STUDY PROTOCOL

REVISED Investigating Pneumonia Etiology Among Refugees and the Lebanese population (PEARL): A study protocol [version 2; peer review: 2 approved, 1 approved with reservations]

Thomas Kesteman[1], Ali Ghassani[2], Crystel Hajjar[3], Valentina Picot[1], Marwan Osman[4], Zahraa Alnajjar[5], Florence Komurian-Pradel[1], Melina Messaoudi[1], Stéphane Pouzol[1], PEARL Study Group, Hicham Ghazi Soulaiman[6], Philippe Vanhems[7], Octavio Ramilo[8], Dolla Karam-Sarkis[9,10], Josette Najjar-Pellet[1], Monzer Hamze[4], Hubert Endtz[1,11]

[1]Fondation Mérieux, Lyon, 69002, France
[2]Amel Association, Beirut, Lebanon
[3]Faculté de Pharmacie, Université Saint-Joseph, Beirut, Lebanon
[4]Laboratoire Microbiologie Santé et Environnement, Lebanese University, Tripoli, Lebanon
[5]Chtoura Hospital, Chtoura, Lebanon
[6]Al-Bashaer Medical Center, Tripoli, Lebanon
[7]Infection Control and Epidemiology Unit, Edouard Herriot Hospital, Hospices Civils de Lyon, Lyon, 69002, France
[8]Nationwide Childrens' Hospital and the Ohio State University College of Medicine, Columbus, OH, 43205, USA
[9]Laboratoire des Agents Pathogènes, Faculté de Pharmacie, Université Saint-Joseph, Beirut, Lebanon
[10]Laboratoire Rodolphe Mérieux, Université Saint-Joseph, Beirut, Lebanon
[11]Erasmus Medical Center, Rotterdam, The Netherlands

v2 **First published:** 18 Apr 2018, **2**:19
https://doi.org/10.12688/gatesopenres.12811.1
Latest published: 13 Jun 2019, **2**:19
https://doi.org/10.12688/gatesopenres.12811.2

Abstract

Background: Community-acquired pneumonia (CAP), a leading cause of mortality, mainly affects children in developing countries. The harsh circumstances experienced by refugees include various factors associated with respiratory pathogen transmission, and clinical progression of CAP. Consequently, the etiology of CAP in humanitarian crisis situations may differ to that of settled populations, which would impact appropriate case management. Therefore, the Pneumonia Etiology Among Refugees and the Lebanese population (PEARL) study was initiated with the objective of identifying the causal pathogenic microorganisms in the respiratory tract of children and adults from both the refugee and host country population presenting with signs of CAP during a humanitarian crisis.

Methods: PEARL, a prospective, multicentric, case-control study, will be conducted at four primary healthcare facilities in Tripoli and the Bekaa valley over 15 months (including two high-transmission seasons/winters). Sociodemographic and medical data, and biological samples will be collected from at least 600 CAP cases and 600

Open Peer Review

Approval Status ✓ ? ✓

	1	2	3
version 2 (revision) 13 Jun 2019	✓ view		✓ view
version 1 18 Apr 2018	? view	? view	

1. **Anne-Laure Page**, Epicentre, Paris, France
2. **Stephen R. Howie**, University of Auckland, Auckland, New Zealand
3. **Carina King**, Karolinska Institutet, Stockholm, Sweden

controls. Nasopharyngeal swabs, sputum, urine and blood samples will be analyzed at five clinical pathology laboratories in Lebanon to identify the bacterial and viral etiological agents of CAP. Transcriptomic profiling of host leukocytes will be performed.
Conclusions: PEARL is an original observational study that will provide important new information on the etiology of pneumonia among refugees, which may improve case management, help design antimicrobial stewardship interventions, and reduce morbidity and mortality due to CAP in a humanitarian crisis.

Keywords
Community-Acquired Pneumonia, Lebanon, Refugees, Etiology, Case-Control Studies, Epidemiology, Prevention & Control, Risk Factors

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Thomas Kesteman (thomas.kesteman@ext.fondation-merieux.org)

Author roles: Kesteman T: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Ghassani A**: Conceptualization, Investigation, Methodology, Resources, Writing – Review & Editing; **Hajjar C**: Formal Analysis, Investigation, Resources, Writing – Review & Editing; **Picot V**: Conceptualization, Data Curation, Funding Acquisition, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Osman M**: Investigation, Methodology, Resources, Writing – Review & Editing; **Alnajjar Z**: Investigation, Methodology, Resources, Writing – Review & Editing; **Komurian-Pradel F**: Conceptualization, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Messaoudi M**: Conceptualization, Methodology, Writing – Review & Editing; **Pouzol S**: Conceptualization, Methodology, Writing – Review & Editing; **Soulaiman HG**: Conceptualization, Investigation, Resources, Writing – Review & Editing; **Vanhems P**: Conceptualization, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Ramilo O**: Conceptualization, Data Curation, Formal Analysis, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Karam-Sarkis D**: Conceptualization, Investigation, Methodology, Resources, Writing – Review & Editing; **Najjar-Pellet J** : Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; **Hamze M**: Conceptualization, Investigation, Methodology, Resources, Writing – Review & Editing; **Endtz H**: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project is supported by the Bill and Melinda Gates Foundation [OPP1168739] and the Fondation Mérieux.

How to cite this article: Kesteman T, Ghassani A, Hajjar C *et al.* **Investigating Pneumonia Etiology Among Refugees and the Lebanese population (PEARL): A study protocol [version 2; peer review: 2 approved, 1 approved with reservations]** Gates Open Research 2019, **2**:19 https://doi.org/10.12688/gatesopenres.12811.2

First published: 18 Apr 2018, **2**:19 https://doi.org/10.12688/gatesopenres.12811.1

REVISED **Amendments from Version 1**

As compared with version 1, and thanks to the accurate remarks of the two reviewers, the version 2 of this manuscript has been stripped of imprecisions, clarified, and complemented with additional details (inclusion criteria, molecular methods, procedures for transcriptomics, conventional microbiology methods), as well as a more comprehensive and nuanced discussion. Changes included the addition of the complete methodology of a triplex PCR that had not been disclosed in a previous scientific article. This prompted us to add M Stéphane Pouzol, who designed and developed that technique, as an author of the second version of the manuscript.

'Erasmus Medical Center, Rotterdam, The Netherlands' has been added as affiliation for co-author Hubert Endtz.

See referee reports

Introduction

Lower respiratory tract infections (LRTI) are the second leading cause of mortality worldwide, accounting for an estimated 2.8 million deaths annually, and mainly affect children in developing countries[1]. Community-acquired pneumonia (CAP) is caused by a variety of bacteria and viruses and is mainly characterized by lobar or broncho-pneumonic changes. However, identification of the etiology of pneumonia is often difficult, and optimal prevention and treatment strategies for CAP critically depend on a full understanding of its etiology. For example, intracellular ("atypical") bacteria (*Chlamydophila, Mycoplasma,* etc.) require treatment with specific antibiotics, while purely viral infections do not. Furthermore, clinical exams, chest radiology and biological tests lack specificity, and blood cultures yield positive results in only 10 to 20% of cases.

Current interventions for CAP are primarily based on etiological studies conducted in the early 1980s[2–4], which indicated bacteria are responsible for almost half of all cases of CAP[5,6]; *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Staphylococcus aureus* were the most commonly identified bacteria. In children < 5-years-old, bacteria were responsible for severe forms of CAP: *S. pneumoniae* and *H. influenzae* accounted for approximately 60% of cases of severe and fatal pneumonia[7,8] and *S. pneumoniae* alone accounted for 11% of mortalities overall in children < 5-years-old[9]. However, other bacteria (e.g. *Mycoplasma pneumoniae, Chlamydophila pneumoniae, Legionella pneumophila, Bordetella pertussis,* etc.*)* may be responsible for a significant proportion of CAP[10]. Moreover, respiratory viruses such as influenza viruses, respiratory syncytial virus (RSV) and human metapneumovirus (hMPV) are known to make significant, often seasonal, contributions to CAP. Finally, superinfections, although poorly understood, are widely accepted to contribute to severe CAP. A clear picture of the etiology of CAP would help to estimate the potential impact of novel public health interventions, such as antiviral therapies in case management or influenza vaccination for prevention in vulnerable populations. In order to provide up-to-date data on the etiology of CAP, several studies were recently conducted, in particular two large studies in the USA in 2010–2012, targeting children and adults[11,12], and two studies that targeted under-fives in low and middle income countries, the GABRIEL Pneumonia Study in 2010–2014, and the Pneumonia Etiology Research for Child Health (PERCH) study in 2011–2014[13]. These studies pointed to the importance of viruses in the etiology of CAP.

LRTI are a major cause of morbidity and mortality in the acute phase of humanitarian crises. Case management (diagnoses, treatment choice and delivery) and public health interventions (immunization strategies) for CAP in humanitarian crisis settings are conducted blindly or based on the assumption that evidence gathered in non-crisis settings applies to displaced populations. However, large knowledge gaps remain in many areas, including the etiology of these infections[14,15]. No studies have been specifically designed to assess the etiology of CAP in populations facing humanitarian crisis.

Evaluation of the etiology of infection

The reference method for identifying the pathogen responsible for an infection is to sample the infected tissues, i.e., the lung, in patients with pneumonia, and screen for pathogenic agents by culture or molecular tools. Such invasive procedures are difficult to set up, even in hospital settings, and virtually impossible in primary health care. To circumvent this issue, the causative agent can be identified using samples from the upper respiratory tract using non-invasive techniques, e.g. nasopharyngeal swabs[11,12]. Due to the existence of healthy carriers of several potential pathogens (e.g. *Streptococcus pneumoniae*) in the population, there is a need to adjust for carriage: the higher the proportion of asymptomatic carriers of a given pathogen, the lesser the chances that this agent is the etiology of CAP when found in the upper respiratory tract.

In the last decade, transcriptomic analyses of the blood of patients infected with different pathogens have revealed gene expression patterns that correlate strongly with individual etiologic agents[16,17]. White blood cells express different genes in response to infection with different agents, and these patterns can be used to distinguish viral and bacterial infections with high accuracy[18]. These transcriptomic patterns can even distinguish between different viral (or bacterial) infections, and also enable evaluation of the severity of infection. RNA sequencing analysis of the gene expression profiles of white blood cells during infection enables differentiation of viral and bacterial infections with higher specificity than white blood cell counts[19]. Therefore, transcriptomics represents an innovative tool that reduces the need to identify pathogens by culturing respiratory tract samples and does not need to be adjusted for asymptomatic carriage. The ideal control samples for transcriptomic analysis are samples from healthy patients, i.e. those attending for vaccination or other routine healthcare visits.

Situation of Syrian refugees in Lebanon

Lebanon is severely affected by a complex, chronic and protracted refugee crisis due to the ongoing war in Syria that began in March 2011. The Bekaa valley is the main entry point for Syrian refugees, followed by the North Lebanon region and Beirut, with refugees settling all over the country. The crisis not only affects Syrian refugees, but also the local Lebanese population and Palestinian refugees who settled decades ago.

The majority of Syrian refugees reside in approximately 4000 informal tented settlements and mobile centers, garages and unfinished buildings, with no officially established access to food, water, sanitary means, health or education.

Respiratory tract infections (RTI) are among the leading causes of morbidity and mortality in children and adults affected by the current humanitarian crisis in Lebanon. According to Médecins Sans Frontières (MSF), RTI accounted for at least 56% of ambulatory or in-patient health care visits in the first six months of 2015, and up to 79% of visits among children under 5-years-old in the healthcare facilities attended mainly by refugees in the Bekaa valley (MSF, personal communication). Unpublished data from primary healthcare facilities in Lebanon suggests that 10–25% of these cases of RTI are CAP (Ali Ghassani, personal communication), though data on the etiology and severity of these cases of CAP cases is lacking.

In Lebanon, CAP peaks between October and April. Specifically, 80% of the primary consultations for children during winter are related to RTI, compared to 20% in the summer (Ali Ghassani, personal communication). The Lebanese surveillance program for pulmonary infections due to *S. pneumoniae* considers the general population, and therefore current data on the incidence, predominant serotypes, and antimicrobial susceptibility of CAP among the refugee population is not available[20]. UNICEF and the Lebanese Ministry of Health started the introduction of a 13-valent pneumococcal vaccine (PCV) as part of a National Program in 2016, among children of Syrian, Palestinian and Lebanese origin.

Objectives

In order to fill the knowledge gaps described above, the Pneumonia Etiology Among Refugees and the Lebanese population (PEARL) study was initiated in 2016 to identify the causal pathogenic microorganisms in the respiratory tract of children and adults presenting with signs of CAP to health facilities run by medical associations in the context of a humanitarian crisis. The first inclusions took place in November 2016, and the study is expected to last until March 2018.

The primary objective of the PEARL study is to estimate the population attributable fractions (PAFs) of specific viral and bacterial pathogens, i.e. the proportion of CAP attributable to each pathogen, in both the refugee population and Lebanese population. The goals behind this objective are (i) to enable local healthcare staff to provide more accurate diagnoses and improved case management and care, (ii) to help designing antimicrobial stewardship interventions, and (iii) to help assessing the impact of PCV, as it will generate baseline data on the burden of CAP caused by *Streptococcus pneumoniae* at the introduction phase of PCV in the national vaccination program.

The secondary objectives of the PEARL study are to: identify *S. pneumoniae* serotypes in nasopharyngeal and blood samples; identify the antimicrobial susceptibility profiles of the pathogenic bacteria isolated from nasopharyngeal and blood samples and compare these profiles with antibiotic prescriptions; examine the association between respiratory viral infections and invasive pneumococcal infections; identify risk factors for CAP in this population, especially those that may be modifiable (crowding, tobacco smoking, domestic sources of smoke produced by cooking or heating, etc.), and thus provide data for prevention programs; provide current data on the incidence and severity of CAP in vulnerable populations in Lebanon; provide a unique transcriptomics dataset as regards the sociodemographic profile of the patients and spectrum of diseases; compare microbiological and transcriptomic methods in estimating viral vs bacterial attributable fractions of LRTI; and assess the operational capacity of rapid, nuclear acid-based point-of-care diagnostic tests in a humanitarian crisis.

Protocol

Study setting and design

Based on refugee density, two main regions were selected for this multicentric prospective case-control study: the Bekaa valley and the city of Tripoli (Figure 1). Both regions have similar demographic composition, dynamics and climate. The population eligible for the study consists of any child (>2-months-old) or adult attending one of four primary healthcare facilities (no hospital facilities) that provide routine health care and immunization: one site located in Tripoli and three sites in the Bekaa valley (Kamed el Loz, Machghara, and Baalbeck El Ain). In addition to refugees, local and vulnerable Lebanese individuals who also attend these health centers (although to a lower extent) will be included for ethical and practical reasons.

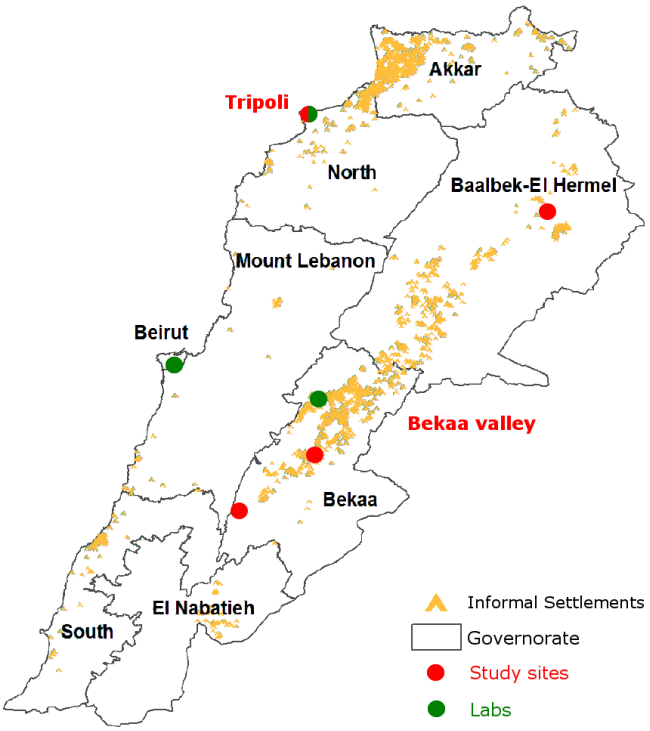

Figure 1. Concentration of Syrian refugees in Lebanon (source: UNHCR[21], image adapted under CC-BY 3.0 license) and location of the PEARL study sites.

Nasopharyngeal, urine and blood samples will be collected from patients with CAP (cases). Adults with CAP able to produce sputum will be asked to provide a sputum sample.

As CAP-related pathogens can also be carried by healthy individuals, control individuals will be recruited from the same health facilities; cases and controls will be matched in a 1:1 ratio by age, season and site. Nasopharyngeal swabs and urine will be collected from all controls; a subset of controls (15%) will also be asked to provide a small blood sample (500–1000 µL) for transcriptomic analysis. Controls will not be asked to provide sputum samples.

After enrolment in the study, patients will be managed as per local guidelines and doctors' recommendations; in particular, the study won't alter procedures for referral for complementary examinations or hospitalization.

Inclusion criteria

Cases. Physicians will evaluate case definitions on the basis of a clinical examination and patient history. The case definition was based on WHO's IMCI[22]. Cases should meet all of the following inclusion criteria: (i) patient aged > 2 months, (ii) with cough or dyspnea, (iii) lower chest wall indrawing (in children ≤ 3 years only) or tachypnea, (iv) no wheezing suggestive of asthma at auscultation, (v) onset of symptoms within the last 14 days, (vi) and informed consent statement signed by the patient, parent, or legal guardian (Supplementary File 1). Tachypnea is defined as > 50 breaths per minute in patients between 2-months and < 1-year-old; > 40 breaths per minute, between 1 and < 5-years-old; 30 breaths per minute, between ≥ 5 and <18-years-old and > 20 breaths per minute, if ≥ 18 years-old. Characteristics of wheezing suggestive of asthma included: expiratory wheezing, high pitched wheezing, wheezing with history of asthma and without fever nor history of fever, and any clinical picture such as the clinician would exclude LRTI and retain asthma as final diagnostic.

Exclusion criteria are (i) any characteristic of healthcare-associated pneumonia, (including hospitalization at an acute care hospital for 2 or more days within 90 days of infection, residence in a nursing home or long-term care facility, recent intravenous antibiotic therapy, or wound care from medical staff within the 30 days prior to the current infection); or (ii) increased risk of lower airway disease, such as immunocompromised status due to underlying disease, including hemodialysis, or immunosuppressant treatment; or (iii) treatment with inhaled corticosteroids or other asthma medications.

The study nurse will check inclusion criteria, and ask the patient or their parent/guardian to provide signed informed consent and fill in the case report form (CRF), which includes a medical history (e.g. HIV infection, tuberculosis, respiratory infections), risk factors for pneumonia, prior medical treatment and immunizations, description of clinical signs that warrant enrolment in the study, current and recent treatment (antibiotic therapy), disease progression, and socio-economic characteristics (Supplementary File 2). Pulse oximetry is used to assess the severity of CAP[23]; pulse oximeters will be made available to all four healthcare facilities for the present study and pulse oximetry will be reported in the CRF. Data will be anonymized by attributing a unique patient identification code number issued at enrolment of each patient.

Physicians may request other tests such as chest radiography without interfering with the study protocol. Antimicrobial treatment may begin immediately after blood sampling, if required. Cases and controls may also undergo clinical and lab tests as required. Chest X-ray is not required as this examination is not available at all health care facilities, especially the three remote healthcare facilities in Bekaa valley; if available, results from chest X-rays will be reported in the CRF.

Severity of pneumonia will be classified according to the IMCI/WHO guidelines[22]. Moreover, in children under five years of age, we will assess the breathing rate twice, since it has been shown that the cut-off of > 50 breaths per minute, assessed twice, is more specific for pneumonia than the standard IMCI recommendations for children 1–5 years old (40 breaths/min)[24], and was successfully adopted in the ALMANACH protocol[25,26].

Controls. An aged-matched control will be recruited for each case by age group (± 1 year for children aged 2 months to 4 years, ± 2 years for children 5–17 years, and ± 5 years for patients aged 18–49 years old and adults aged ≥ 50 years old). Controls will be matched to cases attending the same site in the same calendar month (some flexibility is allowed since controls for cases included at the end of one month may be recruited in the next calendar month). The study nurse will identify next patient consulting for reasons other than respiratory (upper or lower respiratory tract) or gastrointestinal infection, check eligibility, obtain written informed consent from the participant, and fill in the CRF.

Controls should meet all of the following inclusion criteria: (i) patients aged > 2 months attending one of the four sites participating in the study for symptomatic disease or immunization; and (ii) informed consent signed by patient or parent/guardian. Patients will be excluded as controls if they (i) exhibited any symptom of RTI (cough, dyspnea, chest wall indrawing, tachypnea, fever, coryzal symptoms/"cold") or (ii) intestinal infection (watery/bloody diarrhea, abdominal cramps) in the last 5 days. Individuals with gastro-intestinal symptoms were excluded as some viruses (adenovirus, enterovirus, and coronavirus) are proven or suspected to cause either respiratory or gastro-intestinal symptoms[27–29], and including such patients as controls would result in an underestimation of the attributable fraction of these pathogens.

Sample size calculation

As pathogen distributions can vary according to patient age, the univariate and multivariate analyses will be stratified by age group. We calculated the total sample size required for each age group to detect a difference in pathogen prevalence between cases and controls (assuming an equal distribution of samples between the age groups) based on a power of 90% ($\alpha = 0.05$). Pearson's χ^2 test indicated 150 cases and 150 controls were required in each of the four age strata (2 months to 4 years,

5–17 years, 18–49 years, > 50 years) to detect a 15% difference in pathogen frequency between cases and controls (with a carriage prevalence < 10% in controls) or detect a difference in pathogen frequency of 20% (with a carriage prevalence of up to 30% in controls). If no differences between age groups is observed, pooling the data from age groups will increase statistical power.

Thus, the final sample size will be a minimum 600 cases and 600 controls with the aim of including a maximum of 900 cases and 900 controls.

Analysis of biological samples

Biological samples will be collected and stored using standard protocols and transported together with the sample log to the laboratories in a multicentric manner (Table 1). All analytical tests will be performed according to good clinical laboratory practice (GCLP)[30] following standard operating procedures defined for the study. Sample processing for cases and controls are illustrated in Figure 2 and Figure 3, respectively. The total volume of blood taken is 0.5-1.0 mL in controls of all ages and, in cases, 2.5-5.0 mL in infants (<1 year old), 3.5-6.0 mL in children aged one to four, 10.5-17.0 and 6 mL in children aged five or above, and 18.5-19.0 mL in adults (≥18 years old), with a flexibility in the age cut-off depending on the individual's weight and overall health status.

Clinically relevant results will be communicated to the clinicians in a timely manner. Such results include: (i) identification of pathogen bacteria in blood, by conventional microbiology or molecular tools, (ii) identification of clinically relevant bacteria in nasopharyngeal swab or sputum of cases, e.g. detection of atypical bacteria; (iii) identification of RSV in infants (cases or controls) aged ≤12 months, and (iv) positive Binax® test in urine of case. Whenever the lab identifies a pathogen or sensitivity profile of importance for the clinical management of pneumonia, they will provide their etiological diagnosis to the clinical team to ensure appropriate case management. All laboratory data will also be transmitted to the local Fondation Mérieux office in Beirut.

Molecular biology testing of respiratory samples. A trained nurse or lab technician-in-charge will collect two nasopharyngeal samples from all cases and controls[31]: one will be discharged in virus transport medium (VTM) and the other in skim milk-tryptone-glucose-glycerol (STGG) medium. Nasopharyngeal swabs will be processed and analyzed as shown in Table 1 and Figure 2 and Figure 3.

Nucleic acids extracted from 350 µL of VTM using QIAamp® DNA mini kit (Qiagen, Hilden, Germany) will be subjected to a real-time triplex PCR assay targeting *Streptococcus pneumoniae, Staphylococcus aureus* and *Haemophilus influenzae B,* the three most common bacteria detected in CAP. This triplex PCR has been used in previous studies[32,33]. Briefly, 5 µL extracted nucleic acids are added to 9 µL Takyon No ROX Probe 2X MasterMix dTTP (Eurogentec, Seraing, Belgium) and 4 µL of a solution containing 1 µM of each primer and probe. Each PCR

Table 1. Collection, storage, processing and analysis of clinical samples.

Specimen	Collection container	Storage temperature	Assay	Processing	Laboratory
Blood	Blood culture bottle	+4°C	Blood culture and AST	Real time	CH, LMSE
	Tempus	+4°C/-20°C	Transcriptomics	Batch	NCH
	EDTA tube	+4°C	WBC	Real time	EB, CH
		+4°C/-80°C	Triplex PCR	Batch	LRM, LMSE
Urine	Sterile container (or sterile adhesive bags for babies)	+4°C	Antibiotic activity*	Batch	LRM, LMSE
		+4°C	Binax	Real time	CH, LMSE
Nasopharyngeal swab	Viral transport medium	+4°C/-80°C	Triplex PCR	Batch	LRM, LMSE
		+4°C/-80°C	Micro-array (BioFire)*	Real time	
		+4°C/-80°C	*S. pneumo* CAPS-PCR	Batch	
Nasopharyngeal swab	STGG medium	+4°C	Bacterial culture*	Real time	CH, LMSE
		+4°C/-80°C	CAPS-PCR	Batch	LRM, LMSE
Sputum (adults only)	Sterile container	+4°C	Bacterial culture	Real time	CH, LMSE
		+4°C/-80°C	Triplex PCR	Batch	LRM, LMSE
		+4°C/-80°C	Micro-array (BioFire)*	Real time	
		+4°C/-80°C	*S. pneumo* CAPS-PCR	Batch	

*: assay also performed in controls; CH: Chtoura hospital; LMSE: Laboratoire Microbiologie Santé et Environnement; LRM: Laboratoire Rodolphe Mérieux; NCH: Nationwide Children's Hospital; EB: El Bashaer health center.

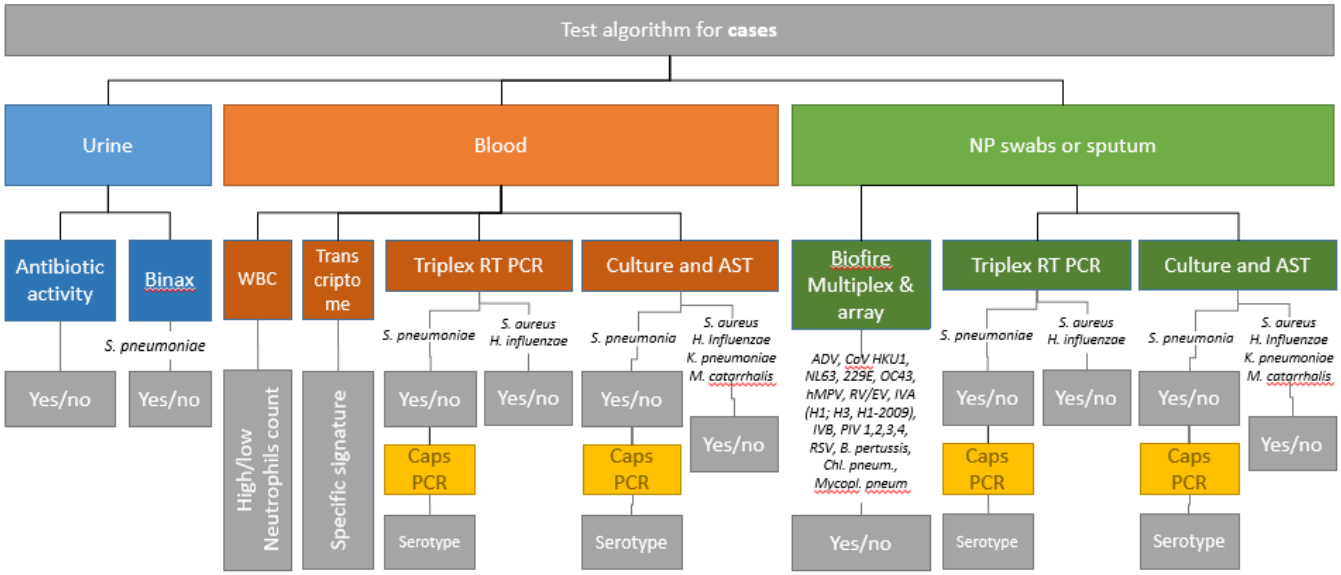

Figure 2. Sample analysis for cases. NP, nasopharyngeal; WBC, white blood cells; AST, Antimicrobial Susceptibility Testing; Caps PCR, capsular antigen serotyping PCR.

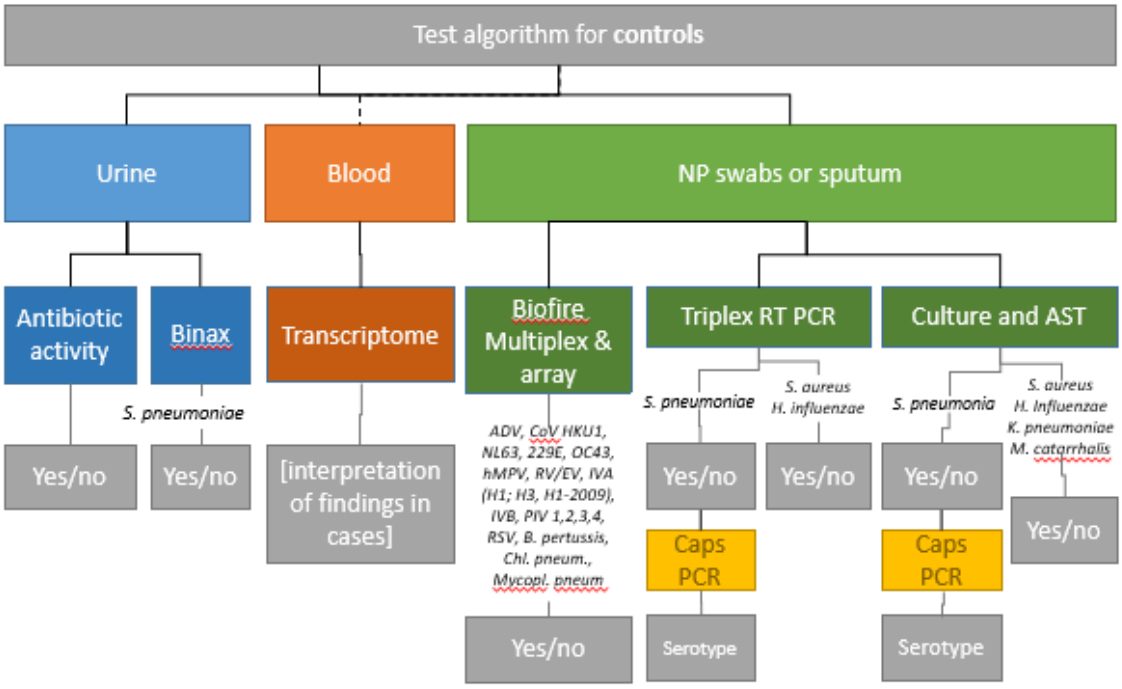

Figure 3. Sample analysis for controls. NP, nasopharyngeal; AST, Antimicrobial Susceptibility Testing; Caps PCR, capsular antigen serotyping PCR.

mixture is submitted to 95°C for 10 min then 40 cycles of 8 sec at 95°C then 34 sec at 60°C. The respective 5'-3' sequences of forward primers, reverse primers and probes are: lytA gene (S. pneumoniae) ACG AAT AAC CAA CCA AAC AAC, CCA GTA GCC AGT GTC ATT C, and tca Atc Gtc Aag Ccg ttc t using HEX as a fluorophore (capital letter in probes indicate locked nucleic acid substitution); vicK gene (S. aureus) GAA GCA GTC TAA CCG TAG TC, GGG ATA TTA TAT ACC CAG ACA CG, and tcc Tta Cca Ccg Cca taa, using FAM as a fluorophore; bexA gene (H. influenzae B) ATT TGA GAA ACG CAA AGA

CC, ATT TGA GAA ACG CAA AGA CC, and agt Ttc Aca Tag Ccc gag t, using Cy5 as a fluorophore. Amplification curves will be examined individually by two independent technicians, at the local laboratory and at the Laboratoire des Pathogènes Emergents (Lyon, France). Cycle thresholds (Ct) value will be manually set so that it intersects the exponential curve at its inflexion point. Any exponential signal observed between 0 and 40 Ct value will be considered as positive.

The FilmArray Respiratory Filmarray (BioFire Diagnostics, Salt Lake City, UT, USA) will be used to identify pathogens that cause pneumonia from VTM[34] and sputum samples. The 17 viruses and three atypical bacteria detected by this assay are adenovirus (ADV), coronavirus (CoV) HKU1, coronavirus NL63, coronavirus 229E, coronavirus OC43, hMPV, human rhinovirus (RV)/enterovirus (EV), influenza virus A (IVA), influenza A/H1, influenza A/H3, influenza A/H1-2009, influenza B, parainfluenza virus (PIV) 1, parainfluenza virus 2, parainfluenza virus 3, parainfluenza virus 4, RSV, *Bordetella pertussis, Chlamydophila pneumoniae,* and *Mycoplasma pneumoniae.*

Samples from cases positive for *Streptococcus pneumoniae* by conventional microbiology or triplex PCR will be serotyped using a Multiplex Real-Time PCR assay[35] that can identify 40 capsular antigen serotypes (1, 2, 3, 4, 5, 6A/B, 6C, 7C, 7F, 8, 9N/L, 9V, 10A, 10F, 11A, 12F, 13, 14, 15A, 15B/C, 16F, 17F, 18C, 19A, 19F, 20, 21, 22F, 23A, 23B, 23F, 24, 31, 33F, 34, 35A, 35B, 35F, 38, 39) and includes an internal control (*LytA*).

Culture of bacteria from respiratory samples. Sputum samples and STGG media will be inoculated onto different selective agar plates according to standard laboratory procedures (at least one blood culture medium, one chocolate agar, and one medium selective for gram-negative bacilli) and incubated under specific conditions (37°C, aerobic atmosphere supplemented with CO_2) to determine the presence of respiratory pathogens and assess antibiotic susceptibility[36]. Quality of sputum and absence of contamination by saliva will be assessed by standard laboratory procedures[36]. Positive cultures will be subjected to Gram staining and examined by light microscopy, and sub-cultured for identification and antimicrobial susceptibility testing if the strain is confirmed to be clinically relevant.

Detection of bacteria in blood. Duplicate aerobic haemoculture assays will be performed for all cases using BacT/ALERT (bioMérieux, Marcy l'Etoile, France) or BACTEC (Becton Dickinson, Franklin Lakes, NJ, USA) automated blood culture systems. The target blood volume for blood culture is 1-3 mL in under-fives and 6 mL in children aged five or above and in adults, following manufacturers' recommendations. If the volume of blood recovered is insufficient to inoculate two bottles, a single culture will be performed. Positive cultures will be examined by light microscopy (Gram staining) and subcultured on agar culture media for identification. Antimicrobial susceptibility testing will be performed for all clinically relevant strains.

Whole blood (200 µL EDTA) samples from all cases will be extracted using QIAamp DNA blood mini kit (Qiagen, Hilden, Germany), and the same semi-quantitative multiplex Real-Time PCR assay as for respiratory samples will be conducted on 5 µL of extracted DNA to identify *Streptococcus pneumoniae, Staphylococcus aureus* and *Haemophilus influenzae B*.

Transcriptomics. Blood samples (0.5–1 mL) for transcriptomics from all cases and 15% of controls will be collected in Tempus tubes containing stabilizing agents for conservation of mRNA, and shipped to the Nationwide Children's Hospital Research Institute, in Columbus, Ohio, USA. Sampling for transcriptomics was initiated in May 2017.

Total RNA will be isolated from whole blood collected in Tempus tube and analyzed for quality using the RNA 6000 Nano Kit (Agilent Technologies). Globin mRNAs will be further removed with GOBINclean kit (Thermo Fisher Scientific). Poly(A)-enriched next-generation sequencing library construction will be performed using the KAPA mRNA Hyper Prep Kit (KAPA Biosystems) with 500 ng of input total RNA and 11 amplification cycles according to the manufacturer's protocol. Individual libraries will be quantitated via quantitative PCR using the KAPA Library Quantification Kit, Universal (KAPA Biosystems) and equimolar pooled. Final pooled libraries will be sequenced on an Illumina HiSeq 2500 with single-end 70-base-pair (bp) read lengths.

Urinalysis. The presence of antibiotics in urine samples from all cases and controls will be assessed using the disk diffusion method[37], and the rapid immunochromatographic Binax® assay (Alere, Orlando, FL, USA) will be used for qualitative detection of *S. pneumoniae* C polysaccharide antigen in urine[38].

Other data collected

All involved healthcare facilities will be required to provide monthly report of the total number of consultations and, whenever available, the number of LRTIs. This will help weighting the monthly results from the study to reflect the actual incidence of CAP. A subgroup of patients or parents/legal guardians, randomly selected, was called back at least one month after inclusion to re-evaluate the outcome of included cases (hospitalization, death).

Study outcomes

The primary outcome of the PEARL study is the etiological distribution of CAP, expressed as PAFs for individual pathogenic agents. PAFs will be calculated for each three-month period of the 15-month study, and by patient age group (2 to 11-months-old and 1 to 4, 5 to 17, 18 to 49, and ≥ 50-years-old), by site location (Tripoli vs. Bekaa valley), subpopulation (Syrian refugees versus Lebanese population), and severity of pneumonia.

Additional indicators will be examined to assist data interpretation and for public health and clinical care purposes, including

the incidence of RTI and other diseases in the populations consulting the four healthcare facilities involved in the study, agent-specific hospitalization and fatality rates, the socio-demographic characteristics of the patients, clinical signs of CAP (e.g. severity criteria), epidemic features if appropriate (e.g. types of influenza viruses circulating during epidemic periods), and care provided (e.g. antimicrobial use).

Data management and analysis

All data (CRF, informed consent forms, data logs) will be transmitted to the local Fondation Mérieux office in Beirut for data entry using EpiInfo version 7.2 (CDC, Atlanta, GA, USA). The data will be anonymized by attributing a unique patient identification code number issued at patient enrolment. Access to the names of participants and their corresponding identification codes will be restricted and forms will be secured in locked cupboards.

Data analyses will include (i) descriptive comparisons of the sociodemographic, clinical, and lab data of cases and controls; (ii) univariate/multivariate logistic regression of the relationship between case/control status and pathogens, adjusted for age, season, site, and pneumonia risk factors; and (iii) computation of PAFs for every pathogen, specified by age group, period, study site, subpopulation (Syrian refugees, or general Lebanese population), and pneumonia severity, as described previously[39]. Ninety-five-percent confidence intervals will be calculated, with regards to the outcomes linked to the primary and secondary objectives. Analyses will be performed using R version 3.3.2 (R Core Team, Vienna, Austria[40]) and/or Stata version 13.0 (StataCorp LP, College Station, TX, USA) software. RNA sequencing data from transcriptomics will be analyzed at the Nationwide Children's Hospital, OH, USA. Quality control of raw reads will be performed with FASTQC. Reads will be aligned to the reference human genome (GRCh38) using hisat2 after quality and adapter trimming by cutadapt. FeatureCounts program will be used to quantify total number of read counts for each gene. The RNA sequencing data analysis will be performed in the R programming language, using DESeq2 R package for size factor, dispersion estimation calculation and differential gene expression analysis. A number of supervised and unsupervised analytical approaches, and modular analyses will be used to identify the pathogen-associated transcriptome profiles as previously described[41–43]. Etiologic fractions will be also computed from transcriptomics data, independently from calculations from microbiological data, and compared with each other.

Dissemination of the study outcomes

The results of the study (microbiological and transcriptomics results) will be published as scientific publications in international peer-reviewed journals. The designation of co-authors will conform to international guidelines governing publications.

The findings of the project, when completed, will be presented to the Ministry of Health of Lebanon, and to the medical and scientific community, such as United Nations agencies and non-governmental organizations, involved in medical care of vulnerable populations in Lebanon. The results will also be communicated via the GABRIEL website (https://www.gabriel-network.org/). The study data relevant to a publication authored by the investigators will be available for review in a public data repository

Ethical considerations

This study will be conducted in accordance with the Declaration of Helsinki[44], the recommendations for Good Ethical Practices in Epidemiology of the Association of French-language Epidemiologists[45], the Good Clinical Practice recommendations from the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) relevant to observational studies[46], and the GCLP guidelines of the WHO[30]. The informed consent statement has been translated into Arabic and was validated during the pilot phase of the study. All adult patients and the parents or legal guardians of children (i.e. ≤18 years) will be asked to sign an informed consent statement prior to enrolment in the study.

As no national ethic committee exists in Lebanon, the study protocol and all other documents related to the trial (informed consent, CRFs, amendments) have been approved by the institutional review boards of the following organizations: El Bashaer Association, Tripoli, Lebanon (5th December 2015); Université Libanaise, Ecole doctorale des Sciences et Technologies, Tripoli, Lebanon (3rd December 2015); Amel Association, Beirut, Lebanon (9th December 2015); and Université Saint-Joseph, Beirut, Lebanon (10th November 2016).

Current status of the study

Patient enrollment began in November 2016 and was completed in March 2018. At the time of submission of the first version of the manuscript, data cleaning and analysis was ongoing.

Discussion/conclusions

The Pneumonia Etiology Among Refugees and Lebanese population (PEARL) study will yield unique data on the etiology of CAP in the context of a humanitarian crisis. This study is not only original in terms of target population (refugees), but also as it will assess patients of all ages and adopt a primary health care approach while previous major studies only targeted hospitalized cases[2–5,7,13]. Therefore, this study may provide a unique and valuable perspective on the etiology of CAP in the context of a humanitarian crisis.

However, assessment of the etiology of CAP in low resource primary health care settings relies on clinical signs for case definition, as chest X-rays are not available at most of the centers in this study. Inclusion criteria for cases of pneumonia are a controversial issue. Several case definitions, based on different clinical, radiological, or biological criteria, have been used in different studies. Opting for a single case definition imposes restrictions on the anatomical and histological levels of the respiratory tract involved, on the severity of disease, and therefore may introduce a bias towards the relative prevalence

or contribution of specific pathogens. This makes it difficult to compare results between studies using different case definitions; the present study does not avoid this limitation.

We also anticipate other limitations. For example, the imperfect sensitivity of the diagnostic tests used to identify certain pathogens may result in an underestimation of certain etiological fractions[47]. Moreover, with regards to surveillance of pneumonia, given that the population covered by the healthcare facilities participating in this study is hard to define, estimation of the precise incidence of CAP in this population might not be possible. Additionally, certain aspects of CAP will not be explored in this study, especially AIDS-associated pneumonia and tuberculosis. We expect individuals with HIV will be excluded from the present study as HIV infection is one type of immunosuppression, which is an exclusion criterion. As the prevalence of HIV is very low (< 0.1%) in both the general Lebanese and Syrian populations[48], inclusion of cases HIV/AIDS-associated CAP would only have a small effect on the results. Patients with tuberculosis are also expected to be excluded, since duration of symptoms >14 days is an exclusion criterion. The incidence rate of tuberculosis among Syrian refugees did not increase in the last couple of years and remains relatively low at close to pre-war levels, i.e. 10-20/100.000 population[49]. Therefore, we consider it unlikely that tuberculosis is a major cause of acute LRTI in our population[50].

Finally, the external validity of our study is limited in time (15 months) and space (four health facilities). A case-control study cannot replace longitudinal surveillance systems, as are currently being set up in Lebanon, e.g. the Lebanese Inter-Hospital Pneumococcal Surveillance Program[20]. Therefore, although the situation of Syrian refugees in Lebanon shares a lot of characteristics with other crises, one should be cautious about applying the results of PEARL to other humanitarian crises, because of the variations of the epidemiology of LRTI.

Despite these limitations, the PEARL study is expected to provide healthcare planners with an empirical basis for the management of CAP in the context of a refugee crisis. Such information may help to guide population-based health interventions, such as immunization strategies for pneumococci, *H. influenzae* and influenza. In particular, analysis of *S. pneumoniae* serotypes in vaccinated and non-vaccinated individuals will provide information on the herd effect following introduction of the PCV in Lebanon and will provide proxy baseline data for evaluation of the success of this vaccination program. This study will also provide practical experience and a methodology for determination of the etiology and involvement of viral and bacterial agents in CAP in other similar humanitarian crisis settings.

Data availability

No data is associated with this article.

Grant information

This project is supported by the Bill and Melinda Gates Foundation [OPP1168739] and the Fondation Mérieux.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

We would like to thank all of the participants in this study, who provided their time and biological samples. We are very grateful to all health care workers, nurses, medical doctors and laboratory technicians at the participating health facilities, in particular: Ahmad Al Hallak, Anas Altabaa, Ali Rida, Asma Allouch, Danielle Chaaya, Hossam Al Nazer, Hussein Madie, Khaled Hallak, Majdeddine Mouzawak, Marianne Abi Fadel, Maryam Menhem, Nahed Elboosh, Souraya Nasser, Taha Abdou, Yasmine Amraoui, Zeina Jbara, and the other individuals who helped to set up this protocol. Khaled S. Aoun, Kamile, and especially Shafiq Bassil from Bioteck company deserve a special note for their kind collaboration in the transportation of biological samples. We thank Rana Hajjeh, Abdullah Brooks and Keith Klugman for their careful review of the protocol. Editorial support was provided by Andrea Devlin of Science Editing Experts.

PEARL Study Group members: *Valentina Abdel Khelek (Amel Association, Lebanon), Mohammad Alabrash (Al-Bashaer Medical Center, Lebanon), Ghadban Al Ghadban (Amel Association), Mohamad Al Zayed (Amel Association), Cynthia Bakkalian (Fondation Mérieux, France), Zeinab Farhat (Amel Association), Ahmad Obeid (Al-Bashaer Medical Center), Pierre Salloum (Bioteck, Lebanon), Haneen Saty (Amel Association).*

Supplementary material

Supplementary File 1: Informed consent form.
Click here to access the data.

Supplementary File 2: Case report form.
Click here to access the data.

References

1. Lozano R, Naghavi M, Foreman K, *et al.*: **Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010.** *Lancet.* 2012; **380**(9859): 2095–128.
PubMed Abstract | Publisher Full Text

2. Shann F: **Etiology of severe pneumonia in children in developing countries.** *Pediatr Infect Dis.* 1986; **5**(2): 247–52.
PubMed Abstract

3. Wall RA, Corrah PT, Mabey DC, *et al.*: **The etiology of lobar pneumonia in the Gambia.** *Bull World Health Organ.* 1986; **64**(4): 553–8.
PubMed Abstract | Free Full Text

4. Ikeogu MO: **Acute pneumonia in Zimbabwe: bacterial isolates by lung aspiration.** *Arch Dis Child.* 1988; **63**(10): 1266–7.
PubMed Abstract | Publisher Full Text | Free Full Text

5. Scott JA, Hall AJ, Muyodi C, *et al.*: **Aetiology, outcome, and risk factors for mortality among adults with acute pneumonia in Kenya.** *Lancet.* 2000; **355**(9211): 1225–30.
PubMed Abstract | Publisher Full Text

6. Scott JA, Brooks WA, Peiris JS, *et al.*: **Pneumonia research to reduce childhood mortality in the developing world.** *J Clin Invest.* 2008; **118**(4): 1291–300.
PubMed Abstract | Publisher Full Text | Free Full Text

7. Rudan I, Boschi-Pinto C, Biloglav Z, *et al.*: **Epidemiology and etiology of childhood pneumonia.** *Bull World Health Organ.* 2008; **86**(5): 408–16.
PubMed Abstract | Publisher Full Text | Free Full Text

8. GBD 2015 LRI Collaborators: **Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the Global Burden of Disease Study 2015.** *Lancet Infect Dis.* 2017; **17**(11): 1133–61.
PubMed Abstract | Publisher Full Text | Free Full Text

9. O'Brien KL, Wolfson LJ, Watt JP, *et al.*: **Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates.** *Lancet.* 2009; **374**(9693): 893–902.
PubMed Abstract | Publisher Full Text

10. Castriota-Scanderbeg A, Popolizio T, Sacco M, *et al.*: **[Diagnosis of mycoplasma pneumonia in children: which is the role of thoracic radiography?].** *Radiol Med.* 1995; **89**(6): 782–6.
PubMed Abstract

11. Jain S, Williams DJ, Arnold SR, *et al.*: **Community-acquired pneumonia requiring hospitalization among U.S. children.** *N Engl J Med.* 2015; **372**(9): 835–45.
PubMed Abstract | Publisher Full Text | Free Full Text

12. Jain S, Self WH, Wunderink RG, *et al.*: **Community-Acquired Pneumonia Requiring Hospitalization among U.S. Adults.** *N Engl J Med.* 2015; **373**(5): 415–27.
PubMed Abstract | Publisher Full Text | Free Full Text

13. Levine OS, O'Brien KL, Deloria-Knoll M, *et al.*: **The Pneumonia Etiology Research for Child Health Project: a 21st century childhood pneumonia etiology study.** *Clin Infect Dis.* 2012; **54 Suppl 2**: S93–101.
PubMed Abstract | Publisher Full Text | Free Full Text

14. Connolly MA, Gayer M, Ryan MJ, *et al.*: **Communicable diseases in complex emergencies: impact and challenges.** *Lancet.* 2004; **364**(9449): 1974–83.
PubMed Abstract | Publisher Full Text

15. Blanchet K, Ramesh A, Frison S, *et al.*: **Evidence on public health interventions in humanitarian crises.** *Lancet.* 2017; **390**(10109): 2287–96.
PubMed Abstract | Publisher Full Text

16. Mejias A, Suarez NM, Ramilo O: **Detecting specific infections in children through host responses: a paradigm shift.** *Curr Opin Infect Dis.* 2014; **27**(3): 228–35.
PubMed Abstract | Publisher Full Text | Free Full Text

17. Zaas AK, Chen M, Varkey J, *et al.*: **Gene expression signatures diagnose influenza and other symptomatic respiratory viral infections in humans.** *Cell Host Microbe.* 2009; **6**(3): 207–17.
PubMed Abstract | Publisher Full Text | Free Full Text

18. Ramilo O, Mejías A: **Shifting the paradigm: host gene signatures for diagnosis of infectious diseases.** *Cell Host Microbe.* 2009; **6**(3): 199–200.
PubMed Abstract | Publisher Full Text

19. Hu X, Yu J, Crosby SD, *et al.*: **Gene expression profiles in febrile children with defined viral and bacterial infection.** *Proc Natl Acad Sci U S A.* 2013; **110**(31): 12792–7.
PubMed Abstract | Publisher Full Text | Free Full Text

20. Hanna-Wakim R, Chehab H, Mahfouz I, *et al.*: **Epidemiologic characteristics, serotypes, and antimicrobial susceptibilities of invasive *Streptococcus pneumoniae* isolates in a nationwide surveillance study in Lebanon.** *Vaccine.* 2012; **30 Suppl 6**: G11–7.
PubMed Abstract | Publisher Full Text

21. UNHCR: **Refugees from Syria: Lebanon**. 2015.
Reference Source

22. World Health Organization, UNICEF: **Integrated Management of Childhood Illness.** Geneva. 2014.
Reference Source

23. Plüddemann A, Thompson M, Heneghan C, *et al.*: **Pulse oximetry in primary care: primary care diagnostic technology update.** *Br J Gen Pract.* 2011; **61**(586): 358–9.
PubMed Abstract | Publisher Full Text | Free Full Text

24. Rambaud-Althaus C, Althaus F, Genton B, *et al.*: **Clinical features for diagnosis of pneumonia in children younger than 5 years: a systematic review and meta-analysis.** *Lancet Infect Dis.* 2015; **15**(4): 439–50.
PubMed Abstract | Publisher Full Text

25. Rambaud-Althaus C, Shao AF, Kahama-Maro J, *et al.*: **Managing the Sick Child in the Era of Declining Malaria Transmission: Development of ALMANACH, an Electronic Algorithm for Appropriate Use of Antimicrobials.** *PLoS One.* 2015; **10**(7): e0127674.
PubMed Abstract | Publisher Full Text | Free Full Text

26. Shao AF, Rambaud-Althaus C, Samaka J, *et al.*: **New Algorithm for Managing Childhood Illness Using Mobile Technology (ALMANACH): A Controlled Non-Inferiority Study on Clinical Outcome and Antibiotic Use in Tanzania.** *PLoS One.* 2015; **10**(7): e0132316.
PubMed Abstract | Publisher Full Text | Free Full Text

27. Wilhelmi I, Roman E, Sánchez-Fauquier A: **Viruses causing gastroenteritis.** *Clin Microbiol Infect.* 2003; **9**(4): 247–62.
PubMed Abstract | Publisher Full Text

28. Zhou HT, Yi HS, Guo YH, *et al.*: **Enterovirus-related diarrhoea in Guangdong, China: clinical features and implications in hand, foot and mouth disease and herpangina.** *BMC Infect Dis.* 2016; **16**: 128.
PubMed Abstract | Publisher Full Text | Free Full Text

29. Romero JR, Rotbart HA: **Enteroviruses.** In: Murray PR, Baron EJ, Jorgensen JH, Pfaller MA, Yolken RH, editors. *Manual of Clinical Microbiology.* 8th edition. Washington, DC: ASM Press; 2003; 1427–38.

30. World Health Organization: **Good clinical laboratory practice (GCLP).** Geneva: World Health Organization on behalf of the Special Programme for Research and Training in Tropical Diseases; 2009.
Reference Source

31. World Health Organization: **WHO guidelines for the collection of human specimens for laboratory diagnosis of avian influenza infection [Internet]**. 2005; Accessed: 2016-03-16. (Archived by WebCite® at http://www.webcitation.org/6g33epwmS).
Reference Source

32. Albrich WC, Madhi SA, Adrian PV, *et al.*: **Pneumococcal colonisation density: a new marker for disease severity in HIV-infected adults with pneumonia.** *BMJ Open.* 2014; **4**(8): e005953.
PubMed Abstract | Publisher Full Text | Free Full Text

33. Albrich WC, Madhi SA, Adrian PV, *et al.*: **Genomic load from sputum samples and nasopharyngeal swabs for diagnosis of pneumococcal pneumonia in HIV-infected adults.** *J Clin Microbiol.* 2014; **52**(12): 4224–9.
PubMed Abstract | Publisher Full Text | Free Full Text

34. Babady NE: **The FilmArray® respiratory panel: an automated, broadly multiplexed molecular test for the rapid and accurate detection of respiratory pathogens.** *Expert Rev Mol Diagn.* 2013; **13**(8): 779–88.
PubMed Abstract | Publisher Full Text

35. Messaoudi M, Milenkov M, Albrich WC, *et al.*: **The Relevance of a Novel Quantitative Assay to Detect up to 40 Major *Streptococcus pneumoniae* Serotypes Directly in Clinical Nasopharyngeal and Blood Specimens.** *PLoS One.* 2016; **11**(3): e0151428.
PubMed Abstract | Publisher Full Text | Free Full Text

36. Thomson RBJ, Miller JM: **Specimen Collection, Transport, and Processing: Bacteriology.** In: Murray PR, Baron EJ, Jorgensen JH, Pfaller MA, Yolken RH, editors. *Manual of Clinical Microbiology.* 8th edition. Washington, DC: ASM Press; 2003; 286–330.

37. Driscoll AJ, Bhat N, Karron RA, *et al.*: **Disk diffusion bioassays for the detection of antibiotic activity in body fluids: applications for the Pneumonia Etiology Research for Child Health project.** *Clin Infect Dis.* 2012; **54 Suppl 2**: S159–64.
PubMed Abstract | Publisher Full Text

38. Smith MD, Derrington P, Evans R, *et al.*: **Rapid diagnosis of bacteremic pneumococcal infections in adults by using the Binax NOW *Streptococcus pneumoniae* urinary antigen test: a prospective, controlled clinical evaluation.** *J Clin Microbiol.* 2003; **41**(7): 2810–3.
PubMed Abstract | Publisher Full Text | Free Full Text

39. Bénet T, Sánchez Picot V, Messaoudi M, *et al.*: **Microorganisms Associated With Pneumonia in Children < 5 Years of Age in Developing and Emerging Countries: The GABRIEL Pneumonia Multicenter, Prospective, Case-Control Study.** *Clin Infect Dis.* 2017; **65**(4): 604–12.
PubMed Abstract | Publisher Full Text

40. R Core Team: **R: A Language and Environment for Statistical Computing.** Vienna, Austria; 2016.

41. Mejias A, Dimo B, Suarez NM, *et al.*: **Whole blood gene expression profiles to assess pathogenesis and disease severity in infants with respiratory syncytial virus infection.** *PLoS Med.* 2013; **10**(11): e1001549.
PubMed Abstract | Publisher Full Text | Free Full Text

42. Suarez NM, Bunsow E, Falsey AR, *et al.*: **Superiority of transcriptional profiling over procalcitonin for distinguishing bacterial from viral lower respiratory tract infections in hospitalized adults.** *J Infect Dis.* 2015; **212**(2): 213–22.
PubMed Abstract | Publisher Full Text | Free Full Text

43. Wallihan RG, Suárez NM, Cohen DM, *et al.*: **Molecular Distance to Health Transcriptional Score and Disease Severity in Children Hospitalized With Community-Acquired Pneumonia.** *Front Cell Infect Microbiol.* 2018; **8**: 382.
PubMed Abstract | Publisher Full Text | Free Full Text

44. World Medical Association: **Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects.** Fortaleza; 2013.
Reference Source

45. Association des Epidémiologistes de Langue Française: **Recommandations de Déontologie et de Bonnes Pratiques en Epidémiologie.** Bordeaux; 2007.
Reference Source

46. **International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use**. Good Clinical Practice. 1997.

47. Hammitt LL, Feikin DR, Scott JAG, *et al.*: **Addressing the Analytic Challenges of Cross-Sectional Pediatric Pneumonia Etiology Data.** *Clin Infect Dis.* 2017; **64**(suppl_3): S197–204.
PubMed Abstract | Publisher Full Text | Free Full Text

48. Bozicevic I, Riedner G, Haghdoost A: **HIV case reporting in the countries of North Africa and the Middle East.** *J Int AIDS Soc.* 2014; **17**(1): 18962.
PubMed Abstract | Publisher Full Text | Free Full Text

49. Ismail MB, Rafei R, Dabboussi F, *et al.*: **Tuberculosis, war, and refugees: Spotlight on the Syrian humanitarian crisis**. Leong JM, editor. *PLOS Pathog.* 2018; **14**(6): e1007014.
PubMed Abstract | Publisher Full Text | Free Full Text

50. Bellos A, Mulholland K, O'Brien KL, *et al.*: **The burden of acute respiratory infections in crisis-affected populations: a systematic review.** *Confl Health.* 2010; **4**: 3.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Peer Review Status: ✔ ? ✔

Version 2

Reviewer Report 13 October 2020

https://doi.org/10.21956/gatesopenres.14036.r29789

✔ **Carina King**

Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden

This presents a clear study protocol, aiming to answer an important (but challenging!) question - of LRTI aetiology in a humanitarian setting.

Background:

- Can you also provide information on vaccine coverage (if it can be disaggregated?), and influenza vaccine policies?
- Indicate what formulation of PCV.

Protocol:

- Why was chest indrawing only considered for children <=3 years?
- Can you provide a bit more information on whether pulse oximetry was new in this setting, and if so, what training was done and quality control?
- It wasn't clear whether you intend to use the averaged RR over two counts, or just one. And whether this will be done for all age groups.
- Will any analyses of multi-pathogen versus single pathogen presentation be done? Thinking of the PERCH findings, where this was an important result?

Is the rationale for, and objectives of, the study clearly described?

Yes

Is the study design appropriate for the research question?

Yes

Are sufficient details of the methods provided to allow replication by others?

Yes

Are the datasets clearly presented in a useable and accessible format?
Not applicable

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Infectious epidemiology, paediatric pneumonia

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 14 January 2020

https://doi.org/10.21956/gatesopenres.14036.r27382

Anne-Laure Page
Epicentre, Paris, France

I have reviewed the response from the authors and believe that the revisions are appropriate.

Is the rationale for, and objectives of, the study clearly described?
Not applicable

Is the study design appropriate for the research question?
Not applicable

Are sufficient details of the methods provided to allow replication by others?
Not applicable

Are the datasets clearly presented in a useable and accessible format?
Not applicable

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 26 November 2018

https://doi.org/10.21956/gatesopenres.13878.r26432

?

Stephen R. Howie
Department of Paediatrics: Child and Youth Health, University of Auckland, Auckland, New Zealand

Thank you for the opportunity to review this study protocol. This protocol describes a multicentre case control study of pneumonia etiology in refugee populations in Lebanon. This is a population at high risk of pneumonia and data on etiology in this group will be welcome. The study has already completed recruitment, and is undergoing analysis. It is no simple task undertaking such a study, and I look forward to seeing its findings.

I have a number of comments and suggestions. The protocol is clear in some of its aspects and unclear in others: I commend the authors for the former and I will focus my comments on the latter, which are points to be addressed.

The authors rightly stress the limitations of 'blindly' applying etiology data from other studies to this population. They should be similarly cautious about applying their findings to those suffering in the midst of humanitarian crisis elsewhere, and even to those in the study settings at other times. Fifteen months is relatively short, and refugee contexts are unstable and prone to epidemics. What happens in that 15 month period may not be what is happening the next year. A case-control study such as this certainly has value, and longitudinal surveillance would add more value. The relationship between existing pneumococcal surveillance systems in Lebanon and this study is not clear from the protocol; adding broader pathogen surveillance to those systems has much to recommend it. I note too the lack of any data about TB in the study protocol, a gap it would be good to fill. Are there plans to address this in the future?

The protocol makes much of the benefits of transcriptomic analysis (including that it 'reduces the need to identify pathogens by culturing respiratory tract samples') but is relatively silent on how this will be undertaken and how data will be used to achieve the primary aims. Biomarkers for pneumonia etiology (whether they be transcriptomic or other), while attractive conceptually, have many limitations, and their practical applicability at this point need careful scrutiny. I wonder whether the transcriptomic aspect of this study is an exploratory add-on (not in itself an invalid thing to do). I see no mention of power calculations with respect to the transcriptomic element, for instance, or how it will be integrated into the analysis as a whole.

The power calculation description is restricted to the primary objective. It would be helpful to show that there is adequate power for all objectives.

The analytic methods planned for the primary objective are fairly restricted. I would encourage the authors to learn from the approaches taken in studies such as PERCH and get the most they

can out of the data. This can still be done even though recruitment has been completed.

There is no mention of using a point-of-care testing approach except at the end of the objectives and I wonder what was meant by that.

The case definition used was not explicitly linked to the WHO definition, but reference was made to using WHO categorisation of severity. Clarifying the relationship between study and WHO definitions would be helpful.

What does the protocol mean when it says that inclusion requires “no wheezing suggestive of asthma at auscultation”? What sort of wheezing is that? Does ‘suggestive’ mean age, history, trial of bronchodilator, or something else? What does the protocol mean when it specifies “particular attention to pneumonia with a breathing rate > 50 breaths per minute’? Was this a de facto entry criterion?

What does the protocol mean when it says that ‘relevant’ results will be communicated to clinicians. Which are these? Did you share NP or sputum PCR results and if so how did you suggest that clinicians interpret them?

The protocol appears to state that 6ml of blood will be taken for blood culture. Even in infants? The sample collection protocols requires clarification here.

Is the rationale for, and objectives of, the study clearly described?
Yes

Is the study design appropriate for the research question?
Partly

Are sufficient details of the methods provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Not applicable

Competing Interests: Dr. Howie has a patent 'Lipocalin-2 as a Biomarker for Pneumococcal Infection Status' pending, which is not of specific relevance to this article but of broad relevance to the subject of pneumonia aetiology. No other competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 01 Mar 2019

Thomas Kesteman, Fondation Mérieux, Lyon, France

We would like to express thanks the reviewer for his useful comments and suggestions. We

carefully reviewed his report and we changed the manuscript according to his propositions. Our answers to his questions and comments are listed in the text below.

Comment: The authors rightly stress the limitations of 'blindly' applying etiology data from other studies to this population. They should be similarly cautious about applying their findings to those suffering in the midst of humanitarian crisis elsewhere, and even to those in the study settings at other times.
Action/Answer: We added a sentence in the paragraph regarding the limitations of the study (*Discussion* chapter): "although the situation of Syrian refugees in Lebanon shares a lot of characteristics with other crises, one should be cautious about applying the results of PEARL to other humanitarian crises, because of the variations of the epidemiology of LRTI."

Comment: Fifteen months is relatively short, and refugee contexts are unstable and prone to epidemics. What happens in that 15 month period may not be what is happening the next year. A case-control study such as this certainly has value, and longitudinal surveillance would add more value. The relationship between existing pneumococcal surveillance systems in Lebanon and this study is not clear from the protocol; adding broader pathogen surveillance to those systems has much to recommend it.
Action/Answer: We added a sentence in the *Discussion,* just before the one above: "Finally, the external validity of our study is limited in time (15 months) and space (four health facilities). A case-control study cannot replace longitudinal surveillance systems, as are currently being set up in Lebanon, e.g. the Lebanese Inter-Hospital Pneumococcal Surveillance Program.".

Comment: I note too the lack of any data about TB in the study protocol, a gap it would be good to fill. Are there plans to address this in the future?
Action/Answer: This was acknowledged as a limitation of our study in the first version of the manuscript: "certain aspects of CAP will not be explored in this study [...]. Patients with tuberculosis are expected to be excluded, since duration of symptoms >14 days is an exclusion criterion." Nevertheless, we added the following sentence: "The incidence rate of tuberculosis among Syrian refugees did not increase in the last couple of years and remains relatively low at close to pre-war levels, i.e. 10-20/100.000 population. Therefore, we consider it unlikely that tuberculosis is a major cause of acute LRTI in our population." and we provided references for this assertion.

Comment: The protocol makes much of the benefits of transcriptomic analysis (including that it 'reduces the need to identify pathogens by culturing respiratory tract samples') but is relatively silent on how this will be undertaken and how data will be used to achieve the primary aims. Biomarkers for pneumonia etiology (whether they be transcriptomic or other), while attractive conceptually, have many limitations, and their practical applicability at this point need careful scrutiny. I wonder whether the transcriptomic aspect of this study is an exploratory add-on (not in itself an invalid thing to do). I see no mention of power calculations with respect to the transcriptomic element, for instance, or how it will be integrated into the analysis as a whole.
Action/Answer: The reviewer is right: transcriptomics is an add-on, as we have clearly stated that "Patient enrollment began in November 2016" and that "Sampling for transcriptomics was initiated in May 2017" (i.e. 7 months after the initiation of the study).

Therefore, the overall study has not been powered to meet this secondary objective.
In order to clarify these points, we added the following sentence in the chapter related to Data management and analysis: "Etiologic fractions will be also computed from transcriptomics data, independently from calculations from microbiological data, and compared with each other."

Comment: The power calculation description is restricted to the primary objective. It would be helpful to show that there is adequate power for all objectives.
Action/Answer: The study has been powered for the primary objective only, and the uncertainty around the values corresponding to secondary objectives will be provided by 95% confidence intervals. We added the following sentence in *Data management and analysis*: "Ninety-five-percent confidence intervals will be calculated, with regards to the outcomes linked to the primary and secondary objectives."

Comment: The analytic methods planned for the primary objective are fairly restricted. I would encourage the authors to learn from the approaches taken in studies such as PERCH and get the most they can out of the data. This can still be done even though recruitment has been completed.
Action/Answer: Methodologies such as partially latent class models (if this is what Dr Howie refers to) have been considered for PEARL study. Nevertheless, they can poorly be applied to our dataset given the restricted range of samples in our study: nasopharyngeal swab mainly, some sputum, and blood cultures -that are expected to have a low yield in primary health care and will therefore be poorly informative. In comparison with PERCH, where information from a much wider range of biological samples taken from severe cases was used, our dataset would be poorly suitable for complex models. Another drawback of PLC models is that they appear as a 'blackbox' while Population Attributable Fraction, resulting from Generalized Linear Models, are much more straightforward and comprehensible.

Comment: There is no mention of using a point-of-care testing approach except at the end of the objectives and I wonder what was meant by that.
Action/Answer: Actually, the PoC tests will be performed in laboratories located nearby but not in the institution where the patients will be included. The results are therefore not become available real time and cannot be considered true PoC tests.

Comment: The case definition used was not explicitly linked to the WHO definition, but reference was made to using WHO categorisation of severity. Clarifying the relationship between study and WHO definitions would be helpful.
Action/Answer: Exact. We added the following sentence in *Inclusion criteria*: "The case definition of CAP was based on the WHO's IMCI. "

Comment: What does the protocol mean when it says that inclusion requires "no wheezing suggestive of asthma at auscultation"? What sort of wheezing is that? Does 'suggestive' mean age, history, trial of bronchodilator, or something else?
Action/Answer: The reviewer is correct. This is indeed not clear. We added the following sentence in *Inclusion criteria*: "Characteristics of wheezing suggestive of asthma included: expiratory wheezing, high pitched wheezing, wheezing with history of asthma and without

fever nor history of fever, and any clinical picture such as the clinician would exclude LRTI and retain asthma as final diagnostic."

Comment: What does the protocol mean when it specifies "particular attention to pneumonia with a breathing rate > 50 breaths per minute'? Was this a de facto entry criterion?
Action/Answer: No, it just means that we measured the breathing rate twice in children, in order to evaluate its pertinence in our own dataset. In the manuscript, we therefore changed the sentence as follows: "Moreover, in children under five years of age, we will assess the breathing rate twice, since it has been shown that the cut-off of > 50 breaths per minute, assessed twice, is more specific for pneumonia than the standard IMCI recommendations for children 1–5 years old (40 breaths/min), and was successfully adopted in the ALMANACH protocol."

Comment: What does the protocol mean when it says that 'relevant' results will be communicated to clinicians. Which are these? Did you share NP or sputum PCR results and if so how did you suggest that clinicians interpret them?
Action/Answer: The criteria for feedback of individual lab data, as specified in the protocol shared with laboratories, included:

- Identification of clinically relevant bacteria in blood (e.g. no bacteria associated to skin contamination, like coagulase-negative *Staphylococci*), by conventional microbiology or molecular tools. Information of identification and AST results will both be reported to the clinician by phone as soon as they will become available.
- Identification of clinically relevant bacteria in nasopharyngeal swab or sputum of CAP cases, by conventional microbiology or molecular tools. This includes detection of atypical bacteria (*Chlamydia pneumoniae* or *Mycoplasma pneumoniae*) by BioFire. Information related to the identification and AST results will be reported by phone to the clinician as soon as they become available.
- Identification of RSV in infants (cases or controls) aged ≤12 months.
- Positive Binax test in urine of case.

This was added to the manuscript, in the chapter *Analysis of biological samples*.

Comment: The protocol appears to state that 6ml of blood will be taken for blood culture. Even in infants? The sample collection protocols requires clarification here.
Action/Answer: No, in infants and small children, the volume of blood taken was lower. The following sentence has been added to clarify this: "The target blood volume for blood culture is 1-3 mL in under-fives and 6 mL in children aged five or above and in adults, following manufacturers' recommendations."

Competing Interests: No competing interests were disclosed.

Reviewer Report 28 September 2018

https://doi.org/10.21956/gatesopenres.13878.r26639

?

Anne-Laure Page
Epicentre, Paris, France

This manuscript describes the study protocol for a study on the etiology of pneumonia among refugees and local population in Lebanon. The study, which has already completed all inclusions, was designed to investigate the population-attributable fractions (PAF) of specific viral and bacterial pathogens by comparing the proportion of pathogens detected by culture, PCR and rapid tests among pneumonia cases and age-matched controls. In addition, it aims at describing the pathogens isolates (antimicrobial susceptibility and serotypes if applicable), identifying risk factors for pneumonia and providing transcriptomics data to identify expression patterns associated with specific pathogens. The manuscript is clear and well written and describes the study adequately, except for some areas, which would require additional information:

1. Describe the recommended volumes of blood collection for blood culture according to the age
2. The study sites are primary healthcare facility (no hospital facility), but some of the inclusion criteria (in particular chest indrawing in children < 3 years of age) would require immediate referral to hospital. Please describe how inclusions are performed for these patients.
3. Describe the rationale for having gastrointestinal infections as exclusion criteria for controls.
4. The sample size calculation is stratified by age group (150 patients for each of four age groups), leading to a final sample size of minimum 600 cases and maximum 900 cases. It would be helpful to know what were the expected proportions of patients in each age group to understand the final sample size calculation. The authors could comment on the fact that finally 1420 patients were included, which is not consistent with the initial sample size calculated.
5. More information could be provided on the culture of bacterial from respiratory samples (sputum samples and nasopharyngeal swabs). Is quality of the sputum and absence of saliva contamination assessed? Which selective media are used?
6. Please provide more information about transcriptomics analysis (at least a reference).
7. The references cited for the semi-quantitative multiplex real-time PCR assay to detect *S. pneumoniae*, *S. aureus* and *H. influenzae* B do not provide sufficient information on the method. Authors should also describe more in details how this assay is performed on EDTA whole blood (volume of blood used for extraction and in the PCR assay). Interpretation of the PCR results from blood, sputum and nasopharyngeal swabs should also be described, in particular if there is a threshold (in Ct or copies/mL) used to consider the result as positive.
8. It is unclear whether the additional indicators described in the study outcomes would come from the study results or from other sources. Some, like clinical signs of CAP, can be analysed from the study data, whereas others, while for others, like agent-specific hospitalization and fatality rates, the corresponding data is not collected in the CRF.
9. The primary objective states that the PAFs will be estimated “using a combination of conventional methods and transcriptomics”, but it is unclear in the analysis plan if or how the transcriptomics data will be used to determine the PAFs and how these transcriptomics data will be analysed.

Is the rationale for, and objectives of, the study clearly described?
Yes

Is the study design appropriate for the research question?
Yes

Are sufficient details of the methods provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Not applicable

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 01 Mar 2019

Thomas Kesteman, Fondation Mérieux, Lyon, France

We would like to express thanks the reviewer for her useful comments and suggestions. We carefully reviewed her report and we changed the manuscript according to her propositions. These changes included adding the complete methodology of a triplex PCR that had not been disclosed in a previous scientific article. Therefore, we added M Stéphane Pouzol, who designed and developed that technique, as an author of the present manuscript. Our answers to her questions and comments are listed in the text below.

Comment: Describe the recommended volumes of blood collection for blood culture according to the age.
Action/Answer: The following sentence has been added to clarify this: "The total volume of blood taken is 0.5-1.0 mL in controls of all ages and, in cases, 2.5-5.0 mL in infants (<1 year old), 3.5-6.0 mL in children aged one to four, 10.5-17.0 and 6 mL in children aged five or above, and 18.5-19.0 mL in adults (≥18 years old), with a flexibility in the age cut-off depending on the individual's weight and overall health status."

Comment: The study sites are primary healthcare facility (no hospital facility), but some of the inclusion criteria (in particular chest indrawing in children < 3 years of age) would require immediate referral to hospital. Please describe how inclusions are performed for these patients.
Action/Answer: We added the following in *Study setting and design*: "After enrolment in the study, patients will be managed as per local guidelines and doctors' recommendations; in particular, the study won't alter procedures for referral for complementary examinations or hospitalization."

Comment: Describe the rationale for having gastrointestinal infections as exclusion criteria for controls.
Action/Answer: We added the following in *Inclusion criteria - Controls* and provided a reference: "Individuals with gastro-intestinal symptoms were excluded as some viruses (adenovirus, enterovirus, and coronavirus) are proven or suspected to cause either respiratory or gastro-intestinal symptoms, and including such patients as controls would result in an underestimation of the attributable fraction of these pathogens."

Comment: The sample size calculation is stratified by age group (150 patients for each of four age groups), leading to a final sample size of minimum 600 cases and maximum 900 cases. It would be helpful to know what were the expected proportions of patients in each age group to understand the final sample size calculation.
Action/Answer: At the moment of protocol design, there were no data available regarding incidence of lower respiratory tract infection in the different age groups. We assume that will would include an approximately equal number of cases in each of the four age groups. Hence, 150*4 = 600.

Comment: The authors could comment on the fact that finally 1420 patients were included, which is not consistent with the initial sample size calculated.
Action/Answer: Actually, this is a result of the study, and that should not have appeared in the protocol. Therefore, we removed that sentence. It will be discussed in the manuscript presenting the results of the study.

Comment: More information could be provided on the culture of bacterial from respiratory samples (sputum samples and nasopharyngeal swabs). Is quality of the sputum and absence of saliva contamination assessed? Which selective media are used?
Action/Answer: We agree with the reviewer's comment. We modified the manuscript as follows: "Sputum samples and STGG media will be inoculated onto different selective agar plates according to standard laboratory procedures (at least one blood culture medium, one chocolate agar, and one medium selective for gram-negative bacilli) and incubated under specific conditions (37°C, aerobic atmosphere supplemented with CO_2) to determine the presence of respiratory pathogens and assess antibiotic susceptibility. Quality of sputum and absence of contamination by saliva will be assessed by standard laboratory procedures."

Comment: Please provide more information about transcriptomics analysis (at least a reference).
Action/Answer: Done. The *Transcriptomics* paragraph of *Analysis of biological samples* chapter and the *Data management and analysis* chapter have been substantially supplemented to include all procedures related to transcriptomics, and we added references to articles where those methods have been used.

Comment: The references cited for the semi-quantitative multiplex real-time PCR assay to detect S. pneumoniae, S. aureus and H. influenzae B do not provide sufficient information on the method.
Action/Answer: We amended references and added the information required for performing the PCR in the *Molecular biology testing of respiratory samples* paragraph of

Analysis of biological samples chapter: sequences of primers, probes, reagents, and temperature cycle.

Comment: Authors should also describe more in details how this assay is performed on EDTA whole blood (volume of blood used for extraction and in the PCR assay).
Action/Answer: Done. The *Detection of bacteria in blood* paragraph of *Analysis of biological samples* chapter has been modified as follows: "Whole blood (200 µL EDTA) samples from all cases will be extracted using QIAamp DNA blood mini kit (Qiagen, Hilden, Germany), and the same semi-quantitative multiplex Real-Time PCR assay as for respiratory samples will be conducted on 5 µL of extracted DNA to identify *Streptococcus pneumoniae, Staphylococcus aureus* and *Haemophilus influenzae* B."

Comment: Interpretation of the PCR results from blood, sputum and nasopharyngeal swabs should also be described, in particular if there is a threshold (in Ct or copies/mL) used to consider the result as positive.
Action/Answer: Done. The following sentence has been added in *Molecular biology testing of respiratory samples* paragraph of *Analysis of biological samples* chapter: "Amplification curves will be examined individually by two independent technicians, at the local laboratory and at the Laboratoire des Pathogènes Emergents (Lyon, France). Cycle thresholds (Ct) value will be manually set so that it intersects the exponential curve at its inflexion point. Any exponential signal observed between 0 and 40 Ct value will be considered as positive."

Comment: It is unclear whether the additional indicators described in the study outcomes would come from the study results or from other sources. Some, like clinical signs of CAP, can be analysed from the study data, whereas others, while for others, like agent-specific hospitalization and fatality rates, the corresponding data is not collected in the CRF.
Action/Answer: Agree. The following sentence has been added in *Other data collected*: "A subgroup of patients or parents/legal guardians, randomly selected, was called back at least one month after inclusion to re-evaluate the outcome of included cases (hospitalization, death)."

Comment: The primary objective states that the PAFs will be estimated "using a combination of conventional methods and transcriptomics", but it is unclear in the analysis plan if or how the transcriptomics data will be used to determine the PAFs and how these transcriptomics data will be analysed.
Action/Answer: Agree. This is indeed misleading. We removed this part of the sentence in the primary objective and added the following secondary objective: "compare microbiological and transcriptomic methods in estimating viral vs bacterial attributable fractions of LRTI".

Competing Interests: No competing interests were disclosed.