Multi-centre EuRopean study of MAjor Infectious Disease Syndromes

Acute Respiratory Infections in Adults

Protocol Version 6.0

15-May-2020

MERMAIDS

PREPARE is funded by the European Commission under grant number 602525

DOCUMENT CONTROL

Document information

	Information
Document Owner	Peter Horby
Issue Date	11-May-2016
File Name	MERMAIDS_Protocol_ARI_V4.docx

Document History

Version	Issue Date	Changes
1.2	7-Aug-2015	Final approved version
2.0	20-Nov-2015	• Minor changes in wording to improve clarity • Changed primary care Day2 visit window to 40-56 hours • Changed Day28 visit window to -3/+7 days • Added assessment of febrile illness at Day28 • All participants to receive gift vouchers at Day28 • Participants will be asked to consent to remain in the study if they lose capacity. Proxy assent will also be sought in these cases For full details see: MERMAIDS_Protocol_ARI_V2.0_summary_of_changes.docx
3.0	11-Mar-2016	• Inclusion criteria: added 'dyspnoea' as an eligible respiratory symptom for both primary and secondary care • Exclusion criteria: Patient is enrolled in an interventional clinical study *of an anti-infective or immunomodulatory therapy* • Delay of Day 2 blood sample collection up to Day 4 will be permissible. • Clarified that participants may be followed up in a hospital, GP or outpatient clinic, at home or at work For full details see: MERMAIDS_Protocol_ARI_V3.0_summary_of_changes.docx
4.0	11-May-2016	• Day 28 visit window extended to -3/+21 days • Added exclusion criterion: previous enrolment in MERMAIDS-ARI
5.0	06-Apr-2020	• Removal of exclusion criteria: "Patient is enrolled in an interventional clinical study of an anti-infective or immunomodulatory therapy" • Addition of section 7.2.3 Study consenting procedures during the SARS-CoV-2 pandemic
6.0	15-May-2020	• Increased duration of symptom onset to admission from 7 days to 14 days

Document Approvals

Role	Name©	Institution	Signature	Date
Coordinating Investigator	Peter Horby	University of Oxford		
Co-Principal Investigator	Menno de Jong	Academic Medical Centre, Amsterdam		
Statistician	Sébastien Matamoros	Academic Medical Centre, Amsterdam		

Study Title:

Multi-centre EuRopean study of MAjor Infectious Disease Syndromes (MERMAIDS) – Acute Respiratory Infections

Short title: MERMAIDS - ARI

Ethics Ref: 15/WM/0254

Study ID: ISRCTN18034878

Coordinating Principal Investigator:	Professor Peter Horby Nuffield Department of Clinical Medicine Centre for Tropical Medicine **Postal address:** Nuffield Department of Medicine Research Building University of Oxford, Old Road Campus Roosevelt Drive Oxford, OX3 7FZ, UK
Co-Principal Investigator	Professor Menno de Jong, Academic Medical Centre, Amsterdam
Investigators:	Professor Marion Koopmans, Erasmus Medical Centre Professor Peter Openshaw, Imperial College London Professor Chris Butler, University of Oxford Professor Theo Verheij, University Medical Centre Utrecht Dr Louise Sigfrid, University of Oxford Cosimo Cristella, Academic Medical Centre, Amsterdam Dr Frank van Someren Gréve, Academic Medical Centre, Amsterdam Brendan Scicluna, Academic Medical Centre, Amsterdam Professor Greet Ieven, University of Antwerp Professor Herman Goossens, University of Antwerp James Lee, University of Oxford Ushma Galal, University of Oxford
Sponsor:	University of Oxford Joint Research Office 1st floor, Boundary Brook House Oxford, OX4 7LE
Funder:	European Commission, FP7 & H2020
Chief Investigator Signature:	

Confidentiality Statement

This document contains confidential information that must not be disclosed to anyone other than the Sponsor, the Investigator Team, host organisation, and members of the Research Ethics Committee, unless authorised to do so.

TABLE OF CONTENTS

1. SYNOPSIS

Study Title	Multi-centre EuRopean study of MAjor Infectious Disease Syndromes (MERMAIDS): Acute Respiratory Infections in Adults
Ref. no. / short title	MERMAIDS – ARI
Study Design	Prospective, observational study
Study Participants	Adults (≥ 18 years old) with a suspected community acquired acute respiratory tract infection, Group 1: attending primary care Group 2: admitted to hospital
Planned Sample Size	Primary Care patients (Group 1): n = 750 Hospitalised Patients (Group 2): n = 2,250
Planned Study Period	60 months

	Objectives	**Outcome measures**
Primary objective	Identify host and pathogen related determinants of severity of community acquired acute respiratory infections (ARI) in adults.	Differentially expressed host genes (≥ 2-fold difference in expression levels) as assessed by RNA transcriptome microarray in hospitalised and primary care managed cases of ARI, stratified by pathogen* and comorbidity**. *Influenza virus, Human Rhinovirus, Respiratory Syncytial Virus, *Streptococcus pneumoniae*; **No comorbidity, chronic pulmonary disease, chronic cardiovascular disease, chronic metabolic disease (diabetes)
Secondary objectives	Describe the aetiology, clinical management and outcomes of adult patients with community acquired ARI, in both primary care and hospital care, across Europe.	<u>In both groups:</u> • Prevalence of detection of putative pathogens in respiratory tract samples • Proportion of cases receiving antibiotics, antivirals, antifungals and/or immunomodulators. • 28 day mortality <u>Additional in group 1 (primary care):</u> • Proportion of cases requiring hospitalisation

	Objectives	Outcome measures
		Additional in group 2 (hospitalised patients): • Severity of illness at enrolment as assessed by Pneumonia Severity Index (PSI) and CURB-65 • Proportion of cases requiring during admission: supplemental oxygen; non-invasive or invasive mechanical ventilation; extra-corporeal life support. • Duration of invasive mechanical ventilation and extra-corporeal life support, if applicable • Proportion of cases requiring Intensive Care Unit (ICU)/High Care Unit (HCU) admission • Hospital - and ICU/HCU length of stay • In-hospital mortality
Tertiary Objectives	To develop and validate prognostic and diagnostic algorithms.	This is an exploratory objective. Some measures that could be included in algorithms are classifier gene sets based on host gene expression profiles, pathogen profiles, demographics, co-morbidities, risk factors, and clinical parameters. The diagnostic/prognostic algorithms will be evaluated for their ability to correctly diagnose infecting pathogen and/or predict adverse outcome in ARI.
	To gain understanding into pathophysiological mechanisms contributing towards development of severe disease.	This is an exploratory objective to increase pathophysiological insights by integrative (systems medicine) analyses of pathogen- and patient characteristics. Measures that will be included in these integrative analyses are host gene expression profiles, clinical data on disease progression/outcome, deep sequencing of pathogen genomes and microbiomes.

2. ABBREVIATIONS

ARI	Acute respiratory infection
CI	Chief Investigator
CRF	Case Record Form
CRO	Contract Research Organisation
CURB-65	Severity Score (Confusion, Urea, Respiratory rate, Blood pressure, Age >65y)
DAVID	Database for Annotation, Visualization, and Integrated Discovery
DM	Data Management
eCRF	Electronic Case Record Form
EDC	Electronic data capture
EDTA	Ethylenediaminetetraacetic acid
FDR	False Discovery Rate
GCP	Good Clinical Practice
GO	Gene Ontologies
GSEA	Gene Set Enrichment Analysis
HRV	Human Rhinovirus
ICH	International Conference of Harmonisation
JC	Julius Centre
MAQC	MicroArray Quality Control
MERMAIDS	Multi-centre EuRopean study of MAjor Infectious Disease Syndromes
NHS	National Health Service
NRES	National Research Ethics Service
PCR	Polymerase Chain Reaction
PI	Principal Investigator
PIL	Participant/ Patient Information Leaflet
PREPARE	Platform for European Preparedness Against (Re-)emerging Epidemics
PSI	Pneumonia Severity Index / Fine Score
PSP	Project Specific Procedures
REC	Research Ethics Committee
RO	Research Online
R&D	NHS Trust R&D Department
REC	Research Ethics Committee
RSV	Respiratory Syncytial Virus

SOG	Study Operations Group
SOP	Standard Operating Procedure
SSC	Study Steering Committee
SST	Serum-separating Tube
WHO	World Health Organisation
WI	Work Instructions
WP3	Work Package 3

3. BACKGROUND AND RATIONALE

3.1 Background

PREPARE (Platform for European Preparedness Against (Re-)emerging Epidemics) is a European Commission funded network for harmonised large-scale clinical research studies on infectious diseases. The long-term aim is to establish a network that is prepared to rapidly respond to any severe infectious disease outbreak, providing real-time evidence for clinical management of patients and for informing public health responses. For more information on PREPARE please visit our website (www.prepare-europe.eu).

This study is part of the Multi-centre EuRopean study of MAjor Infectious Disease Syndromes (MERMAIDS), which comprises work package 3 (WP3) of the PREPARE platform.

3.2 Current limitations in understanding the pathogenesis of acute respiratory infections

Based on historic and recent pandemics or pandemic threats, as well as on knowledge of transmissibility and epidemiology, emerging or (re-)emerging pathogens causing acute respiratory infections (ARI) are considered the most likely candidates to cause the next pandemic.

There is a substantial body of knowledge on clinical risk factors for disease severity and outcome of ARIs. Validated clinical risk stratification tools, such as the Pneumonia Severity Index (PSI) (1) and CURB-65 score (2) predict outcomes and help to guide clinical decision making. Patient groups at risk for developing severe disease are well known, such as the elderly, patients with chronic pulmonary, cardiovascular or metabolic disease or immunocompromised patients. In national guidelines for prevention and treatment of ARI these risk groups are often combined.

However, the underlying pathophysiologic processes that determine the severity of ARI across the diversity of risk groups and pathogens likely differ and are not well understood. This means that within broad phenotypic 'risk groups' it is currently not possible to predict who is at increased risk of becoming severely ill. Consequently, there are no opportunities to tailor preventive and therapeutic interventions to individual risk. Personalised or precision medicine is becoming a reality in some areas of oncology, but has yet to enter infectious diseases (3). In patients that do become moderately or severely ill, there is an assumption that the underlying pathophysiological processes are the same, hence all patients will benefit equally from the same intervention, such as antivirals and immunomodulators. Increased insight into differences in underlying pathophysiological processes across clinical backgrounds and pathogens will open avenues for the development of predictive algorithms that anticipate severity at individual patient levels and direct strategies for individualized therapeutic interventions, hence improving clinical outcomes. Additionally, there may be

commonalities in host- and pathogen-related factors that determine outcome which could be used in generic diagnostic (e.g. distinguishing viral versus bacterial infection(4,5)) and therapeutic approaches. This could be of relevance to the next emerging novel respiratory pathogen, hence contributing to clinical preparedness for a new epidemic threat.

Analysis of the host gene expression profile (transcriptome) provides a high-resolution insight into host responses and pathophysiology, providing opportunities to understand the specific contributions of risk factors, comorbidities and pathogen-specific traits. As an example, human challenge studies using influenza virus, respiratory syncytial virus (RSV) and human rhinovirus (HRV) have indicated that peripheral blood gene expression profiles distinguish between specific viral aetiologies (6), and symptomatic and asymptomatic influenza virus infection (7,8), thus providing valuable insight into the development of clinical disease as well as avenues towards host-targeted approaches of aetiological diagnostics.

3.3 Rationale for study design

While experimental research remains essential to provide the foundations of our understanding of disease pathogenesis, detailed patient-oriented studies across the spectrum of clinical disease and aetiologies are needed to understand the complexities of 'real-life' pathogenesis of human infection, including the combined complex influences of age, comorbidities and co-infections with other pathogens in development of severe disease. This approach requires adequately sized and well defined cohorts with regards to age, presence of comorbidities and microbial aetiologies. To date, such patient-oriented research in ARI has been limited to relatively small and heterogeneous groups of patients, in which data from patients of different ages with a diverse spectrum of comorbidities were usually combined (9–11). This makes identification of clinical background- or pathogen-specific pathophysiological pathways problematic if not impossible. In addition, most studies to date have focused on the severe end of the clinical spectrum (11–13). Although there are many genes that have increased expression regardless of the pathogen (14,15) differences are generally observed if the samples are taken early enough in the disease (16). There is also limited data on primary care patients and outpatients. Finally, many studies have investigated a single time point (11,13,17–21); analysing more time points is required to fully understand the dynamic nature of the host transcriptional profile (7).

Therefore, given the heterogeneity of clinical backgrounds and causative pathogens, the generation of meaningful insights into the relative contributions of host and pathogen factors to severe disease will require adequately sized cohorts of ARI patients across disease severities. The aim of the MERMAIDS ARI study is to establish such a cohort by enrolling sufficient numbers of patients across the most prevalent comorbidities (chronic pulmonary/cardiovascular/metabolic disease) and most prevalent pathogens (Influenza A, HRV, RSV and *Streptococcus pneumoniae*) with mild (community care) or moderately severe (hospitalised) ARI to have adequate statistical power to analyse differences in host response between these predefined subgroups. A small subset of ARI patients will progress to severe disease, as defined by the need for mechanical ventilation and/or death.

3.4 Study Aims

3.4.1 Determinants of ARI disease severity

The primary aim of this study is to identify host and pathogen related determinants of disease severity of ARI. In this study we will include a prospectively defined and sufficiently sized patient cohort that will allow investigations of relative contributions of host and pathogen factors in development of mild, moderate and severe ARI. The sample size of this study is determined to provide adequate power to identify differences in RNA expression profiles between groups of patients with ARI stratified into groups by major categories of comorbidity (chronic pulmonary/cardiovascular/metabolic disease), and

the major causative pathogens (Influenza A, HRV, RSV and *Streptococcus pneumoniae*). Setting a high bar with regards to sample size, by basing our power calculation on the ability to detect differentially expressed genes between these predefined groups, will ensure the groups will be large enough to also compare less complex determinants of ARI severity, e.g. local and systemic innate and adaptive immune markers, and pathogen markers (e.g. pathogen load).

Our analysis will identify differentially expressed genes between patients with different comorbidities within and across ARI aetiologies and severities by combining a fold change assessment (>2-fold nominal change expression levels) combined with a $p<0.05$ correcting for multiple comparisons using the Benjamini-Hochberg False Discovery Rate (FDR) < 0.05 in accordance with guidelines from the MicroArray Quality Control (MAQC) group (22).

3.5 Aetiology, clinical management and outcomes of ARI across Europe

This study will be the largest comprehensive study to investigate and compare aetiology, pneumonia severity scores, clinical management and outcomes of adult ARI across Europe, in both primary and hospital care.

3.6 Diagnostic and prognostics markers

Analysis of transcriptional profiles will give us unique insight into determinants of disease severity, and it may also provide diagnostic and prognostic biomarkers. These may include classifier gene sets for aetiological diagnosis and prognosis, as well as composite algorithms using genetic markers, cytokine profiles, demographics, clinical backgrounds, and clinical parameters. Furthermore, we will be able to validate biomarkers and biomarker profiles that have already been determined in other studies, such as cytokine profiles from the MOSAIC study in patients with severe influenza infection (http://www1.imperial.ac.uk/mosaic/about/).

3.7 Pathophysiological mechanisms

Using a systems medicine approach, integrative analysis of pathogen and patient characteristics (e.g. clinical data on disease progression, respiratory microbiota, pathogen genotypes etc.) will be performed in relation to RNA transcriptional profiles. With pathway analyses and gene network clustering and correlations, we aim to identify commonalities in host- and pathogen-related factors. Such commonalities, if identified, can inform management in case of a future (severe) ARI or other epidemics of unknown origin.

4. STUDY DESIGN

International, multicentre, prospective observational study.

This study is designed to enrol sufficient subjects in each class of major pathogen (influenza virus, human rhinovirus, respiratory syncytial virus and *Streptococcus pneumoniae*) and comorbidity (no comorbidity, chronic pulmonary disease, chronic cardiovascular disease, and chronic metabolic disease) combination, in order to be adequately powered to detect 2 fold differences in host RNA transcriptome profiles between severity levels within each pathogen-comorbidity class.

5. OBJECTIVES AND OUTCOME MEASURES

	Objectives	Outcome measures
Primary objective	Identify host and pathogen related determinants of severity of community acquired acute respiratory infections (ARI) in adults.	Differentially expressed host genes (≥ 2-fold difference in expression levels) as assessed by RNA transcriptome microarray in hospitalised and primary care managed cases of ARI, stratified by pathogen* and comorbidity**. *Influenza virus, Human Rhinovirus, Respiratory Syncytial Virus, *Streptococcus pneumoniae*; **No comorbidity, chronic pulmonary disease, chronic cardiovascular disease, chronic metabolic disease (diabetes)
Secondary objectives	Describe the aetiology, clinical management and outcomes of adult patients with community acquired ARI, in both primary care and hospital care, across Europe.	<u>In both groups:</u> • Prevalence of detection of putative pathogens in respiratory tract samples. • Proportion of cases receiving antibiotics, antivirals, antifungals and/or immunomodulators. • 28 day mortality <u>Additional in group 1 (primary care):</u> • Proportion of cases requiring hospitalisation <u>Additional in group 2 (hospitalised patients):</u> • Severity of illness at enrolment as assessed by Pneumonia Severity Index (PSI) and CURB-65 • Proportion of cases requiring during admission: supplemental oxygen; non-invasive or invasive mechanical ventilation; extra-corporeal life support. • Duration of invasive mechanical ventilation and extra-corporeal life support, if applicable • Proportion of cases requiring Intensive Care Unit (ICU)/High Care Unit (HCU) admission • Hospital - and ICU/HCU length of stay • In-hospital mortality
Tertiary Objectives	To develop and validate prognostic and diagnostic algorithms.	This is an exploratory objective. Some measures that could be included in algorithms are classifier gene sets based on host gene expression profiles, pathogen profiles, demographics, co-morbidities, risk factors, and clinical parameters. The diagnostic/prognostic algorithms will be evaluated for their ability to correctly diagnose infecting pathogen and/or predict adverse outcome in ARI.

	Objectives	Outcome measures
	To gain understanding into pathophysiological mechanisms contributing towards development of severe disease.	This is an exploratory objective to increase pathophysiological insights by integrative (systems medicine) analyses of pathogen- and patient characteristics. Measures that will be included in these integrative analyses are host gene expression profiles, clinical data on disease progression/outcome, deep sequencing of pathogen genomes and microbiomes.

6. PARTICIPANT IDENTIFICATION

6.1 Study participants

Group 1: Adults attending primary care with a new episode of community acquired respiratory tract illness

Group 2: Adults admitted to hospital with a new episode of community acquired respiratory tract illness.

6.2 Inclusion criteria Group 1 – primary care patients

1) Age ≥ 18 years
2) Patient is self-attending to primary care (i.e. not a home visit)
3) Clinical suspicion of a new episode of acute respiratory tract infection
4) Onset of the following symptoms within the last 14 days:
    i. Sudden onset of self-reported fever OR temperature of ≥ 38°C at presentation

    AND

    ii. At least one respiratory symptom (cough, sore throat, runny or congested nose, dyspnoea)

    AND

    iii. At least one systemic symptom (headache, muscle ache, sweats or chills or tiredness).

6.3 Exclusion criteria Group 1 – primary care patients

1) Patient lacks capacity to provide informed consent
2) No informed consent is provided by patient
3) The attending primary care physician decided to send patient to the hospital for assessment and possible hospital admission
4) Patient has been previously enrolled in the MERMAIDS-ARI study

6.4 Inclusion criteria Group 2 – hospitalised patients

1) Age ≥ 18 y

2) Clinical suspicion of a new episode of acute respiratory tract infection
3) Patient is admitted to hospital
4) Primary reason for hospital admission is clinical suspicion of a new episode of ARI
5) Onset of the following symptoms within the last 14 days:
    i. Sudden onset of self-reported fever OR temperature of ≥ 38°C at presentation

    AND

    ii. At least one respiratory symptom (cough, sore throat, runny or congested nose, dyspnoea)

    AND

    iii. At least one systemic symptom (headache, muscle ache, sweats or chills or tiredness).

6.5 Exclusion criteria Group 2 – hospitalised patients

1) Patient lacks capacity to provide informed consent
2) No informed consent is provided by patient
3) Patient has been transferred from another hospital
4) Patient has been previously enrolled in the MERMAIDS-ARI study

7. STUDY PROCEDURES

7.1 Screening and eligibility assessment

Potential patients for the study will be identified by local PREPARE co-ordinators, clinicians, nurses, microbiologists or virologists and assessed against the eligibility criteria listed above and their eligibility recorded in the site study protocol. A 'Screening Log' will be maintained of all the patients who undergo screening regardless of whether they decide to participate in the study. If there is a delay between introduction of the study/initial eligibility assessment and consent, then the recruiter must confirm all eligibility criteria are still met.

7.2 Recruitment and informed consent

The participant must sign and date the latest approved version of the Informed Consent form before any study specific procedures are performed.

Written and verbal versions of the Participant Information Sheet and Informed Consent will detail the exact nature of the study; the implications and constraints of the protocol; and any risks involved in taking part. It will be clearly stated that the participant is free to withdraw from the study at any time for any reason without prejudice to future care, and with no obligation to give the reason for withdrawal. Participant Information Sheets will be available in the common local language.

Written Informed Consent will be confirmed by the dated signatures of the participant and by the person who presented and obtained the informed consent. The person obtaining consent must be suitably qualified and experienced, and be authorised to do so by the Principal Investigator. A copy of the signed Informed Consent will be given to the participants. The original signed form will be retained at the study site.

7.2.1 Primary care

Participants will be identified by practice staff as they present naturally to primary care with an acute respiratory infection. Alternatively, advertisements may be used to invite potentially eligible participants to contact a recruiting primary care site to enquire about joining the study. In the primary care practice, a GCP-trained doctor or research nurse will assess the participants' eligibility and explain the study verbally to potential participants. They will also be given a Participant Information Sheet and allowed ample time to read it and ask questions. Consent will be taken by a GCP-trained doctor or research nurse.

7.2.2 Secondary care

Potential participants may be identified by a member of their immediate care team and asked to give a verbal consent to be approached by research staff who are not members of their care team. If they agree to this then the verbal consent will be recorded in the participant's medical record. Written and verbal versions of the Participant Information Sheet and Informed Consent will then be presented by GCP-trained study staff. Consent will be taken by a GCP-trained doctor or research nurse.

7.2.2.1 Study consenting procedures during the SARS-CoV-2 pandemic

During the SARS-CoV-2 / COVID-19 pandemic, study procedures are adjusted to minimalize interference with clinical care and reduce the burden on the patient, especially with regard to shortages of personal protection materials. The following adjustments will be applied:

Patients who are eligible for the study will be approached by a member of their healthcare team. Verbal information on the study will be provided. If the patient gives verbal informed consent to participate in the study, the study procedures will be executed. Verbal informed consent will be recorded in the patient notes. Once the patient is deemed non-infectious, a member of the healthcare team will undertake the informed consent procedure as stated in section 7.2.2.

7.3 Withdrawal from study

Patients are free to withdraw consent at any time without providing a reason. Patients who wish to withdraw consent for the study will have anonymised data collected up to the point of that withdrawal of consent included in the analyses. The patient will not contribute further data to the study. Data up to the time of withdrawal will be included in the analyses unless the patient explicitly states that this is not their wish.

7.4 Co-enrolment

Subjects may be co-enrolled in other studies if the study co-ordinators have been informed and have given their approval, to ensure the other study would not interfere with the results of this study or compromise patient welfare.

7.5 Potential risks and benefits

Since this is an observational study, we do not foresee any risks involved in participating in the study, and during the informed consent process the opportunity will be given to understand the objectives and any inconveniences of the study. Participating in the study will not directly benefit the participant, but will provide information on the burden, aetiology, risk factors, medical management and outcomes of the diseases and inform prevention, management and future research and management of emerging or (re-)emerging infections with epidemic potential.

7.6 Baseline assessments and subsequent visits

Following provision of consent, baseline data will be collected from patient medical charts and/or patient interview. Biological samples and information on treatment and care from subjects attending primary care (group 1) will be collected according to the schedule in Table 1a. They will have clinical baseline data and samples taken on day 0 and be invited for a follow up appointment on day 2 for sampling.

For subjects admitted to hospital (group 2), samples and data will be collected according to Table 1b.

Patients in both groups will be followed up at discharge and on day 28 for mortality. In order to avoid loss to follow up, in exceptional circumstances where a participant may not be able to attend hospital or their GP surgery/outpatient clinic, the day 28 follow up visit may take place in the participant's home or their place of work.

Table 1a. Schedule of procedures for subjects included in group 1 (primary care)

Day	0	1	2	28†
Eligibility screen	X			
Informed consent	X			
Baseline data	X			
Nasopharyngeal Swab	X			X
Nasopharyngeal RNA	X			X
Physical examination	X			
EDTA blood	X		X	
RNA blood	X		X	X
SST blood (serum)	X			X
Patient symptom scores	X			X
Concomitant meds	X	X	X	X
Follow up of outcomes				X

Day 0 = at initial presentation
Day 2 = 2 days after initial presentation. Delay of Day 2 blood sample collection up to Day 4 will be permissible.
† = -3/+21days

Table 1b. Schedule of procedures for subjects included in group 2 (hospital admissions)

Day	0	1	2	3 to discharge	At discharge^	28†
Eligibility screen	X					
Informed consent	X					
Baseline data	X					
Nasopharyngeal Swab	X				X	X
Nasopharyngeal RNA	X				X	X
Tracheal aspirate#	-	-	-	-		
Bronchoalveolar lavage fluid (BALF)#	-	-	-	-		
Physical examination	X					
EDTA blood	X		X		X	
RNA blood	X		X		X	X
SST blood (serum)	X					X
Clinical observations	X	X	X	X	X	X
Concomitant meds	X	X	X	X	X	X
Clinical management	X	X	X	X	X	
Follow up for outcome						X

Day 0 = within 24h of hospital admission
Day 2 = 2 days after hospital admission. Delay of Day 2 blood sample collection up to Day 4 will be permissible.
= if available at any time from clinical diagnostic procedure
^if discharged before or on day 2, then discharge samples will replace the day 2 samples
† = -3/+21 days

7.7 Clinical parameters

7.7.1 From all subjects

- Patient demographics and medical history, at enrolment
- Description of clinical signs and symptoms, in the 14 days before enrolment and at enrolment (Day 0)
- Physical examination, at enrolment (including respiratory rate and oximetry - see CRF for complete list)
- Discharge/final diagnosis as determined by attending physician
- Presence of febrile illness at day 28
- Mortality at day 28

7.7.2 Additional clinical parameters from Group 1

- in the 14 days before enrolment to day 28:
  - All relevant laboratory results as ordered by routine care
  - All relevant prescribed medication (e.g. antivirals, antibiotics, antifungals, anti-inflammatory drugs, steroids, other immune modulators, statins, ACE-inhibitors/angiotensin receptor blockers).
  - Results of all relevant microbiological testing as ordered for routine care

- Percentage of patients admitted to hospital by day 28

7.7.3 Additional clinical parameters from Group 2

- Pneumonia Severity Index (PSI) at admission
- CURB-65 severity score, at enrolment
- Requirement of (at any time during admission):
  - supplemental oxygen
  - non-invasive mechanical ventilation
  - invasive mechanical ventilation
  - extra-corporeal life support
- If applicable, duration of invasive mechanical ventilation and extra-corporeal life support
- In the 14 days before admission to day 28:
  - All relevant laboratory results as ordered by routine care
  - All relevant prescribed medication (e.g. antivirals, antibiotics, antifungals, anti-inflammatory drugs, steroids, other immune modulators, statins, ACE-inhibitors/angiotensin receptor blockers).
  - Results of all relevant microbiological testing as ordered for routine care
- Date of hospital admission and discharge, transfer or death
- Date of high care unit/intensive care unit admission and discharge, transfer or death, if applicable
- Clinical outcome: hospital discharge, transfer or in-hospital death

7.8 Definition of co-morbid phenotypes

A. Chronic pulmonary disease

Definition:

- Patients with confirmed COPD (chronic bronchitis or emphysema) OR possible COPD diagnosed by a Physician or the study team from data recorded in the clinical notes
- Patients diagnosed with asthma by a Physician and on regular (or seasonal) treatment with inhaled corticosteroids
- Patients with a diagnosis of Cystic Fibrosis or other bronchiectasis conditions
- Patients on regular medical or supportive (O2) treatment for non-infectious Interstitial lung disease (e.g. pulmonary fibrosis, asbestosis, autoimmune)
- Patients with neuromuscular pulmonary disease (e.g. Amyotrophic Lateral Sclerosis, Muscular Dystrophy, Myasthenia Gravis)

B. Chronic cardiovascular disease[1]

Patients diagnosed by a physician AND on regular medication or with pacemaker/VAD for any of the below:

[1] WHO http://www.euro.who.int/en/health-topics/noncommunicable-diseases/cardiovascular-diseases/cardiovascular-diseases2/definition-of-cardiovascular-diseases

- Coronary Artery Disease (e.g. Angina, Chronic Heart Disease, ischaemic or atherosclerotic heart disease)
- Chronic / congestive heart failure
- Cerebrovascular disease (ischaemic or haemorrhagic stroke, or transient ischaemic attack)
- Congenital heart disease
- Cardiomyopathy

C. Metabolic disease - diabetes

Patients diagnosed with Type 1 or Type 2 diabetes by a Physician or Nurse specialist AND requiring daily insulin or oral diabetes medication to control their diabetes.

7.9 Laboratory assessments

Laboratory samples to be analysed by the reference laboratories in Antwerp and Amsterdam, as described in Table 2. Laboratory samples will be labelled with a unique code allocated to each patient at enrolment. No patient identifiable information will be shared with the reference laboratory. Results will be reported back to study sites at the end of the study.

Table 2: Laboratory assessments, ARI observational study

Specimen	Collection container	Assay	Archive sample in biobank	Centralised/ Decentralised testing
Flocked Nasopharyngeal swab	UTM Viral Transport Media	Identification and quantification of putative respiratory pathogens	Yes	Centralised
Nasopharyngeal RNA	RNA conserving NP swab kit	Microarray analysis of host transcriptome	Yes	Centralised
Blood	EDTA vacutainer	Innate and adaptive immune markers	Yes	Centralised
Blood	Tempus™ Blood RNA tube	Microarray analysis of host transcriptome	Yes	Centralised
Blood	Serum-separating vacutainer	Soluble inflammatory mediators and serum bank (PREPARE Task 3.2)	Yes	Centralised
Tracheal aspirate/ BALF#	Sterile container	Pathogen identification, pathogen shedding	Yes	Centralised

= if available from routine clinical diagnostics

7.10 Sample handling

Sample handling will adhere to the PREPARE Multi-centre EuRopean study of MAjor Infectious Disease Syndromes (MERMAIDS) guidelines for sample handling.

7.11 Standard of care for sampling from patients

Samples taken for the purpose of medical management will, at all times, have priority over samples requested for research reasons. Research samples will, wherever possible, be collected at a time coinciding with samples taken for routine care.

7.12 Future use of samples

Samples collected will be used for the purpose of this study as stated in the protocol and stored for future analyses that are relevant to the study questions. The standard consent form will request consent from subjects for sample storage and/or export of samples to collaborating institutions for investigations that cannot be performed locally. Any proposed plans to use samples other than for those relevant to the study question will be submitted to the relevant ethics committees prior to any testing.

Any database will only identify participants by a participant number. Participant names or any other identifying details will NOT be included in the central database. Data may be used alone or in combination with data from related studies in secondary analyses.

7.13 Medical management and safety reporting

Medical management will be according to standard of care at the treating site and not a part of this research protocol. The research only includes collection of clinical information and specimens and therefore adverse event reporting is not applicable as there is no intervention.

7.14 Discontinuation/withdrawal of participants from study

Subjects can leave the study at any time for any or no reason if they wish to do so without any consequences. The investigator can decide to withdraw a subject from the study for urgent medical reasons.

7.15 Definition of end of study

The end of the study will be the date when all enrolled patients have completed their 28 day follow up.

8. STATISTICS AND ANALYSES

8.1 Statistical considerations regarding power calculation

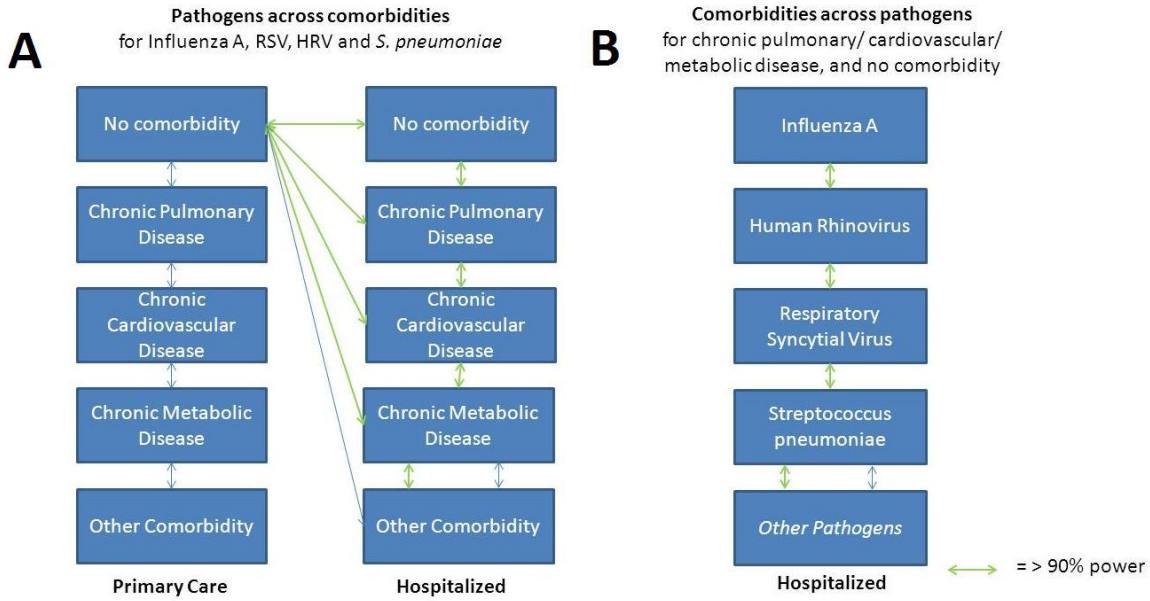
8.1.1 General approach to power calculation

The primary aim of this study is to detect differentially expressed genes as assessed by RNA transcriptome, across all groups of major pathogens and major comorbidity groups. To determine the total number of subjects required to answer this question, these groups are firstly defined (8.1.2). Then the minimally required group size necessary to detect differentially expressed genes between two groups is determined (8.2). Based on the estimated prevalence of the pathogens and comorbidities, the power calculation is performed to determine the minimal amount of subjects in group 2 (8.3.1) and group 1 (8.3.2).

8.1.2 Subgroup comparisons

We have performed sample size calculation on the ability to detect differentially expressed genes with a statistical power of at least 90% between all groups across the major pathogens (influenza virus, human rhinovirus, respiratory syncytial virus and *Streptococcus pneumoniae*), and major comorbidity groups (no comorbidity, chronic pulmonary disease, chronic cardiovascular disease, and chronic metabolic disease in hospitalised patients, and between primary care patients without comorbidities (Figure 1). Basing our power calculation on these subgroups will subsequently provide enough statistical power to detect differentially expressed genes in a number of other aetiologies across these comorbidities as well as for our secondary study objectives.

Figure 1. Predefined relevant subgroups for power calculation



A. This figure represents the predefined subgroups for each ARI aetiology (viruses and bacteria). The sample size estimation is based on achieving a statistical power of >90% for detection of differentially expressed genes in predefined subgroups (e.g. Influenza, HRV, RSV, *Streptococcus pneumoniae*). The statistical power to detect differentially expressed genes in other comorbidities in the above mentioned pathogens will vary (e.g. cancer/ immunosuppressive condition/obesity approx. 80-90%). **B.** This figure displays the subgroups in which we will analyse the differences and commonalities in the transcriptional profiles of hospitalised patients with a specific comorbidity across the different pathogens. The statistical power to investigate these comorbidities in other pathogens will vary (e.g. Parainfluenza virus >90% in chronic pulmonary disease and no comorbidity, coronaviruses approx. 50-80%).

8.2 Estimation of minimally required group sizes

We assume that the expression of each gene is normally distributed on the log scale and that the variance is the same in the different groups and thus the only difference between the groups is the mean. At an alpha of 0.05 and at an expected standard deviation of the gene intensity measurements on the base-two logarithmic scale of 0.7 (realistic for genes that are expressed at moderate to high levels), we expect to be able to detect 90% of >2 fold differentially expressed genes (i.e. an absolute effect size of 1 on a base-two logarithmic scale) between groups of 11 patients. These estimations are likely to be on the conservative side as we chose to err on the side of caution with regard to the chosen value for standard deviation. The calculation was based on the Affymetrix HG U133plus2 microarray platform, which contains 54,515 probe-sets.

Table 3. Results from power calculations

	effect size in fold change		
power	**2**	**1.75**	**1.5**
60%	5	8	15
70%	7	10	18
80%	8	12	23
90%	11	16	31
95%	13	20	38

8.3 Power calculation

8.3.1 Number of hospitalised subjects needed

Based on the demographic and clinical data from 9 large clinical studies (24–32) in patients hospitalized with ARI we can expect about 4.9% of ARI patients to be infected with Influenza A, 4.6% with RSV, 3.9% with HRV. If we include 1500 ARI patients we expect to find group sizes of at least 11 in all groups across the major pathogens and major comorbidity groups (as shown in figure 1). This power calculation is based on detecting individual differentially expressed genes, which requires more power than detecting differentially/up- or down regulated pathways or clusters of genes.

Table 4. Projected prevalence of pathogens in hospitalised ARI patients based on available literature

PATHOGENS	mean (range) %
Influenza A	4.9 (4.8-5.0)
RSV	4.6(2.0-7.2)
HRV	3.9 (3.9-3.9)
Streptococcus pneumoniae	6.9 (0.8-13.3)

Table 5. Projected prevalence of comorbidities in hospitalised ARI patients based on available literature

COMORBIDITIES	**mean (range) %**
Chronic pulmonary disease	31.6 (19.4-42.0)
Chronic cardiovascular disease	20.9 (10.0-41.9)
Chronic metabolic disease	23.0 (14.0-40.9)
No comorbidity	35.8 (18.3-66.7)

8.3.2 Number of primary care subjects needed

Based on the demographic and clinical data from the GRACE cohort (Genomics to combat Resistance against Antibiotics in Community-acquired LRTI in Europe, http://www.grace-lrti.org/) in primary care patients presenting with ARI, we can expect about 9.9% of patients with influenza A infection, 4.6% with RVS and 20.1% with HRV. We expect 50% (range 45-65%) of ARI patients in primary care not to have comorbidities. If we include 500 primary care patients we expect to find group sizes of at least 11 in patients without comorbidities infected with influenza virus A, HRV, RSV or *Streptococcus pneumoniae*.

Table 6. Prevalence of pathogens in GRACE cohort

PATHOGENS	**GRACE data (%)**
Influenza A	9.9
RSV	4.6
HRV	20.1
Streptococcus pneumoniae	5.5

8.4 Statistical considerations regarding outcome measures

8.4.1 RNA expression profiles

Correction for Multiple Comparisons

To control for multiple testing and reduce the false positive rate (FDR), stringent statistical criteria will be used to identify nominal 2-fold differentially expressed genes (i.e. effect size of 1 in absolute terms on a base-two logarithmic scale) with raw P-value < 0.05 and Benjamini-Hochberg False Discovery Rate <0,05.

Dimension reduction

Dimension reduction analysis of expression levels aims at identifying genes that vary the most amongst subjects and phenotypic classes, using methods such as Principal Component Analysis or Multi-Dimensional Scaling.

Class Discovery

Class discovery, or unsupervised cluster analysis aims at grouping genes or samples or both, using solely measures of similarity of expression levels, using methods such as hierarchical clustering, K-means clustering and bi-clustering in combination with resampling and bootstrapping.

Supervised Classification Methods

Supervised classification methods use information on existing, labelled classes to build a classifier, using methods such as linear discriminant analysis. Some methods combine class discovery and classification (33).

Multiple testing

Correction with multiple testing will rely on the False Discovery Rate method (34). Analyses will be performed using standard packages in R, version 3.1 (35).

Gene regulatory networks

A major task in transcriptome analysis is to derive the interaction between genes from the expression levels measured. Relations between genes are represented as a network representing genes as nodes and interactions as edges. Inference of such network is an active field of research and numerous methods exist and are being developed. Analysis of these networks allows the identification of highly connected genes and novel interactions (36–38).

8.4.2 Pathway and network analysis

Functional Annotation and Pathway Enrichment analysis

One approach for functional grouping of genes identified as relevant from a statistical viewpoint is realized by utilizing gene ontologies (GO (39)). Statistical tools exist to identify enriched or depleted categories for a list of genes or proteins of interest. We will use Database for Annotation, Visualization, and Integrated Discovery (40) (DAVID) or BiNGO(41), and pathway databases like KEGG (42) and Reactome (43) to complement the functional ontologies.

Gene Set Enrichment Analysis

This type of analysis ignores what we already know about the biology of a system, and identifies enriched terms without regard for how they are related to one another. Gene Set Enrichment Analysis operates over the whole of the data from a microarray experiment, so no data points are lost prior to functional analysis. It also makes uses of prior knowledge to group sets of genes by related function, meaning that over-representation of a group of functions is considered, as opposed to isolated individual functions. The most commonly used tool for Gene Set Enrichment Analysis is GSEA (44).

Network clustering and correlation network

Each gene is a node and each edge indicates co-expression of two genes under the conditions of the experiment. The use of correlation networks for gene expression data is supported by Cytoscape (45) using either the MCODE (46) or ClusterMaker (47) plugin.

8.5 Statistical analysis of secondary outcome measures

Statistical analysis of secondary endpoints will be descriptive in nature and include estimates of rates and proportions. Hypothesis testing is not envisaged to be a part of the statistical analysis of secondary endpoints.

All participants with evaluable samples and informed consent will be included in the analysis.

Proportions will be estimated for the prevalence of each pathogen in respiratory tract samples; the proportion of cases receiving antibiotics, antivirals, antifungals and/or immunomodulators; and proportion not surviving to 28 days post enrolment. Estimates will be presented for primary and hospital care groups separately, with 95% confidence intervals estimated using the binomial exact method. Additional binary endpoints will be analysed in the same manner.

Time to event variables such as time to discharge (length of stay) or time on ventilation will be described using the Kaplan-Meier method.

9. DATA MANAGEMENT

9.1 Source data

Source documents are original documents, data, and records from which participants' CRF data are obtained. These include, but are not limited to, general practice and hospital medical records.

CRF entries will be considered source data if the CRF is the site of the original recording (e.g. there is no other written or electronic record of data). In this study the CRF will be used as the source document for the documentation of inclusion and exclusion criteria, and baseline assessment information, which will include, but not be limited to, comorbidities and severity of symptoms.

All documents will be stored safely in confidential conditions. On all study-specific documents, other than the signed consent and contact details form, the participant will be referred to by the study participant ID, not by name, including on all laboratory results.

9.2 Access to data

Direct access will be granted to authorised representatives from the Sponsor, host institution and the regulatory authorities to permit study related monitoring, audits and inspections.

9.3 Data recording and record keeping

Research Online (RO) is an electronic data capture (EDC) system that will be used for data collection. Web-based case report forms (eCRF) are implemented into the system to facilitate the study specific data collection. These forms can easily be accessed by all standard web browsers. Multiple validation and range checks will be programmed in the eCRF to assure complete and high quality data. Data that does not comply with these rules or ranges will generate a query that must be resolved immediately or at a later stage. Electronic workflows as multiple skip and jump rules will ensure that only information that is applicable to the patient will appear. After the data of last subject is entered, the database can rapidly be closed and data made available for further analysis and publication purposes. This data will not contain any patient identifiable information.

RO meets all requirements according to ICH-GCP standards for electronic data entry with respect to safeguarding data integrity and data security regulations. Users will have role based access to the system logging in using their personal username and password. The system will log all data entry steps with timestamps and user information. The role based access to the system will avoid unauthorised data access and prevents users from performing actions that they do not have authorisation for.

Project management of the study is facilitated by the integrated real live study progress reports. RO data traffic over the internet is encrypted using secured data communication protocols. Dedicated databases and web servers are hosted in a secure data centre, the database (PostgreSQL) is backed up on a daily basis.

The participants will be identified by a unique study specific number and/or code in any database. The name and any other identifying detail will not be included in any study data electronic file.

9.4 Standard procedures

To assure high quality the data management (DM) department of the Julius Centre (JC), Utrecht, who will be responsible for the data management within the trial, works according to a Quality Management System. All work is carried out in accordance with our written Standard Operating Procedures (SOP) and Work instructions (WI).

10. QUALITY ASSURANCE PROCEDURES

The study will be conducted in accordance with the current approved protocol, GCP, relevant regulations and MERMAIDS Standard Operating Procedures. The monitoring will be performed by the MERMAIDS Study Manager or delegate; these tasks may be delegated to a qualified Organisation. The level of monitoring required will be informed by a risk assessment. All investigators and study related site staff will receive training in study procedures and GCP. Regular study monitoring will be conducted as below to ensure that the rights and wellbeing of human participants are protected during the course of the study and that the data collected is credible and accurate.

Governance of the study will be through the Study Steering Committee (SSC) and the Study Operation Group (SOG).

The SSC will be the primary decision making body, with membership from partner institutions, a senior representative from each participating country, and other independent experts. Decisions of the Committee will be by consensus. Terms of reference for the SSC will set out the full details of the membership, operations and remit.

The SOG will be led from the MERMAIDS Coordination Centre in Oxford and will report to the SSC. The SOG team will provide operational oversight and day-to-day management of the study to ensure compliance to follow up and critical factors such a retention and data quality. The team will comprise those with responsibility for the operational activities and so include study operations, data management and central laboratory.

11. ETHICAL AND REGULATORY CONSIDERATIONS

11.1 Declaration of Helsinki

The Investigator will ensure that this study is conducted in accordance with the principles of the Declaration of Helsinki (64th WMA General Assembly, Fortaleza, Brazil, October 2013).

11.2 Guidelines for good clinical practice

The Investigators will ensure that this study is conducted in accordance with relevant regulations and with Good Clinical Practice.

11.3 Approvals

The protocol, informed consent form, participant information sheet and any proposed advertising material will be submitted to an appropriate Research Ethics Committee (REC), and host institution(s) for written approval. This will be done in collaboration with Work Package 1 of the PREPARE Programme to ensure that the correct regulatory approvals are gained in each participating country.

The Investigator will submit and, where necessary, obtain approval from the above parties for all substantial amendments to the original approved documents.

11.4 Reporting

The CI shall submit once a year throughout the study, or on request, an Annual Progress report to the REC Committee, host organisation and Sponsor. In addition, an End of Study notification and final report will be submitted to the same parties.

11.5 Participant confidentiality

The study staff will ensure that the participants' anonymity is maintained. The participants will be identified only by a study ID number on the CRF and any electronic database. All documents will be stored securely and only accessible by study staff and authorised personnel. The study will comply with the Data Protection Act, which requires data to be anonymised as soon as it is practical to do so.

11.6 Expenses and benefits

Participants will be reimbursed for their time and travel expenses for all study visits in additional to normal care. This reimbursement will be made in the form of gift card(s) or voucher(s) given once the day 28 visit has been completed. No cash payments will be made. The maximum total reimbursement per participant for all study visits will be 70 Euros (or equivalent in a different currency).

11.7 Other ethical considerations

This observational study will not exclude pregnant women or older participants. We are aiming to include as broad a range of samples as possible to be representative of the general population.

It is expected that 2-5% of hospitalised participants might lose capacity at some point during the study. Should this occur, these participants would not be withdrawn from the study. In the UK, and other countries with similar ethical requirements, participants will be asked for their prior consent to continue the collection of study data and samples should they later lose capacity. This will be explained fully during the consent process and consent for use of data and samples will be reconfirmed verbally when/if they regain capacity. The prior consent is optional so participants may opt-out and would be withdrawn from the study should they become incapacitated. Proxy assent will also be sought in order to continue collecting study samples on participants who have lost capacity.

12. FINANCE AND INSURANCE

12.1 Funding

The study is funded by the European Commission.

12.2 Insurance

The University has a specialist insurance policy in place, which would operate in the event of any participant suffering harm as a result of their involvement in the research (Newline Underwriting Management Ltd, at Lloyd's of London). NHS indemnity operates in respect of the clinical treatment, which is provided.

13. PUBLICATION POLICY

The Investigators will be involved in reviewing drafts of the manuscripts, abstracts, press releases and any other publications arising from the study. Authors will acknowledge that the study was funded by the European Commission. Authorship will be determined in accordance with the ICMJE guidelines and other contributors will be acknowledged.

14. REFERENCES

1. Auble TE, Hsieh M, Gardner W, Cooper GF, Stone RA, McCausland JB, et al. A prediction rule to identify low-risk patients with heart failure. Acad Emerg Med. **2005**; 12:514–521.

2. Lim WS, Eerden MM van der, Laing R, Boersma WG, Karalus N, Town GI, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax. **2003**; 58(3):377–382.

3. Chen R, Snyder M. Promise of personalized omics to precision medicine. Wiley Interdiscip. Rev. Syst. Biol. Med. 2013. p. 73–82.

4. Flacher V, Bouschbacher M, Verronese E, Massacrier C, Sisirak V, Berthier-Vergnes O, et al. Human Langerhans Cells Express a Specific TLR Profile and Differentially Respond to Viruses and Gram-Positive Bacteria. J Immunol [Internet]. **2006** [cited 2014 May 30]; 177(11):7959–7967. Available from: http://www.jimmunol.org/cgi/doi/10.4049/jimmunol.177.11.7959

5. Ramilo O, Allman W, Chung W, Mejias A, Ardura M, Glaser C, et al. Gene expression patterns in blood leukocytes discriminate patients with acute infections. Blood [Internet]. **2007** [cited 2014 May 30]; 109(5):2066–77. Available from: http://bloodjournal.hematologylibrary.org/content/109/5/2066.short

6. Zaas AK, Chen M, Varkey J, Veldman T, Hero AO, Lucas J, et al. Gene expression signatures diagnose influenza and other symptomatic respiratory viral infections in humans. Cell Host Microbe [Internet]. **2009** [cited 2013 Nov 13]; 6(3):207–17. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2852511&tool=pmcentrez&rendertype=abstract

7. Huang Y, Zaas AK, Rao A, Dobigeon N, Woolf PJ, Veldman T, et al. Temporal dynamics of host molecular responses differentiate symptomatic and asymptomatic influenza a infection. PLoS Genet [Internet]. **2011** [cited 2014 May 30]; 7(8):e1002234. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3161909&tool=pmcentrez&rendertype=abstract

8. Woods CW, McClain MT, Chen M, Zaas AK, Nicholson BP, Varkey J, et al. A host transcriptional signature for presymptomatic detection of infection in humans exposed to influenza H1N1 or

H3N2. PLoS One [Internet]. **2013** [cited 2014 Aug 12]; 8(1):e52198. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3541408&tool=pmcentrez&rendertype=abstract

9. Lole KS, Arya RP, Tripathy AS, Ramdasi AY, Chadha MS, Sangle S a, et al. Role of host immune response and viral load in the differential outcome of pandemic H1N1 (2009) influenza virus infection in Indian patients. PLoS One [Internet]. **2010** [cited 2014 May 30]; 5(10). Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2948498&tool=pmcentrez&rendertype=abstract

10. Simmons CP, Popper S, Dolocek C, Chau TNB, Griffiths M, Dung NTP, et al. Patterns of host genome-wide gene transcript abundance in the peripheral blood of patients with acute dengue hemorrhagic fever. J Infect Dis [Internet]. **2007** [cited 2013 Nov 13]; 195(8):1097–107. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17357045

11. Berdal J-E, Mollnes TE, Wæhre T, Olstad OK, Halvorsen B, Ueland T, et al. Excessive innate immune response and mutant D222G/N in severe A (H1N1) pandemic influenza. J Infect [Internet]. **2011** [cited 2014 Aug 13]; 63(4):308–16. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21781987

12. Bermejo-Martin JF, Martin-Loeches I, Rello J, Antón A, Almansa R, Xu L, et al. Host adaptive immunity deficiency in severe pandemic influenza. Crit Care [Internet]. **2010**; 14(5):R167. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3219262&tool=pmcentrez&rendertype=abstract

13. Kongchanagul A, Suptawiwat O, Boonarkart C, Kitphati R, Puthavathana P, Uiprasertkul M, et al. Decreased expression of surfactant protein D mRNA in human lungs in fatal cases of H5N1 avian influenza. J Med Virol [Internet]. **2011** [cited 2014 Jul 21]; 83(8):1410–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21678446

14. Tang BMP, McLean AS, Dawes IW, Huang SJ, Cowley MJ, Lin RCY. Gene-expression profiling of gram-positive and gram-negative sepsis in critically ill patients. Crit Care Med [Internet]. **2008** [cited 2014 May 29]; 36(4):1125–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18379237

15. Nau GJ, Richmond JFL, Schlesinger A, Jennings EG, Lander ES, Young R a. Human macrophage activation programs induced by bacterial pathogens. Proc Natl Acad Sci U S A [Internet]. **2002**; 99(3):1503–8. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=122220&tool=pmcentrez&rendertype=abstract

16. Waterer GW. Community-acquired pneumonia: genomics, epigenomics, transcriptomics, proteomics, and metabolomics. Semin Respir Crit Care Med [Internet]. **2012** [cited 2014 May 30]; 33(3):257–65. Available from: http://www.nejm.org/doi/full/10.1056/NEJM199512143332408

17. Greninger AL, Chen EC, Sittler T, Scheinerman A, Roubinian N, Yu G, et al. A metagenomic analysis of pandemic influenza A (2009 H1N1) infection in patients from North America. PLoS

One [Internet]. **2010** [cited 2014 May 24]; 5(10):e13381. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2956640&tool=pmcentrez&rendertype=abstract

18. Liu L, Li R, Pan Y, Chen J, Li Y, Wu J, et al. High-throughput screen of protein expression levels induced by cyclooxygenase-2 during influenza a virus infection. Clin Chim Acta [Internet]. Elsevier B.V.; **2011** [cited 2014 May 30]; 412(11-12):1081–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21354125

19. Yu S-Y, Hu Y-W, Liu X-Y, Xiong W, Zhou Z-T, Yuan Z-H. Gene expression profiles in peripheral blood mononuclear cells of SARS patients. World J Gastroenterol [Internet]. **2005**; 11(32):5037–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16124062

20. Reghunathan R, Jayapal M, Hsu L-Y, Chng H-H, Tai D, Leung BP, et al. Expression profile of immune response genes in patients with Severe Acute Respiratory Syndrome. BMC Immunol [Internet]. **2005** [cited 2014 Aug 13]; 6:2. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=546205&tool=pmcentrez&rendertype=abstract

21. Parnell G, McLean A, Booth D, Huang S, Nalos M, Tang B. Aberrant cell cycle and apoptotic changes characterise severe influenza A infection--a meta-analysis of genomic signatures in circulating leukocytes. PLoS One [Internet]. **2011** [cited 2014 May 30]; 6(3):e17186. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3050844&tool=pmcentrez&rendertype=abstract

22. Shi L, Campbell G, Jones WD, Campagne F, Wen Z, Walker SJ, et al. The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. Nat Biotechnol. **2010**; 28(8):827–838.

23. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis. **1987**; 40:373–383.

24. Spoorenberg SMC, Bos WJW, Heijligenberg R, Voorn PGP, Grutters JC, Rijkers GT, et al. Microbial aetiology, outcomes, and costs of hospitalisation for community-acquired pneumonia; an observational analysis. BMC Infect Dis. **2014**; 14(1):335.

25. Reed C, Chaves SS, Perez A, D'Mello T, Daily Kirley P, Aragon D, et al. Complications among adults hospitalized with influenza: a comparison of seasonal influenza and the 2009 H1N1 pandemic. Clin Infect Dis [Internet]. **2014** [cited 2015 Feb 10]; 59(2):166–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24785230

26. Pletz MW, Terkamp C, Schumacher U, Rohde G, Schütte H, Welte T, et al. Vitamin D deficiency in community-acquired pneumonia: low levels of 1,25(OH)2 D are associated with disease severity. Respir Res [Internet]. **2014** [cited 2015 Feb 10]; 15:53. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4046524&tool=pmcentrez&rendertype=abstract

27. Aliberti S, Brambilla AM, Chalmers JD, Cilloniz C, Ramirez J, Bignamini A, et al. Phenotyping community-acquired pneumonia according to the presence of acute respiratory failure and severe sepsis. Respir Res. Respiratory Research; **2014**; 15(1):27.

28. Pereira JM, Teixeira-Pinto A, Basílio C, Sousa-Dias C, Mergulhão P, Paiva JA. Can we predict pneumococcal bacteremia in patients with severe community-acquired pneumonia? J Crit Care. Elsevier Inc.; **2013**; 28(6):970–4.

29. Lee N, Hui DSC, Zuo Z, Ngai KLK, Lui GCY, Wo SK, et al. A prospective intervention study on higher-dose oseltamivir treatment in adults hospitalized with influenza a and B infections. Clin Infect Dis. **2013**; 57(11):1511–9.

30. Pasquale M Di, Ferrer M, Esperatti M, Crisafulli E, Giunta V, Li Bassi G, et al. Assessment of severity of ICU-acquired pneumonia and association with etiology. Crit Care Med [Internet]. **2014** [cited 2015 Feb 10]; 42(2):303–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23989176

31. Louie JK, Acosta M, Samuel MC, Schechter R, Vugia DJ, Harriman K, et al. A novel risk factor for a novel virus: obesity and 2009 pandemic influenza A (H1N1). Clin Infect Dis [Internet]. **2011** [cited 2015 Jan 14]; 52(3):301–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21208911

32. Oliveira W, Carmo E, Penna G, Kuchenbecker R, Santos H, Araujo W, et al. Pandemic H1N1 influenza in Brazil: analysis of the first 34,506 notified cases of influenza-like illness with severe acute respiratory infection (SARI). Euro Surveill. **2009**; 14(42):1–6.

33. Qiu P, Plevritis SK. Simultaneous class discovery and classification of microarray data using spectral analysis. J Comput Biol A J Comput Mol Cell Biol. **2009**; 16(7):935–944.

34. Cheng C, Pounds S. False discovery rate paradigms for statistical analyses of microarray gene expression data. Bioinformation. **2007**; 1(10):436–446.

35. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL. 2013. p. http://www.R–project.org/.

36. Baitaluk M, Kozhenkov S, Ponomarenko J. An integrative approach to inferring gene regulatory module networks. PLoS One. **2012**; 7(12):e52836.

37. Schadt EE, Lamb J, Yang X, Zhu J, Edwards S, Guhathakurta D, et al. An integrative genomics approach to infer causal associations between gene expression and disease. Nat Genet. **2005**; 37(7):710–717.

38. Emmert-Streib F, Glazko G V., Altay G, Matos Simoes R de. Statistical inference and reverse engineering of gene regulatory networks from observational expression data. Front Genet. **2012**; 3:8.

39. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. **2000**; 25(1):25–29.

40. Dennis Jr G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: database for annotation, visualization, and integrated discovery. Genome Biol. **2003**; 4(5):P3.

41. Maere S, Heymans K, Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics. **2005**; 21(16):3448–3449.

42. Kotera M, Hirakawa M, Tokimatsu T, Goto S, Kanehisa M. The KEGG databases and tools facilitating omics analysis: latest developments involving human diseases and pharmaceuticals. Methods Mol Biol. **2012**; 802:19–39.

43. Joshi-Tope G, Gillespie M, Vastrik I, D'Eustachio P, Schmidt E, Bono B de, et al. Reactome: a knowledgebase of biological pathways. Nucleic Acids Res. **2005**; 33(Database issue):D428–432.

44. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. **2005**; 102(43):15545.

45. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. **2003**; 13(11):2498–2504.

46. Bader GD, Hogue CW V. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. **2003**; 4:2.

47. Morris JH, Apeltsin L, Newman AM, Baumbach J, Wittkop T, Su G, et al. clusterMaker: a multi-algorithm clustering plugin for Cytoscape. BMC Bioinformatics. **2011**; 12:436.