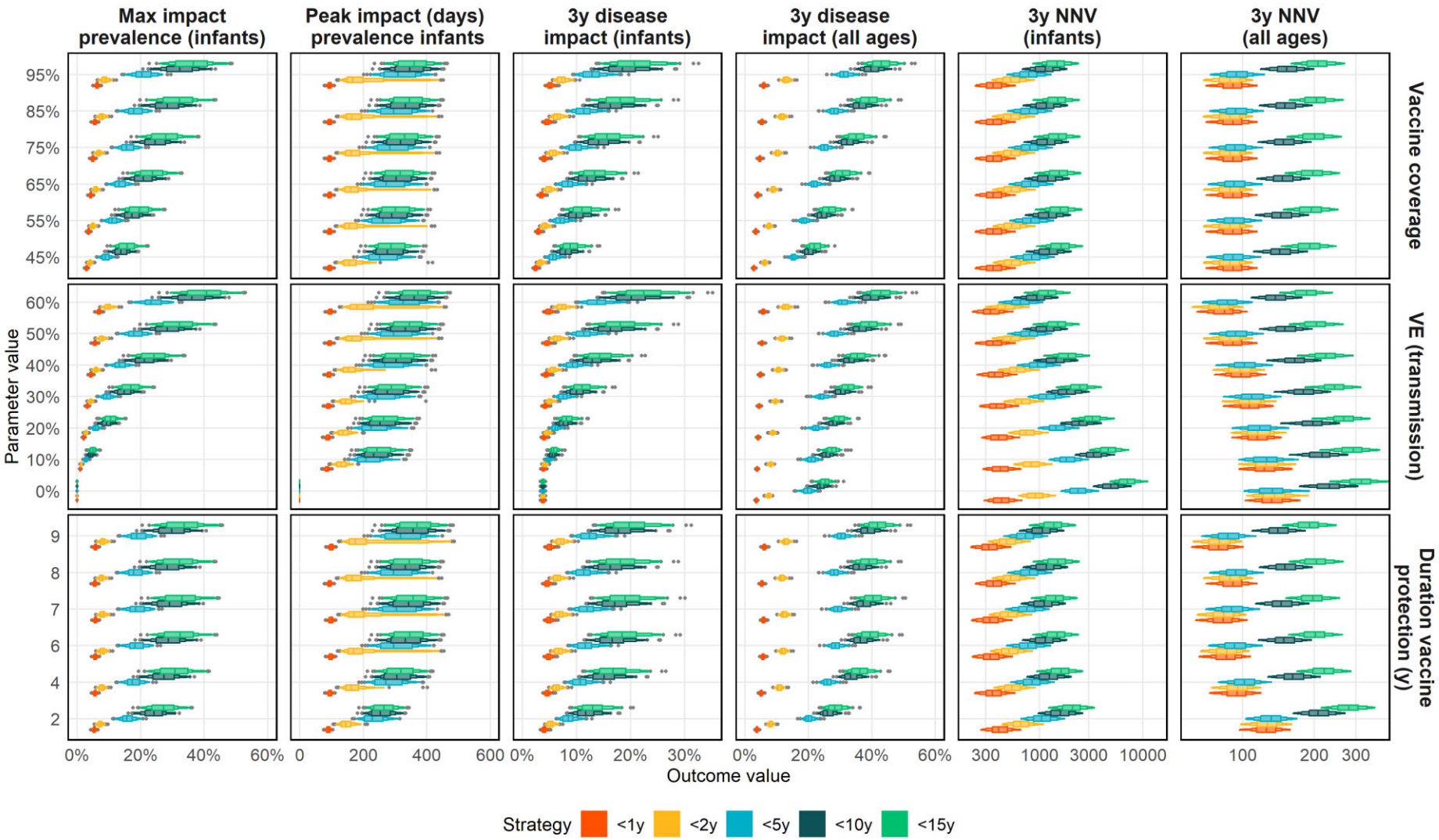
Unsurprisingly, reduced VE against transmission would result in substantially reduced impact of a PCV campaign. This by itself highlights the importance of the indirect effects on pneumococcal transmission. If PCVs would not protect against transmission, there would be no added benefit of widening the age targeting of a campaign on infant severe pneumococcal disease cases. However, as not all cases occur in infants, an <5y campaign would still prevent 20% (16 – 23) of cases over a 3y period, whereas an <10y and <15y campaign would increase this to 24% (20 – 27) and 25% (21 – 28). This highlights the importance of direct protection, even for these older age groups.

However, the assumed VE against carriage of 50% increases the impact of an <5y campaign on all severe pneumococcal disease cases over 3y to 28% (25 – 32) and increases the impact of an <15y campaign to 39% (34 – 46). The indirect protection from the wider age targeting even increases the impact on infant cases over 3y from 4% (3 – 4) for both vaccine strategies to 11% (9 – 16) and 18% (13 – 26) for the <5y and <15y campaigns.

To summarise, our sensitivity analyses show that our modelled results of the impact of PCV campaigns are especially sensitive to factors that affect the indirect effect of the PCV campaign: the assumed contact rates with the (unvaccinated) host population, vaccine coverage, and vaccine efficacy against transmission. Wider age targeting could mitigate the reduced impact from any of these factors, but this may come at a cost of reduced efficiency. The impact is also affected by the migration rate and the duration of vaccine protection, but to lower degree than the other factors. These especially affect the longevity of the impact of the PCV campaign, beyond a 3y period. Finally, our results are not affected by the assumed increased acquisition rate in malnourished individuals, primarily because we lack the data to accurately parameterize this variable.



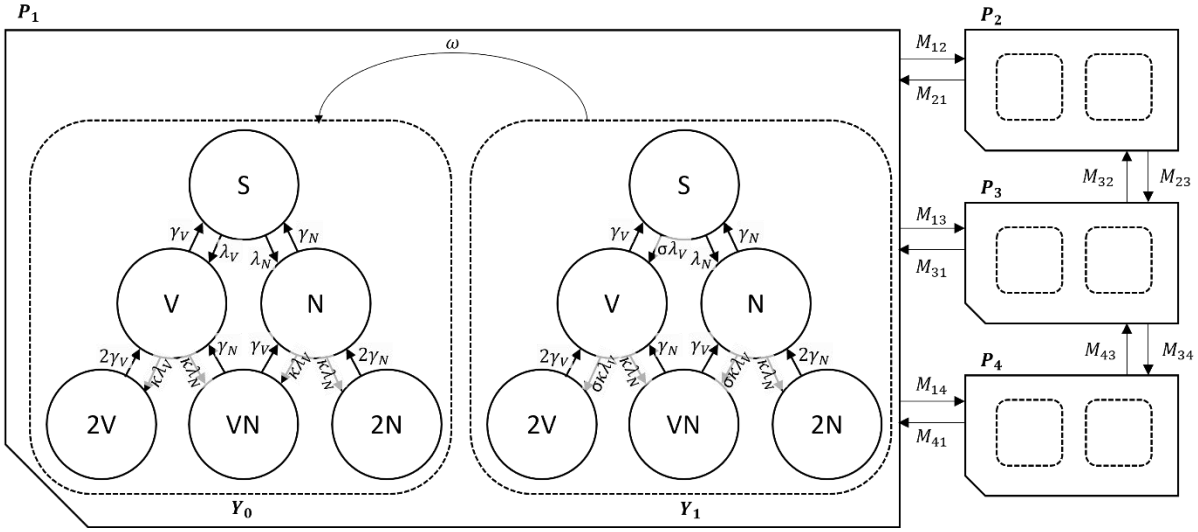
Supplemental Figure D1. Parametric sensitivity analyses for the mixing with the host population, migration rate, and relative susceptibility of malnourished individuals. Parameters are shown in facet rows, with parameter values on the y-axes. Parameter values denote i) the proportion of contacts made outside of the crisis-affected population, ii) the average number of years individuals remain resident in the crisis-affected population, and iii) the relative increase in susceptibility to acquisition of pneumococci in malnourished individuals. Parameter values used in the main model (1.7%, 7.7 years, and 1.2) are highlighted with an asterisk. Different outcomes are shown in facet columns, with outcome values on the x-axes. Outcome values denote i) the maximum relative reduction on infant VT prevalence, ii) the time (days) after the PCV campaign at which i occurs, iii) the cumulative impact on severe pneumococcal disease over three years following the PCV campaign in all age groups and infants, and iv) the NNV for these campaigns. Boxen plots show the distribution of the outcome values (x-axes) of 100 model runs for vaccine strategies targeting children <1y (red), <2y (yellow), <5y (blue), <10y (dark green), and <15y (light green).



Supplemental Figure D2. Parametric sensitivity analyses for the vaccine coverage, duration of vaccine protection, and vaccine efficacy. Parameters are shown in facet rows, with parameter values on the y-axes. Parameter values denote i) the vaccine coverage in targeted age groups, ii) the average duration of vaccine protection in years, and the vaccine efficacy against transmission. Parameter values used in the main model (85%, 8 years, and 50%) are highlighted with an asterisk. Different outcomes are shown in facet columns, with outcome values on the x-axes. Outcome values denote i) the maximum relative reduction on infant VT prevalence, ii) the time (days) after the PCV campaign at which i occurs, iii) the cumulative impact on severe pneumococcal disease over three years following the PCV campaign in all age groups and infants, and iv) the NNV for these campaigns. Boxen plots show the distribution of the outcome values (x-axes) of 100 model runs for vaccine strategies targeting children <1y (red), <2y (yellow), <5y (blue), <10y (dark green), and <15y (light green).

Alternative model structure

Our main model structure assumes no superinfections of serotypes of the same group (VT or NVT), and only promotes coexistence of serotypes arbitrarily. We explored the structural uncertainty of our model to this assumption by assessing our results in an alternative structurally neutral model that does allow for superinfection of serotypes of the same group (Supplemental Figure D3).



Supplemental Figure D3. Compartmental model structure of the pneumococcal transmission model. Pneumococcal serotypes defined as those included in the PCV vaccine (vaccine-type, V), or not included in the vaccine (non-vaccine type, N). At any time *t*, Susceptible individuals (S) acquire serotypes in either group at rates $\lambda_V(t)$ and $\lambda_N(t)$, and immediately become infectious carriers themselves. Carriers of VTs or NVTs remain partially susceptible to infection with an additional serotype, and may develop a superinfection (2V, VN, 2N) at rates $\kappa\lambda_V(t)$ and $\kappa\lambda_N(t)$, where κ represents competition between serotypes. Carriers are assumed to clear their infection of all serotypes in either group at rates γ_V for VTs and γ_N for NVTs. Individuals may be without vaccine derived protection (Y_0), or with full vaccine derived protection (Y_1). Those with vaccine protection acquire VT at a reduced rate σ, and lose their level of protection at rate ω. Different populations P_1, P_2, and P_3, P_4 represent malnourished and non-malnourished strata of a displaced and host population. Individuals of population *p* migrate to population *q* at rate M_{qp}. Age groups, ageing, and transmission between vaccine protection strata, populations, and age groups are not shown.

Epidemiological transitions

In this alternative model structure, the values for epidemiological compartment $X_a^{p,y}$, where $X \in \{S, V, N, 2V, VN, 2V\}$ (i.e. $S_a^{p,y}$, $V_a^{p,y}$, $N_a^{p,y}$, $2V_a^{p,y}$, $VN_a^{p,y}$, $2N_a^{p,y}$), again hold the proportion of people in age group *a* in population *p*, that are in compartment *X* in vaccinated stratum *y*.

Values for model compartments within a single age group in a model population sum to 1, i.e.

$$\sum_{y=0}^{N_{\text{vac}}} \left(S_a^{p,y} + V_a^{p,y} + N_a^{p,y} + 2V_a^{p,y} + VN_a^{p,y} + 2N_a^{p,y} \right) = 1 \qquad 30$$

The demographic changes in the model compartments, $\Delta D_{X_a^{p,y}}(t)$, remain the same as in the base model structure listed in equation 4. The epidemiological transitions are altered as shown in the following set of equations. For ease of reading, we have again omitted the time notations, and gave the notations for the age group, population, and vaccinated stratum a lighter colour:

$$\Delta E_{S_a^{p,y}}(t) = -\left(\sigma_a^{p,y} {\lambda_V}_a^p + {\lambda_N}_a^p\right)\overline{S}_a^{p,y} + \gamma_{V\,a}\overline{V}_a^{p,y} + \gamma_{N\,a}\overline{N}_a^{p,y}$$

$$\Delta E_{V_a^{p,y}}(t) = \sigma_a^{p,y} {\lambda_V}_a^p \overline{S}_a^{p,y} - \left(\sigma_a^{p,y}\kappa{\lambda_V}_a^p + \kappa{\lambda_N}_a^p + \gamma_{V\,a}\right)\overline{V}_a^{p,y} + \gamma_{N\,a}\overline{VN}_a^{p,y} + 2\gamma_{V\,a}\overline{2V}_a^{p,y}$$

$$\Delta E_{N_a^{p,y}}(t) = {\lambda_N}_a^p \overline{S}_a^{p,y} - \left(\kappa{\lambda_N}_a^p + \sigma_a^{p,y}\kappa{\lambda_V}_a^p + \gamma_{N\,a}\right)\overline{N}_a^{p,y} + \gamma_{V\,a}\overline{VN}_a^{p,y} + 2\gamma_{N\,a}\overline{2N}_a^{p,y}$$

$$\Delta E_{2V_a^{p,y}}(t) = \sigma_a^{p,y}\kappa{\lambda_V}_a^p\overline{V}_a^{p,y} - 2\gamma_{V\,a}\overline{2V}_a^{p,y}$$

$$\Delta E_{VN_a^{p,y}}(t) = \kappa{\lambda_N}_a^p\overline{V}_a^{p,y} + \sigma_a^{p,y}\kappa{\lambda_V}_a^p\overline{N}_a^{p,y} - \left(\gamma_{V\,a} + \gamma_{N\,a}\right)\overline{VN}_a^{p,y}$$

$$\Delta E_{2N_a^{p,y}}(t) = \kappa{\lambda_N}_a^p\overline{N}_a^{p,y} - 2\gamma_{N\,a}\overline{2N}_a^{p,y}$$

$${\lambda_V}_a^p = \beta_a^{p,V}\sum_{b=1}^{N_{\text{age}}}\sum_{q=1}^{N_{\text{pop}}}\left(c_{ba}^p\left(\delta_{q,p}(1-o^p) + o^p T_{qp}\right)\sum_{y=0}^{N_{\text{vac}}}\left(\overline{V}_b^{q,y} + \overline{2V}_b^{q,y} + \overline{VN}_b^{q,y}\right)\right)$$

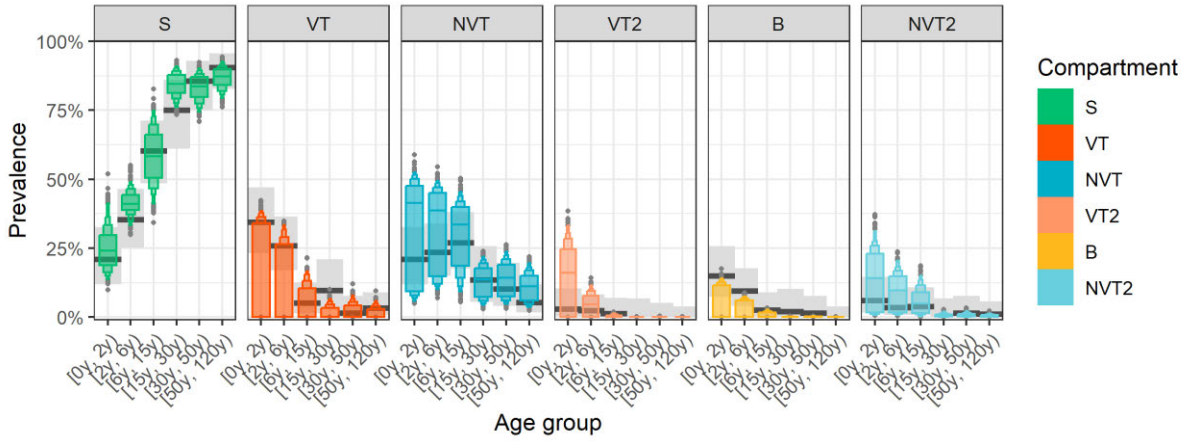
$${\lambda_N}_a^p = \beta_a^{p,N}\sum_{b=1}^{N_{\text{age}}}\sum_{q=1}^{N_{\text{pop}}}\left(c_{ba}^p\left(\delta_{q,p}(1-o^p) + o^p T_{qp}\right)\sum_{y=0}^{N_{\text{vac}}}\left(\overline{N}_b^{q,y} + \overline{2N}_b^{q,y} + \overline{VN}_b^{q,y}\right)\right) \qquad 31$$

The parameter definitions and values are the same as in the main model. We then combine the demographic and epidemiological transitions, and calculate the change for each model compartment at time *t* as:

$$\frac{dX_a^{p,y}}{dt} = \Delta D_{X_a^{p,y}}(t) + \Delta E_{X_a^{p,y}}(t) \qquad 32$$

Model fit

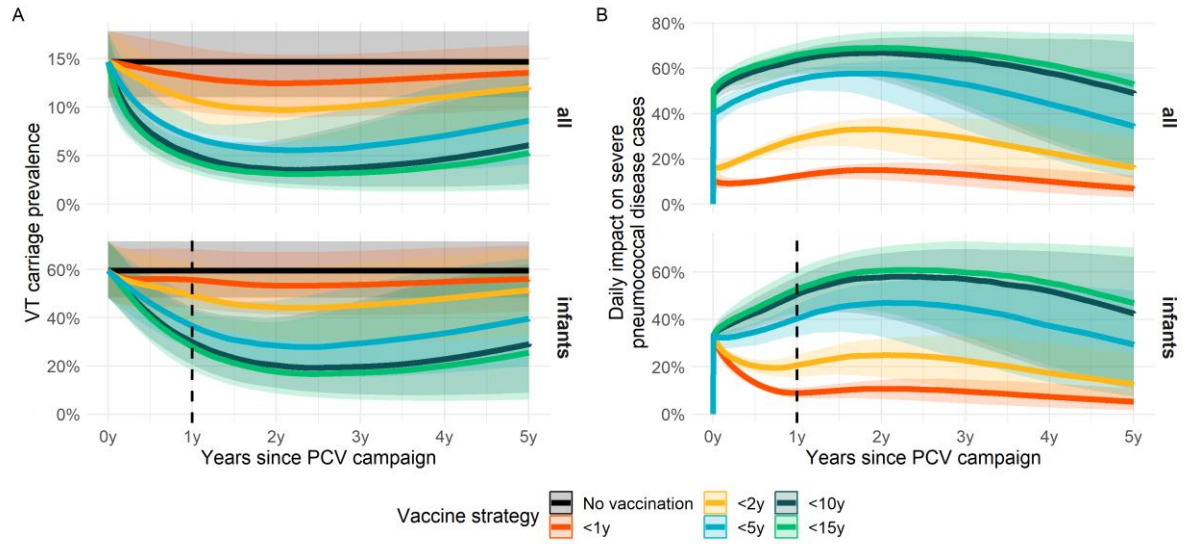
Supplemental Figure D4 shows the fit of the alternative model compared to the observed data. We got an overall good fit to the data, with the exception of prevalence in <2y in the 2VT compartment, which was overestimated in the model.



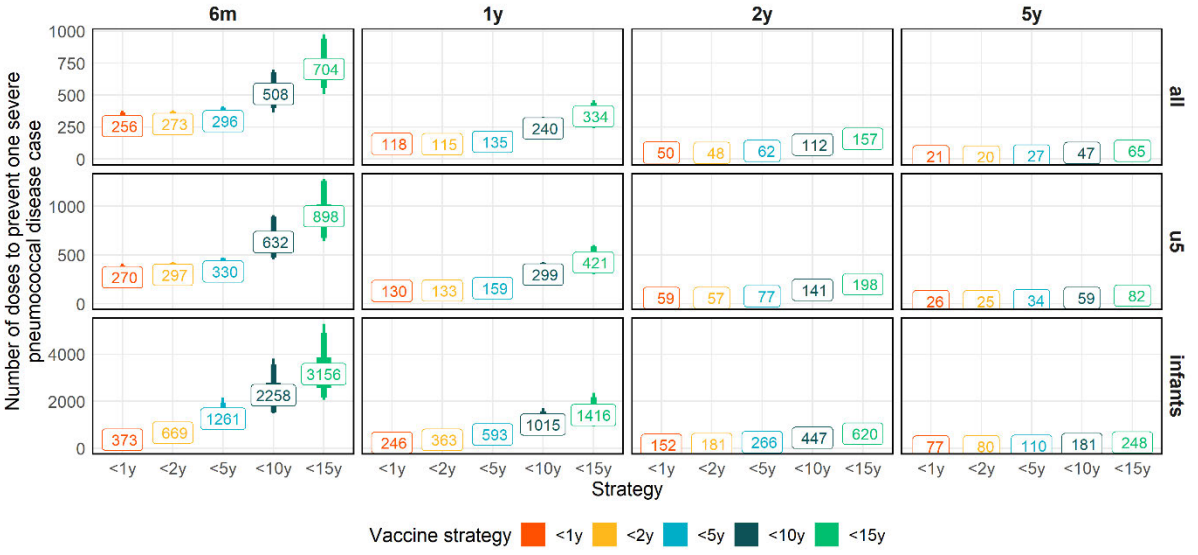
Supplemental Figure D4. Model fit to observed data. Prevalence by age is shown for S (susceptibles), VT (vaccine-type carriers), NVT (non-vaccine-type carriers), VT2 (carriers of 2 or more VTs, but no NVTs), B (both VT and NVT carriers), and NVT2 (carriers of 2 or more NVTs, but no VTs). Observed estimates are shown as black bars with error bars indicating their 95% confidence intervals. Coloured boxen plots show the distribution of modelled prevalence from 1000 model runs sampling from the joint-posterior distribution.

Model outcomes

Supplemental Figure D5 shows the impact of different PCV campaigns on VT prevalence and severe pneumococcal disease incidence over time. Supplemental Figure D6 shows the NNV to prevent one severe pneumococcal disease case for different age groups and periods since the PCV campaign.



Supplemental Figure D5. Impact of PCV mass vaccination campaigns on prevalence and incidence. A: VT carriage prevalence (VT + 2VT + B) in all age groups and infants only. B: reduction in daily severe pneumococcal disease cases compared to no vaccination, in all age groups and infants only. In all plots, thick lines show the median estimates and shaded areas corresponding 95% credible intervals from 500 model posteriors. A dotted vertical line is plotted for infants 1 year after the PCV campaign. Infants to the right of this line have been born after the PCV campaign and are thus all unvaccinated. The reductions in these birth cohorts result from indirect vaccine protection.



Supplemental Figure D6. Number needed to vaccinate to prevent one severe pneumococcal disease case. In different time periods (facet columns) and age groups (facet rows). Estimates show median and 95% CrI from 500 posterior samples.

Our alternative model projected overall much greater impact of all PCV campaigns, with much longer duration. However, qualitatively, this did not alter the relative rank of each

vaccination strategy in terms of impact and efficiency compared to our base model. The base model structure has been used to effectively explain pneumococcal transmission dynamics before and after PCV introduction in several settings[1,11,15,16], while the neutral model has predominantly been used to understand coexistence of biologically similar strains.[17,18] We therefore kept this as the main model in our analysis.

Section E. Global demand forecast

Rationale

In order to more fully illuminate the case for PCV use by humanitarian actors, and to understand the unmet need, we used our identified vaccination strategies to estimate the potential demand for PCV over the next five years.

Methods

We extracted estimates of the number of displaced individuals from the last 5 years from the child displacement dataset from Unicef, which aggregates data from UNHCR refugee statistics, the Global Internal Displacement Database, and UNRWA.[19–22] We assumed that it would be unlikely that PCV campaigns would be needed in high income settings, where refugees would likely be covered by routine immunization schedules, and restricted the dataset to countries classified as low or lower-middle income countries by the World Bank.[23] As the Unicef displacement dataset does not provide age-specific data, we took population estimates for the included countries from UN WPP, assumed the age distributions were representative for the displaced population, and used these values to estimate the age-specific proportions of the displaced populations, e.g. the total number of children <5y.[24]

We then assumed that all children <5y or <10y in currently displaced populations would receive a single dose of PCV, followed by annual catch-up campaigns in children <1y born after the initial MVC. We assumed that the average increasing trend in displaced populations would continue, and that each newly displaced population would as well receive an initial <5y or <10y vaccination campaign, followed by annual catch-up campaigns. We assumed a vaccination coverage of 100% in our calculations, without any adjustment for wasting of vaccine doses.

Results

In 2022, an estimated 112 million people were forcibly displaced in low and lower middle-income countries, which has on average increased by 16 million people annually within the past five years. On average, 14% and 27% of these populations are aged <5y and <10y.

40 million doses of PCV would be needed if an <5y campaign would be implemented in all these populations, followed by annual catch-up campaigns. If <10y campaigns would be implemented, 64 million doses of PCV would be needed. If two doses would be administered in infants, the need would increase to 58 and 87 million doses.

Discussion

While these are only initial estimates assuming PCVs would be used in all displaced populations, they highlight the substantial need there would be in order to make PCVs available to crisis-affected populations.

References used in supplemental material

1 Flasche S, Ojal J, Le Polain de Waroux O, *et al.* Assessing the efficiency of catch-up campaigns for the introduction of pneumococcal conjugate vaccine: a modelling study based on data from PCV10 introduction in Kilifi, Kenya. *BMC Medicine* 2017; **15**. DOI:10.1186/s12916-017-0882-9.

2 Awori JO, PERCH Study Group. Association between malnutrition and pneumococcal pneumonia in the PERCH study. 2018.

3 von Mollendorf C, Cohen C, de Gouveia L, *et al.* Risk Factors for Invasive Pneumococcal Disease Among Children Less Than 5 Years of Age in a High HIV Prevalence Setting, South Africa, 2010 to 2012. *The Pediatric Infectious Disease Journal* 2015; **34**: 27.

4 von Mollendorf C, von Gottberg A, Tempia S, *et al.* Increased Risk for and Mortality From Invasive Pneumococcal Disease in HIV-Exposed but Uninfected Infants Aged <1 Year in South Africa, 2009–2013. *Clinical Infectious Diseases* 2015; **60**: 1346–56.

5 Verhagen LM, Hermsen M, Rivera-Olivero IA, *et al.* Nasopharyngeal carriage of respiratory pathogens in Warao Amerindians: significant relationship with stunting. *Tropical Medicine & International Health* 2017; **22**: 407–14.

6 Verhagen LM, Gómez-Castellano K, Snelders E, *et al.* Respiratory infections in Eñepa Amerindians are related to malnutrition and *Streptococcus pneumoniae* carriage. *J Infect* 2013; **67**: 273–81.

7 Gebre T, Tadesse M, Aragaw D, *et al.* Nasopharyngeal Carriage and Antimicrobial Susceptibility Patterns of *Streptococcus pneumoniae* among Children under Five in Southwest Ethiopia. *Children (Basel)* 2017; **4**: 27.

8 County Department of Health. Kilifi County SMART Survey Report. 2016. https://www.nutritionhealth.or.ke/wp-content/uploads/SMART%20Survey%20Reports/Kilifi%20County%20SMART%20Survey%20Report%20November2016.pdf.

9 Hammitt LL, Etyang AO, Morpeth SC, *et al.* Effect of ten-valent pneumococcal conjugate vaccine on invasive pneumococcal disease and nasopharyngeal carriage in Kenya: a longitudinal surveillance study. *The Lancet* 2019; **393**: 2146–54.

10 Silaba M, Ooko M, Bottomley C, *et al.* Effect of 10-valent pneumococcal conjugate vaccine on the incidence of radiologically-confirmed pneumonia and clinically-defined pneumonia in Kenyan children: an interrupted time-series analysis. *The Lancet Global Health* 2019; **7**: e337–46.

11 Ojal J, Griffiths U, Hammitt LL, *et al.* Sustaining pneumococcal vaccination after transitioning from Gavi support: a modelling and cost-effectiveness study in Kenya. *The Lancet Global Health* 2019; **7**: e644–54.

12 Lipsitch M, Abdullahi O, D'Amour A, *et al.* Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal serotypes in Kenya with a Markov transition model. *Epidemiology* 2012; **23**: 510–9.

13 Melegaro A, Choi YH, George R, Edmunds WJ, Miller E, Gay NJ. Dynamic models of pneumococcal carriage and the impact of the Heptavalent Pneumococcal Conjugate Vaccine on invasive pneumococcal disease. *BMC Infectious Diseases* 2010; **10**: 90.

14 Ter Braak CJF, Vrugt JA. Differential Evolution Markov Chain with snooker updater and fewer chains. *Stat Comput* 2008; **18**: 435–46.

15 Le Polain De Waroux O, Edmunds WJ, Takahashi K, *et al.* Predicting the impact of pneumococcal conjugate vaccine programme options in Vietnam. *Hum Vaccin Immunother* 2018; **14**: 1939–47.

16 Choi YH, Jit M, Flasche S, Gay N, Miller E. Mathematical Modelling Long-Term Effects of Replacing Prevnar7 with Prevnar13 on Invasive Pneumococcal Diseases in England and Wales. *PLOS ONE* 2012; **7**: e39927.

17 Gjini E, Valente C, Sá-Leão R, Gomes MGM. How direct competition shapes coexistence and vaccine effects in multi-strain pathogen systems. *J Theor Biol* 2016; **388**: 50–60.

18 Lipsitch M, Colijn C, Cohen T, Hanage WP, Fraser C. No coexistence for free: Neutral null models for multistrain pathogens. *Epidemics* 2009; **1**: 2–13.

19 United Nations High Commissioner for Refugees. UNHCR - Refugee Statistics. https://www.unhcr.org/refugee-statistics/ (accessed March 2, 2023).

20 Global Internal Displacement Database. IDMC - Internal Displacement Monitoring Centre. https://www.internal-displacement.org/database (accessed Dec 19, 2023).

21 UNRWA Registered Population Dashboard. UNRWA. https://www.unrwa.org/what-we-do/relief-and-social-services/unrwa-registered-population-dashboard (accessed Dec 19, 2023).

22 Child Displacement and Refugees. UNICEF DATA. https://data.unicef.org/topic/child-migration-and-displacement/displacement/ (accessed Dec 19, 2023).

23 World Bank Country and Lending Groups – World Bank Data Help Desk. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups (accessed Dec 19, 2023).

24 United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects - Population Division - United Nations. 2019. https://population.un.org/wpp/Download/Standard/Population/ (accessed April 19, 2020).

Chapter 5

Discussion

This PhD thesis aimed to identify strategies for pneumococcal conjugate vaccine use in humanitarian crises. I hypothesised that an effective strategy would require vaccination of the age groups driving transmission in settings without routine vaccination, outlined in detail in chapters 0 and 1. Using the combined contact and carriage data in a camp for internally displaced people in Somaliland, presented in chapters 2 and 3, I identified children aged 2-4 and 5-14 years as the main age-groups exposing children <5 years to pneumococci. Reducing VT transmission from these age groups though vaccination could therefore, in theory, be an effective approach to prevent pneumococcal disease in this population. The modelling study in chapter 4 formally tested this approach, and indeed identified single-dose PCV campaigns vaccinating children <5 and <10 as effective and efficient interventions to prevent pneumococcal disease cases for up to three years. Crucially, these campaigns were able to partially prevent cases in children born or migrated into the population after the vaccination campaign, which a campaign in children <1 and <2 years could not achieve. The current WHO guidance is to use PCV in humanitarian settings in children under 1 year of age, and to consider its use in children under 5 years of age, using 3 doses in infants (children aged <12 months), a single dose in children older than 2, and 1 or 2 doses in children aged 1-2 years[1]. This guidance is directly implemented in specific guidance of Save the Children and Médecins Sans Frontières (MSF) for the use of PCV - leading humanitarian actors delivering vaccination services to beneficiaries in crisis-affected populations.[2,3] However, the evidence generated in this thesis suggests that wider age-targeting may be needed for improved control of pneumococcal disease in crises.

Summary of research findings

A detailed discussion of the specific research findings is presented within the different research papers. This section summarises key findings in relation to the specific objectives of this PhD in more detail.

I reviewed the literature on the pneumococcal burden and current use of pneumococcal conjugate vaccines in humanitarian crises, the result of which is presented in chapter 1. There is a clear research gap and our current understanding of the pneumococcal disease burden in humanitarian crises is limited. Other reviews have shown that acute respiratory infections and pneumonia specifically are leading causes of morbidity and mortality in crises,[4,5] but the aetiologic contribution of pneumococci in these settings is largely unknown. This is in part due to the difficulty in determining the cause of respiratory infections.[6] Rapid urinary antigen tests are used to determine whether *Streptococcus pneumoniae* is the causative agent in community-acquired pneumonia in adults[7], but their specificity in children

is limited due to high rates of pneumococcal colonization.[6] Additionally, there is general a lack of health research in humanitarian crises, in part because working in these settings is uniquely challenging[8], and may even be deemed unethical if the research is done at the expense of humanitarian action.[9] A substantial pneumococcal burden may be inferred from the high pneumococcal burden observed in other settings before the widespread use of PCVs,[10] and the high prevalence in crisis settings of pneumococcal risk factors that can increase the risk of carriage, disease, and deaths from pneumococci.[11] Pneumococcal carriage is a precursor to disease, and can give important insights in the circulating pneumococci and their age distribution, but there is a general lack of data on pneumococcal carriage in humanitarian settings as well.

Despite improvements in more affordable access for PCVs such as the Humanitarian Mechanism and the development of the PNEUMOSIL vaccine in recent years, the use of PCVs in humanitarian settings has been limited.[12,13] The underutilization of PCVs is likely multifactorial. Use of other vaccines such as the oral cholera vaccine is often affected by the available stockpile[14,15], but this does not affect PCVs. The cost of PCV remains relatively high compared to other vaccines, but other studies have shown that PCV campaigns in humanitarian crises are cost-effective at the WHO threshold.[16,17] While pneumococcal outbreaks have been reported[18,19], they are less visible than those of measles, polio, or cholera, which may result in the prioritisation of other vaccines. Critically, there is a lack of evidence on the effectiveness of PCVs in crises, especially on the effectiveness of mass-vaccination campaigns with alternative age targeting and with reduced doses. These are commonly used for other vaccines in humanitarian settings,[20] especially reactively for outbreak control. The narrative review built a case for informing vaccination strategies through a combination of mathematical modelling and primary data collection to populate key model parameters with realistic, hitherto unavailable estimates.

I designed a study to collect data on social mixing, demography, pneumococcal carriage prevalence and related risk factors in a population affected by a humanitarian crisis. We partnered with Save the Children and MSF to find a suitable field site, and conducted the survey in Digaale, a camp for internally displaced people, in Somaliland in 2019. The study design is described in more detail in chapters 1 and 2. Questions on demography, mortality, migration, and malnutrition (in children aged 6 weeks to 5 years) were based on the Standardized Monitoring and Assessment of Relief and Transitions (SMART) methodology, which was developed to be able to rapidly assess nutritional status of children under five and the mortality rate of the population in a humanitarian crisis.[21] SMART surveys are regularly conducted, which allowed improved comparability of the population living in Digaale to that

of populations in other crises. The social contact survey questions were largely based on similar surveys conducted in other populations.[22–25]

A total of 464 households participated in the survey. Individual contact and risk-factor surveys were completed for 509 participants, while nasopharyngeal swabs were collected from 453 participants. The average household size of 4.5 in the displaced population was lower than the national Somaliland average of 6[26], and slightly lower than average household sizes in IDP populations in Somalia.[27] A high proportion of contacts involved direct touch, most contacts were made at home, and very few contacts were reported to be made outside of the IDP camp. Overall, contacts were mostly age-assortative, in particular in children, while more intergenerational contacts were reported in adults. No other contact data has been collected in other crisis-affected populations or in Somaliland, and surveys in other settings have used different definitions of a contact, which makes comparing contact estimates difficult. However, the average number of direct contacts was similar to those observed in a rural setting in Uganda[22], and lower than those in townships in South Africa and Zimbabwe.[23,28] Overall, the contact data gave important insights to mixing between age groups in an IDP camp. Mixing between age groups was not substantially different from that observed in other settings, but did differ substantially in the location where these contacts occurred. The survey also showed a high prevalence of risk factors for lower respiratory tract infections in the IDP camp, including poor shelter quality, exposure to indoor air pollution, crowing, and acute malnutrition and stunting.

This was also the first study to assess pneumococcal carriage prevalence in an IDP camp or in Somaliland, and only the second study to my knowledge to assess pneumococcal carriage prevalence in a crisis-affected population, with the other instance a survey in Mae La refugee camp in Thailand.[29] Carriage prevalence in Somaliland was 36% in all age groups, 70% in children under 5, and 80% in children under 2. This is similar to prevalence estimates of the Mae La refugee camp and other high transmission settings in East-Africa[29–32], but not as high as has been observed in rural Gambia.[33] Carriage prevalence was associated with the number of household members under 5, recent respiratory symptoms, the total number of physical contacts, and stunting. Most pneumococcal exposure in children under 5 years was due to contact with carriers aged 2-5 and 6-14 years, but not carriers under 2 years. This indicates that the pre-school and school aged children are the main transmitters of pneumococci in this population. Around 60% of circulating serotypes in children under 5 were covered by the different PCVs, though this increased based on the valency of the vaccines. A substantial proportion of circulating serotypes that are likely causing pneumococcal disease are covered by the different PCVs. The proportion of serotypes included in the vaccines (VT) was similar for the two ten-valent vaccines

PNEUMOSIL and Synflorix. Overall, these findings show that any of the PCVs could potentially have a substantial impact in this setting.

I then constructed a pneumococcal transmission model, parameterized using these data, to simulate and assess the effectiveness of different PCV vaccination campaigns in Digaale and other crisis-affected populations. Results suggest that single-dose PCV campaigns targeting children 6 weeks to 1 year only would have a limited impact. However, extending the age-eligibility to include children up to the age of 4 or older, who were identified as the main pneumococcal transmitters in Digaale in a previous analysis and also have a high burden of disease, achieves substantially better impact for up to two or three years. A <5 campaign may prevent up to a third of all IPD cases in a two-year period. Crucially, infants born after the PCV campaign would remain partially protected in this period, highlighting the importance of indirect vaccine protection in achieving this impact. Parametric sensitivity analyses show that increased rates mixing or migration with an unvaccinated host population would reduce the strategy's effectiveness, and broader age targeting may be needed to achieve a similar effect. This was also reflected in other modelled settings, Bentiu camp in South Sudan in 2015 and the city of Maiduguri in Nigeria in 2016. The efficiency of these campaigns was assessed by the number needed to vaccinate to prevent one IPD case. An <5y campaign was the most efficient use of PCV, as the additional doses administered to children 2-4y were more effective in preventing IPD than doses administered to children <2y alone. Due to a lack of costing data, we did not formally assess the cost-effectiveness of these strategies. However, PCV campaigns in Somalia and South Sudan were previously shown to be cost-effective at WHO thresholds using a static model that does not account for indirect vaccination effects.[16,17] Given the importance of the indirect protection, the simulated PCV campaigns would likely be cost-effective as well. There are no other dynamic transmission models that assessed the effect of PCV campaigns in humanitarian settings. However, modelling studies assessing oral cholera vaccine dose administration found similar results, namely that a single dose strategy vaccinating a broader age group is more effective than providing two doses in the youngest ages.[34,35] Like these cholera models, my pneumococcal model highlights the importance of maximizing indirect protection in settings where sufficient vaccination coverage using routine vaccination is not feasible, which may require a shift in the way decision makers approach vaccination strategies. Overall, the modelled data suggest that single-dose PCV campaigns that achieve high vaccination coverage in children up to 4 years or older can be effective in preventing a substantial proportion of IPD for 2-3 years during a humanitarian crisis. This is likely a good proxy for the overall impact on pneumococcal disease and death. If PCV campaigns would be

implemented at a large scale, an estimated 40 to 64 doses may be needed for humanitarian use globally over the next five years.

Limitations and challenges

Primary data collection

A substantial proportion of this thesis describes the cross-sectional survey conducted in Digaale IDP camp in 2019. While this survey generated unique data from a displaced population, it is not without limitations. First, we cannot assess any temporal variation in estimates as we only collected data at a single time point. Migration rates, contact rates, and pneumococcal carriage[36,37] may all vary over time, and we do not know how representative the estimates are for different periods in the year. Pneumococcal serotypes also vary in their frequency, duration of carriage, and invasiveness.[38,39] For example, serotypes 1, 5, and 7 are highly invasive, but rarely identified in carriage data. While we used microarray to serotype pneumococci, which has a good sensitivity in detecting minor serotypes in co-colonizing pneumococci,[40] we may have missed circulating pneumococci that have a shorter duration of carriage by only taking a single swab.

We also did not collect any data in the host population of Hargeisa, and cannot quantify how carriage or contact rates differ to the displaced population. We could only conduct surveys during daylight hours, and men of working age would already have left for jobs in the city of Hargeisa or to tend to their animals. Although I post-stratified estimates to partially correct for this, estimates may be affected by selection bias. This likely only minimally affects pneumococcal carriage prevalence, which is low in these age groups, but these individuals may differ in their contact behaviour which would be unobserved. Many of the collected data were self-reported by participants, and contacts of young children were reported by their parents, and recall bias may have affected their responses. For example, we used local event calendars to determine ages and when asking people whether deaths in their household occurred in the six months preceding the survey. We also made sure to explain the definitions for a household, for a contact, and for self-reported morbidities as much as possible. However, we cannot exclude that errors in data collection may have affected some results, such as the high mortality rates.

Our final sample size was relatively small, with only 509 participants in the contact and risk factor survey, and only 453 participants from whom a nasopharyngeal swab was collected. Our target sample size was 700 people, which we believed would enable us to detect pneumococcal carriage within a single age group with 2-10% precision, and contact rates

between age groups with a precision of 0.7 contacts (Appendix C). One reason for the small sample sized was a smaller overall population size than expected. We were able to use finite population correction to adjust for the smaller population size to calculate population level estimates, which limits the effect on study power for the primary outcomes of the survey.[41] However, we did not apply this correction to assess the relationship between carriage and certain risk factors such as malnutrition, and may be limited in power to detect any such associations.

Data collection was conducted on electronic tablets using Open Data Kit (ODK). While ODK allows automated verification of answers and is easy to use, it is limited in that it is not possible to link previously collected data to a survey form.[42] In principle, this technology is only needed for longitudinal surveys, but we designed our study so that we had multiple contact moments with households and participants, where data was collected using different survey forms. Human error occasionally resulted in errors in the unique household, participant, or swab ID entered in the form. While it was able to retrieve most of these links using manual matching (Chapter 3 Supplemental Material Section A), 53 records for collected swabs could not be completely matched. We have resolved this issue for additional surveys conducted after the initial 2019 survey, not presented in this thesis, by using ODK-X, which does allow for the required linkage of data. We have also used printed barcodes on swab labels, which eliminates the need of manual entry.

Digaale is a relatively stable setting and small IDP camp. It was established in 2014 as a result of a large influx of displaced people in Somaliland following an acute food insecurity crisis. While we found a high prevalence of risk factors common in humanitarian settings, including high rates of mortality, migration, overcrowding, and malnutrition, it may only be representative for a subset of crisis-affected populations and not for other settings, such as large scale IDP or refugee camps like Bentiu or Cox's Bazar during the acute phase of a crisis. This is in part related to the difficulties in conducting research in humanitarian settings.[43] We needed to conduct the study in an established setting, where we know we would have safe access to the population and where we would have minimal resources (a team of data collectors and a cold-chain facility to store nasopharyngeal swabs). In addition, we needed the local support of a humanitarian organization with sufficient capacity to host the research, which we eventually found in Save the Children Somaliland that has been present since 1951. Yet, it may have been helpful to conduct contact and carriage surveys in multiple crisis typologies, not just camp displacement, to understand whether mixing and pneumococcal transmission may differ in other settings. We did not have the resources to study several populations, and the Covid-19 pandemic prohibited us from conducting a

survey in a rural area in South Sudan. Data from other settings, specifically other crisis typologies, would have increased the robustness of our model projections.

Data collection focussed on estimating pneumococcal carriage prevalence and social mixing between age groups. It may have been useful to alter data collection in order to refine assumptions on other important parameters, such as the disease burden or case-carrier ratio of different serotypes in crises. It is challenging to assess pneumococcal disease with high specificity, and existing surveillance systems are rarely adequate in crisis settings.[44]

Sample storage and shipment

We had substantial challenges with sample storage and shipment, which significantly delayed shipment and analysis of nasopharyngeal swab samples. Nasopharyngeal swabs need to be stored at ultra-low temperatures (ULT) for pneumococci to remain viable. We were fortunate to be able to store samples in a ULT freezer at the local tuberculosis hospital, which was the only known ULT freezer in the country at the time. The freezer broke down for two weeks, during which samples had to be placed back in a -20C freezer. Somaliland is not internationally recognized as an independent country. Australia prohibits air cargo originating directly from Somalia,[45] including Somaliland, which required us to ship via a third country. Some shipping companies also placed an internal embargo on Somalia, prohibiting them from working with us. We tried to ship our samples during the Covid-19 pandemic, during which international transport options -especially for biological shipments - were substantially limited. Local airlines that were accepting cargo did not allow the use of dry ice, so we identified phase change materials (PCMs) as a passive cold-chain method to keep samples below -15C during shipment.

Transmission model

The transmission model combines all available data, and allowed simulation of different PCV campaigns. However, there are several limitations to the model. First, some simplifying structural assumptions were made, including grouping all serotypes together as a non-vaccine type (NVT) or vaccine-type, depending on their inclusion in the PCV of interest. This is necessary for computational efficiency to keep the number of model compartments to a manageable level, but ignores specific differences between individual serotypes. In addition, the diamond-shaped model structure artificially promotes co-existence of different serotypes.[46] The model also does not explicitly model acquisition of immunity, but implicitly incorporates this by incorporating empirical data on the age- and serotype specific duration

of carriage,[38] and by reducing the probability of an effective contact for older ages using a fitted parameter. Similar model structures have successfully been used to describe pneumococcal transmission and disease pre- and post-PCV use in a range of different settings.[47–51] However, it is not clear whether the assumptions of this model, specifically around naturally acquired immunity and serotype replacement, hold in populations where PCV is introduced but use is then ceased.

There is also uncertainty surrounding model parameters. While we collected key primary data in Digaale IDP camp, we had to extrapolate these data to represent pneumococcal transmission in other settings. In the absence of local disease data, we applied case-carrier ratios from a study in Kenya to calculate likely IPD cases.[47] We adjusted acquisition and invasiveness rates for malnourished individuals, but the effect of malnutrition on in particular pneumococcal transmission is not understood sufficiently.[11] The effect of malnutrition could have been more explicitly accounted for, e.g. by incorporating different degrees of malnutrition, differentiating between stunting and acute malnutrition, accounting for malnutrition-related mortality, or by incorporating changes in malnutrition prevalence over time. Data were fitted to pre-vaccination estimates from Digaale IDP camp only. The total effect and the longevity of PCV impact is highly dependent on assumptions about the PCV. We modelled PCV campaigns using PNEUMOSIL, and had to assume vaccine efficacy (VE) against VTs is similar to that of other PCVs. There is some uncertainty around the VE[52,53] and especially the duration of vaccine protection[54] of reduced dose schedules, the latter of which especially affects the modelled longevity PCV impact. We conducted parametric sensitivity analyses to better understand the sensitivity and uncertainty of our modelled estimates to these assumptions.

We also assumed that none of the modelled populations had ongoing or previous routine PCV coverage. This may limit the representativeness of the findings to settings where PCV vaccination is or was ongoing. In these settings, some level of herd immunity will already be in place, and narrower age targeting with catch-up vaccination limited to age cohorts that missed their PCV doses may be sufficient to control transmission of VTs without unnecessary wasting of PCV in older children. Implementing a PCV campaign with narrower age targeting may be a good option in settings with good reliable data on existing PCV vaccination coverage. Strategies were assessed on their effect on pneumococcal prevalence and IPD cases, and their efficiency using the NNV. Other outcomes, such as the effect on pneumococcal pneumonia, mortality, or cost per DALY averted could have been used to compare strategies, and guide policy makers when designing PCV campaigns. The modelled outcomes were presented as relative impact, and the impact on other outcomes may be proportional to that on IPD.

Areas for future research

Maintenance strategies

An aspect of PCV use in humanitarian crises not assessed in this PhD is what should be done after a vaccination campaign has been conducted. Ideally, routine vaccination can be (re)established, but that is unfortunately not the reality in the majority of settings. In the absence of subsequent vaccinations, our model projects that VT carriage will eventually return back to the original baseline when vaccine protection has fully waned. One option could be to conduct additional PCV campaigns, potentially only in narrow age groups vaccinating birth-cohorts born after a previous PCV campaign. Such campaigns could be conducted every 24, 12, or 6 months, and compared in terms of their effectiveness and efficiency. Ideally, this would be coupled with cost-effectiveness analyses, as every PCV campaign has substantial fixed costs associated with it. It may also allow exploring the feasibility of offering booster doses at an older age in more detail, and assess whether different maintenance strategy may be needed in different settings, e.g. due to differences in migration.

Evaluating the impact of a PCV campaign

The modelled results suggest substantial impact of PCV campaigns, but ideally these would be validated by empirical evidence. We received funding from the Bill and Melinda Gates Foundation for the Evaluating the Effectiveness of a Pneumococcal Immunization Campaign in a Crisis-affected population (EEPICC) study (https://clinicaltrials.gov/study/NCT04945681). In this study, which is currently ongoing and was directly informed by the work presented in this thesis, we will estimate the impact of a PCV campaign using PNEUMOSIL in children <5y in the Digaale IDP camp on pneumococcal carriage prevalence. Figure 1 shows a timeline of this study, including the timing of the initial survey and modelling work presented in this thesis in relation to this follow-up study. We gave 2 doses of PNEUMOSIL to all consenting infants aged 6 weeks to 1 year, spaced 4 weeks apart, and a single dose to all participating children aged 1 to 4. A change in carriage prevalence will be assessed 6 months, 1 year, and 2 years after the PCV campaign, compared to a new baseline collected just before the PCV campaign. We also conducted a carriage survey in the host population in the city of Hargeisa, which will allow us to understand potential differences in carriage between the two populations.

This is not the only study assessing the impact of a PCV campaign in a humanitarian setting. MSF Epicentre assessed the effect of administering a single full or fractional dose of

PNEUMOSIL in children aged 1 to 9 years, compared to a control population where no campaign was conducted using a cluster-randomized trial in the Maradi region in Niger.[55] They found a substantial reduction in VT carriage in both intervention arms. Fractional dose campaigns could be useful to reduce the cost of the still relatively expensive PCVs.

A third study is also evaluating the impact of a PCV campaign in the Sahel region in Burkina Faso. The empirical evidence of these studies will substantially contribute to the modelled evidence already generated in this thesis, and can be used to validate the modelled results.

Improving model parameters

The modelled results were sensitive to the assumption around the degree of mixing between populations (chapter 4). Research to improve certain modelled estimates is already ongoing.

I received a Robert Austrian Research Award at the 2022 meeting of the International Society of Pneumonia and Pneumococcal Diseases (ISPPD) to quantify pneumococcal transmission rates between the displaced population in Digaale and the non-displaced population in Hargeisa, informed by DNA sequencing of pneumococci collected in the 2022 carriage surveys as part of the EEPICC study. The results of this analysis will improve the robustness of the transmission model to the assumption of mixing in the Digaale population.[56]

The high modelled impact of PCV campaigns is dependent on the assumed VE of single-dose vaccination. A large cluster-randomized trial assessing the non-inferiority of reduced dose PCV strategies (1p+1 and 0p+1) against three dose schedules (3p+0 and 2p+1) on VT carriage has recently been concluded in Nha Trang, Vietnam. It will provide additional insights in the VE against transmission of different vaccine doses. A follow-up carriage survey conducted one year after ceasing vaccination as part of the trial will provide additional insights into the speed of resurgence of VTs. A similar trial, the PVS study, comparing the effect of a 1p+1 to a 3p+0 schedule in the high-transmission setting in the Gambia is also concluding, and will provide additional insights in the VE of reduced doses.[57]

Other interventions

As is the case for PCV, other vaccines such as those protecting against rotavirus, human papillomavirus, malaria, and respiratory syncytial virus are also underutilized. In general, there is a lack of evidence-based guidance for interventions in humanitarian settings.[58,59] The evidence generation pathway combining modelling with minimal data collection of key

parameters has been effective to assess the likely impact of PCV vaccination campaigns, and may be used as a reference when attempting to improve the evidence base for other interventions. Ideally, the effect and cost-effectiveness of PCV would be directly compared to other interventions, as a stand-alone intervention or as a multi-intervention strategy, to assist policy makers in deciding which services to implement.

Conclusions

In my PhD research, I designed and conducted a cross-sectional survey in a camp for displaced people in Somaliland and collected novel data on social mixing and nasopharyngeal carriage prevalence. The lack of data from humanitarian settings remains a challenge. Researchers should aim to include crisis-affected populations when designing studies, ideally in collaborations with INGOs and local partners. My modelling work shows that current guidance on the use of PCV in crisis settings is insufficient: for a PCV campaign to be effective, the main transmitters in addition to those at highest risk of severe disease should be vaccinated for a substantial impact on pneumococcal transmission and disease that lasts for 2 to 3 years. Single-dose campaigns, which are operationally much easier to implement, can be effective, given a high enough vaccination coverage and broad enough age targeting. Vaccinating children up to the age of 4 is an effective and efficient use of PCV, while extending the age range can further improve the effectiveness of the vaccination campaign. Broader age targeting may be needed to mitigate suboptimal impact in settings with high degree of mixing with an unvaccinated population, or high migration rates. The number of people affected by humanitarian crises is large and growing, and likely experience high burdens of pneumococcal disease. Single-dose mass vaccination PCV campaigns are effective, and should be considered by decision makers.

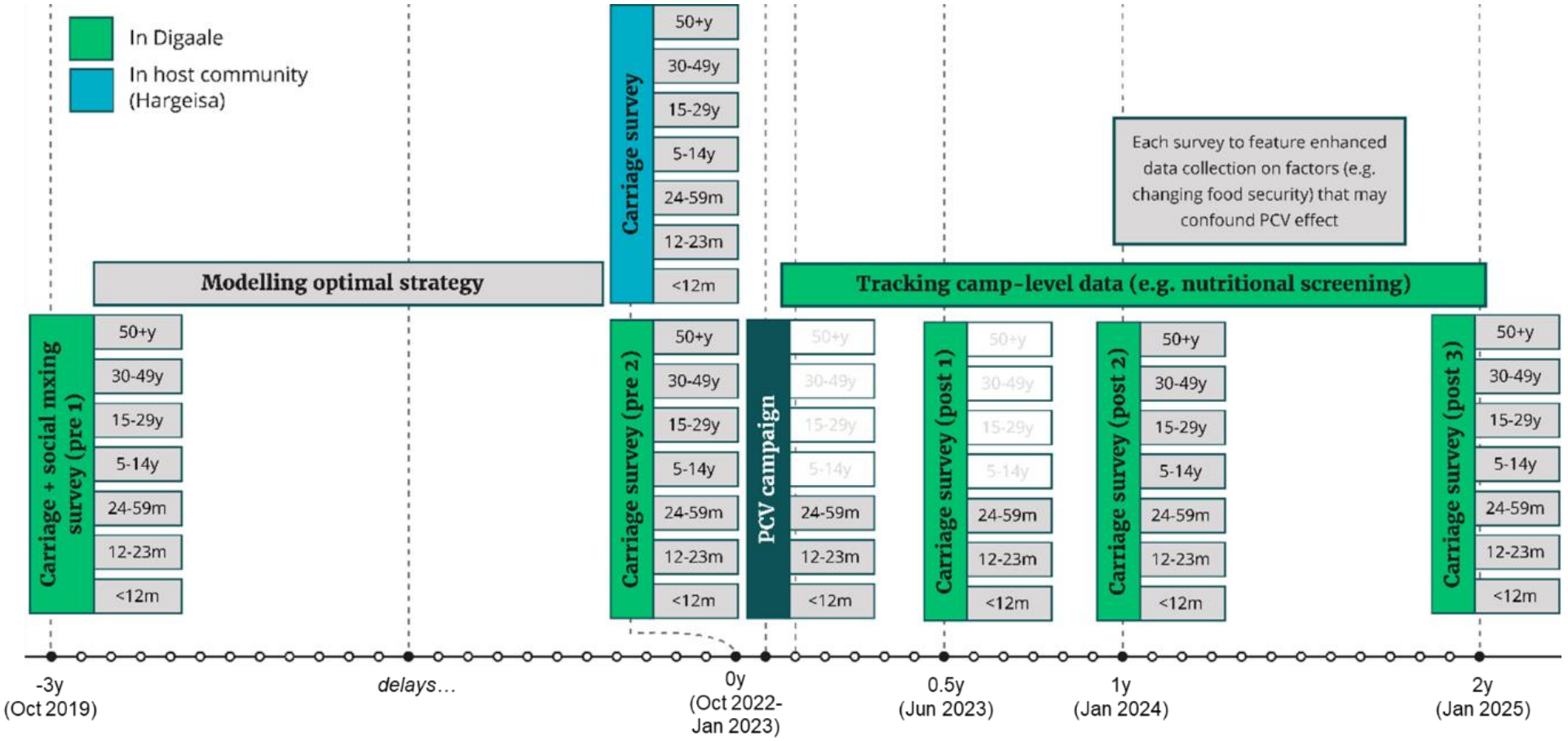


Figure 1. Study timeline of the EEPICC study: Evaluating the Effectiveness of a Pneumococcal Immunization Campaign in a Crisis-affected population.

Reference list

1 World Health Organization. Pneumococcal conjugate vaccines in infants and children under 5 years of age: WHO position paper – February 2019. *Wkly Epidemiol Rec* 2019; **94**: 85–104.

2 Médecins Sans Frontières. Pneumococcal Conjugate Vaccine (PCV). MSF Medical Guidelines. https://medicalguidelines.msf.org/en/viewport/EssDr/english/pneumococcal-conjugate-vaccine-pcv-39851029.html (accessed May 20, 2024).

3 Save the Children. Mass vaccination: Pneumococcal conjugate vaccine (PCV). COMPASS for Health, Nutrition & WASH. https://compass.savethechildren.org.uk/ (accessed May 20, 2024).

4 Bellos A, Mulholland K, O'Brien KL, Qazi SA, Gayer M, Checchi F. The burden of acute respiratory infections in crisis-affected populations: a systematic review. *Conflict and Health* 2010; **4**: 3.

5 Chen SJ, Walker PJ, Mulholland K, Graham HR. Childhood pneumonia in humanitarian emergencies in low- and middle-income countries: A systematic scoping review. *J Glob Health*; **12**: 10001.

6 Rueda ZV, Aguilar Y, Maya MA, *et al.* Etiology and the challenge of diagnostic testing of community-acquired pneumonia in children and adolescents. *BMC Pediatr* 2022; **22**: 169.

7 Viasus D, Calatayud L, McBrown MV, Ardanuy C, Carratalà J. Urinary antigen testing in community-acquired pneumonia in adults: an update. *Expert Rev Anti Infect Ther* 2019; **17**: 107–15.

8 Kohrt BA, Mistry AS, Anand N, Beecroft B, Nuwayhid I. Health research in humanitarian crises: an urgent global imperative. *BMJ Global Health* 2019; **4**: e001870.

9 Pringle JD, Cole DC. Health Research in Complex Emergencies: A Humanitarian Imperative. *J Acad Ethics* 2009; **7**: 115–23.

10 Troeger C, Blacker B, Khalil IA, *et al.* Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Infectious Diseases* 2018; **18**: 1191–210.

11 van Zandvoort K, Checchi F, Diggle E, *et al.* Pneumococcal conjugate vaccine use during humanitarian crises. *Vaccine* 2019; **37**: 6787–92.

12 Leach K, Checchi F. The utilisation of vaccines in humanitarian crises, 2015–2019: A review of practice. *Vaccine* 2022; **40**: 2970–8.

13 Unicef. Update on Humanitarian Mechanism. 2018. https://www.unicef.org/supply/files/SESSION_7_2018_VIC_Update_on_Humanitarian_Mechanism.pdf.

14 Yen C, Hyde TB, Costa AJ, *et al.* The development of global vaccine stockpiles. *Lancet Infect Dis* 2015; **15**: 340–7.

15 Pawar P. The world's stockpile of cholera vaccine is empty—but relief is on the way. Science | AAAS. 2024. https://www.science.org/content/article/world-s-stockpile-cholera-vaccine-empty-relief-way (accessed May 7, 2024).

16 Gargano LM, Hajjeh R, Cookson ST. Pneumonia prevention: Cost-effectiveness analyses of two vaccines among refugee children aged under two years, *Haemophilus influenzae* type b-containing and pneumococcal conjugate vaccines, during a humanitarian emergency, Yida camp, South Sudan. *Vaccine* 2017; **35**: 435–42.

17 Gargano LM, Hajjeh R, Cookson ST. Pneumonia Prevention during a Humanitarian Emergency: Cost-effectiveness of *Haemophilus Influenzae* Type B Conjugate Vaccine and Pneumococcal Conjugate Vaccine in Somalia. *Prehosp Disaster Med* 2015; **30**: 402–11.

18 Coldiron ME, Touré O, Frank T, Bouygues N, Grais RF. Outbreak of Pneumococcal Meningitis, Paoua Subprefecture, Central African Republic, 2016–2017. *Emerg Infect Dis* 2018; **24**: 1720–2.

19 Crellen T, Rao VB, Piening T, Zeydner J, Siddiqui MR. Seasonal upsurge of pneumococcal meningitis in the Central African Republic. *Wellcome Open Research* 2018; **3**: 134.

20 Lam E, McCarthy A, Brennan M. Vaccine-preventable diseases in humanitarian emergencies among refugee and internally-displaced populations. *Human Vaccines & Immunotherapeutics* 2015; **11**: 2627–36.

21 SMART. Standardized Monitoring and Assessment for Relief and Transitions Manual 2.0. 2017. https://smartmethodology.org/wp-content/uploads/2018/02/SMART-Manual-2.0_Final_January-9th-2017-for-merge-3.pdf.

22 le Polain de Waroux O, Cohuet S, Ndazima D, *et al.* Characteristics Of Human Encounters And Social Mixing Patterns Relevant To Infectious Diseases Spread By Close Contact: A Survey In Southwest Uganda. *BMC Infect Dis* 2017; **18**. DOI:10.1101/121665.

23 Melegaro A, Del Fava E, Poletti P, *et al.* Social Contact Structures and Time Use Patterns in the Manicaland Province of Zimbabwe. *PLoS ONE* 2017; **12**: e0170459.

24 Horby P, Thai PQ, Hens N, *et al.* Social Contact Patterns in Vietnam and Implications for the Control of Infectious Diseases. *PLOS ONE* 2011; **6**: e16965.

25 Mossong J, Hens N, Jit M, *et al.* Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLOS Medicine* 2008; **5**: e74.

26 Central Statistics Department, Ministry of Planning and National Development, Somaliland Government, Ministry of Planning and National Development, Somaliland Government. The Somaliland Health and Demographic Survey 2020. 2020.

27 Hanmer LC, Rubiano-Matulevich E, Santamaria J. Differences in Household Composition: Hidden Dimensions of Poverty and Displacement in Somalia. The World Bank, 2021 DOI:10.1596/1813-9450-9818.

28 Johnstone-Robertson SP, Mark D, Morrow C, *et al.* Social Mixing Patterns Within a South African Township Community: Implications for Respiratory Disease Transmission and Control. *Am J Epidemiol* 2011; **174**: 1246–55.

29 Turner P, Turner C, Jankhot A, *et al.* A Longitudinal Study of *Streptococcus pneumoniae* Carriage in a Cohort of Infants and Their Mothers on the Thailand-Myanmar Border. *PLOS ONE* 2012; **7**: e38271.

30 Nackers F, Cohuet S, le Polain de Waroux O, *et al.* Carriage prevalence and serotype distribution of *Streptococcus pneumoniae* prior to 10-valent pneumococcal vaccine introduction: A population-based cross-sectional study in South Western Uganda, 2014. *Vaccine* 2017; **35**: 5271–7.

31 Heinsbroek E, Tafatatha T, Phiri A, *et al.* Pneumococcal carriage in households in Karonga District, Malawi, before and after introduction of 13-valent pneumococcal conjugate vaccination. *Vaccine* 2018; **36**: 7369–76.

32 Abdullahi O, Karani A, Tigoi CC, *et al.* The prevalence and risk factors for pneumococcal colonization of the nasopharynx among children in Kilifi District, Kenya. *PLoS One* 2012; **7**: e30787.

33 Hill PC, Akisanya A, Sankareh K, *et al.* Nasopharyngeal Carriage of *Streptococcus pneumoniae* in Gambian Villagers. *Clin Infect Dis* 2006; **43**: 673–9.

34 Havumaki J, Meza R, Phares CR, Date K, Eisenberg MC. Comparing alternative cholera vaccination strategies in Maela refugee camp: using a transmission model in public health practice. *BMC Infectious Diseases* 2019; **19**: 1075.

35 Leung T, Eaton J, Matrajt L. Optimizing one-dose and two-dose cholera vaccine allocation in outbreak settings: A modeling study. *PLOS Neglected Tropical Diseases* 2022; **16**: e0010358.

36 Bojang A, Jafali J, Egere UE, *et al.* Seasonality of Pneumococcal Nasopharyngeal Carriage in Rural Gambia Determined within the Context of a Cluster Randomized Pneumococcal Vaccine Trial. *PLOS ONE* 2015; **10**: e0129649.

37 Numminen E, Chewapreecha C, Turner C, *et al.* Climate induces seasonality in pneumococcal transmission. *Sci Rep* 2015; **5**. DOI:10.1038/srep11344.

38 Lipsitch M, Abdullahi O, D'Amour A, *et al.* Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal serotypes in Kenya with a Markov transition model. *Epidemiology* 2012; **23**: 510–9.

39 Løchen A, Truscott JE, Croucher NJ. Analysing pneumococcal invasiveness using Bayesian models of pathogen progression rates. *PLoS Comput Biol* 2022; **18**: e1009389.

40 Satzke C, Dunne EM, Porter BD, Klugman KP, Mulholland EK, PneuCarriage project group. The PneuCarriage Project: A Multi-Centre Comparative Study to Identify the Best Serotyping Methods for Examining Pneumococcal Carriage in Vaccine Evaluation Studies. *PLOS Medicine* 2015; **12**: e1001903.

41 Lumley T. Complex Surveys: A Guide to Analysis Using R. John Wiley & Sons, 2011.

42 ODK-X. Selecting the Appropriate Tool Suite. ODK-X Documentation. 2023. https://docs.odk-x.org/select-tool-suite/ (accessed May 7, 2024).

43 Smith J, Blanchet K. Research Methodologies in Humanitarian Crises. 2019. https://www.elrha.org/wp-content/uploads/2020/02/R2HC-Research-Methodologies-in-Humanitarian-Crises-new.pdf (accessed May 28, 2024).

44 Abuzerr S, Zinszer K, Assan A. Implementation challenges of an integrated One Health surveillance system in humanitarian settings: A qualitative study in Palestine. *SAGE Open Medicine* 2021; **9**: 20503121211043038.

45 Australian Government Department of Home Affairs. Aviation Transport Security (Prohibited Cargo - Somalia) Instrument 2015. 2015. https://www.legislation.gov.au/F2015L02057/asmade (accessed May 7, 2024).

46 Lipsitch M, Colijn C, Cohen T, Hanage WP, Fraser C. No coexistence for free: Neutral null models for multistrain pathogens. *Epidemics* 2009; **1**: 2–13.

47 Flasche S, Ojal J, Le Polain de Waroux O, *et al.* Assessing the efficiency of catch-up campaigns for the introduction of pneumococcal conjugate vaccine: a modelling study based on data from PCV10 introduction in Kilifi, Kenya. *BMC Medicine* 2017; **15**. DOI:10.1186/s12916-017-0882-9.

48 Ojal J, Griffiths U, Hammitt LL, *et al.* Sustaining pneumococcal vaccination after transitioning from Gavi support: a modelling and cost-effectiveness study in Kenya. *The Lancet Global Health* 2019; **7**: e644–54.

49 Choi YH, Jit M, Flasche S, Gay N, Miller E. Mathematical Modelling Long-Term Effects of Replacing Prevnar7 with Prevnar13 on Invasive Pneumococcal Diseases in England and Wales. *PLOS ONE* 2012; **7**: e39927.

50 Le Polain De Waroux O, Edmunds WJ, Takahashi K, *et al.* Predicting the impact of pneumococcal conjugate vaccine programme options in Vietnam. *Hum Vaccin Immunother* 2018; **14**: 1939–47.

51 Masala GL, Lipsitch M, Bottomley C, Flasche S. Exploring the role of competition induced by non-vaccine serotypes for herd protection following pneumococcal vaccination. *J R Soc Interface* 2017; **14**. DOI:10.1098/rsif.2017.0620.

52 Temple B, Tran HP, Dai VTT, *et al.* Efficacy against pneumococcal carriage and the immunogenicity of reduced-dose (0 + 1 and 1 + 1) PCV10 and PCV13 schedules in Ho Chi Minh City, Viet Nam: a parallel, single-blind, randomised controlled trial. *The Lancet Infectious Diseases* 2023; **23**: 933–44.

53 Berman-Rosa M, O'Donnell S, Barker M, Quach C. Efficacy and Effectiveness of the PCV-10 and PCV-13 Vaccines Against Invasive Pneumococcal Disease. *Pediatrics* 2020; **145**: e20190377.

54 Le Polain De Waroux O, Flasche S, Prieto-Merino D, Goldblatt D, Edmunds WJ. The Efficacy and Duration of Protection of Pneumococcal Conjugate Vaccines Against Nasopharyngeal Carriage: A Meta-regression Model. *The Pediatric Infectious Disease Journal* 2015; **34**: 858–64.

55 Coldiron M, Soumana I, Langendorf C, *et al.* OA-240 Determining whether mass vaccination campaigns with fractional-dose PCV10 (PNEUMOSIL) could accelerate group protection against pneumococcal transmission in sub-Saharan Africa. *BMJ Global Health* 2023; **8**. DOI:10.1136/bmjgh-2023-EDC.20.

56 van Zandvoort K, Hassan AI, Ibrahim S, *et al.* Quantifying pneumococcal transmission rates between an internally displaced and non-displaced population in Somaliland. The 13th International Symposium on Pneumococci and Pneumococcal Diseases (ISPPD), Cape Town, South Africa, 2024.

57 Mackenzie GA, Osei I, Salaudeen R, *et al.* A cluster-randomised, non-inferiority trial of the impact of a two-dose compared to three-dose schedule of pneumococcal conjugate vaccination in rural Gambia: the PVS trial. *Trials* 2022; **23**: 71.

58 Turner T, Green S, Harris C. Supporting Evidence-based Health Care in Crises: What Information Do Humanitarian Organizations Need? *Disaster Medicine and Public Health Preparedness* 2011; **5**: 69–72.

59 Blanchet K, Ramesh A, Frison S, *et al.* Evidence on public health interventions in humanitarian crises. *The Lancet* 2017; **390**: 2287–96.

Appendix A

Ethical approvals from the Research Ethical Committees at the London School of Hygiene and Tropical Medicine, and the Republic of Somaliland Ministry of Health Development

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
United Kingdom
Switchboard: +44 (0)20 7636 8636

www.lshtm.ac.uk

Observational / Interventions Research Ethics Committee

Mr Kevin van Zandvoort
LSHTM

10 July 2019

Dear Kevin,

Study Title: ESPICC | Cross-sectional study to estimate social contacts and nasophayngeal carriage of pneumococcus in Digale IDP camp near Hargeysa, Somaliland

LSHTM Ethics Ref: 16577

Thank you for responding to the Observational Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Conditions of the favourable opinion

Approval is dependent on local ethical approval having been received, where relevant.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document Type	File Name	Date	Version
Information Sheet	Information sheet	15/05/2019	1
Information Sheet	Informed Consent_Participant with impartial witness_Final_23.09.16	15/05/2019	1
Protocol / Proposal	questionnaire_paper_household	15/05/2019	1
Protocol / Proposal	questionnaire_paper_contacts	15/05/2019	1
Information Sheet	Informed Assent_Cild_and Representative	21/05/2019	1
Investigator CV	KvZandvoort	21/05/2019	1
Investigator CV	SFlasche	21/05/2019	1
Investigator CV	FChecchi	21/05/2019	1
Investigator CV	Dr. Mohamed Bobe CV (003)	21/05/2019	1
Investigator CV	CV_Emma Diggle	21/05/2019	1
Investigator CV	K. Mulholland CV - Sept 2018	21/05/2019	1
Investigator CV	OlePolain	21/05/2019	1
Investigator CV	BRao	21/05/2019	1
Protocol / Proposal	LSHTM Protocol Template - ESPICC	31/05/2019	1
Covering Letter	response_LSHTM_ethics_committee	09/07/2019	1

After ethical review

The Chief Investigator (CI) or delegate is responsible for informing the ethics committee of any subsequent changes to the application. These must be submitted to the Committee for review using an Amendment form. Amendments must not be initiated before receipt of written favourable opinion from the committee.

The CI or delegate is also required to notify the ethics committee of any protocol violations and/or Suspected Unexpected Serious Adverse Reactions (SUSARs) which occur during the project by submitting a Serious Adverse Event form.

An annual report should be submitted to the committee using an Annual Report form on the anniversary of the approval of the study during the lifetime of the study.

At the end of the study, the CI or delegate must notify the committee using an End of Study form.

All aforementioned forms are available on the ethics online applications website and can only be submitted to the committee via the website at: http://leo.lshtm.ac.uk

Additional information is available at: www.lshtm.ac.uk/ethics

Yours sincerely,

[redacted]

Professor Jimmy Whitworth
Chair

ethics@lshtm.ac.uk
http://www.lshtm.ac.uk/ethics/

Improving health worldwide

Jamhuuriyadda Somaliland | Republic of Somaliland
Wasaaradda Horumarinta Caafimaadka | Ministry of Health Development
Agaasimaha Guud | The Director General

Ref:MOHD/DG:3/13075/2019 Date:28/8/2019

TO: SAVE THE CHILDREN AND LONDON SCHOOL OF HYGIENE AND TROPICAL MEDICINE

Cc: Director of Planning Policy and strategic information MOHD

Cc: Vice Minister of MOHD

Cc: Minister of MOHD

SUBJECT: EVALUATING STRATEGIES FOR PNEUMOCOCCAL IMMUNIZATION SURVEY

The ministry of health Somaliland, under the process of the department of planning/policy and strategic information, unit of research has reviewed and discussed your proposal documents of "Evaluating Strategies for Pneumococcal Immunization Survey in Digaale IDP located Hargeisa Somaliland' has been approved to conduct and implement.

1	Name of Principal Investigator	SAVE THE CHILDREN AND LONDON SCHOOL OF HYGIENE AND TROPICAL MEDICINE
2	Name of the document	Evaluating Strategies for Pneumococcal Immunization Survey
3	Questionnaires	Accepted/approved
4	Study proposal/protocol	Accepted and approved

We approved the study to be conducted in the presented form, the ministry of health Somaliland also expects to be informed about any changes occurring during the study and any revision in the protocol and tools, the copy of the final report (soft as well as hard copy will also be submitted to the ministry of Health Development)

NB: Custodial of the data is owned by the ministry Development

[redacted]

Director General of MOHD

Tell: +252 4466061 Email:herrgeye@gmail.com

Appendix B

Study protocol for cross-sectional survey

ESPICC Digale

Evaluating Strategies for Pneumococcal Immunisation Campaigns in Crises

Cross-sectional study Digale IDP camp

Version 1.2, 21/5/2019

SPONSOR: London School of Hygiene & Tropical Medicine
FUNDERS: Elhra's Research for Health in Humanitarian Crises (R2HC) Programme and the Global Challenges Research Fund

Protocol authorised by:

Name:	[redacted]e	**Role:**	Co-Principal Investigator
Signature:	[redacted]	**Date:**	31/05/2019
Name:	**Francesco Checchi**	**Role:**	Co-Principal Investigator
Signature:	[redacted]	**Date:**	31 May 2019

Main Contacts

Co-Principal Investigator: Dr Stefan Flasche; Stefan.Flasche@lshtm.ac.uk
Co-Principal Investigator: Prof Francesco Checchi; Francesco.Checchi@lshtm.ac.uk
Co-investigator: Mr Kevin van Zandvoort; Kevin.Van-Zandvoort@lshtm.ac.uk

For general queries, supply of documentation, and collection of data, please contact:

Co-investigator: Mr Kevin van Zandvoort
Address: Keppel Street, WC1E 7HT, London, UK
E-mail: Kevin.Van-Zandvoort@lshtm.ac.uk

Funder
This study is funded by Elhra's Research for Health in Humanitarian Crises (R2HC) Programme, which aims to improve health outcomes by strengthening the evidence base for public health interventions in humanitarian crises. The R2HC programme is funded by the UK Government (DFID), the Wellcome Trust, and the UK National Institute for Health Research (NIHR). This study receives additional funding from the Global Challenges Research Fund (GCRF).

Table of Contents

GLOSSARY OF ABBREVIATIONS

ARI	Acute respiratory tract infections
GAM	Global Acute Malnutrition
IDP	Internally Displaced People
LMIC	Lower and Middle Income Countries
MCMC	Markov Chain Monte Carlo
MCRI	Murdoch Children's Research Institute
MUAC	Middle-Upper Arm Circumference
NP	Nasopharyngeal
NVT	Non vaccine-targeted serotypes
ODK	Open Data Kit
PCV	Pneumococcal Conjugate Vaccine
SAM	Severe Acute Malnutrition
SMART	Standardized Monitoring and Assessment for Relief and Transitions
STGG	Skim milk tryptoneglucose-glycerol
VT	Vaccine-targeted serotypes
WHO	World Health Organization

STUDY SUMMARY

TITLE Evaluate Strategies for Pneumococcal Immunisation Campaigns in Crises

DESIGN Cross-sectional study design.

AIMS The primary aim for the ESPICC study is to identify effective and efficient pneumococcal conjugate vaccine (PCV) strategies during humanitarian crises. The aim for this cross-sectional study is to collect relevant data to parametrize mathematical models, and use these models in order to identify optimal strategies for PCV use in humanitarian crises.

OUTCOME MEASURES The outcomes for this cross-sectional study are: 1) the age specific point-prevalence of pneumococcal carriage; 2) the distribution of pneumococcal serotypes; 3) the average number of age-specific daily social contacts; and 4) the point-prevalence of malnutrition in children aged 6-59 months.

POPULATION Internally Displaced People (IDP) and refugees living in Digale IDP camp.

ELIGIBILITY All individuals living in Digale IDP camp will be eligible for study participation, except if they have head or facial injuries that would contraindicate swabbing of the nasopharynx.

DURATION Data will be collected over a 4-6 week period.

1. INTRODUCTION

1.1 BACKGROUND

Nearly 1% of the world's population was forcibly displaced due to insecurity and war in 2017.[1] In the same year, hundreds of millions were affected by armed conflicts,[2,3] and almost 100 million were impacted by natural disasters.[4] Crises substantially affect people's lives, and can dramatically increase premature mortality.[5–7] In most crises, these excess deaths are often attributable to the indirect effects of crisis-emergent factors, such as the breakdown of public health services, food insecurity, inadequate water and sanitation, and overcrowding; factors that increase both the incidence and severity of disease.[8,9] Infectious diseases are of particular concern, and require specific control measures such as vaccines.

Streptococcus pneumoniae (the pneumococcus) is a human commensal that commonly resides in the nasopharynx, and occasionally causes disease (e.g. pneumonia, meningitis, and sepsis), especially in young children and people with weakened immune systems.[10] The pneumococcal disease burden in crises is largely unknown, but likely substantial. Pneumococcal meningitis outbreaks have occasionally been reported in humanitarian settings,[11,12] and pneumococcal pneumonia is a major concern. During crises, acute respiratory tract infections (ARI) and diarrhoeal disease make up the top two causes of morbidity in all age groups, with ARIs alone accounting for 20-35% of mortality in children younger than five years of age.[13] The exact aetiology of these ARIs remains unknown, but more than half of all ARI-related deaths worldwide were caused by pneumococci in the pre-pneumococcal conjugate vaccination era.[14] Risk factors that are commonly exacerbated in crises, such as malnutrition, indoor air pollution, and overcrowding, can increase pneumococcal carriage, transmission, disease, and mortality (Table 1), which likely amplifies this burden in crises. In addition, the displacement and crowding of people from a range of different communities may expose them to a range of circulating serotypes that they have not seen before, increasing the risk of disease and probably extending the risk even more into older age groups.

Table 1. Crisis-emergent risk factors that can plausibly affect the pneumococcal burden.

Risk factor	Increased transmission (carriage)	Increased probability that carriage leads to disease	Increased case-fatality ratio	Selected references
Acute malnutrition	**++***	**+++***	**+++***	15,16
Measles outbreaks and other viral respiratory tract infections	**++**	**++**	**++**	17–19
Overcrowding and altered social contact patterns	**+++***	**-**	**-**	15,16,20,21
Disrupted routine pneumococcal conjugate vaccine use	**+[i]**	**+++**	**-**	22–24
Low access to curative care	**-**	**+**	**+++**	25,26
Smoke inhalation	**-**	**+**	**-**	27,28
Inadequate water and sanitation	**++**	**+**	**-**	29,30

* potential shift in the age-specific risk (younger average acquisition and increased carriage and disease among all age groups).
i increase in carriage of vaccine-targeted serotypes, but not in overall carriage.
- no effect on outcome; + small effect on outcome; ++ medium effect on outcome; +++ large effect on outcome

To date, only a small subset of licensed vaccines that are routinely used in most stable settings is commonly used in humanitarian crises. These usually include measles, polio, and (recently) cholera, with context-specific threats such as meningococcal disease or yellow fever infrequently addressed.[31] However, the prioritisation of pathogens targeted by these vaccines may not comprehensively address the local anticipated preventable disease burden. More recent additions to the vaccine portfolio, such as vaccines protecting against *Streptococcus pneumoniae*, have rarely been used in humanitarian settings.

Pneumococcal conjugate vaccines (PCVs) effectively protect against pneumococcal disease.[10] There are currently two PCV products available, protecting against 10 (PCV10) or 13 (PCV13) of more than 90 known pneumococcal serotypes. In contrast to (unconjugated) pneumococcal polysaccharide vaccines,[32]

PCVs are recommended for use in children and, in addition to the direct protection against pneumococcal disease, also elicit indirect protection through interrupted transmission of vaccine-targeted serotypes (VT).[10] Although their impact is partially dampened by replacement colonisation of the nasopharynx by non-vaccine serotypes, these serotypes are generally less likely to cause severe disease, resulting in a net benefit.[33] PCVs have now been introduced in the routine childhood immunisation programmes of the majority of countries.[34] In most places where PCVs are used at high coverage, the marked reduction in VT transmission has expanded the benefit beyond vaccinees alone.[35–38]

Insufficient evidence on the causes underlying the disease burden during crises and limited guidance on vaccine priorities for humanitarian decision-makers may partly explain the hitherto narrow uptake of vaccine interventions. In an attempt to improve this situation the World Health Organization (WHO) introduced a Framework for Decision-Making on Vaccination in Humanitarian Emergencies in 2012, which was updated in 2017.[39] This three-step framework aims to implement the most appropriate vaccination interventions in each crisis, given the local epidemiology, vaccine characteristics, and other context-specific considerations. The framework emphasises expanding the range of vaccines offered to crisis-affected populations, but also recommends adapted vaccination strategies, including expanded age ranges and reduced-dose regimens.

Although the WHO Framework lists PCVs as one of the vaccines to be considered for use in crises,[39] and despite a likely high preventable pneumococcal disease burden, they have rarely been used during crises.[40–43] The rationale for integrated PCV vaccination strategies in crises is clear: mass vaccination campaigns delivered as part of the initial package of interventions in the acute emergency phase of new crises could rapidly establish direct and indirect protection when vulnerability due to malnutrition, congestion of unplanned settlements, and lack of curative health services is likely to peak. These campaigns should ideally be multi-antigen interventions (e.g. bundling measles and cholera) or multi-interventional (e.g. bed nets or micronutrient supplementation).

A PCV-specific barrier to vaccination in crises has long been its price. While PCVs have been prohibitively expensive, a "Humanitarian Mechanism" sponsored in 2017 by the WHO, Unicef, Médecins Sans Frontières and Save the Children now guarantees more affordable PCV procurement by humanitarian actors and affordable expedited delivery.[44] Although some 600 000 doses of PCV have been delivered through this mechanism to date,[42] this only covers a small proportion of crises affected populations at risk.

1.2 RATIONALE FOR CURRENT STUDY

A key barrier that has not yet been addressed is the insufficient evidence on optimal PCV deployment strategies via mass vaccination campaigns and their expected impact in crises.[45,46] In places where they have been used, they have been administered through different strategies targeted at different age groups.[47–50] The impact of those alternative approaches has not been assessed.

The best evidence of vaccine impact comes from cluster-randomised controlled trials (cRCT). However, these are resource-intensive and exceptionally challenging to conduct during crises, with additional ethical concerns related to randomisation of vulnerable populations to potentially less protected trial arms.[51] Moreover, only a small subset of many possible combinations of potentially viable dosing strategies and age ranges can be investigated.

We propose an alternative evidence generation pathway, consisting of primary data collection in combination with mathematical modelling. Mathematical models are increasingly used to synthesize a multitude of evidence for vaccine decision making, particularly if indirect vaccine effects form a key part of the desired impact.[22,52,53] If adequately parameterised, these models are useful to simulate the pneumococcal epidemiology of a specific setting and predict PCV impact under various vaccination strategies.[22,54]

A limitation to the use of modelling to inform PCV use in crises-affected populations is the lack of context-specific data for model parameterisation. The key drivers of pneumococcal transmission are social contact behaviour (a proxy for disease transmission routes) and the pre-PCV prevalence of nasopharyngeal carriage that helps identifying pockets of the population driving pneumococcal transmission. Studies have measured both in a multitude of settings,[55,56] but few have been done in LMICs and evidence from crisis-

affected populations is entirely absent. The main drivers of transmission are often children, due to the nature and frequency of their contacts in combination with high prevalence of pneumococcal carriage.[57,58] However, in displaced populations, both social contact patterns and pneumococcal carriage may be considerably altered from their pre-crisis baseline (see Table 1), so primary data is needed to construct meaningful models for hypothesis generation.

Therefore, we aim to collect this data using a cross-sectional study in a crisis-setting, specifically Digale Internally Displaced People (IDP) camp near Hargeysa, the capital of the autonomous region of Somaliland in Somalia. Digale IDP camp is a camp with an estimated 800-1000 households. The population consists of IDPs from within Somaliland and refugees from other places in Somalia. It has an outpatient health facility on site. Save the Children International supports nutrition and child protection services.

Digale is one of a total of 9 camps in or near Hargeysa where Save the Children International provide some of the services. The estimated combined population size across all camps is ~ 60 000 IDPs, of which Digale makes up 10 – 15%. The camp is located 30mins south of Hargeysa, and therefore relatively well accessible by NGOs and the host population.

Data collection will entail measurements of social contacts through specifically designed surveys, and measurement of nasopharyngeal carriage will take place through microbiological analysis of nasopharyngeal swabs. In addition, we propose to take anthropometric measures of children aged 6-59month old in order to assess the association between malnutrition and pneumococcal carriage, and to conduct a demographic household survey to get a broader understanding of the characteristics of this population.

This cross-sectional study is part of a larger study. In this larger study, we will use the results generated in this cross-sectional study to parametrize mathematical models. These models will be used to simulate and identify optimal pneumococcal vaccination strategies for this population, and can also be used to identify vaccination strategies in a range of different crisis-affected scenarios. Specifically, we aim to identify pneumococcal vaccination strategies that can effectively be employed in a crisis situation as described by the characteristics of Digale IDP camp.

2. STUDY OBJECTIVES

Primary objectives

A. To collect robust estimates of the following data in Digale IDP camp:
    1) age specific point-prevalence of pneumococcal carriage in the following age strata: <1y; 1y; 2-5y; 6-14y; 15-29y; 30-49y; 49+y;
    2) average number of age specific daily contacts in the following age strata: <1y; 1y; 2-5y; 6-14y; 15-29y; 30-49y; 49+y
    3) prevalence of malnutrition in children aged 6-59 months old.
B. Process the data collected in objectives A1-3 to enable parametrization of mathematical models.
C. Use mathematical models to simulate and predict the impact of alternative vaccination strategies in this and similar populations.

Secondary objectives

The following secondary objectives will lead to more robust estimates in objective C.

D. To assess the distribution of pneumococcal serotypes in the following age strata: <1y; 1y; 2-5y; 6-14y; 15-29y; 30-49y; 49+y;
E. To assess the association between malnutrition and pneumococcal carriage.
F. To assess the population death rate and proportional mortality over a retrospective period.

3. STUDY DESIGN

We will conduct a cross-sectional population-based study. Data will be collected over a 4 week period. A retrospective period of 3 years will be used to collect data for objective E.

We aim to recruit 700 participants, 100 each in the following 7 age groups:

- Children under 1 year of age;
- Children 1 year of age;
- Children 2 to 5 years of age;
- Children 6 to 14 years of age;
- Children and adults 15 to 29 years of age;
- Adults 30 to 49 years of age;
- Adults over 49 years of age.

Assuming that the population distribution of Somalia in 2019 is a good proxy for the population distribution in Digale IDP camp at the time of the study, this implies that children under 2 years of age and people over 49 years of age will be oversampled (Figure 1). A large sample is needed in order to generate robust estimates of pneumococcal carriage in these age groups, as i. they usually experience the highest burden of pneumococcal disease, and ii. young children usually the main drivers of pneumococcal transmission.

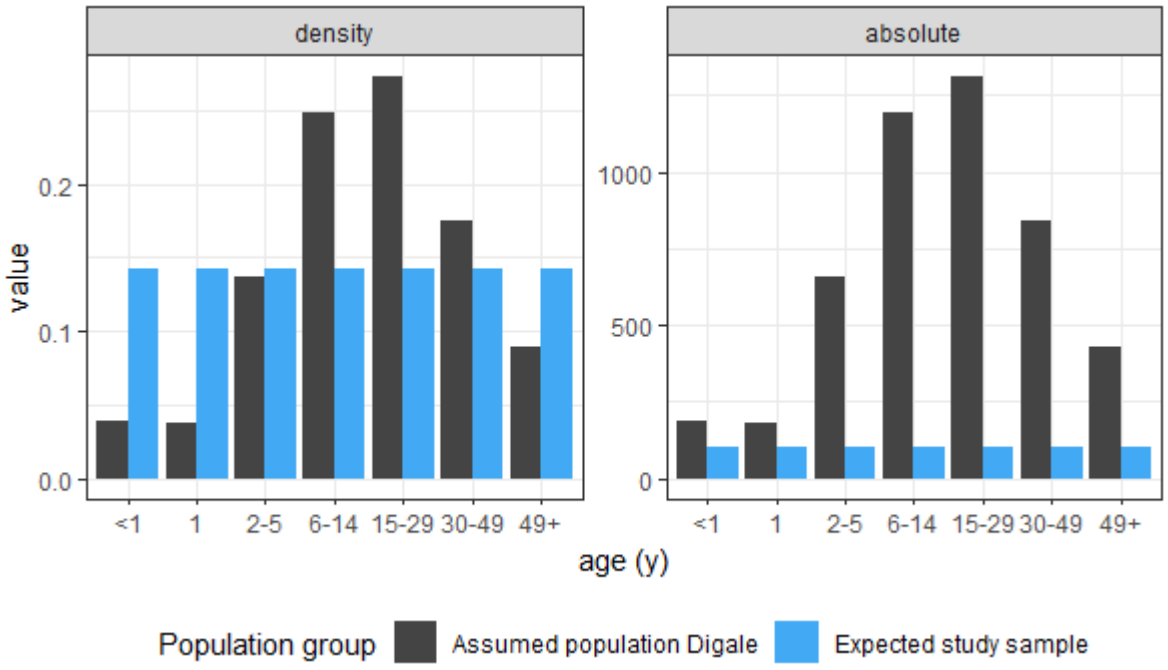


Figure 1. Comparison between the age-specific distribution of population density and absolute population size in the (assumed) population in Digale and expected study sample.

The study will be conducted in Digale IDP camp, situated 6km from Hargeisa. Hargeisa is the capital of the semi-autonomous region of Somaliland in Somalia. Around 800-1000 households are resident in Digale. Our operational partner for this study will be Save the Children. Save the Children also implement Education, Health, Nutrition and WASH programs in the other centres in the region. PCVs have never been introduced in Somalia, whereas the general national routine immunization coverage is slightly above 50%.

The study will end when a sample of 700 participants is reached, or when the 4 week period has ended.

3.1 STUDY OUTCOME MEASURES

3.1.1 PRIMARY OUTCOMES

- Age specific point-prevalence of pneumococcal carriage in the nasopharynx. This will be measured as the proportion of study participants with presence of *Streptococcus pneumoniae* in the primary culture of a nasopharyngeal swab.

- Frequency and duration and type of age-specific person to person contacts within a single day (24 hour period).
- Point-prevalence of (acute) malnutrition in children aged 6 to 59 months old.

3.1.2 SECONDARY OUTCOMES

- Distribution of pneumococcal serotypes. This will be measured as the proportion of individuals colonized with one or more specific serotypes.
- Average number of age specific daily social contacts.

3.2 RISKS AND BENEFITS

There are no immediate individual risks or benefits associated with any of the measurements taken. Children may benefit from an updated assessment of malnutrition, as children identified as severely malnourished will be referred to the local health facility. Individuals may experience discomfort from the nasopharyngeal swab.

4. SELECTION AND WITHDRAWAL OF PARTICIPANTS

No specific intake tests need to be done before a participant can enter the study. All individuals living in Digale IDP camp are eligible to participate in this study. Individuals will only be excluded if they have head or facial injuries that would contraindicate swabbing of the nasopharynx. Any other injury or morbidity will not exclude from participation. Participants will not be included if they or their representative either refuses or is not able to provide informed consent to participate.

Participants can withdraw their consent at all times. Any data that has already been collected will then be destroyed.

5. ENROLMENT PROCEDURE

5.1 REGISTRATION PRACTICALITIES

We will employ quota sampling to select 100 individuals in each of the 7 age strata (<1y; 1y; 2-5y; 6-14y; 15-29y; 30-49y; 49+y).

We aim to visit all households in Digale. With 15 - 20% nonresponse, it should be possible to include, on average, at least one participant from 700 households in total. A household is defined as: “a person or a group of persons, related or unrelated, who live together and share a common source of food”.

Because social contacts will be correlated within a household (i.e. the contact between a child and its mother is the same as the contact between the mother and her child), we initially aim to select only a single participant in each household. However, children are expected to be clustered within households (eg some households have many children, some have none or few). Therefore, it may occur that more than one individual is selected in each household, if this would fulfil the set quota. This will be done by randomly revisiting households to fill up any remaining quota once all have been visited.

First, the purpose of the study will be explained to the household head. An information sheet will help explanation of the study. The team will strongly emphasize the social benefit and confidentiality of the results. Consent will be asked for all elements of the study: the household survey, social contact survey, anthropometric survey (if applicable given the age of the participant), and swabbing procedure. If the household head refuses to participate, teams should not persist on participation, but make a note and move on to the next household.

All individuals who are members of the selected household will be eligible for study participation, except if they have head or facial injuries that would contraindicate swabbing of the nasopharynx. Participants will not be excluded because of any other illness, frailty, or condition, in order to reach a representative sample of the population. The age-group for the member of the household chosen to participate in the study will be selected through quota sampling. If multiple participants of the required age-group are present in the household, the participant will be randomly chosen amongst all those eligible. If the selected participant is not available when the study team is there, the team shall revisit the household later.

Once the participant is selected from the household, the nurse/nursing assistant will carefully introduce the study to the participant and to his/her caregiver if the participant is underage (<18y). Because swabbing of the nasopharynx is a harmless, but invasive and uncomfortable procedure, the field team will make sure that this is clear to the participant. This will be done by showing them a swab and explaining the procedure, assisted by a visual image on the information sheet given to the participant.

6. MEASUREMENTS

All electronic data will be collected using the Open Data Kit (ODK) software on electronic tablets. This software can be programmed to automatically skip irrelevant questions and perform plausibility checks, thereby increasing the internal validity of the study. Tablets will not require any internet access in the

field, but will need to be charged overnight. A limited number of paper-based alternatives will be available in case of technological failure.

Questionnaires will be developed in English, then translated to Somali and back-translated to English to ensure internal validity.

6.1 HOUSEHOLD SURVEY

We will conduct a household survey according to SMART (Standardized Monitoring and Assessment for Relief and Transitions)[59] guidelines to retrospectively assess mortality and demography in the study population. This will provide estimates for birth rates, mortality rates, and migration rates, and help to create a realistic representation of population dynamics in our model. Moreover, information on causes of death will provide a better understanding about proportional mortality in the population.

A recall period of 3 years from the interview date, starting at a memorable date, will be used.

The survey will specifically ask: household size, housing facilities, the ages and number of people who entered the household (in migrants and live births), and ages and number of people who left the household (out migrants and deaths) since the start of the recall period. For those that died, verbal autopsy according to WHO questionnaires will be performed at a separate time to establish the probable cause of death. The analysis of verbal autopsy questionnaires is now automated through InterVal software. We will also collect data on exposure to indoor air pollution, as this is a known confounder for pneumococcal carriage.

6.2 INDIVIDUAL SURVEY

6.2.1 SOCIAL CONTACT DATA

To allow for comparability with other social contact surveys, we will conduct a face-to-face survey. The survey is based on a similar survey conducted by Le Polain et al in rural Uganda, which should increase the external validity.[56] We will obtain data on the frequency, duration, and intimacy of contacts within and outside of household settings, for a single day.

Such questionnaires have been shown to perform poorly if participants are not notified in advance to remember their contacts over a 24 hour period. Therefore, we will ask participants to remember their contacts in the study period in order to minimize recall bias at recruitment. This study period will be defined as: the time between waking up on the day before data-collection and waking up on the day of data-collection. Social contacts will be measured on every day of the week, in order to get an estimate of the average number of social contacts that is representative for the entire week. A caregiver will answer the questionnaire as a proxy for young children.

Social contact surveys require a good approximation of age. In some rural areas, age is not always known. If the age is unknown, it will be approximated as follows: For children, if the year and month of birth is not known by the parent, it will be registered from the child's immunisation card. Otherwise, we will ask whether a child is older or younger than a neighbour's child with a known age. Alternatively, we will use a local-event calendar listing important events in the past five years can be reported, where the mother can be asked in what month the child was born. A local-event calendar will be created together with the field-team in Digale.

The social contact survey will require a reasonable estimate of the age of the respondent and the contact. If an age is not known, and cannot be retrieved via other measures, one can make a best guess. The method by which age was approximated will be noted.

Participants will be asked to list all places he or she has visited during the study period (e.g. their household, neighbour, market, mosque, clinic, etc). For each place visited, we will then ask about all persons encountered in this place. For each of the encounters listed, the participants will be asked to recall:

1) Whether the encounter was a contact. A contact will be defined as: A person with whom you spent 5 minutes or more, and with whom you exchanged at least three words.

2) The approximate duration of contacts (<5 minutes; 5-14 minutes; 15-59 minutes; 1-4 hours; >4 hours). If the same person is encountered multiple times, the sum of the duration of these encounters will be recorded.
3) Whether physical contact has occurred. A physical contact will be defined as a contact with who you had physical contact. This includes hand shaking, sharing a bike, kissing, embracing, but also sharing a glass or other utensils that are passed directly from mouth to mouth.
4) Relationship to the contact (i.e. household member, relative, friend, etc).
5) Age of the contact (exact or approximated).

We will also ask to give a rough approximation of the amount of people that were met, but that cannot be defined as a contact.

6.2.2 NASOPHARYNGEAL SWAB

Nasopharyngeal (NP) swabs will be taken to estimate the prevalence of pneumococcus in the study population. Swabbing will be done according to WHO guidelines[60]. Sampling from the nasopharynx is recommended over the oropharynx in infants and children. Swabs should be made from nylon or Dacron materials, as these are preferred to maintain sensitivity in further molecular analyses.

The standard method will be used to administer nasopharyngeal swabs. Data-collectors, nurses or nursing assistants, will be trained in the correct and standardized way of collecting swabs:

"Hold the infant or young child's head securely. Tip their head backwards slightly and pass the swab directly backwards, parallel to the base of the NP passage. The swab should move without resistance until reaching the nasopharynx, located about one-half to two-thirds the distance from the nostril to ear lobe.
If resistance occurs, remove the swab and attempt again to take the sample entering through the same or the other nostril. Failure to obtain a satisfactory specimen is often due to the swab not being fully passed into the nasopharynx. Once the swab is in location, rotate the swab 180◦, or leave in place for 5 s to saturate the swab tip; remove the swab slowly. All swabs should be processed; however, to assist with interpreting the results, investigators should record whether the procedure was acceptable or suboptimal. Recording if secretions are present on the swab and whether the swab was potentially contaminated (e.g. touched by the investigator or dropped on the ground) may also be helpful in interpretation."

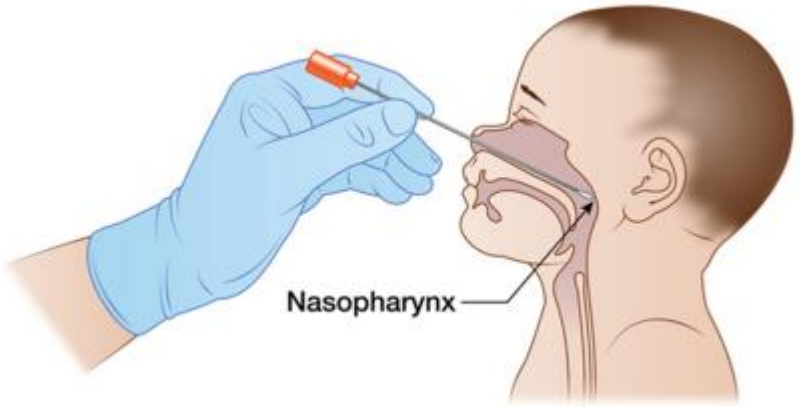


Figure 2. Recommended method to take nasopharyngeal swabs.

NP swabs are to be aseptically placed in the storage medium immediately after collection. It is inserted in the bottom of the STGG medium, raised slightly, after which the shaft is to be cut off using sterile scissors and the lid closed. The tube is then placed in a cool box for storage during the day. At the end of each day, swabs will be transferred to low temperatures.

Ideally, swabs are to be placed in ultra low temperatures (ULT; < -80C). However, this requires either an adequate freezer or an appropriate alternative such as dry ice. Our partners at the Murdoch Children's Research Institute (MCRI) performed additional analyses showing that samples remain viable when stored in in -20C. We will use dry ice as provided by the Somaliland MoH to store our sampels. However, given possible logistic difficulties of procuring dry ice in Hargeisa, a domestic freezer (-20C) will be available as an alternative to store samples during the study period.

No further processing, such as aliquotting, needs to be done on the field site, as this will all be done at the MCRI in Melbourne, Australia. Therefore, no local laboratory facilities are needed for this study.

The swabs will be stored and transported in skim milk tryptoneglucose-glycerol (STGG), which is the recommended medium to do so. STGG needs to remain cold (at +4C) between preparation and data-collection, so domestic refrigerators are required locally. The medium can be at room-temperature at the time of data collection, so does not need to stay cold in the field. Swabs will be placed on wet ice directly after collection in the field, and frozen within 24hrs.

Upon completion of the study, samples will be shipped to MCRI in Melbourne, Australia in cryoboxes under negative col chain as UN3373 (Category B – biological substance).

After transport to MCRI, nasopharyngeal swabs will be tested for pneumococcal serotypes using Microarray with an additional culture amplification step. This had shown to have 100% sensitivity to detect the major pneumococcal serotype present in the swab (95 – 100), and 99% sensitivity to detect minor pneumococcal serotypes (95 – 100). It has a positive predictive value (PPV) of 100% (98 – 100).[61]

Information on the history of oral antibiotic treatment and respiratory symptoms in the two weeks preceding the study will be collected in the individual survey, as these may affect carriage rates.

6.2.3 ANTHROPOMETRY

Anthropometric measurements will be collected according to SMART guidelines[59] from all 6 – 59 month old children (300 in our proposed sample). The resulting estimates of malnutrition (prevalence of stunting, wasting, and oedema) will be scientifically beneficial as they can be used to assess the association between malnutrition and carriage of pneumococcal serotypes, and to assess how malnutrition may affect a vaccination strategy. They will also benefit the operational partner (Save the Children), as they provide an updated estimate of the level of malnutrition in the population.

As part of the survey, weight, height, age, middle-upper-arm circumference (MUAC) and presence of oedema will be recorded. Weight will be measured using an electronic scale, whilst height (or length for children under 2 years of age) will be measured using a measuring board. MUAC strips will be used to estimate MUAC, whilst a trained worker will assess the presence of oedema. If a child is physically disabled, this will be reported (i.e. missing limbs may bias weight estimates).

7. STATISTICS AND DATA ANALYSIS

7.1 SAMPLE SIZE CALCULATIONS

We will sample 700 participants, 100 each in the following 7 age groups:

- Children under 1 year of age;
- Children 1 year of age;
- Children 2 to 5 years of age;
- Children 6 to 14 years of age;
- Children and adults 15 to 29 years of age;
- Adults 30 to 49 years of age;
- Adults over 49 years of age.
-

A sample of 100 persons per age group will allow to estimate the point-prevalence of pneumococcal carriage within a single stratum (age group) with a precision of 2-10%, depending on the actual pneumococcal prevalence and using a confidence level of 95%. In a study conducted in Gambia before introduction of PCVs in the routine immunization programme, the prevalence of pneumococcal carriage was found to be 90% in infants. If similar rates would be present in Digale, we would be able detect this with a precision of 6%.

The precision around the average number of daily contacts in any age group will depend on the actual average number of daily contacts and degree of overdispersion in the distribution of the number of contacts. In Uganda, the mean number of contacts across all age groups was found to be 7, with an

overdispersion parameter of 10.6. If similar numbers would be present in Digale, we would be able to detect this with a precision of 0.7 contacts.

Ultimately, the study needs not to be powered to detect a difference in carriage or contacts between age-groups. Rather, we need sufficient precision in order to accurately parametrize mathematical models for pneumococcal transmission. We constructed a mathematical model, simulating a sample size of 100 individuals in each age group, the same carriage rates as observed in Gambia[62], and social contact rates as detect by le Polain et al in Uganda[56]. This allowed us to adequately parametrize transmission rates in a simple Susceptible – Infected – Susceptible (SIS) model, shown below (Figure 3).

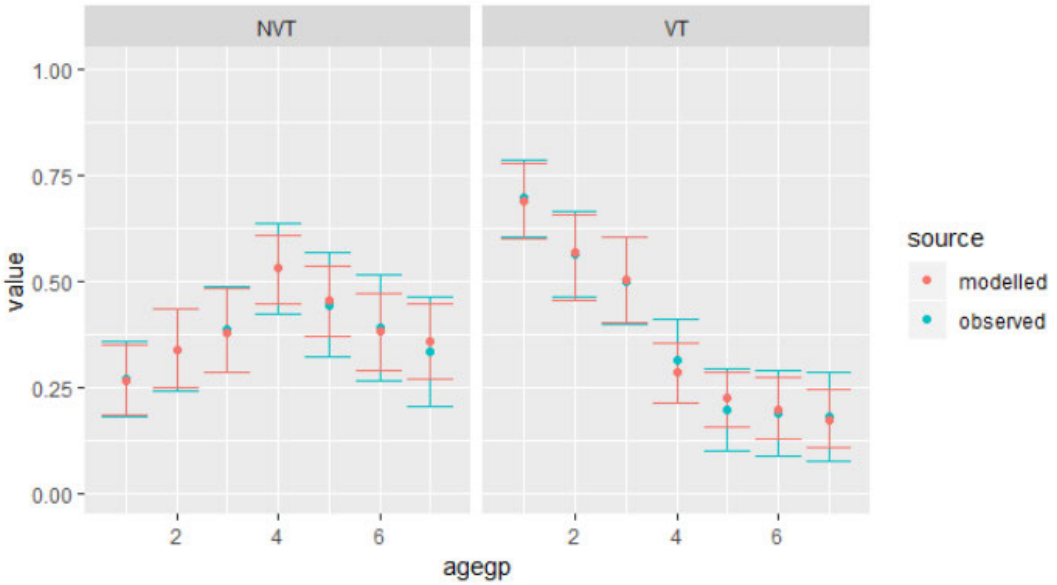


Figure 3. Age stratified compartmental SIS model, using contact patterns as observed in Uganda with s sample size of 100 individuals in each age-group. Model is fit to VT and NVT prevalence data as observed in Gambia, assuming that 100 individuals would have been sampled in each age group. Blue points show observed estimates with corresponding 95% confidence intervals if sample would have been 100 individuals in each age group. Red points show results of the fitted model with corresponding 95% credible intervals.

7.2 STATISTICAL ANALYSIS

All data analyses will be conducted in R software.

1) Nasopharyngeal carriage

   Prevalence of nasopharyngeal *Streptococcus pneumoniae* carriage will be presented as proportions and 95% confidence intervals. We will present prevalence overall, and for each age group specific.

   We will also present the distribution of circulating serotypes in each age group. Serotypes will be classified as vaccine- or non-vaccine-type, where vaccine-type (VT) are the 10 or 13 serotypes to which PCV10 or PCV13 offer protection.

   We will use binomial regression techniques to assess the association between pneumococcal carriage and multiple covariates, such as age, sex, and malnutrition.

2) Social contact matrix

   We will first undertake a descriptive analysis of the data, summarizing the number of physical and nonphysical contacts by age, sex, area of residence, and others.

   We will use count models such as zero-inflated negative binomial regression techniques to assess the age-specific number of daily contacts, and control for confounding factors such as sex and household size. We will perform a stratified analysis by contact type (all contacts, only physical contacts, only nonphysical contacts) and setting (e.g. household, market, place of worship).

   Established methods using negative binomial regression models[55,63] will be used to construct social contact matrices, which can be used to parametrize mathematical models. Social contact

matrices will be based on the average number of age-specific contacts, type of contact, and duration of contact.

3) Anthropometric measures

Standardized methods will be used to analyse anthropometric measures[59]. We will calculate weight-for height, height for age, and weight for age.

Classification of acute malnutrition:
Severe acute malnutrition as all children with a weight for height (WFH) z-score <-3, and all children with a WFH z-score >-3 and with oedema.
Moderate acute malnutrition as all children with a WFH z-score between -2 and -3.
Global acute malnutrition as all children with a WFH z-score < -2.

The relationship between malnutrition and nasopharyngeal carriage of *Streptococcus pneumoniae* will be assessed using binomial regression models.

4) Modelling

Data on nasopharyngeal carriage and social contact patterns will feed into mathematical models of pneumococcal transmission. We will use a realistic age-structured deterministic compartmental Susceptible – Infected – Susceptible (SIS) model to model pneumococcal transmission. This model structure is commonly used to model pneumococcal transmission rates[23,54,64].

The model will be stratified by age, colonization status, and if required, degree of malnutrition. Pneumococcal serotypes will be classified by the protection offered by the PCVs (i.e. serotypes to which PCVs protect are classified as vaccine-type, others as non-vaccine type). The model will allow for colonization with only vaccine-types (VT), only non-vaccine types (NVT), or both, allowing for co-colonization of VTs and NVTs.

The model will be fitted to nasopharyngeal carriage estimates using a Markov Chain Monte Carlo (MCMC) algorithm.

We will explore the short and long term effects of vaccine impact by simulating vaccination strategies using different assumptions. This can inform on the age groups to target, number of doses to administer to each age group, vaccination coverage in each age group, and frequency at which strategies should be employed.

8. REGULATORY ISSUES

8.1 ETHICS APPROVAL

We will apply for ethics approval at LSHTM Research Ethics Committee, as well as at the Research Department of the Directorate of Planning, Policy & Strategic Information at the Ministry of Health Development of the Republic of Somaliland.

8.2 COMMUNITY INVOLVEMENT

Local staff and partners will be crucial in helping to carry out the study. We will meet community representatives in advance of the study, and ensure that the survey will respect community beliefs and social values. We will design appropriate community sensitisation with input from local staff and community representatives.

8.3 CONSENT

Informed consent will be obtained prior to the participant undergoing procedures that are specifically for the purposes of the study. We will first explain the study to each household head, and ask for written consent for the household survey and individual survey. We will then ask written consent from each participant for the individual survey as well. Consent will be sought from each participant only after a full explanation has been given, an information leaflet offered and time allowed for consideration.

An information sheet is attached as an appendix. It will be read and explained to each eligible adult and to the representative of each child (less than 18 years of age), as well as to child themselves. If the study participant of representative is illiterate, consent will be given through a mark or fingerprint. Each consent from an illiterate person will also be signed by an adult legal witness. Children aged 8-18y will be asked to provide assent through a signature, mark, or fingerprint.

Consent forms and information sheets will be translated to Somali and back translated to English.

The right of the participant to refuse to participate without giving reasons will be respected at all times. All participants are free to withdraw at any time from the protocol treatment without giving reasons and without prejudicing further treatment.

The PI is responsible for ensuring that all vulnerable participants are protected and participate voluntarily in an environment free from coercion or undue influence.

8.4 CONFIDENTIALITY

All procedures will be conducted in a location that ensures privacy.

Any participants' identifiable data collected by the Study Coordination Centre will be stored securely and their confidentiality protected in accordance with the Data Protection Act 1998. All data will be stored encrypted on passport protected electronic tablets. This data will be synchronized with a password protected database on a password protected encrypted server. Data on electronic tablets will be deleted upon completion of the survey. Data will be stored on the encrypted server for a minimum of 10 years. Data will not be stored with person identifiers.

Informed consent forms will be kept separated from other study documents.

9. DISSEMINATION OF FINDINGS

Findings will be published in scientific journals and presented at scientific conferences. Co-authorship will be offered based on International Committee of Medical Journal Editors criteria. Other individuals who made important contributions will be acknowledged.

Findings will also be disseminated to the community in Digale, internally within Save the Children and Médecins Sans Frontières, and at regional relief hubs.

10. REFERENCES

1 United Nations High Commissioner for Refugees. UNHCR Global Trends 2017. 2018. https://www.unhcr.org/statistics/unhcrstats/5b27be547/unhcr-global-trends-2017.html (accessed Nov 22, 2018).

2 Sundberg R, Melander E. Introducing the UCDP Georeferenced Event Dataset. *J Peace Res* 2013; **50**: 523–32.

3 Croicu M, Sundberg R. UCDP GED Codebook version 18.1. 2017.

4 Université catholique de Louvain CRED. EM-DAT: The Emergency Events Database. 2019. http://www.emdat.be.

5 Toole M, Waldman R. The public health aspects of complex emergencies and refugee situations. *Annu Rev Public Health* 1997; **18**: 283–312.

6 Heudtlass P, Speybroeck N, Guha-Sapir D. Excess mortality in refugees, internally displaced persons and resident populations in complex humanitarian emergencies (1998–2012) – insights from operational data. *Confl Health* 2016; **10**: 15.

7 Thomas SL, Thomas SD. Displacement and health. *Br Med Bull* 2004; **69**: 115–27.

8 Checchi F, Warsame A, Treacy-Wong V, Polonsky J, Ommeren M van, Prudhon C. Public health information in crisis-affected populations: a review of methods and their use for advocacy and action. *The Lancet* 2017; **390**: 2297–313.

9 Checchi F, Gayer M, Grais RF, Mills EJ. Public health in crisis-affected populations: a practical guide for decision-makers. London: Humanitarian Practice Network, Overseas Development Institute, 2007.

10 Klugman KP, Dagan R, Malley R, Whitney CG. Pneumococcal conjugate vaccine and pneumococcal common protein vaccines. In: Plotkin's Vaccines, 7th edn. Philadelphia: Elsevier, 2018.

11 Coldiron ME, Touré O, Frank T, Bouygues N, Grais RF. Outbreak of Pneumococcal Meningitis, Paoua Subprefecture, Central African Republic, 2016–2017. *Emerg Infect Dis* 2018; **24**: 1720–2.

12 Crellen T, Rao VB, Piening T, Zeydner J, Siddiqui MR. Seasonal upsurge of pneumococcal meningitis in the Central African Republic. *Wellcome Open Res* 2018; **3**: 134.

13 Bellos A, Mulholland K, O'Brien KL, Qazi SA, Gayer M, Checchi F. The burden of acute respiratory infections in crisis-affected populations: a systematic review. *Confl Health* 2010; **4**: 3.

14 Troeger C, Blacker B, Khalil IA, *et al.* Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect Dis* 2018; **18**: 1191–210.

15 Fonseca Lima E, Mello MJG, Albuquerque MFPM, *et al.* Risk factors for community-acquired pneumonia in children under five years of age in the post-pneumococcal conjugate vaccine era in Brazil: a case control study. *BMC Pediatr* 2016; **16**. DOI:10.1186/s12887-016-0695-6.

16 Selwyn BJ. The epidemiology of acute respiratory tract infection in young children: comparison of findings from several developing countries. Coordinated Data Group of BOSTID Researchers. *Rev Infect Dis* 1990; **12 Suppl 8**: S870-888.

17 Small C, Shaler CR, McCormick S, *et al.* Influenza infection leads to increased susceptibility to subsequent bacterial superinfection by impairing NK cell responses in the lung. *J Immunol Baltim Md 1950* 2010; **184**: 2048–56.

18 Zhou H, Haber M, Ray S, Farley MM, Panozzo CA, Klugman KP. Invasive Pneumococcal Pneumonia and Respiratory Virus Co-infections. *Emerg Infect Dis* 2012; **18**: 294–7.

19 Akramuzzaman SM, Cutts FT, Wheeler JG, Hossain MJ. Increased Childhood Morbidity After Measles Is Short-term in Urban Bangladesh. *Am J Epidemiol* 2000; **151**: 723–35.

20 Dueger EL, Asturias EJ, Matheu J, Gordillo R, Torres O, Halsey N. Increasing penicillin and trimethoprim-sulfamethoxazole resistance in nasopharyngeal *Streptococcus pneumoniae* isolates from Guatemalan children, 2001-2006. *Int J Infect Dis IJID Off Publ Int Soc Infect Dis* 2008; **12**: 289–97.

21 Petrosillo N, Pantosti A, Bordi E, *et al.* Prevalence, Determinants, and Molecular Epidemiology of *Streptococcus pneumoniae* Isolates Colonizing the Nasopharynx of Healthy Children in Rome. *Eur J Clin Microbiol Infect Dis* 2002; **21**: 181–8.

22 Ojal J, Griffiths U, Hammitt LL, *et al.* The merits of sustaining pneumococcal vaccination after transitioning from Gavi support - a modelling and cost-effectiveness study for Kenya. *bioRxiv* 2018; : 369603.

23 Choi YH, Jit M, Flasche S, Gay N, Miller E. Mathematical Modelling Long-Term Effects of Replacing Prevnar7 with Prevnar13 on Invasive Pneumococcal Diseases in England and Wales. *PLOS ONE* 2012; **7**: e39927.

24 Le Polain De Waroux O, Flasche S, Prieto-Merino D, Goldblatt D, Edmunds WJ. The Efficacy and Duration of Protection of Pneumococcal Conjugate Vaccines Against Nasopharyngeal Carriage: A Meta-regression Model. *Pediatr Infect Dis J* 2015; **34**: 858–64.

25 Singer M, Nambiar S, Valappil T, Higgins K, Gitterman S. Historical and Regulatory Perspectives on the Treatment Effect of Antibacterial Drugs for Community-Acquired Pneumonia. *Clin Infect Dis* 2008; **47**: S216–24.

26 File TM, Schentag JJ. What Can We Learn from the Time Course of Untreated and Partially Treated Community-Onset *Streptococcus pneumoniae* Pneumonia? A Clinical Perspective on Superiority and Noninferiority Trial Designs for Mild Community-Acquired Pneumonia. *Clin Infect Dis* 2008; **47**: S157–65.

27 Muthumbi E, Lowe BS, Muyodi C, Getambu E, Gleeson F, Scott JAG. Risk factors for community-acquired pneumonia among adults in Kenya: a case-control study. *Pneumonia Nathan Qld* 2017; **9**: 17.

28 Dherani M, Pope D, Mascarenhas M, Smith KR, Weber M, Bruce N. Indoor air pollution from unprocessed solid fuel use and pneumonia risk in children aged under five years: a systematic review and meta-analysis. *Bull World Health Organ* 2008; **86**: 390-398C.

29 Reisman J, Rudolph K, Bruden D, Hurlburt D, Bruce MG, Hennessy T. Risk Factors for Pneumococcal Colonization of the Nasopharynx in Alaska Native Adults and Children. *J Pediatr Infect Dis Soc* 2014; **3**: 104–11.

30 Collins DA, Hoskins A, Snelling T, *et al.* Predictors of pneumococcal carriage and the effect of the 13-valent pneumococcal conjugate vaccination in the Western Australian Aboriginal population. *Pneumonia* 2017; **9**. DOI:10.1186/s41479-017-0038-x.

31 Bellos A. Vaccination in Humanitarian Emergencies: Literature review and case studies. 2012. http://www.who.int/immunization/sage/meetings/2012/april/2_SAGE_WGVHE_SG1__Lit_Review_CaseStudies.pdf (accessed Nov 21, 2018).

32 Grabenstein JD, Musher DM. Pneumococcal Polysaccharide Vaccines. In: Plotkin's Vaccines, 7th edn. Philadelphia: Elsevier, 2018.

33 Feikin DR, Kagucia EW, Loo JD, *et al.* Serotype-specific changes in invasive pneumococcal disease after pneumococcal conjugate vaccine introduction: a pooled analysis of multiple surveillance sites. *PLoS Med* 2013; **10**: e1001517.

34 World Health Organization. Official country reported coverage estimates time series. 2018. http://apps.who.int/immunization_monitoring/globalsummary/timeseries/tscoveragebcg.html (accessed Feb 27, 2019).

35 Klugman KP. Herd protection induced by pneumococcal conjugate vaccine. *Lancet Glob Health* 2014; **2**: e365–6.

36 Loughlin AM, Hsu K, Silverio AL, Marchant CD, Pelton SI. Direct and indirect effects of PCV13 on nasopharyngeal carriage of PCV13 unique pneumococcal serotypes in Massachusetts' children. *Pediatr Infect Dis J* 2014; **33**: 504–10.

37 Hammitt LL, Akech DO, Morpeth SC, *et al.* Population effect of 10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of *Streptococcus pneumoniae* and non-typeable *Haemophilus influenzae* in Kilifi, Kenya: findings from cross-sectional carriage studies. *Lancet Glob Health* 2014; **2**: e397–405.

38 Simonsen L, Taylor RJ, Young-Xu Y, Haber M, May L, Klugman KP. Impact of Pneumococcal Conjugate Vaccination of Infants on Pneumonia and Influenza Hospitalization and Mortality in All Age Groups in the United States. *mBio* 2011; **2**: e00309-10.

39 World Health Organization. Vaccination in Acute Humanitarian Emergencies. A framework for decision making. 2017.

40 Gargano LM, Hajjeh R, Cookson ST. Pneumonia Prevention during a Humanitarian Emergency: Cost-effectiveness of *Haemophilus Influenzae* Type B Conjugate Vaccine and Pneumococcal Conjugate Vaccine in Somalia. *Prehospital Disaster Med* 2015; **30**: 402–11.

41 Gargano LM, Hajjeh R, Cookson ST. Pneumonia prevention: Cost-effectiveness analyses of two vaccines among refugee children aged under two years, *Haemophilus influenzae* type b-containing and pneumococcal conjugate vaccines, during a humanitarian emergency, Yida camp, South Sudan. *Vaccine* 2017; **35**: 435–42.

42 Unicef. Update on Humanitarian Mechanism. 2018. https://www.unicef.org/supply/files/SESSION_7_2018_VIC_Update_on_Humanitarian_Mechanism.pdf.

43 Pereira A de L, Southgate R, Ahmed H, O'Connor P, Cramond V, Lenglet A. Infectious Disease Risk and Vaccination in Northern Syria after 5 Years of Civil War: The MSF Experience. *PLOS Curr Disasters* 2018; published online Feb 2. DOI:10.1371/currents.dis.bb5f22928e631dff9a80377309381feb.

44 World Health Organization. The Humanitarian Mechanism. 2017. http://www.who.int/immunization/programmes_systems/sustainability/The_Humanitarian_Mechanism_ToRs.pdf?ua=1.

45 Blanchet K, Ramesh A, Frison S, *et al.* Evidence on public health interventions in humanitarian crises. *The Lancet* 2017; **390**: 2287–96.

46 Médecins Sans Frontières. Médecins Sans Frontières (MSF) briefing on provisional agenda item 20.1: Global vaccine action plan, Document A71/39. 2018. https://msfaccess.org/sites/default/files/2018-05/WHO-WHA-71_MSF-briefing-GVAP-Agenda-Item-20-1.pdf.

47 Ravelo JL. MSF bends donation policy for pneumonia vaccine. Devex. 2015. https://www.devex.com/news/sponsored/msf-bends-donation-policy-for-pneumonia-vaccine-85317 (accessed Feb 27, 2019).

48 Peyraud N, Quéré M, Duc G, *et al.* A post-conflict vaccination campaign, Central African Republic. *Bull World Health Organ* 2018; **96**: 540–7.

49 Médecins Sans Frontières. Nigeria: 'There were hardly any children under 5 years of age'. 2016. https://www.msf.ie/article/nigeria-there-were-hardly-any-children-under-5-years-age (accessed Feb 27, 2019).

50 Médecins Sans Frontières. Vaccination activities in Médecins Sans Frontières. 2015. https://www.who.int/immunization/sage/meetings/2015/october/MSF_SAGE_October2015_vf.pdf.

51 Leaning J. Ethics of research in refugee populations. *The Lancet* 2001; **357**: 1432–3.

52 Baguelin M, Flasche S, Camacho A, Demiris N, Miller E, Edmunds WJ. Assessing Optimal Target Populations for Influenza Vaccination Programmes: An Evidence Synthesis and Modelling Study. *PLoS Med* 2013; **10**: e1001527.

53 Flasche S, Jit M, Rodríguez-Barraquer I, *et al.* The Long-Term Safety, Public Health Impact, and Cost-Effectiveness of Routine Vaccination with a Recombinant, Live-Attenuated Dengue Vaccine (Dengvaxia): A Model Comparison Study. *PLOS Med* 2016; **13**: e1002181.

54 Flasche S, Ojal J, Le Polain de Waroux O, *et al.* Assessing the efficiency of catch-up campaigns for the introduction of pneumococcal conjugate vaccine: a modelling study based on data from PCV10 introduction in Kilifi, Kenya. *BMC Med* 2017; **15**. DOI:10.1186/s12916-017-0882-9.

55 Mossong J, Hens N, Jit M, *et al.* Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLoS Med* 2008; **5**. DOI:10.1371/journal.pmed.0050074.

56 le Polain de Waroux O, Cohuet S, Ndazima D, *et al.* Characteristics Of Human Encounters And Social Mixing Patterns Relevant To Infectious Diseases Spread By Close Contact: A Survey In Southwest Uganda. *BMC Infect Dis* 2017; **18**. DOI:10.1101/121665.

57 Althouse BM, Hammitt LL, Grant L, *et al.* Identifying transmission routes of *Streptococcus pneumoniae* and sources of acquisitions in high transmission communities. *Epidemiol Infect* 2017; **145**: 2750–8.

58 Weinberger DM, Pitzer VE, Regev-Yochay G, Givon-Lavi N, Dagan R. Association Between the Decline in Pneumococcal Disease in Unimmunized Adults and Vaccine-Derived Protection Against Colonization in Toddlers and Preschool-Aged Children. *Am J Epidemiol* 2019; **188**: 160–8.

59 SMART. Standardized Monitoring and Assessment for Relief and Transitions Manual 2.0. 2017.

60 Satzke C, Turner P, Virolainen-Julkunen A, *et al.* Standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*: Updated recommendations from the World Health Organization Pneumococcal Carriage Working Group. *Vaccine* 2013; **32**: 165–79.

61 Satzke C, Dunne EM, Porter BD, Klugman KP, Mulholland EK, PneuCarriage project group. The PneuCarriage Project: A Multi-Centre Comparative Study to Identify the Best Serotyping Methods for Examining Pneumococcal Carriage in Vaccine Evaluation Studies. *PLOS Med* 2015; **12**: e1001903.

62 Hill PC, Akisanya A, Sankareh K, *et al.* Nasopharyngeal Carriage of *Streptococcus pneumoniae* in Gambian Villagers. *Clin Infect Dis* 2006; **43**: 673–9.

63 Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am J Epidemiol* 2006; **164**: 936–44.

64 Le Polain De Waroux O, Edmunds WJ, Takahashi K, *et al.* Predicting the impact of pneumococcal conjugate vaccine programme options in Vietnam. *Hum Vaccines Immunother* 2018; **14**: 1939–47.

Appendix C

Paper-based questionnaires used in the survey

0	General information	
0.0	interviewer ID	
0.1	Date (YYYY – MM – DD)	Time (HH – MM)
	- -	-
0.2	Household ID	

1	Housing facilities	
1.0	How many separate rooms do you have?	
1.1	How many of these separate rooms are places to sleep?	
1.2	What is the main source of drinking water for members of your household? only select a SINGLE answer	
	piped water	***water from spring***
	piped into dwelling ☐	protected spring ☐
	piped into compound, yard, or plot ☐	unprotected spring ☐
	piped into neighbour ☐	rainwater collection ☐
	public tap/standpipe ☐	tanker truck ☐
	tubewell/borehole ☐	cart with small tank/drum ☐
		surface water (e.g. river, dam, lake, pond, stream) ☐
	dug well	
	protected well ☐	bottled water ☐
	unprotected well (e.g. shallow well) ☐	
		don't know ☐
		refuse to answer ☐
		other
1.3	Is there water leakage from the walls or roof when it rains?	
	yes ☐	don't know ☐
	no ☐	refuse to answer ☐
1.4	Is there a draft inside the house?	
	yes ☐	don't know ☐
	no ☐	refuse to answer ☐
1.5	What sort of fuel is used inside the house? Can select MULTIPLE answers	

	firewood	☐	diesel	☐
	charcoal	☐	electricity	☐
	LPG	☐		
	processed solid fuel	☐	don't know	☐
	kerosene ethanol	☐	refuse to answer	☐
	torches	☐		
	solar	☐	other	
1.6	Is ventilation present in the cooking room?			
	outside	☐	don't know	☐
	yes	☐	refuse to answer	☐
	no	☐		
1.7	When did your family settle in Digale?			
	<1 year ago	☐	don't know	☐
	1-2 years ago	☐	refuse to answer	☐
	2-3 years ago	☐		
	>3 years ago	☐		

2 Demography and mortality

2.0 Including yourself, could you list all individuals who are currently living in this household?

	Name or initials	Sex		Age (y)	Born during recall period?		Joined during recall period?	
		Male	Female		Yes	No	Yes	No
1		☐	☐		☐	☐	☐	☐
2		☐	☐		☐	☐	☐	☐
3		☐	☐		☐	☐	☐	☐
4		☐	☐		☐	☐	☐	☐
5		☐	☐		☐	☐	☐	☐
6		☐	☐		☐	☐	☐	☐
7		☐	☐		☐	☐	☐	☐
8		☐	☐		☐	☐	☐	☐
9		☐	☐		☐	☐	☐	☐
10		☐	☐		☐	☐	☐	☐

2.1 If there are any, can you list any household members who have left to live somewhere else during the recall period?

	Name or initials	Sex		Age (y)	Born during recall period?		Joined during recall period?	
		Male	Female		Yes	No	Yes	No
1		☐	☐		☐	☐	☐	☐
2		☐	☐		☐	☐	☐	☐
3		☐	☐		☐	☐	☐	☐
4		☐	☐		☐	☐	☐	☐
5		☐	☐		☐	☐	☐	☐

2.2 If there are any, can you list any household members who have died during the recall period?

	Name or initials	Sex		Age (y)	Born during recall period?		Joined during recall period?		Location of death	Cause of death
		M	F		Y	N	Y	N		
1		☐	☐		☐	☐	☐	☐		
2		☐	☐		☐	☐	☐	☐		
3		☐	☐		☐	☐	☐	☐		

	4		☐	☐			☐	☐	☐	☐			
	5		☐	☐			☐	☐	☐	☐			

2.3	Does anyone who currently lives in the household smoke, uses khat, or uses snuff?		
	smoke	☐	don't know ☐
	khat	☐	refuse to answer ☐
	snuff	☐	
			other
	none	☐	

0	General information	
0.0	interviewer ID	
0.1	Date (YYYY – MM – DD)	Time (HH – MM)
	- -	-
0.2	Household ID	
0.3	Participant id	
0.4	Answers given by	
	self ☐ parent (proxy) ☐	other

1	Sociodemographic characteristics		
1.0	**Is the participant a young child (<5yo)?**		
	yes	☐	go to 1.0.2
	no	☐	go to 1.0.1
1.0.1	**What is your age in years?**		
	☐☐		go to 1.1
1.0.2	**Can you tell me your date of birth?**		
	yes	☐	
	no	☐	go to 1.0.2.2
1.0.2.1	**Birth date (DD – MM – YYYY)**		
	☐☐ - ☐☐ - ☐☐☐☐		go to 1.1
1.0.2.2	**What is your age in months**		
	☐☐		
1.0.3	**Is exact age known?**		
	yes	☐	go to 1.1
	no	☐	
1.0.4	**How was age estimated?**		
	local event calendar	☐	
	proxy age other person	☐	
	guess	☐	
1.1	**What is your sex?**		
	male	☐	
	female	☐	
1.2	**What is your primary occupation or daily activity?**		
	pre-school child	☐	
	school/college/university student	☐	
	office worker	☐	
	shop worker	☐	
	manual worker	☐	
	agriculture	☐	
	market salesman	☐	
	household	☐	
	unemployed	☐	
	retired	☐	
	don't know	☐	
	refuse to answer	☐	
		☐	
	other		

2	Animal contact			
2.0	**Do you live close to any of the following animals?**			
	Chickens, ducks, gees	☐	none	☐
	Cows, goats, sheep	☐		
	Rodents	☐	don't know	☐
	Primates	☐	refuse to answer	☐
	Bats	☐		
	Antelope	☐	other	
2.1	**Do you touch the following animals at least once per week?**			
	Chickens, ducks, gees	☐	none	☐
	Cows, goats, sheep	☐		
	Rodents	☐	don't know	☐
	Primates	☐	refuse to answer	☐
	Bats	☐		
	Antelope	☐	other	
2.2	**In the last month, have you been bitten, scratched, or cut by any of these animals?**			
	Chickens, ducks, gees	☐	none	☐
	Cows, goats, sheep	☐		
	Rodents	☐	don't know	☐
	Primates	☐	refuse to answer	☐
	Bats	☐		
	Antelope	☐	other	
2.3	**In the last month, have you killed, butchered, or cooked any of these animals?**			
	Chickens, ducks, gees	☐	none	☐
	Cows, goats, sheep	☐		
	Rodents	☐	don't know	☐
	Primates	☐	refuse to answer	☐
	Bats	☐		
	Antelope	☐	other	

3 Travel

3.0 How far and how often do you travel outside of Digale?

Place	How often do you visit this place?							When you go to this place, how much time do you spend there?						
	Most days of the week	At least once per week	At least once per month	Less than once per month	Never	Don't know	Refuse to answer	<1h	1-2h	Half a day	A whole day	More than one day	Don't know	Refuse to answer
<5km away (<1 hour walk)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
5-10km away (1-2hr walk; i.e. Hargeissa)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
>10km away	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

3.1 What is the name of the village, town, or settlement of the furthest place you went to last week?

Digale	☐	don't know	☐
Hargeissa	☐	refuse to answer	☐

other

3.2 **We are now interested in all the places where you have spent time yesterday. Could you first list all the places you visited, starting from when you woke up yesterday until when you woke up today? For each place, could you tell us whether it was inside or outside of Digale? Could you also estimate how much time you spent there? Can you also estimate the number of people who you saw for only a very short period of time (<5mins), and with who you exchanged at least three words in a face-to-face conversation? (examples are greeting someone on your way, seeing someone in a shop, or seeing a few children in the playground).**

	Type of setting													Location			Time spent here						Number casual contacts					
Setting ID	Home	Another house	Work	School	Mosque	Transport	Leisure	Shop	Market	Water source	Garden	Other	Name other	In Digale	Outside Digale	Name of location	<15mins	15mins - <1h	1h - <2h	2h - <4h	>4h	Don't know	0	1-2	3-5	6-10	>10	Don't know
S01	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		☐	☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
S02	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		☐	☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
S03	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		☐	☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
S04	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		☐	☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
S05	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		☐	☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
S06	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		☐	☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
S07	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		☐	☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
S08	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		☐	☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
S09	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		☐	☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
S10	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐		☐	☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

4	Human contacts
4.0	We will now ask you to remember who you was in contact with yesterday, **between the time of waking up yesterday and the time of waking up today**. We will ask this for each of the settings you listed in the previous question. We are only interested in your **direct contacts.** These are people with who you spent 5mins or longer, and with whom you exchanged at least three words. For each of them , we would like to know: - their age in years; - sex; - your relationship to the contact; - whether this contact lives in Digale; - how often you contact this person in general; - the total time spent with that person in that place; - We would also like to know whether contact was physical or nonphysical. **Nonphysical contact** happens when you haven't touched the person. **Physical contact** includes hand shaking, embracing, kissing, and sharing a bike, and also sharing a glass or other utensils passed directly from mouth to mouth. We only ask for their contact initials or name to help in remembering the contact, but we will delete that information at the end of the interview. If you had contact with the same person in multiple settings, we will record this multiple times, once for each setting. If you don't know their exact age, you can take a best guess.

Contact initials or name	Contact ID	Setting ID	Age	Sex		Contact type		relationship to person							Does contact live in Digale?				How often do you have contact with this person in general?							How much time did you spend with this person in this place?					Match contact id
				male	female	Physical	nonphysical	household member	other relative	coworker or schoolmate	friend	other	don't know	refuse to answer	yes	no	don't know	refuse to answer	daily or almost daily	at least once per week	at least once per month	less than once per month	never met before	don't know	refuse to answer	5-15min	15min-1h	1h-2h	2h-4h	>4h	
	C	S																													C
	C	S																													C
	C	S																													C
	C	S																													C
	C	S																													C
	C	S																													C
	C	S																													C
	C	S																													C
	C	S																													C
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

5	Health questions			
	We are almost at the end of this questionnaire, but have a few remaining health related questions.			
5.0	**Have you been feeling ill or had any of the following symptoms in the past two weeks?**			
	cough	☐	none	☐
	sore throat	☐		
	runny nose	☐	don't know	☐
	sneezing	☐	refuse to answer	☐
	diarrhoea	☐		
	headache	☐	other	
	wheezing	☐		
	difficulty breathing	☐		
	fever with chills like malaria	☐		
	fever without chills	☐		
5.1	**Have you received any antibiotics in the past two weeks?**			
	Yes	☐	don't know	☐
	No	☐	refuse to answer	☐
5.2	**Have you ever been diagnosed with any of the following illnesses by a healthcare worker?**			
	Pneumonia	☐	none	☐
	Diabetes mellitus	☐		
	Sickle cell disease	☐	don't know	☐
	HIV	☐	refuse to answer	☐
	Chronic Obstructive Pulmonary Disease (COPD)	☐		
	Asthma	☐		
5.3	**Do you use any of the following substances?**			
	Tobacco (cigarettes or cigars)	☐	none	☐
	Tobacco (snuff)	☐		
	Khat	☐	don't know	☐
	Alcohol	☐	refuse to answer	☐

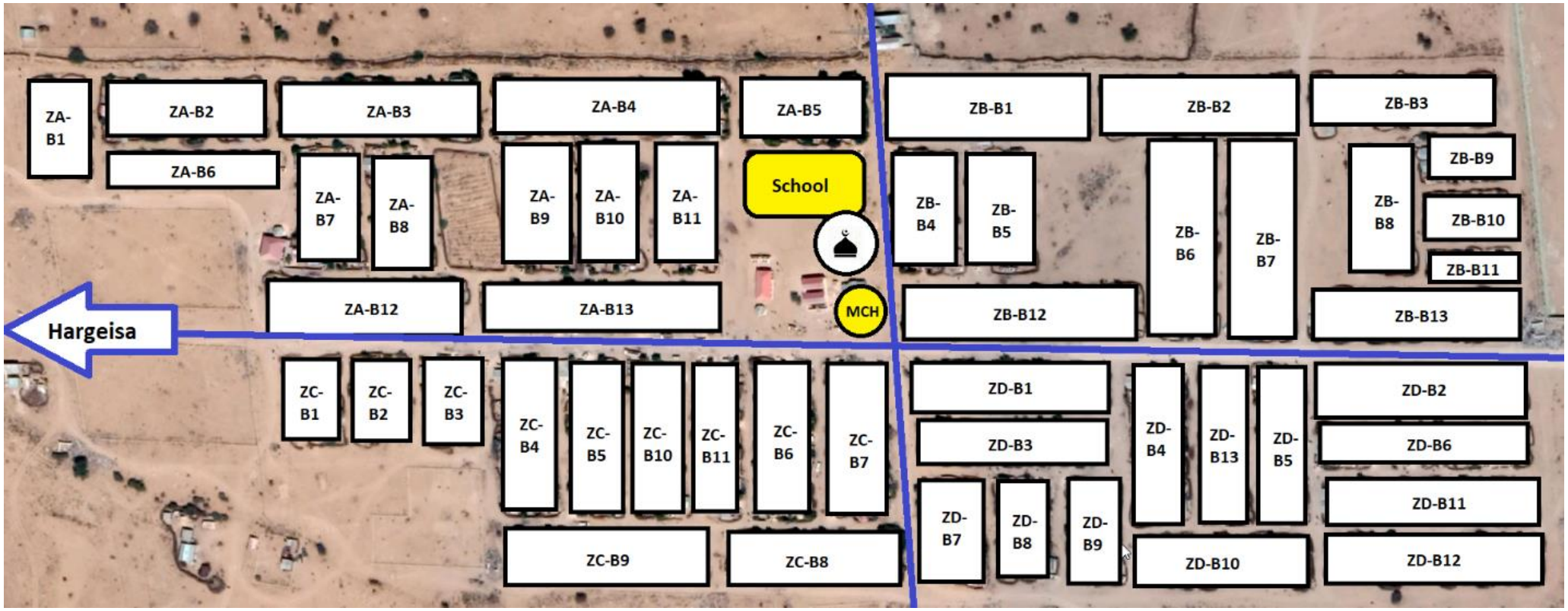
6	Nutritional survey intake
	Thank you very much for your participation in this contact survey. The names you provided earlier are now destroyed. **[if 6mo > age > 59mo]** **go to 7.0** **[else]** **go to 8.0**

7	Nutritional survey			
	We will now take a few anthropometric measures. This allows us to assess the general health of your child.			
7.0	**Weight (in kgs, round to nearest 100 gram)**			
	☐☐.☐			
7.1	**Height/length in centimetres**			
	☐☐☐			
7.2	**Is height or length measured?**			
	height ☐			
	length ☐			
7.3	**MUAC in centimetres, to nearest mm**			
	☐☐.☐			
7.4	**Oedema**			
	Absent ☐		Not able to assess ☐	
	Mild: both feet and ankles ☐			
	Moderate: both feet, plus lower legs, hands or lower arms ☐			
	Severe: generalized oedema including both feet, legs, hands, arms, and facce ☐			
7.5	**Any issue that may affect results?**			
	child has missing limb ☐		none ☐	
	technical issue with measurement equipment ☐	go to 7.5.1	other	
	child was (partially) clothed ☐	go to 7.5.2		
7.5.1	**What equipment has issues?**			
	weight scale ☐			
	measuring board ☐			
	MUAC tape ☐			
7.5.2	**What clothes did the child wear?**			
	only undergarments ☐		other	
	skirt/trousers ☐			
	shirt ☐			
	shoes ☐			

8	Nasopharyngeal swab		
	We will now take the nasopharyngeal swab. This will only take a moment.		
8.0	**Is the swab collected?**		
	yes ☐	go to 8.0.1	
	no ☐	go to 8.0.6	
8.0.1	**Time swab collected (HH – MM)**		
	☐☐ - ☐☐		
8.0.2	**Copy code from vial**		
	☐☐☐☐☐☐		
8.0.3	**Are secretions present on the swab?**		
	yes ☐		
	no ☐		
8.0.4	**Possible contamination of the swab?**		
	touched by anybody ☐	none ☐	
	dropped on the ground ☐	other	
8.0.5	**An adverse events?**		
	none ☐	yes	
	go to 8.1	go to 8.1	
8.0.6	**Why was the swab not collected?**		
	participant refused ☐	other	
	no swab available ☐		
8.1	**This was the end of the questionnaire. Thank you very much for your cooperation. Do you have any last comments or remarks for us?**		
	none ☐	yes	

Appendix D

Map of Digaale IDP camp in 2019


ZA-B1
ZA-B2
ZA-B3
ZA-B4
ZA-B5
ZA-B6
ZA-B7
ZA-B8
ZA-B9
ZA-B10
ZA-B11
ZA-B12
ZA-B13
School
MCH
Hargeisa
ZB-B1
ZB-B2
ZB-B3
ZB-B4
ZB-B5
ZB-B6
ZB-B7
ZB-B8
ZB-B9
ZB-B10
ZB-B11
ZB-B12
ZB-B13
ZC-B1
ZC-B2
ZC-B3
ZC-B4
ZC-B5
ZC-B10
ZC-B11
ZC-B6
ZC-B7
ZC-B9
ZC-B8
ZD-B1
ZD-B3
ZD-B7
ZD-B8
ZD-B9
ZD-B4
ZD-B13
ZD-B5
ZD-B2
ZD-B6
ZD-B11
ZD-B10
ZD-B12