BMJ Open

Serological surveillance of influenza and other infections in an English sentinel network: pilot study protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-024285
Article Type:	Protocol
Date Submitted by the Author:	18-May-2018
Complete List of Authors:	de Lusignan, Simon; University of Surrey, Department of Clinical and Experimental Medicine; St. Georges University of London, Division of Population Health Sciences and Education Borrow, Ray; Public Health England, Vaccine Evaluation Unit, Manchester Royal Infirmary Tripathy, Manasa ; University of Surrey, Department of Clinical and Experimental Medicine Linley, Ezra; Public Health England, Vaccine Evaluation Unit, Manchester Royal Infirmary Zambon, Maria; Public Health England, Hoschler, Katja; Public Health England Ferreira, Filipa; University of Surrey, Department of Clinical and Experimental Medicine Andrews, Nick; Public Health England, Modelling and Economics Department Yonova, Ivelina; University of Surrey, Department of Clinical and Experimental Medicine; Royal College of General Practitioners, CIRC Hriskova, Mariya ; University of Surrey, Department of Clinical and Experimental Medicine; Royal College of General Practitioners, CIRC Rafi, Imran; Royal College of General Practitioners, Pebody, Richard; Public Health England, Centre for Infectious Disease Surveillance and Control
Keywords:	Medical Records System, Computerized, Serology, Data Collection, Population Surveillance, Records as topic, Influenza, Human

Serological surveillance of influenza and other infections in an English sentinel network: pilot study protocol

Simon de Lusignan[1,2]
Professor of Primary Care and Clinical Informatics
Head of Department of Clinical and Experimental Medicine
s.lusignan@surrey.ac.uk
SdeL

Ray Borrow[4]
Head of the Vaccine Evaluation Unit
Ray.Borrow@phe.gov.uk
RB

Manasa Tripathy[1]
Practice Liaison Officer
m.tripathy@surrey.ac.uk
MT

Ezra Linley[4]
Deputy Head of Vaccine Evaluation Unit
Ezra.Linley@phe.gov.uk
EL

Maria Zambon[3]
Head of Reference Microbiology
Maria.Zambon@phe.gov.uk
MZ

Katja Hoschler[3]
Senior Scientist, Respiratory Virus Unit
Katja.hoschler@phe.gov.uk
KH

Filipa Ferreira[1]
Project Manager
f.ferreira@surrey.ac.uk
FF

Nick Andrews[3]
Senior statistician
nick.andrews@phe.gov.uk
NA

Ivelina Yonova[1,2]
Practice Liaison Officer
i.yonova@surrey.ac.uk
IY

Mariya Hriskova[1,2]
Practice Liaison Officer
m.hriskova@surrey.ac.uk
MH

Imran Rafi[2]
Chair of Clinical Innovation and Research Centre
Irafi@rcgp.org.uk
IR

Richard Pebody [3]
Head of Flu Surveillance
richard.pebody@phe.gov.uk
RP

[1] Department of Clinical and Experimental Medicine, University of Surrey, GUILDFORD, GU2 7XH, UK
[2] Royal College of General Practitioners Research and Surveillance Centre, Euston Square, LONDON,
[3] Public Health England, London, UK
[4] Vaccine Evaluation Unit, Public Health England, Manchester Royal Infirmary, Manchester, UK

Correspondence to:
Professor Simon de Lusignan
s.lusignan@surrey.ac.uk
T: 01483 683089
Section of Clinical Medicine & Ageing
Department of Clinical and Experimental Medicine
University of Surrey
GUILDFORD, GU2 7XH, UK

Word Count: 5847

Serological surveillance of influenza and other infections in an English sentinel network: pilot study protocol

ABSTRACT

Introduction

Background:
Rapidly undertaken age-stratified serology studies can produce valuable data about a new emerging infection including back-ground population immunity and seroincidence during an influenza pandemic. Traditionally seroepidemiology studies have used surplus laboratory sera with little or no clinical information or have been expensive detailed population-based studies. We propose collecting population-based sera from the Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC), a sentinel network with extensive clinical data.

Aim:
To pilot a method for provision of nationally representative serum samples and associated patient data to measure sero-positivity and seroincidence due to seasonal influenza and other infections of public health importance, and create a population- based serology bank for investigation of other important infections.

Methods and Analysis:

Setting and Participants: We will recruit 6 RCGP RSC practices already taking nasopharyngeal virology swabs. Patients who attend for a scheduled blood test will be consented to donate additional blood samples. Approximately 100-150 residual blood samples will be collected from each of the following age-bands - 18-29, 30-39, 40-49, 50-59, 60-69, 70+ years.

Methods: We will send the samples to the Public Health England (PHE) Seroepidemiology Unit (SEU) for processing and storage. These samples will be tested for influenza antibodies, using haemagglutination inhibition assays (HAI). Serology results will be pseudonymised, sent to the RCGP RSC and combined using existing processes at the RCGP RSC secure hub. The influenza seroprevalence results from the RCGP cohort will be compared against those from the annual PHE influenza residual serosurvey.

Ethics and dissemination

Ethical approval was granted by the Proportionate Review Sub- Committee of the London – Camden & Kings Cross on 6 February 2018. This study received approval from Health Research Authority on 7 February 2018. On completion the results will be made available via peer- reviewed journals.

Strengths and Limitations of this study

Strengths:

- The Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) is one of the oldest sentinel networks in Europe, it has completed its 51st season of surveillance.
- Practices in RCGP take microbiological samples, including influenza virology specimens which are linked at individual level.
- We have the potential through this network to link vaccine exposure, virologically confirmed influenza and serology.

Limitations:

- This is a pilot study to demonstrate we can collect samples across pre-defined adult age-groups - we will not collect samples from children, who may be exposed to live attenuated intranasal vaccine (LAIV) in recent years.
- Pharmacist and other non-GP vaccination can lead to missing information on the computerised medical record system at General Practice.

Keywords:

Medical records systems, computerized
Population Surveillance
Serology
Influenza, human
Data collection
Records as topic
Seroepidemiologic Studies
Primary Health Care
Pandemics

INTRODUCTION

Rationale for the pilot study and background

There have been calls for a World Serology Bank as serology could tell us much about the susceptibility of the population to infectious disease.[1] This pilot study explores the potential to establish a serology bank, based on a sentinel network, and focussed in the first instance on influenza and respiratory disease.

Serological data potentially allow the assessment of the severity of a new influenza strain by providing the capability to detect asymptomatic and mild infections and thus determine the symptomatic proportion.[2] The number of infections can be determined if the age-specific prevalence of immunity prior to and then during and after the pandemic are known. Thus, the number of people infected (and therefore no longer susceptible) can be calculated. If these data are available early (particularly the background population immunity and the symptomatic proportion), they can be used to adjust planning assumptions and to help predict the impact of the pandemic on health care services and optimal intervention strategies.

Serum archives would help assess the severity of a novel influenza virus and allow modification of local, national and international pandemic plans.[3] This requirement was a lesson from the 2009 A(H1N1) pandemic and its importance is recognised by the World Health Organization (WHO), the European Centre for Disease Control (ECDC), the UK Department of Health (DH), and the Joint Committee on Vaccination and Immunisation (JCVI).

Despite its apparent merit, the use of serology for seasonal influenza surveillance has several limitations. Seasonal influenza infections occur regularly and individuals are re-infected throughout life with related strains. At present, two serum samples have to be analysed in order to evaluate an increase of antibody between acute and convalescent samples. The prior (lifetime) exposure of individuals to circulating influenza viruses in previous years and a high degree of cross-reactivity between antibodies to different seasonal influenza virus strains can lead to results which are difficult to interpret. Distinguishing between recent infection and recent vaccination also presents further difficulties. Consequently, seroepidemiological studies have not often been used for the investigation of seasonal influenza at a population level. However, when analysing immunologically naïve individuals, such as young children or a naïve population at the beginning of a pandemic, seroepidemiology can provide very valuable information on infection susceptibility in the population due to the absence of prior infection with an emerging virus and the effect of vaccination.[2]

The UK undertook a series of influenza serosurveys during the 2009 pandemic based on residual blood samples from the Public Health England (PHE) National Seroepidemiology Programme.[4] These are samples submitted to PHE and National Health Service (NHS) laboratories for routine diagnostic purposes. Although this work delivered critical information on background population seroprevalence and seroincidence, several issues were raised in post-pandemic reviews. These reviews highlighted that, although this information was gathered and published earlier than almost any other country, even earlier availability of this intelligence would have been critical to inform important national policy decisions. The following key recommendations have been made in relation to influenza seroepidemiology:

- The Science and Technology Committee (3rd report 2010-12) stated that seroepidemiological data need to be available earlier in the time course of a future pandemic to help with risk assessment (i.e. the likely spread of the disease across the population);[5]
- The Chief Medical Officer (CMO)-Statistics Legacy Group (SLG) determined that serosurveillance is critical to determine population immunity and community infection rates. These data cannot be obtained from other sources and are vital to making modelling predictions of the pandemic.[6]
- The 2011 UK Influenza Pandemic Preparedness Strategy includes seroepidemiology as a key surveillance initiative that will be required at the start of any pandemic, and states that work should be underway to enhance capability to respond, based on the H1N1 (2009) influenza pandemic.[7]

- The Scientific Pandemic Influenza Advisory Committee (SPI), Subgroup on Modelling (SPI-M) group recommended that the PHE strengthen population-based influenza seroepidemiology, including collection of key epidemiological information on vaccination status and underlying risk status.[8]
- Finally ECDC has highlighted the importance of influenza seroepidemiology, as has WHO as part of the Fineberg report into the pandemic response and the lack of preparedness including the need for a proper assessment of severity at national and subnational levels early in a pandemic.[9]

The RCGP RSC and PHE, and its predecessor organisations have an over 50 year history of collaboration in influenza and respiratory disease surveillance and vaccine effectiveness studies.[10] This study builds on this long term collaboration.

The RCGP RSC provides a suitable venue for collecting influenza serology data. This sentinel network is one of the longest established and has a nationally representative network of practices.[11] The network has a long history of feedback about data quality, particularly in the areas of influenza like illness (ILI) and other respiratory infections. The network also collects virology specimens to detect influenza in season – allowing matching in individuals between swab report and serology.[12] The RCGP RSC also collects data about vaccine exposure. Combining information about ILI, virology results and vaccine exposure enables the estimation of influenza vaccine effectiveness.[13] Its dashboard capability provides a method for near real time feedback to practices about sample collection.[14]

This pilot study tests whether collecting serology data from a sentinel network and linking it to virology, and clinical record data at the individual level would provide a low cost way of creating a high quality sero-epidemiological resource.[15]

Aim:

To pilot a mechanism to undertake population-based surveys that collect serological specimens that will be linked with key epidemiological information at strategic time points after each influenza season as a resource to be deployed in a future pandemic, for seasonal influenza and potentially other infections of public health importance.

Objectives:

- Establish a system that allows volunteer patients to provide a serology sample. The result of that sample will be linked to that patient's pseudonymised record and any influenza virology swab data if patient participated in the nasopharyngeal swabbing project. This will provide high quality data about vaccine exposure, and any medically reported influenza like illnesses (ILI) or other condition of scientific interest, included within an approved scientific protocol, or meeting public health needs in a pandemic.
- Pilot laboratory programme for the processing and storage of these specimens
- Pilot linkage of these biological specimens to the Public Health England Respiratory Virus Unit (RVU) analysis programme for the detection of influenza antibodies
- Link the results to epidemiological data in particular vaccination history, age and underlying clinical risk factor status and any swab results. The serology results can be linked to vaccine brand.
- Link the pilot results to the statistical and mathematical modelling work that is underway and compare to results from the existing residual sera programme to enable rapid analysis and interpretation of available epidemiological and laboratory data and in particular derive seroprevalence, seroincidence and case-severity end-points and see the feasibility of using seroprofile for predictive modelling of seasonal influenza
- At the end of the project, evaluate the feasibility of the work programme and make recommendations about how best to provide influenza seroepidemiology work at scale and in a future pandemic and apply these approaches to other vaccine preventable infections

- Estimate the costs of providing a national seroepidemiology service based on samples from a subset of RCGP RSC practices.

METHODS AND ANALYSIS

Study Design

Study setting and population:

The project will consist of four interlinked work-packages:
1) Population sampling and collection of biological specimens
2) Laboratory analysis
3) Data management
4) Statistical and modelling analysis.

1) Population sampling and collection of biological specimens

The RCGP RSC practices will collect the biological specimens:
The RCGP RSC will be used to collect the biological specimens. Practices who participated in the virology swabbing scheme will be invited.

The project proposes to pilot a population-based seroprevalence survey, involving 100-150 individuals across the following age range: 18-29, 30-39, 40-49, 50-59, 60-69, 70+ years, following the 2017/18 influenza season. Patients, who attend their pilot sentinel network practice for routine blood test during the study period will be asked to also provide an additional blood sample for serology. This will provide information on seroprevalence to set in context other measures of impact of an influenza season/epidemic in a population, and provide the most accurate measures of population exposure. This new approach to serology banking represents a compromise between using residual serum samples from laboratories where less is known about the patient's medical and immunisation history and formal surveys that can collect such data, but can have non- response bias.

Analysis carried out using the blood collected in RCGP RSC practices and sent to the Seroepidemiology Unit (SEU) archive. The SEU archive is an opportunistic collection of residual serum samples from routine microbiological testing, submitted voluntarily each year from laboratories throughout England. SEU archive sera are stored at the PHE North West regional laboratory in Manchester and are anonymised and permanently unlinked from any patient identifying information, with only age, gender, date of collection (if available) and contributing laboratory retained.

	All practices (5-6 practices) during the study period	
Age Band	100 sample threshold crossed after the following number of weeks	150 sample threshold crossed after the following number of weeks
18-29	3	5
30-39	3	5
40-49	2	2
50-59	1	2
60-69	1	2
70+	1	1

Table 1: **Blood test plan for the pilot study. This table represents the week in which sample collection would be complete, if all patients consented. The pilot study will take place as**

soon as ethical approval is achieved and would involve the collection of a total of up to 150 samples from approximately six practices over the study period.

Serology samples will be analysed at the Respiratory Virus Unit (RVU) at the PHE Colindale using HAI with representative vaccine strains. Each sample will be tested once and not in duplicate, the result will have strong identifiers pseudonymised (in accordance with current best practice) and be returned to the RCGP RSC hub's encrypted server, decrypted and linked to individual patients pseudonymised data.

The RCGP RSC will recruit six practices (depending on practice list size) to ascertain the feasibility and establish the approach. We will aim to collect an evenly distributed set of sera across all age bands that are set out in Table 1. Older people have more chronic disease and hence have higher number of blood tests performed than younger people; so the target number of samples would be achieved sooner. There will be no attempt to select patients for serology on the basis of whether they have had an influenza immunisation, or not or had a virology specimen taken for influenza. The number of samples is based on a combined list size of 50,000 across the participating practices.

The proposed method would involve feeding back to practices as each age band reaches the target number of samples. (minimum 100, maximum 150 per age band). University of Surrey has developed methods for RCGP RSC to give practice specific feedback about vaccine exposure and data quality - these methods could be used to provide feedback to practices participating in serum collection.

Patient selection:

Inclusion criteria: All patients of age 18 and over who visit their practice for a routine blood test and provide another sample for serology are eligible for inclusion in the analysis. The main inclusion criteria for practices is that practices are within our influenza swabbing practices' list with quota sampling according to the table above.

Exclusion criteria: Patients who have explicitly opted out of data sharing will be excluded from the analysis. We will identify these patients using the opt-out codes within GP information systems where the patients have made an explicit choice to opt out; patients will be informed of their option to opt-out via posters in the practices and information sheets.

2) Laboratory analysis

Samples will be submitted to the PHE Manchester laboratory in practice batches of clotted vacutainer bottles, and will be accompanied by the standard request form. This form will be generated by existing ICE pathology request software used by the GP practices.

The standard method for sending specimens to the laboratory is via a pathology request software such as an ICE system. When a patient requires a blood test, the GP can print a test request form via their ICE system. To send samples to PHE's SEU/VEU:

- The practice will need to print out an additional test request form for routine blood sampling
- The University of Surrey will provide the practices with detailed guidance on sample collection and postage to PHE's SEU/VEU

Upon receiving consent, practices will be able to send specimens to PHE's SEU/VEU lab via pre-paid envelopes. A Material Transfer Agreement (MTA) will be put in place for the transport of blood samples from GP practices to Manchester SEU/VEU.

We suggest a single Read code[16] is allocated to mark that the serology specimen has been consented to and sent. This will facilitate one-to-one matching of specimens with the consenting patient/subjects clinical history.

Whole blood samples will be received, processed to obtain sera, catalogued and stored in archive freezers within the Serum Archives section of the Vaccine Evaluation Unit (VEU) at Manchester according to existing SOPs, modified where necessary. It is envisioned that sera collected will be permanently transferred to the SEU archive once this study is complete.

Sera will be processed to undergo analysis at the Respiratory Virus Unit (RVU) of the National Infection Service of PHE (London, UK) using haemagglutination inhibition (HAI) assays and other measures of functional antibody status according to established protocols to detect antibody levels against relevant circulating and/or vaccine influenza strains.[17] Antigens will be grown in-house in egg and cell-culture and Influenza B antigens will be diethyl ether extracted as previously described.[18] Briefly, for the HAI assay, sera will be treated to remove non-specific inhibitors using receptor destroying enzyme (RDEII) and then twofold serially diluted starting at a 1:10 dilution, followed by mixing with an equal volume (25µl) of PBS containing 4HA units of each of the strains. Turkey red blood cells (RBC) will be used for the influenza A(H1N1)pdm09 and influenza B components, and Guinea pig RBC for H3N2. HAI titres will be expressed as the reciprocal of the last serum dilution that results in complete inhibition of agglutination. Additionally, analysis using influenza neutralisation and/or Neuraminidase Antibody Inhibition assays in the format of the Enzyme-linked Lectin Assay[16] may be considered to investigate the functionality of the antibody further.

Following analysis, the results will be reported in form of an Excel table (containing titres against each antigen/influenza virus for every sample) to the data manager of the project. The data (with pseudonymised strong identifiers) will be linked to RCGP RSC data to allow analysis of relevant data including vaccine exposure and previous influenza.

Using the UK laboratory bounded code list

UK laboratories are currently obligated to use the Pathology Bounded Code List (PBCL), a subset of NHS Digital Read codes, when electronically reporting pathology results to GPs, but are freely allowed to choose which PBCL codes are used for each test. A PBCL code assigned to the test (Table 2) will allow one week in areas for the RCGP RSC team to follow the collection of the samples by age-band.

V2_READ_CODE	V2_TERM
4JDb.	Influenza (A&B) serology
43L..	Sample serology

Table 2: Read codes to flag that a specimen has been taken

3) Data extraction and data management

Data collection from volunteer RCGP RSC practice

Data will be extracted from RCGP RSC databases. The RCGP RSC and PHE have worked together to provide state of the art National influenza surveillance for over 50 years.[10]

These databases store data received from participating RCGP practices. A UK general practice is a registration based system where all citizens can register with a single General Practice (GP) of their choice. Practices are computerised, and data entered into computerised medical record systems either as coded data, or free text. We will extract the coded data, and our results will be based on this element of the record. We will extract all coded data, pseudonymising as close to sources as possible. Where patients have a range of codes inserted in their record suggesting they opt out of record sharing we will not analyse their data.

The RCGP RSC will only extract coded data, e.g. where the GP or other health professional codes a disease or symptom into their computerised medical record system, and will process these data as required for this project.[16,19] Coded routine data from UK primary care has been widely used in research.[20]

Data collection methods will follow best practice, including pseudonymisation

These data extractions will be conducted in accordance with best practice, using the Clinical Informatics and Health Outcomes Research Group's standard operating procedures for data extraction, pseudonymisation, and transfer. The method and governance procedure has been developed by the University of Surrey, using an approved provider.

Pseudonymisation is the standard approach for protecting patient's privacy. It is a process that involves the removal of all personal identifiers from data – such as name, date of birth, etc. However, there is a risk that if data are linked to other data a person might be identified.[21,22] Therefore although all identifiers are removed we keep data encrypted during transfer and on a secure network that meets NHS Information Governance standards to minimise the risk of re-identification. A legally binding definition of pseudonymisation has been introduced into European law on the recommendation of the European Data Protection Supervisor (EDPS).[19]

We "pseudonymise" strong identifiers (in this study NHS number) so that we can link further data to the same individual's record. For this study we need, for example, to be able to link whether there have been any immune changes to the individual that had been vaccinated (and with the specific brand and batch number). Pseudonymisation allows us to do this without knowing any of the strong personal identifiers of that individual.

All data processing and analysis in the present proposed study will be conducted within the secure IT environment of the Clinical Informatics and Health Outcomes Research Group, at the University of Surrey. The information security policies and procedures of the Research Group have been approved by the NHS Digital as meeting the Information Governance Toolkit (IGT) standards.[23]

The following routinely collected patient data will be collected for the study:

- Demographic information: age, gender, ethnicity, registered date.
- Lower Super Output Area (LSOA): full postcodes will be automatically and immediately transformed into Lower Super Output Area (LSOA) which can be used for calculating deprivation scores, using the Index of Multiple Deprivation (IMD), within GP computer systems upon extraction. This would provide information about any inequities in access according to level of social deprivation using geographical information system (GIS) methods.
- Influenza vaccination – including date of vaccination and brand/lot
- Primary care consultations following vaccination, any other markers of health care utilisation, and referral to further care.
- Reactogenicity outcomes of seasonal influenza vaccination as listed in the research literature and any contemporary EU guidance.
- Life-style/risk factors – e.g. Body Mass Index (BMI), smoking status.
- Records of other diseases and long term conditions – e.g. chronic respiratory disease, chronic heart disease, chronic kidney disease, chronic liver disease, chronic neurological disease, diabetes, immunosuppression, pneumonia, etc.
- Pregnancy.
- PHE results of influenza HAI testing will be linked to this data-set at the University of Surrey according to established information governance (IG) procedures

Practices will receive weekly feedback via our dashboard system about progress to target within each age-band (Table 1).[14]

Data tables for analysis:
RCGP RSC will prepare tables for analysis as set out in the section below on statistical modelling and analysis. A schedule for reports would be created, as is currently carried out for reporting influenza vaccine effectiveness studies, and blank tables prepared in advance as outlined in this protocol.

4) Statistical and modelling analysis.
PHE will lead on this analysis and follow the outline plan set out below.

The proof-of-concept study will consist of a total of up to 100-150 participants per age-band from whom blood samples have been taken during the survey period.

The evaluation will include the following elements:

- A comparison of representativeness of the recruited study population will be undertaken in relation to national surveys.
- A comparison of the influenza seroprevalence results from the RCGP cohort against those from the annual PHE influenza residual serosurvey
- Demographic, epidemiological and lab data will be checked for completeness, errors and inconsistencies
- Samples with titres >40 by HAI will be considered seropositive to report prevalence of influenza antibody detected by HAI.[4] In addition, geometric mean titres and reverse cumulative distribution (RCD) curves will be calculated for some preliminary, exploratory comparisons of groups such as those vaccinated versus unvaccinated – although conclusions will be limited by the samples available.
- Acceptability to participate amongst practices and patients ascertained through measures such as response rate
- Feasibility of extending the pilot approach to a wider number of practices and collection of additional sample types e.g. oral fluid ascertained

University of Surrey will report the extent to which we have samples from common households – this might provide information about shared infections/levels of immunity within households.

We will produce a weekly report on progress to target by age-band; and provide feedback to practices and to patients about the utility of the data provided. This will be a simple table – based on Table 1, explaining the anticipated number of weeks to complete the sampling by age-band.

Project Management
The pilot project will be a collaborative project lead by Prof S de Lusignan at University of Surrey, with RCGP and PHE as collaborators. The RCGP-PHE scientific committee will oversee the project in collaboration with the principal investigator.

This research and information governance framework for RCGP RSC sits within the University of Surrey's formal frameworks for information and research governance. In addition, all externally funded projects and collaborative projects with external partners are supported and guided by the University's Research and Enterprise Support (RES) service. RES ensures that university-supported projects are financially viable, and that legal issues of knowledge transfer and intellectual properties are addressed. The project team is supported by IT services dedicated to the Faculty and to the Department of Clinical and Experimental Medicine. Our secure analysis servers are optimised for routine healthcare data processing, to provide faster deliveries for our projects.

ETHICS AND DISSEMINATION

Ethical Approval

PHE has ethical approval (05/Q0505/45) for the collection and use of unlinked and anonymised residual serum samples in cross-sectional antibody prevalence studies for the surveillance of population immunity to vaccine preventable diseases of public health importance and the collection has been extensively used for this purpose.

We will seek to collect serum samples from a cohort of patients that are registered with one of the RCGP RSC practices. Potential participants will be attending a pre-scheduled blood appointment, where the healthcare professional treating them will inform them about the study and seek their consent about whether they would be interested in donating an additional blood sample as part of the study.

Blood serum is acellular and not considered a material subject to the Human Tissue Act 2004.[24] However, practices will all need to put in place a Material Transfer Agreement with Public Health England prior to starting their surveillance.[25]

Information Governance

The Clinical Informatics and Health Outcomes Research Group at the University of Surrey has worked with routinely collected healthcare data in a number of research and evaluation projects for over 20 years.[26] The Research Group works within the research and Information Governance frameworks for health and social care in the United Kingdom, and is compliant with the University's best practice standards. The University of Surrey is registered with the Information Commissioner's Office Data Protection Register, and is compliant with the Data Protection Act, and other legislations.

In addition, the Research Group reviewed its departmental information governance policies and procedures, against the requirements of the NHS Information Governance Toolkit (IGT) for Hosted Secondary Use Team/ Project, Version 14.1.[20]

Dissemination and Public Register Disclosure

The outputs from the research will be disseminated primarily through peer review papers in high impact journals within the domains of primary care, surveillance, vaccines, and infectious diseases.[27,28] We will present findings at relevant seminars and conferences. The University of Surrey, in accordance with PHE policy, will post a summary of the study protocol and results within 12 months of study completion.

DISCUSSION

Strengths:

The strengths of this application are that it uses an established sentinel network, it builds on PHE expertise in serological analysis, and pilots a much lower cost method of establishing a serology bank.

The RCGP RSC is a sentinel network that collects and monitors data from primary care, particularly influenza and other respiratory illnesses, with some practices in the network that have been providing data for decades.[11] These practices have data quality that is as good as it gets in primary care and the practices are used to taking specimens.

PHE has expertise in serology, but many of serosurveys used residual serum samples from diagnostic blood tests[4] and hence the clinical information available with these is very limited. Detailed serology surveys are also conducted, but these are extremely expensive.

In this pilot the serology sampling has been designed to provide almost real time results which can be linked to pseudonymised patient data extracted by RCGP RSC. Conducting the sampling alongside routine blood tests reduces overall time for practices and costs overall. We will create a database with individual level serology, virology specimens confirming influenza or other respiratory disease diagnoses, details of past medical history and vaccine exposure data.

Limitations:

This pilot is limited by its design as a pilot within the available resource envelope and data quality. The limitations of this study are: its small size; limitation to adult specimens only; no specific targeting of those who have had previous virology or a specific points within the annual cycle of vaccination or of influenza infection; and data quality particularly of out-of-practice vaccine exposure.

The study is not powered to detect differences between groups, instead to demonstrate our ability to collect samples across all age groups. The ability to collect across all age-groups have been questioned – both by the original authors of the call for a World Serology Bank,[29] and set as a principal challenge by potential funders if the pilot is a success. Hence the focus of the pilot is on sample collection across age-groups.

We have not included children or young people under 18 years in this pilot. We feel this is appropriate for the first pilot of this type. In a substantive study this group may be important because children are important vectors of disease[30] and the UK is one of the few countries to systematically immunise the population using live attenuated intranasal vaccination (LAIV).[31]

This pilot is not targeting those who have had virology specimens, nor to be aligned with a particular point in the vaccination cycle. Generally vaccination takes place in the early autumn in the UK, with seasonal influenza starting to circulate around the year end. In a major study it may be possible to look at immunity in the population, exposed and unexposed to vaccine, and at residual immunity to the circulating strain of influenza.

Patients vaccinated against influenza outside of General Practice may not have information coded into their computerised medical record system at General Practice.[32] Vaccination data are sometimes missing, or sometimes incomplete as the standard reporting form from pharmacist to practice only indicate that the person has been vaccinated against flu, not which brand of vaccine or batch.

CONCLUSION

This pilot test our ability to collect samples across all adult age-bands within a sentinel network. If successful we will pursue resources to expand this into a larger study.

ACKNOWLEDGEMENTS

Patients for allowing their data to be used for surveillance and research. Practices who have agreed to be part of the RCGP RSC and allow us to extract and used health data for surveillance and research. Other members of the Clinical Informatics and Health Outcomes Research Group at University of Surrey. Apollo Medical Systems for data extraction. Collaboration with EMIS, TPP, In-Practice and Micro-test CMR supplier for facilitating data extraction. Colleagues at Public Health England.

LIST OF ABBREVIATIONS

BMI	Body Mass Index
CHPR	Centre for Health Protection Research
ECDC	European Centre for Disease Control
GIS	Geographical information system
GP	General Practitioner – A family physician providing NHS care to a registered list of patients
HAI	Haemagglutination inhibition assays
IG	Information Governance
IGT	Information Governance Toolkit – the NHS standard required for securely holding individual patient level daata
IMD	Index of Multiple Deprivation – a measure of socioeconomic status
JCVI	Joint Committee on Vaccination and Immunisation
LAIV	Live attenuated intranasal vaccination
LSOA	Lower Super Output Area
NHS	National Health Service
NIHR	National Institute for Health Research
PBCL	Pathology Bounded Code List
PHE	Public Health England
RCD	Reverse Cumulative Distribution
RCGP	Royal College of General Practitioners
RSC	Research and Surveillance Centre (within RCGP)
RVU	Respiratory Virus Unit
SEU	Seroepidemiology Unit
WHO	World Health Organization

REFERENCES

[1] Metcalf CJ, Farrar J, Cutts FT, Basta NE, Graham AL, Lessler J, Ferguson NM, Burke DS, Grenfell BT. Use of serological surveys to generate key insights into the changing global landscape of infectious disease. *Lancet*. 2016 Aug 13;388(10045):728-30. doi: 10.1016/S0140-6736(16)30164-7.

[2] Li ZN, Lin SC, Carney PJ, Li J, Liu F, Lu X, Liu M, Stevens J, Levine M, Katz JM, Hancock K. IgM, IgG, and IgA antibody responses to influenza A(H1N1)pdm09 hemagglutinin in infected persons during the first wave of the 2009 pandemic in the United States. *Clin Vaccine Immunol*. 2014 Aug;21(8):1054-60. doi: 10.1128/CVI.00129-14.

[3] Laurie KL, Huston P, Riley S, Katz JM, Willison DJ, Tam JS, Mounts AW, Hoschler K, Miller E, Vandemaele K, Broberg E, Van Kerkhove MD, Nicoll A. Influenza serological studies to inform public health action: best practices to optimise timing, quality and reporting. *Influenza Other Respir Viruses*. 2013;7(2):211-24. doi: 10.1111/j.1750-2659.2012.0370a.x.

[4] Hoschler K, Thompson C, Andrews N, Galiano M, Pebody R, Ellis J, Stanford E, Baguelin M, Miller E, Zambon M. Seroprevalence of influenza A(H1N1)pdm09 virus antibody, England, 2010 and 2011. *Emerg Infect Dis*. 2012 Nov;18(11):1894-7. doi: 10.3201/eid1811.120720.

[5] House of Commons Science and Technology Committee. Scientific advice and evidence in emergencies. Third Report of Session 2010–11. London, House of Commons, 2011. URL: http://www.publications.parliament.uk/pa/cm201011/cmselect/cmsctech/498/498.pdf (accessed 16 May 2018).

[6] Chief Medical Officer's Statistical Legacy Group (CMO-SLG) pandemic influenza preparedness programme. London; Department of Health 2010. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/215871/dh_122754.pdf (accessed 16 May 2018).

[7] DH Pandemic Influenza Preparedness Team . UK Influenza Pandemic Preparedness Strategy. London; Department of Health 2011. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213717/dh_131040.pdf (accessed 16 May 2018).

[8] Scientific Pandemic Influenza Advisory Committee (SPI-M) Modelling Summary. The Scientific Pandemic Influenza Advisory Committee, London, Department of Health, 2013. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/624155/SPI-M_influenza_modelling.pdf (accessed 16 May 2018).

[9] Fineberg HV. Pandemic preparedness and response--lessons from the H1N1 influenza of 2009. *N Engl J Med.* 2014 Apr 3;370(14):1335-42. doi:10.1056/NEJMra1208802.

[10] de Lusignan S, Correa A, Smith GE, Yonova I, Pebody R, Ferreira F, Elliot AJ, Fleming D. RCGP Research and Surveillance Centre: 50 years' surveillance of influenza, infections, and respiratory conditions. *Br J Gen Pract.* 2017 Oct;67(663):440-441. doi: 10.3399/bjgp17X692645.

[11] Correa A, Hinton W, McGovern A, van Vlymen J, Yonova I, Jones S, de Lusignan S. Royal College of General Practitioners Research and Surveillance Centre (RCGP RSC) sentinel network: a cohort profile. *BMJ Open*. 2016 Apr 20;6(4):e011092. doi:10.1136/bmjopen-2016-011092.

[12] de Lusignan S, Correa A, Pathirannehelage S, Byford R, Yonova I, Elliot AJ, Lamagni T, Amirthalingam G, Pebody R, Smith G, Jones S, Rafi I. RCGP Research and Surveillance Centre Annual Report 2014-2015: disparities in presentations to primary care. *Br J Gen Pract.* 2017 Jan;67(654):e29-e40.doi: 10.3399/bjgp16X688573.

[13] Pebody R, Warburton F, Ellis J, Andrews N, Potts A, Cottrell S, Johnston J, Reynolds A, Gunson R, Thompson C, Galiano M, Robertson C, Byford R, Gallagher N, Sinnathamby M, Yonova I, Pathirannehelage S, Donati M, Moore C, de Lusignan S, McMenamin J, Zambon M. Effectiveness of seasonal influenza vaccine for adults and children in preventing laboratory-confirmed influenza in primary care in the United Kingdom: 2015/16 end-of-season results. *Euro Surveill*. 2016 Sep 22;21(38). doi: 10.2807/1560-7917.ES.2016.21.38.30348.

[14] Pathirannehelage S, Kumarapeli P, Byford R, Yonova I, Ferreira F, de Lusignan S. Uptake of a Dashboard Designed to Give Realtime Feedback to a Sentinel Network About Key Data Required for Influenza Vaccine Effectiveness Studies. *Stud Health Technol Inform.* 2018;247:161-165.

[15] de Lusignan S, Correa A. Opportunities and challenges of a World Serum Bank. *Lancet*. 2017 Jan 21;389(10066):250-251. doi: 10.1016/S0140-6736(17)30046-6.

[16] de Lusignan S. Codes, classifications, terminologies and nomenclatures: definition, development and application in practice. *Inform Prim Care.* 2005;13(1):65-70.

[17] Hoschler K, Andrews NJ, Faust SN, Finn A, Pollard AJ, Snape MD, et al. Administration of AS03B-adjuvanted A(H1N1)pdm09 vaccine in children aged <3 years enhances antibody response to H3 and B viruses following a single dose of trivalent vaccine one year later. *Clin Infect Dis.* 2014 Jan;58(2):181-7. doi: 10.1093/cid/cit692.

[18] Höschler K, Southern J, Thompson C, Warburton F, Andrews NJ, Miller E, Zambon M. Responses to live attenuated influenza vaccine in children vaccinated previously with Pandemrix (ASO3(B) adjuvanted pandemic A/H1N1pdm09). *Vaccine*. 2018 May 17;36(21):3034-3040. doi: 10.1016/j.vaccine.2018.04.017.

[19] de Lusignan S, Metsemakers JF, Houwink P, Gunnarsdottir V, van der Lei J. Routinely collected general practice data: goldmines for research? A report of the European Federation for Medical Informatics Primary Care Informatics Working Group (EFMI PCIWG) from MIE2006, Maastricht, The Netherlands. *Inform Prim Care* 2006;14(3):203-9.

[20] Kousoulis AA, Rafi I, de Lusignan S. The CPRD and the RCGP: building on research success by enhancing benefits for patients and practices. *Br J Gen Pract*. 2015 Feb;65(631):54-5. doi: 10.3399/bjgp15X683353.

[21] De Meyer F, De Moor G, Reed-Fourquet L. Privacy Protection through pseudonymisation in eHealth. *Stud Health Technol Inform*. 2008;141:111-8.

[22] de Lusignan S. Effective pseudonymisation and explicit statements of public interest to ensure the benefits of sharing health data for research, quality improvement and health service management outweigh the risks. *Inform Prim Care.* 2014;21(2):61-3. doi: 10.14236/jhi.v21i2.68.

[23] NHS Digital. Information Governance Toolkit (IGT). URL: https://www.igt.hscic.gov.uk/ (accessed 16 May 2018).

[24] List of materials considered to be 'relevant material' under the Human Tissue Act 2004. Human Tissue Authority (HTA). URL: https://www.hta.gov.uk/policies/list-materials-considered-be-%E2%80%98relevant-material%E2%80%99-under-human-tissue-act-2004 (accessed 16 May 2018).

[25] Mascalzoni D, Dove ES, Rubinstein Y, Dawkins HJ, Kole A, McCormack P, Woods S, Riess O, Schaefer F, Lochmüller H, Knoppers BM, Hansson M. International Charter of principles for sharing bio-specimens and data. *Eur J Hum Genet.* 201 Jun;23(6):721-8. doi: 10.1038/ejhg.2014.197. Epub 2014 Sep 24. Erratum in: *Eur J Hum Genet.* 2016 Jul;24(7):1096.

[26] Clinical Informatics and Health Outcomes Research Group. URL: www.clininf.eu (accessed 16 May 2018).

[27] MRC Policy and Guidance on Sharing of Research Data from Population and Patient Studies (Sept 2011) URL: https://mrc.ukri.org/publications/browse/mrc-policy-and-guidance-on-sharing-of-research-data-from-population-and-patient-studies/ (accessed 16 May 2018).

[28] Economic Co-operation and Development (OECD). "Promoting Access to Public Research Data for Scientific, Economic and Social Development" 2007 URL: http://www.oecd.org/sti/sci-tech/38500813.pdf (accessed 16 May 2018).

[29] Metcalf CJ, Mina MJ, Winter AK, Grenfell BT. Opportunities and challenges of a World Serum Bank - Authors' reply. *Lancet.* 2017 Jan 21;389(10066):252. doi:10.1016/S0140-6736(17)30054-5.

[30] Gerlier L, Lamotte M, Dos Santos Mendes S, Damm O, Schwehm M, Eichner M. Estimates of the Public Health Impact of a Pediatric Vaccination Program Using an Intranasal Tetravalent Live-Attenuated Influenza Vaccine in Belgium. *Paediatr Drugs*. 2016 Aug;18(4):303-18. doi: 10.1007/s40272-016-0180-6.

[31] Pebody RG, Green HK, Andrews N, Boddington NL, Zhao H, Yonova I, Ellis J, Steinberger S, Donati M, Elliot AJ, Hughes HE, Pathirannehelage S, Mullett D, Smith GE, de Lusignan S, Zambon M. Uptake and impact of vaccinating school age children against influenza during a season with circulation of drifted influenza A and B strains, England, 2014/15. *Euro Surveill*. 2015;20(39). doi:10.2807/1560-7917.ES.2015.20.39.30029.

[32] de Lusignan S, Hoghton M, Rafi I. Flu vaccination by pharmacists leads to suboptimal medical records. *BMJ*. 2017 Nov 9;359:j5084. doi: 10.1136/bmj.j5084.

AUTHOR'S CONTRIBUTIONS

SdeL

Simon de Lusignan is the lead author who developed the study design and wrote the first draft, edited and contributed to subsequent versions of this protocol.

MT

Manasa Tripathy liaised with the participating practices, edited and contributed to the writing.

RB, EL, KH

Ray Borrow, Ezra Linley and Katja Hoschler helped develop methods, edited and contributed to subsequent versions of the protocol.

FF

Filipa Ferreira manages this study and extensively reviewed the protocol, and contributed to the writing.

MZ:

Maria Zambon reviewed the protocol and contributed to the writing

NA:

Nick Andrews reviewed the protocol and contributed to the writing

IY

Ivelina Yonova helped develop the practice recruitment methods, edited and contributed to the writing.

MH

Mariya Hriskova helped develop the practice recruitment methods, edited and contributed to the writing.

IR

Imran Rafi reviewed the protocol and contributed to the writing

RP

Richard Pebody helped develop methods and contributed to the subsequent versions of the protocol.

FUNDING STATEMENT

This work was supported by the University of Surrey. The funding was obtained as part of the Higher Education Funding Council for England's (HEFCE) Industrial Strategy funding allocated to the University.

COMPETING INTERESTS STATEMENT

Simon de Lusignan has received grant funding through University of Surrey from GSK to report vaccine adverse events, and attended advisory boards for Sanofi and Seqirus.

COMPONENTS OF THE ROSES-I STATEMENT FOR STANDARDIZATION OF THE REPORTING OF SEROEPIDEMIOLOGIC STUDIES

	Item No	STROBE Items	ROSES- I items	Reported on Page
Title and abstract	1	(a). Indicate the study's design with a commonly used term in the title or the abstract (b). Provide in the abstract an informative and balanced summary of what was done and what was found	ROSES-I 1.1: The term "seroepidemiologic," "seroepidemiology," "seroprevalence," or "seroincidence" should be applied to the study in the title or abstract, and the medical subject heading "Seroepidemiologic Studies" be used when the report is of a population-based serological survey.	Pages 3, 4
Introduction	2	Explain the scientific background and rationale for the investigation being reported	ROSES-I 2.1: State what is known about the kinetics of antibody rise, decay, and persistence following infection for the particular virus being studied and the justification for threshold antibody titers or changes in titers used to define evidence of infection ROSES-I 2.2: State what is known about the sensitivity and specificity of the antibody detection assay being used	Pages 5-11
	3	State specific objectives, including any prespecified hypotheses	ROSES-I 3.1: State the specific measure of occurrence that is being estimated, for example, point seroprevalence, cumulative incidence of infection, secondary infection risk	Page 6, 11
EPIDEMIOLOGIC METHODS				
Study design	4	Present key elements of study design early in the paper	ROSES-I 4.1: State which specific seroepidemiologic study design was chosen and why (see Table 1)	Cross- Sectional seroprevalance Page 7

	Item No	STROBE Items	ROSES- I items	Reported on Page
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, methods for sampling, and data collection	ROSES-I 5.1: Describe the timing of the biological sampling in relation to the disease epidemiology in the study population (the beginning, peak, and end of virus transmission) ROSES-I 5.2: Where known, describe the timing of biological sampling in individuals in relation to disease onset and to exposures of interest ROSES-I 5.3: State the interval between sequential biological samples (serial cross-sectional or longitudinal studies), or specify whether only a single sample was collected (cross-sectional study)	NA Methods described on pages 7-11 NA Single Samples Taken
Participants	6	(a). Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case–control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants (b). Cohort study—For matched studies, give matching criteria and the number of exposed and unexposed Case–control study—For matched studies, give matching criteria and the number of controls per case	ROSES-I 6.1: For case-ascertained transmission studies, describe the method of case ascertainment and criteria for defining a “case” ROSES-I 6.2: For household-or institution-based transmission studies, describe the definition of a household or the institution ROSES-I 6.3: For outbreak investigations involving serologic sampling, describe the setting in which the cases were identified, for example, village/ residential setting, occupational workplace ROSES-I 6.4: To aid the interpretation of seroepidemiologic studies of novel influenza A virus subtypes, the results from exposed populations should be compared with the results from unexposed populations. Efforts to validate the assay in virologically confirmed cases should be reported	NA

	Item No	STROBE Items	ROSES- I items	Reported on Page
Variables	7	Clearly define all outcomes, exposures, predictors, potential risk factors, and effect modifiers. Give diagnostic criteria, if applicable	ROSES-I 7.1 The median age and range for each exposure group should be reported ROSES-I 7.2: Describe the potential for immunization (specify vaccine and timing of vaccination in relationship to collection of serum), if applicable, to affect the outcome measures ROSES-I 7.3: Describe any known or potential immunological cross-reactivity that may bias the outcome measures ROSES-I 7.4: Describe illness definitions and methods for ascertaining the presence or absence of clinical illness in subjects	Page 7 NA Page 5 NA
Data sources/ measurement biases	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	ROSES-I 8.1: If relevant, describe measures taken to identify and record immunization history	Page 9-10
Bias	9	Describe any efforts to address the potential sources of bias	ROSES-I 9.1: If relevant, describe efforts to control for the potential effect of immunization on estimates of outcomes	Page 9-10
Study Size	10	Explain how the study size was arrived at	ROSES-I 10.1: Describe the baseline estimated seroprevalence at given antibody titers or incidence of infection and cite published literature to support these estimates	Pages 7-11

	Item No	STROBE Items	ROSES- I items	Reported on Page
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	ROSES-I 11.1: Describe the serological assay's limit of detection and how this limit is defined or calculated. Describe how samples with a result below or on the borderline of the limit were handled in the analysis ROSES-I 11.2: Describe and justify the titer or other result used to define "seropositivity," or the antibody titer change or change in other assay result used to define "seroconversion." Avoid the term "seroconversion" unless referring to change from undetectable to detectable antibody level. Otherwise report the fold-rise in titer. Avoid the term "infection" but report "seroprevalence at a titer of …." ROSES-I 11.3: If statements or inferences are made about protection from infection, describe what is known about the correlation between the assay results and protection from infection and illness	No Results available yet as this study is ongoing
Statistical methods	12	(a). Describe all statistical methods, including those used to control for confounding (b). Describe any methods used to examine subgroups and interactions (c). Explain how missing data were addressed (d). Cohort study—If applicable, explain how loss to follow up was addressed Case–control study—If applicable, explain how matching of cases and controls was addressed Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy (e). Describe any sensitivity analyses	ROSES-I 12.1: if relevant, state how the non-independence of data was managed ROSES-I 12.2: if relevant, report methods used to account for the probability of seropositivity or seroconversion if infected, and to account for decay in antibody titers over time	No Results available yet as this study is ongoing

	Item No	STROBE Items	ROSES- I items	Reported on Page
Laboratory methods	12a		ROSES-I 12a.1: Describe the sample type—serum or plasma. If plasma is used, specify the anticoagulant used (heparin, sodium citrate, EDTA, etc.) ROSES-I 12a.2: Describe the specimen storage conditions (4°C, −20 °C, −80 °C). If frozen prior to the analysis, describe the time to freezing and the number of freeze/thaw cycles prior to testing Serological assays ROSES-I 12a.3: Specify the assay type (e.g., hemagglutination inhibition; virus neutralization/microneutralization; ELISA; other) and methods used to determine the endpoint titer ROSES-I 12a.4: Reference a previously published, CONSISE consensus serologic assay or WHO protocol if used, and any modifications of the protocol. If a previously published protocol is not used, provide full details in supplementary materials ROSES-I 12a.5: State what is known about the determinants of the variability of the antibody detection assay being used ROSES-I 12a.6: Specify the antigen(s) used in the assay, including virus strain name, subtype, lineage or clade, with standardized nomenclature and reference; specify whether live virus or inactivated virus was used (where applicable) ROSES-I 12a.7: Report if antigen(s) from potentially cross-reactive pathogens/strains were used in order to identify cross-reactivity, and specify which antigen was used, including virus name, subtype, strain, lineage and clade, with standardized nomenclature and reference ROSES-I12a.8: If red blood cells were used for a hemagglutinin inhibition assay, specify the animal species from which they were obtained and concentration (v/v) used ROSES-I 12a.9: Describe positive and negative controls used ROSES-I 12a.10: Describe starting and end dilutions ROSES-I 12a.11: Specify laboratory biosafety conditions ROSES-I 12a.12: Specify whether replication was performed, and if so, the acceptable replication parameters ROSES-I 12a.13: Specify whether a confirmatory assay was performed and all specifics of this assay, at the same level of detail ROSES-I 12a.14: Specify international standards used, if appropriate	Page 8-9
RESULTS				
Participants	13	(a). Report the numbers of individuals at each stage of the study—the numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed (b). Give reasons for non-participation at each stage (c). Consider use of a flow diagram	See STROBE item	No Results available yet as this study is Ongoing

	Item No	STROBE Items	ROSES- I items	Reported on Page
Descriptive data	14	a). Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential risk factors (b). Indicate the number of participants with missing data for each variable of interest (c). Cohort study—summarize follow-up time (e.g., average and total amount)	See STROBE item	Page 7
Outcome data	15*	Cohort study—report the numbers of outcome events or summary measures over time Case–control study—report the numbers in each exposure category, or summary measures of exposure Cross-sectional study—report the numbers of outcome events or summary measures	See STROBE item	NA
Main results	16	(a). Give unadjusted estimates and, if applicable, risk factor-adjusted estimates andtheir precision (e.g., 95% confidence interval). Make clear which risk factors were adjusted for and why they were included (b). Report category boundaries when continuous variables were categorized (c). If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	ROSES-I 16.1: Report unadjusted estimates of distribution of titers by age Group ROSES-I 16.2: Report methods to standardize the results from the study sample to the target population	No results available yet as this study is ongoing
Other analyses	17	Report other analyses performed—analyses of subgroups and interactions, and sensitivity analyses	See STROBE item	NA
DISCUSSION				
Key results	18	Summarize key results with reference to study objectives	See STROBE item	NA
Interpretation	19	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	ROSES-I 20.1: Discuss the interpretation of the results in the context of known or potential cross-reactivity	NA

	Item No	STROBE Items	ROSES- I items	Reported on Page
Generalizability	21	Discuss the generalizability (external validity) of the study results	See STROBE item	NA
Other information **Funding**	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	ROSES-I 22.1: Specify if institutional review board approval was received; if not, specify reason (e.g., public health outbreak response/non-research designation)	Page 18

BMJ Open

Serological surveillance of influenza in an English sentinel network: pilot study protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-024285.R1
Article Type:	Protocol
Date Submitted by the Author:	22-Dec-2018
Complete List of Authors:	de Lusignan, Simon; University of Surrey, Department of Clinical and Experimental Medicine; St. Georges University of London, Division of Population Health Sciences and Education Borrow, Ray; Public Health England, Vaccine Evaluation Unit, Manchester Royal Infirmary Tripathy, Manasa ; University of Surrey, Department of Clinical and Experimental Medicine Linley, Ezra; Public Health England, Vaccine Evaluation Unit, Manchester Royal Infirmary Zambon, Maria; Public Health England, Hoschler, Katja; Public Health England Ferreira, Filipa; University of Surrey, Department of Clinical and Experimental Medicine Andrews, Nick; Public Health England, Modelling and Economics Department Yonova, Ivelina; University of Surrey, Department of Clinical and Experimental Medicine; Royal College of General Practitioners, CIRC Hriskova, Mariya ; University of Surrey, Department of Clinical and Experimental Medicine; Royal College of General Practitioners, CIRC Rafi, Imran; Royal College of General Practitioners, Pebody, Richard; Public Health England, Centre for Infectious Disease Surveillance and Control
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Public health, Global health, General practice / Family practice, Infectious diseases
Keywords:	Medical Records System, Computerized, Serology, Data Collection, Population Surveillance, Records as topic, Influenza, Human

Serological surveillance of influenza in an English sentinel network: pilot study protocol

Simon de Lusignan[1,2]
Professor of Primary Care and Clinical Informatics
Head of Department of Clinical and Experimental Medicine
s.lusignan@surrey.ac.uk
SdeL

Ray Borrow[4]
Head of the Vaccine Evaluation Unit
Ray.Borrow@phe.gov.uk
RB

Manasa Tripathy[1]
Practice Liaison Officer
m.tripathy@surrey.ac.uk
MT

Ezra Linley[4]
Deputy Head of Vaccine Evaluation Unit
Ezra.Linley@phe.gov.uk
EL

Maria Zambon[3]
Head of Reference Microbiology
Maria.Zambon@phe.gov.uk
MZ

Katja Hoschler[3]
Senior Scientist, Respiratory Virus Unit
Katja.hoschler@phe.gov.uk
KH

Filipa Ferreira[1]
Project Manager
f.ferreira@surrey.ac.uk
FF

Nick Andrews[3]
Senior statistician
nick.andrews@phe.gov.uk
NA

Ivelina Yonova[1,2]
Practice Liaison Officer
i.yonova@surrey.ac.uk
IY

Mariya Hriskova[1,2]
Practice Liaison Officer
m.hriskova@surrey.ac.uk
MH

Imran Rafi[2]
Chair of Clinical Innovation and Research Centre
Irafi@rcgp.org.uk
IR

Richard Pebody[3]
Head of Flu Surveillance
richard.pebody@phe.gov.uk
RP

[1] Department of Clinical and Experimental Medicine, University of Surrey, GUILDFORD, GU2 7XH, UK
[2] Royal College of General Practitioners Research and Surveillance Centre, Euston Square, LONDON,
[3] Public Health England, London, UK
[4] Vaccine Evaluation Unit, Public Health England, Manchester Royal Infirmary, Manchester, UK

Correspondence to:
Professor Simon de Lusignan
s.lusignan@surrey.ac.uk
T: 01483 683089
Section of Clinical Medicine & Ageing
Department of Clinical and Experimental Medicine
University of Surrey
GUILDFORD, GU2 7XH, UK

Word Count: 5847

Serological surveillance of influenza in an English sentinel network: pilot study protocol

ABSTRACT

Introduction

Background:
Rapidly undertaken age-stratified serology studies can produce valuable data about a new emerging infection including background population immunity and seroincidence during an influenza pandemic. Traditionally seroepidemiology studies have used surplus laboratory sera with little or no clinical information or have been expensive detailed population based studies. We propose collecting population based sera from the Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC), a sentinel network with extensive clinical data.

Aim:
To pilot a method for provision of nationally representative serum samples and associated patient data to measure sero-positivity and seroincidence due to seasonal influenza and other infections of public health importance, and create a population based serology bank for investigation of other important infections.

Methods and Analysis:

Setting and Participants: We will recruit 6 RCGP RSC practices already taking nasopharyngeal virology swabs. Patients who attend for a scheduled blood test will be consented to donate additional blood samples. Approximately 100-150 residual blood samples will be collected from each of the following age-bands - 18-29, 30-39, 40-49, 50-59, 60-69, 70+ years.

Methods: We will send the samples to the Public Health England (PHE) Seroepidemiology Unit (SEU) for processing and storage. These samples will be tested for influenza antibodies, using haemagglutination inhibition assays (HAI). Serology results will be pseudonymised, sent to the RCGP RSC and combined using existing processes at the RCGP RSC secure hub. The influenza seroprevalence results from the RCGP cohort will be compared against those from the annual PHE influenza residual serosurvey.

Ethics and dissemination

Ethical approval was granted by the Proportionate Review Sub- Committee of the London – Camden & Kings Cross on 6 February 2018. This study received approval from Health Research Authority on 7 February 2018. On completion the results will be made available via peer- reviewed journals.

Strengths and Limitations of this study

Strengths:

- The Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) is one of the oldest sentinel networks in Europe, it has completed its 51st season of surveillance.
- Practices in RCGP take microbiological samples, including influenza virology specimens which are linked at individual level.
- We have the potential through this network to link vaccine exposure and serology.

Limitations:

- This is a pilot study to demonstrate we can collect samples across pre-defined adult age-groups - we will not collect samples from children, who may be exposed to live attenuated intranasal vaccine (LAIV) in recent years.
- Pharmacist and other non-GP vaccination can lead to missing information on the computerised medical record system at General Practice.

Keywords:

Medical records systems, computerized
Population Surveillance
Serology
Influenza, human
Data collection
Records as topic
Seroepidemiologic Studies
Primary Health Care
Pandemics

INTRODUCTION

Rationale for the pilot study and background

There have been calls for a World Serology Bank as serology could tell us much about the susceptibility of the population to infectious disease.[1] This pilot study explores the potential to establish a serology bank, based on a sentinel network, and focussed on influenza.

Serological data potentially allow the assessment of the severity of a new influenza strain by providing the capability to detect asymptomatic and mild infections and thus determine the symptomatic proportion.[2] The number of infections can be determined if the age-specific prevalence of immunity prior to and then during and after the pandemic are known. Thus, the number of people infected (and therefore no longer susceptible) can be calculated. If these data are available early (particularly the background population immunity and the symptomatic proportion), they can be used to adjust planning assumptions and to help predict the impact of the pandemic on health care services and optimal intervention strategies.

Serum archives would help assess the severity of a novel influenza virus and allow modification of local, national and international pandemic plans.[3] This requirement was a lesson from the 2009 A(H1N1) pandemic and its importance is recognised by the World Health Organization (WHO), the European Centre for Disease Control (ECDC), the UK Department of Health (DH), and the Joint Committee on Vaccination and Immunisation (JCVI).

Despite its apparent merit, the use of serology for seasonal influenza surveillance has several limitations. Seasonal influenza infections occur regularly and individuals are re-infected throughout life with related strains. The prior (lifetime) exposure of individuals to circulating influenza viruses in previous years and a high degree of cross-reactivity between antibodies to different seasonal influenza virus strains can lead to results which are difficult to interpret. Distinguishing between recent infection and recent vaccination also presents further difficulties. Consequently, seroepidemiological studies have not often been used for the investigation of seasonal influenza at a population level.

The UK undertook a series of influenza serosurveys during the 2009 pandemic based on residual blood samples from the Public Health England (PHE) National Seroepidemiology Programme.[4] These are samples submitted to PHE and National Health Service (NHS) laboratories for routine diagnostic purposes. Although this work delivered critical information on background population seroprevalence and seroincidence, several issues were raised in post-pandemic reviews. These reviews highlighted that, although this information was gathered and published earlier than almost any other country, even earlier availability of this intelligence would have been critical to inform important national policy decisions. The following key recommendations have been made in relation to influenza seroepidemiology:

- The Science and Technology Committee (3rd report 2010-12) stated that seroepidemiological data need to be available earlier in the time course of a future pandemic to help with risk assessment (i.e. the likely spread of the disease across the population);[5]
- The Chief Medical Officer (CMO)-Statistics Legacy Group (SLG) determined that serosurveillance is critical to determine population immunity and community infection rates. These data cannot be obtained from other sources and are vital to making modelling predictions of the pandemic.[6]
- The 2011 UK Influenza Pandemic Preparedness Strategy includes seroepidemiology as a key surveillance initiative that will be required at the start of any pandemic, and states that work should be underway to enhance capability to respond, based on the H1N1 (2009) influenza pandemic.[7]
- The Scientific Pandemic Influenza Advisory Committee (SPI), Subgroup on Modelling (SPI-M) group recommended that the PHE strengthen population based influenza seroepidemiology, including collection of key epidemiological information on vaccination status and underlying risk status.[8]

- Finally ECDC has highlighted the importance of influenza seroepidemiology, as has WHO as part of the Fineberg report into the pandemic response and the lack of preparedness including the need for a proper assessment of severity at national and subnational levels early in a pandemic.[9]

The RCGP RSC and PHE, and its predecessor organisations have an over 50 year history of collaboration in influenza and respiratory disease surveillance and vaccine effectiveness studies.[10] This study builds on this long term collaboration.

The RCGP RSC provides a suitable venue for collecting influenza serology data. This sentinel network is one of the longest established and has a nationally representative network of practices.[11] The network has a long history of feedback about data quality, particularly in the areas of influenza like illness (ILI) and other respiratory infections.[12] The RCGP RSC also collects data about vaccine exposure. Combining information about ILI, virology results and vaccine exposure enables the estimation of influenza vaccine effectiveness.[13] Its dashboard capability provides a method for near real time feedback to practices about sample collection.[14]

This pilot study tests whether collecting serology data from a sentinel network and linking it to clinical record data at the individual level would provide a low cost way of creating a high quality sero-epidemiological resource.[15]

Aim:

To pilot a mechanism to undertake population based surveys that collect serological specimens that will be linked with key epidemiological information at strategic time points after each influenza season as a resource to be deployed in a future pandemic, for seasonal influenza and potentially other infections of public health importance.

Objectives:

- Establish a system that allows volunteer patients from practices to provide a serology sample during a routine blood test. The result of that sample will be linked to that patient's pseudonymised record which will provide high quality data about vaccine exposure, and any medically reported influenza like illnesses (ILI) or other condition of scientific interest, included within an approved scientific protocol, or meeting public health needs in a pandemic.
- Measure the feasibility of collecting 100-150 samples from each of the following age-bands - 18-29, 30-39, 40-49, 50-59, 60-69, 70+ years.
- Pilot laboratory programme for the processing and storage of these specimens
- Pilot linkage of these biological specimens to the Public Health England Respiratory Virus Unit (RVU) analysis programme for the detection of influenza antibodies
- Link the results to epidemiological data in particular vaccination history, age and underlying clinical risk factor status and any swab results. The serology results can be linked to vaccine brand.
- Estimate the costs of providing a national seroepidemiology service based on samples from a subset of RCGP RSC practices.

METHODS AND ANALYSIS

Study Design

Study setting and population:

The project will consist of four interlinked work-packages:
1) Population sampling and collection of biological specimens

2) Laboratory analysis
3) Data management
4) Statistical and modelling analysis.

1) Population sampling and collection of biological specimens

The RCGP RSC practices will collect the biological specimens:
The RCGP RSC will be used to collect the biological specimens. Practices who participated in the virology swabbing scheme will be invited.

The project proposes to pilot a population-based seroprevalence survey, involving 100-150 individuals across the following age range: 18-29, 30-39, 40-49, 50-59, 60-69, 70+ years, following the 2017/18 influenza season. Patients, who attend their pilot sentinel network practice for routine blood test during the study period will be asked to also provide an additional blood sample for serology. This will provide information on seroprevalence to set in context other measures of impact of an influenza season/epidemic in a population, and provide the most accurate measures of population exposure. This new approach to serology banking represents a compromise between using residual serum samples from laboratories where less is known about the patient's medical and immunisation history and formal surveys that can collect such data, but can have non- response bias.

Analysis will be carried out using the blood collected in RCGP RSC practices and sent to the Seroepidemiology Unit (SEU) archive. The SEU archive is a collection of anonymised residual serum samples from routine microbiological testing, submitted voluntarily each year from laboratories throughout England. SEU archive sera are stored at the PHE North West regional laboratory in Manchester and are anonymised and permanently unlinked from any patient identifying information, with only age, gender, date of collection (if available) and contributing laboratory retained.

	All practices (5-6 practices) during the study period	
Age Band	100 sample threshold crossed after the following number of weeks	150 sample threshold crossed after the following number of weeks
18-29	3	5
30-39	3	5
40-49	2	2
50-59	1	2
60-69	1	2
70+	1	1

Table 1: Blood test plan for the pilot study. This table represents the week in which sample collection would be complete, if all patients consented. The pilot study will take place as soon as ethical approval is achieved and would involve the collection of a total of up to 150 samples from approximately six practices over the study period.

Serology samples will be analysed at the Respiratory Virus Unit (RVU) at the PHE Colindale using HAI with representative vaccine strains. Each sample will be tested once and not in duplicate, the result will have strong identifiers pseudonymised (in accordance with current best practice) and be returned to the RCGP RSC hub's encrypted server, decrypted and linked to individual patients pseudonymised data.

The RCGP RSC will recruit six practices (depending on practice list size) to ascertain the feasibility and establish the approach. We will aim to collect an evenly distributed set of sera across all age bands that are set out in Table 1. Older people have more chronic disease and hence have higher number of blood tests performed

than younger people; so the target number of samples would be achieved sooner. There will be no attempt to select patients for serology on the basis of whether they have had an influenza immunisation, or not or had a virology specimen taken for influenza. The number of samples is based on a combined list size of 50,000 across the participating practices.

The proposed method would involve feeding back to practices as each age band reaches the target number of samples. (minimum 100, maximum 150 per age band). University of Surrey has developed methods for RCGP RSC to give practice specific feedback about vaccine exposure and data quality - these methods could be used to provide feedback to practices participating in serum collection.

Patient selection:

Inclusion criteria: All patients of age 18 and over who visit their practice for a routine blood test and provide another sample for serology are eligible for inclusion in the analysis. The main inclusion criteria for practices is that practices are within our influenza swabbing practices' list with quota sampling according to the table above.

Exclusion criteria: Patients who have explicitly opted out of data sharing will be excluded from the analysis. We will identify these patients using the opt-out codes within GP information systems where the patients have made an explicit choice to opt out; patients will be informed of their option to opt-out via posters in the practices and information sheets.

2) Laboratory analysis

Samples will be submitted to the PHE Manchester laboratory in practice batches of clotted vacutainer bottles, and will be accompanied by the standard request form. This form will be generated by existing ICE pathology request software used by the GP practices.

General Practices use a pathology request software, such as the ICE system to send biological specimens to their local laboratory for testing. When a patient requires a blood test, the GP can print a test request form via their ICE system. To send samples to PHE's SEU/VEU:

- The practice will need to print out an additional test request form for routine blood sampling
- The University of Surrey will provide the practices with detailed guidance on sample collection and postage to PHE's SEU/VEU

Upon receiving consent, practices will be able to send specimens to PHE's SEU/VEU lab via pre-paid envelopes. A Material Transfer Agreement (MTA) will be put in place for the transport of blood samples from GP practices to Manchester SEU/VEU.

We suggest a single Read code[16] is allocated to mark that the serology specimen has been consented to and sent. This will facilitate one-to-one matching of specimens with the consenting patient/subjects clinical history.

Whole blood samples will be received, processed to obtain sera, catalogued and stored in -80° C archive freezers within the Serum Archives section of the Vaccine Evaluation Unit (VEU) at Manchester according to existing SOPs, modified where necessary. It is envisioned that sera collected will be permanently transferred to the SEU archive once this study is complete.

Sera will be processed to undergo analysis at the Respiratory Virus Unit (RVU) of the National Infection Service of PHE (London, UK) using haemagglutination inhibition (HAI) assays and other measures of functional antibody status according to established protocols to detect antibody levels against relevant circulating and/or vaccine influenza strains.[17] Antigens will be grown in-house in egg and cell-culture and Influenza B antigens

will be diethyl ether extracted as previously described.[18] Briefly, for the HAI assay, sera will be treated to remove non-specific inhibitors using receptor destroying enzyme (RDEII) and then twofold serially diluted starting at a 1:10 dilution, followed by mixing with an equal volume (25µl) of PBS containing 4HA units of each of the strains. Turkey red blood cells (RBC) will be used for the influenza A(H1N1)pdm09 and influenza B components, and Guinea pig RBC for H3N2. HAI titres will be expressed as the reciprocal of the last serum dilution that results in complete inhibition of agglutination. Where sample volumes after completion of the HAI permit, additional analysis using influenza neutralisation and/or Neuraminidase Antibody Inhibition assays in the format of the Enzyme-linked Lectin Assay [16] may be considered to investigate the functionality of the antibody further.

Following analysis, the results will be reported in form of an Excel table (containing titres against each antigen/influenza virus for every sample) to the data manager of the project. The data (with pseudonymised strong identifiers) will be linked to RCGP RSC data to allow analysis of relevant data including vaccine exposure and previous influenza.

Using the UK laboratory bounded code list

UK laboratories are currently obligated to use the Pathology Bounded Code List (PBCL), a subset of NHS Digital Read codes, when electronically reporting pathology results to GPs, but are freely allowed to choose which PBCL codes are used for each test. A PBCL code assigned to the test (Table 2) will allow one week in areas for the RCGP RSC team to follow the collection of the samples by age-band.

V2_READ_CODE	V2_TERM
4JDb.	Influenza (A&B) serology
43L..	Sample serology

Table 2: Read codes to flag that a specimen has been taken

3) Data extraction and data management

Data collection from volunteer RCGP RSC practice

Data will be extracted from RCGP RSC databases which store pseudonymised data received from participating RCGP practices.[10] A UK general practice is a registration based system where all citizens can register with a single General Practice (GP) of their choice. Practices are computerised, and data entered into computerised medical record systems either as coded data, or free text.

Practices will code 43L.. (Sample serology) into patient record when a patient consents to providing a sample for Serology. We will extract the coded data, and our results will be based on this element of the record. We will extract all coded data, pseudonymising as close to sources as possible. Where patients have a range of codes inserted in their record suggesting they opt out of record sharing we will not analyse their data.

The RCGP RSC will only extract coded data, e.g. where the GP or other health professional codes a disease or symptom into their computerised medical record system, and will process these data as required for this project.[16,19] Coded routine data from UK primary care has been widely used in research.[20]

Data collection methods will follow best practice, including pseudonymisation

These data extractions will be conducted in accordance with best practice, using the Clinical Informatics and Health Outcomes Research Group's standard operating procedures for data extraction, pseudonymisation, and transfer. The method and governance procedure has been developed by the University of Surrey, using an approved provider.

Pseudonymisation is the standard approach for protecting patient's privacy. It is a process that involves the removal of all personal identifiers from data – such as name, date of birth, etc. However, there is a risk that if

data are linked to other data a person might be identified.[21,22] Therefore although all identifiers are removed we keep data encrypted during transfer and on a secure network that meets NHS Information Governance standards to minimise the risk of re-identification. A legally binding definition of pseudonymisation has been introduced into European law on the recommendation of the European Data Protection Supervisor (EDPS).[19]

We "pseudonymise" strong identifiers (in this study NHS number) so that we can link further data to the same individual's record. For this study we need, for example, to be able to link whether there have been any immune changes to the individual that had been vaccinated (and with the specific brand and batch number). Pseudonymisation allows us to do this without knowing any of the strong personal identifiers of that individual.

All data processing and analysis in the present proposed study will be conducted within the secure IT environment of the Clinical Informatics and Health Outcomes Research Group, at the University of Surrey. The information security policies and procedures of the Research Group have been approved by the NHS Digital as meeting the Information Governance Toolkit (IGT) standards.[23]

The following routinely collected patient data will be collected for the study:

- Demographic information: age, gender, ethnicity, registered date.
- Lower Super Output Area (LSOA): full postcodes will be automatically and immediately transformed into Lower Super Output Area (LSOA) which can be used for calculating deprivation scores, using the Index of Multiple Deprivation (IMD), within GP computer systems upon extraction. This would provide information about any inequities in access according to level of social deprivation using geographical information system (GIS) methods.
- Influenza vaccination – including date of vaccination and brand/lot
- Primary care consultations following vaccination, any other markers of health care utilisation, and referral to further care.
- Reactogenicity outcomes of seasonal influenza vaccination as listed in the research literature and any contemporary EU guidance.
- Life-style/risk factors – e.g. Body Mass Index (BMI), smoking status.
- Records of other diseases and long term conditions – e.g. chronic respiratory disease, chronic heart disease, chronic kidney disease, chronic liver disease, chronic neurological disease, diabetes, immunosuppression, pneumonia, etc.
- Pregnancy.
- PHE results of influenza HAI testing will be linked to this data-set at the University of Surrey according to established information governance (IG) procedures

Practices will receive weekly feedback via our dashboard system about progress to target within each age-band (Table 1).[14]

Data tables for analysis:

RCGP RSC will prepare tables for analysis as set out in the section below on statistical modelling and analysis. A schedule for reports would be created, as is currently carried out for reporting influenza vaccine effectiveness studies, and blank tables prepared in advance as outlined in this protocol.

4) Statistical and modelling analysis.

PHE will lead on this analysis and follow the outline plan set out below.

The proof-of-concept study will consist of a total of up to 100-150 participants per age-band from whom blood samples have been taken during the survey period.

The evaluation will include the following elements:

- A comparison of representativeness of the recruited study population will be undertaken in relation to national surveys.
- A comparison of the influenza seroprevalence results from the RCGP cohort against those from the annual PHE influenza residual serosurvey
- Demographic, epidemiological and lab data will be checked for completeness, errors and inconsistencies
- Samples with titres >40 by HAI will be considered seropositive to report prevalence of influenza antibody detected by HAI.[4] In addition, geometric mean titres and reverse cumulative distribution (RCD) curves will be calculated for some preliminary, exploratory comparisons of groups such as those vaccinated versus unvaccinated – although conclusions will be limited by the samples available.
- Acceptability to participate amongst practices and patients ascertained through measures such as response rate
- Feasibility of extending the pilot approach to a wider number of practices

University of Surrey will report the extent to which we have samples from common households – this might provide information about shared infections/levels of immunity within households.

We will produce a weekly report on progress to target by age-band; and provide feedback to practices about the utility of the data provided. This will be a simple table – based on Table 1, explaining the anticipated number of weeks to complete the sampling by age-band.

5) Patient and Public Involvement

No patients or public were involved in the development of the research question or design of this study.

Project Management

The pilot project will be a collaborative project lead by Prof S de Lusignan at University of Surrey, with RCGP and PHE as collaborators. The RCGP-PHE scientific committee will oversee the project in collaboration with the principal investigator.

This research and information governance framework for RCGP RSC sits within the University of Surrey's formal frameworks for information and research governance. In addition, all externally funded projects and collaborative projects with external partners are supported and guided by the University's Research and Enterprise Support (RES) service. RES ensures that university-supported projects are financially viable, and that legal issues of knowledge transfer and intellectual properties are addressed. The project team is supported by IT services dedicated to the Faculty and to the Department of Clinical and Experimental Medicine. Our secure analysis servers are optimised for routine healthcare data processing, to provide faster deliveries for our projects.

ETHICS AND DISSEMINATION

Ethical Approval

PHE has ethical approval (05/Q0505/45) for the collection and use of unlinked and anonymised residual serum samples in cross-sectional antibody prevalence studies for the surveillance of population immunity to vaccine preventable diseases of public health importance and the collection has been extensively used for this purpose.

We will seek to collect serum samples from a cohort of patients that are registered with one of the RCGP RSC practices. Potential participants will be attending a pre-scheduled blood appointment, where the healthcare professional treating them will inform them about the study and seek their consent about whether they would be interested in donating an additional blood sample as part of the study.

Ethical approval was granted by the Proportionate Review Sub- Committee of the London – Camden & Kings Cross on 6 February 2018. This study received approval from Health Research Authority on 7 February 2018.

Blood serum is acellular and not considered a material subject to the Human Tissue Act 2004.[24] However, practices will all need to put in place a Material Transfer Agreement with Public Health England prior to starting their surveillance.[25]

Information Governance

The Clinical Informatics and Health Outcomes Research Group at the University of Surrey has worked with routinely collected healthcare data in a number of research and evaluation projects for over 20 years.[26] The Research Group works within the research and Information Governance frameworks for health and social care in the United Kingdom, and is compliant with the University's best practice standards. The University of Surrey is registered with the Information Commissioner's Office Data Protection Register, and is compliant with the Data Protection Act, and other legislations.

In addition, the Research Group reviewed its departmental information governance policies and procedures, against the requirements of the NHS Information Governance Toolkit (IGT) for Hosted Secondary Use Team/ Project, Version 14.1.[20]

Dissemination and Public Register Disclosure

The outputs from the research will be disseminated primarily through peer review papers in high impact journals within the domains of primary care, surveillance, vaccines, and infectious diseases.[27,28] We will present findings at relevant seminars and conferences. The University of Surrey, in accordance with PHE policy, will post a summary of the study protocol and results within 12 months of study completion.

DISCUSSION

Strengths:

The strengths of this application are that it uses an established sentinel network and builds on PHE expertise in serological analysis to establish a serology bank.

The RCGP RSC is a sentinel network that collects and monitors data from primary care, particularly influenza and other respiratory illnesses, with some practices in the network that have been providing data for decades.[11] These practices have data quality that is as good as it gets in primary care and the practices are used to taking specimens.

PHE has expertise in serology, but many of serosurveys used residual serum samples from diagnostic blood tests[4] and hence the clinical information available with these is very limited. Detailed serology surveys are also conducted, but these are extremely expensive.

In this pilot the serology sampling has been designed to provide almost real time results which can be linked to pseudonymised patient data extracted by RCGP RSC. Conducting the sampling alongside routine blood tests reduces overall time for practices and costs overall.

Collection from the same household may results in selection bias as it is likely that they will have had similar exposures, however it is also a possible strength of the network in trying to understand more about transmission within households or communal establishments, such as old peoples' homes.

We will create a database with individual level serology, virology specimens (if any) confirming influenza or other respiratory disease diagnoses, details of past medical history and vaccine exposure data.

Limitations:

This pilot is limited by its design as a pilot within the available resource envelope and data quality. The limitations of this study are: its small size; limitation to adult specimens only; no specific targeting of those who have had previous virology or at specific points within the annual cycle of vaccination or of influenza infection; and data quality particularly of out-of-practice vaccine exposure.

We set an arbitrary collection strategy across adult age-bands. However, this distribution is not representative of our practices populations' age distribution. Any full scale study would set out to represent the age-sex profile of the population and its geographical distribution.

The study is not powered to detect differences between groups, instead to demonstrate our ability to collect samples across all age groups. The ability to collect across all age-groups have been questioned – both by the original authors of the call for a World Serology Bank,[29] and set as a principal challenge by potential funders if the pilot is a success. Hence the focus of the pilot is on sample collection across age-groups.

We have not included children or young people under 18 years in this pilot. We feel this is appropriate for the first pilot of this type. In a substantive study this group may be important because children are important vectors of disease[30] and the UK is one of the few countries to systematically immunise the population using live attenuated intranasal vaccination (LAIV).[31]

This pilot is not targeting those who have had virology specimens, nor to be aligned with a particular point in the vaccination cycle. Generally vaccination takes place in the early autumn in the UK, with seasonal influenza starting to circulate around the year end. In a major study it may be possible to look at immunity in the population, exposed and unexposed to vaccine, and at residual immunity to the circulating strain of influenza.

Patients vaccinated against influenza outside of General Practice may not have information coded into their computerised medical record system at General Practice.[32] Vaccination data are sometimes missing, or sometimes incomplete as the standard reporting form from pharmacist to practice only indicate that the person has been vaccinated against flu, not which brand of vaccine or batch.

CONCLUSION

This pilot tests our ability to collect samples across all adult age-bands within a sentinel network. If successful we will pursue resources to expand this into a larger study.

ACKNOWLEDGEMENTS

Patients for allowing their data to be used for surveillance and research. Practices who have agreed to be part of the RCGP RSC and have allowed us to extract and use health data for surveillance and research. Other members of the Clinical Informatics and Health Outcomes Research Group at University of Surrey. Apollo Medical Systems for data extraction. Collaboration with EMIS, TPP, In-Practice and Micro-test CMR supplier for facilitating data extraction. Patient advisers at practices and colleagues at Public Health England.

LIST OF ABBREVIATIONS

BMI	Body Mass Index
CHPR	Centre for Health Protection Research
ECDC	European Centre for Disease Control
GIS	Geographical information system
GP	General Practitioner – A family physician providing NHS care to a registered list of patients
HAI	Haemagglutination inhibition assays
IG	Information Governance
IGT	Information Governance Toolkit – the NHS standard required for securely holding individual patient level daata
IMD	Index of Multiple Deprivation – a measure of socioeconomic status
JCVI	Joint Committee on Vaccination and Immunisation
LAIV	Live attenuated intranasal vaccination
LSOA	Lower Super Output Area
NHS	National Health Service
NIHR	National Institute for Health Research
PBCL	Pathology Bounded Code List
PHE	Public Health England
RCD	Reverse Cumulative Distribution
RCGP	Royal College of General Practitioners
RSC	Research and Surveillance Centre (within RCGP)
RVU	Respiratory Virus Unit
SEU	Seroepidemiology Unit
WHO	World Health Organization

REFERENCES

[1] Metcalf CJ, Farrar J, Cutts FT, Basta NE, Graham AL, Lessler J, Ferguson NM, Burke DS, Grenfell BT. Use of serological surveys to generate key insights into the changing global landscape of infectious disease. *Lancet*. 2016 Aug 13;388(10045):728-30. doi: 10.1016/S0140-6736(16)30164-7.

[2] Li ZN, Lin SC, Carney PJ, Li J, Liu F, Lu X, Liu M, Stevens J, Levine M, Katz JM, Hancock K. IgM, IgG, and IgA antibody responses to influenza A(H1N1)pdm09 hemagglutinin in infected persons during the first wave of the 2009 pandemic in the United States. *Clin Vaccine Immunol*. 2014 Aug;21(8):1054-60. doi: 10.1128/CVI.00129-14.

[3] Laurie KL, Huston P, Riley S, Katz JM, Willison DJ, Tam JS, Mounts AW, Hoschler K, Miller E, Vandemaele K, Broberg E, Van Kerkhove MD, Nicoll A. Influenza serological studies to inform public health action: best practices to optimise timing, quality and reporting. *Influenza Other Respir Viruses*. 2013;7(2):211-24. doi: 10.1111/j.1750-2659.2012.0370a.x.

[4] Hoschler K, Thompson C, Andrews N, Galiano M, Pebody R, Ellis J, Stanford E, Baguelin M, Miller E, Zambon M. Seroprevalence of influenza A(H1N1)pdm09 virus antibody, England, 2010 and 2011. *Emerg Infect Dis*. 2012 Nov;18(11):1894-7. doi: 10.3201/eid1811.120720.

[5] House of Commons Science and Technology Committee. Scientific advice and evidence in emergencies. Third Report of Session 2010–11. London, House of Commons, 2011. URL: http://www.publications.parliament.uk/pa/cm201011/cmselect/cmsctech/498/498.pdf (accessed 16 May 2018).

[6] Chief Medical Officer's Statistical Legacy Group (CMO-SLG) pandemic influenza preparedness programme. London; Department of Health 2010. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/215871/dh_122754.pdf (accessed 16 May 2018).

[7] DH Pandemic Influenza Preparedness Team . UK Influenza Pandemic Preparedness Strategy. London; Department of Health 2011. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213717/dh_131040.pdf (accessed 16 May 2018).

[8] Scientific Pandemic Influenza Advisory Committee (SPI-M) Modelling Summary. The Scientific Pandemic Influenza Advisory Committee, London, Department of Health, 2013. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/624155/SPI-M_influenza_modelling.pdf (accessed 16 May 2018).

[9] Fineberg HV. Pandemic preparedness and response--lessons from the H1N1 influenza of 2009. *N Engl J Med.* 2014 Apr 3;370(14):1335-42. doi:10.1056/NEJMra1208802.

[10] de Lusignan S, Correa A, Smith GE, Yonova I, Pebody R, Ferreira F, Elliot AJ, Fleming D. RCGP Research and Surveillance Centre: 50 years' surveillance of influenza, infections, and respiratory conditions. *Br J Gen Pract.* 2017 Oct;67(663):440-441. doi: 10.3399/bjgp17X692645.

[11] Correa A, Hinton W, McGovern A, van Vlymen J, Yonova I, Jones S, de Lusignan S. Royal College of General Practitioners Research and Surveillance Centre (RCGP RSC) sentinel network: a cohort profile. *BMJ Open*. 2016 Apr 20;6(4):e011092. doi:10.1136/bmjopen-2016-011092.

[12] de Lusignan S, Correa A, Pathirannehelage S, Byford R, Yonova I, Elliot AJ, Lamagni T, Amirthalingam G, Pebody R, Smith G, Jones S, Rafi I. RCGP Research and Surveillance Centre Annual Report 2014-2015: disparities in presentations to primary care. *Br J Gen Pract.* 2017 Jan;67(654):e29-e40.doi: 10.3399/bjgp16X688573.

[13] Pebody R, Warburton F, Ellis J, Andrews N, Potts A, Cottrell S, Johnston J, Reynolds A, Gunson R, Thompson C, Galiano M, Robertson C, Byford R, Gallagher N, Sinnathamby M, Yonova I, Pathirannehelage S, Donati M, Moore C, de Lusignan S, McMenamin J, Zambon M. Effectiveness of seasonal influenza vaccine for adults and children in preventing laboratory-confirmed influenza in primary care in the United Kingdom: 2015/16 end-of-season results. *Euro Surveill*. 2016 Sep 22;21(38). doi: 10.2807/1560-7917.ES.2016.21.38.30348.

[14] Pathirannehelage S, Kumarapeli P, Byford R, Yonova I, Ferreira F, de Lusignan S. Uptake of a Dashboard Designed to Give Realtime Feedback to a Sentinel Network About Key Data Required for Influenza Vaccine Effectiveness Studies. *Stud Health Technol Inform.* 2018;247:161-165.

[15] de Lusignan S, Correa A. Opportunities and challenges of a World Serum Bank. *Lancet*. 2017 Jan 21;389(10066):250-251. doi: 10.1016/S0140-6736(17)30046-6.

[16] de Lusignan S. Codes, classifications, terminologies and nomenclatures: definition, development and application in practice. *Inform Prim Care.* 2005;13(1):65-70.

[17] Hoschler K, Andrews NJ, Faust SN, Finn A, Pollard AJ, Snape MD, et al. Administration of AS03B-adjuvanted A(H1N1)pdm09 vaccine in children aged <3 years enhances antibody response to H3 and B viruses following a single dose of trivalent vaccine one year later. *Clin Infect Dis.* 2014 Jan;58(2):181-7. doi: 10.1093/cid/cit692.

[18] Höschler K, Southern J, Thompson C, Warburton F, Andrews NJ, Miller E, Zambon M. Responses to live attenuated influenza vaccine in children vaccinated previously with Pandemrix (ASO3(B) adjuvanted pandemic A/H1N1pdm09). *Vaccine*. 2018 May 17;36(21):3034-3040. doi: 10.1016/j.vaccine.2018.04.017.

[19] de Lusignan S, Metsemakers JF, Houwink P, Gunnarsdottir V, van der Lei J. Routinely collected general practice data: goldmines for research? A report of the European Federation for Medical Informatics Primary Care Informatics Working Group (EFMI PCIWG) from MIE2006, Maastricht, The Netherlands. *Inform Prim Care* 2006;14(3):203-9.

[20] Kousoulis AA, Rafi I, de Lusignan S. The CPRD and the RCGP: building on research success by enhancing benefits for patients and practices. *Br J Gen Pract*. 2015 Feb;65(631):54-5. doi: 10.3399/bjgp15X683353.

[21] De Meyer F, De Moor G, Reed-Fourquet L. Privacy Protection through pseudonymisation in eHealth. *Stud Health Technol Inform*. 2008;141:111-8.

[22] de Lusignan S. Effective pseudonymisation and explicit statements of public interest to ensure the benefits of sharing health data for research, quality improvement and health service management outweigh the risks. *Inform Prim Care.* 2014;21(2):61-3. doi: 10.14236/jhi.v21i2.68.

[23] NHS Digital. Information Governance Toolkit (IGT). URL: https://www.igt.hscic.gov.uk/ (accessed 16 May 2018).

[24] List of materials considered to be 'relevant material' under the Human Tissue Act 2004. Human Tissue Authority (HTA). URL: https://www.hta.gov.uk/policies/list-materials-considered-be-%E2%80%98relevant-material%E2%80%99-under-human-tissue-act-2004 (accessed 16 May 2018).

[25] Mascalzoni D, Dove ES, Rubinstein Y, Dawkins HJ, Kole A, McCormack P, Woods S, Riess O, Schaefer F, Lochmüller H, Knoppers BM, Hansson M. International Charter of principles for sharing bio-specimens and data. *Eur J Hum Genet.* 201 Jun;23(6):721-8. doi: 10.1038/ejhg.2014.197. Epub 2014 Sep 24. Erratum in: *Eur J Hum Genet.* 2016 Jul;24(7):1096.

[26] Clinical Informatics and Health Outcomes Research Group. URL: www.clininf.eu (accessed 16 May 2018).

[27] MRC Policy and Guidance on Sharing of Research Data from Population and Patient Studies (Sept 2011) URL: https://mrc.ukri.org/publications/browse/mrc-policy-and-guidance-on-sharing-of-research-data-from-population-and-patient-studies/ (accessed 16 May 2018).

[28] Economic Co-operation and Development (OECD). "Promoting Access to Public Research Data for Scientific, Economic and Social Development" 2007 URL: http://www.oecd.org/sti/sci-tech/38500813.pdf (accessed 16 May 2018).

[29] Metcalf CJ, Mina MJ, Winter AK, Grenfell BT. Opportunities and challenges of a World Serum Bank - Authors' reply. *Lancet.* 2017 Jan 21;389(10066):252. doi:10.1016/S0140-6736(17)30054-5.

[30] Gerlier L, Lamotte M, Dos Santos Mendes S, Damm O, Schwehm M, Eichner M. Estimates of the Public Health Impact of a Pediatric Vaccination Program Using an Intranasal Tetravalent Live-Attenuated Influenza Vaccine in Belgium. *Paediatr Drugs*. 2016 Aug;18(4):303-18. doi: 10.1007/s40272-016-0180-6.

[31] Pebody RG, Green HK, Andrews N, Boddington NL, Zhao H, Yonova I, Ellis J, Steinberger S, Donati M, Elliot AJ, Hughes HE, Pathirannehelage S, Mullett D, Smith GE, de Lusignan S, Zambon M. Uptake and impact of vaccinating school age children against influenza during a season with circulation of drifted influenza A and B strains, England, 2014/15. *Euro Surveill*. 2015;20(39). doi:10.2807/1560-7917.ES.2015.20.39.30029.

[32] de Lusignan S, Hoghton M, Rafi I. Flu vaccination by pharmacists leads to suboptimal medical records. *BMJ*. 2017 Nov 9;359:j5084. doi: 10.1136/bmj.j5084.

AUTHOR'S CONTRIBUTIONS

SdeL
Simon de Lusignan is the lead author who developed the study design and wrote the first draft, edited and contributed to subsequent versions of this protocol.

MT
Manasa Tripathy liaised with the participating practices, edited and contributed to the writing.

RB, EL, KH
Ray Borrow, Ezra Linley and Katja Hoschler helped develop methods, edited and contributed to subsequent versions of the protocol.

FF
Filipa Ferreira manages this study and extensively reviewed the protocol, and contributed to the writing.

MZ:
Maria Zambon reviewed the protocol and contributed to the writing

NA:
Nick Andrews reviewed the protocol and contributed to the writing

IY
Ivelina Yonova helped develop the practice recruitment methods, edited and contributed to the writing.

MH
Mariya Hriskova helped develop the practice recruitment methods, edited and contributed to the writing.

IR
Imran Rafi reviewed the protocol and contributed to the writing

RP
Richard Pebody helped develop methods and contributed to the subsequent versions of the protocol.

FUNDING STATEMENT

This work was supported by the University of Surrey. The funding was obtained as part of the Higher Education Funding Council for England's (HEFCE) Industrial Strategy funding allocated to the University.

COMPETING INTERESTS STATEMENT

Simon de Lusignan has received grant funding through University of Surrey from GSK to report vaccine adverse events, and attended advisory boards for Sanofi and Seqirus.

COMPONENTS OF THE ROSES-I STATEMENT FOR STANDARDIZATION OF THE REPORTING OF SEROEPIDEMIOLOGIC STUDIES

	Item No	STROBE Items	ROSES- I items	Reported on Page
Title and abstract	1	(a). Indicate the study's design with a commonly used term in the title or the abstract (b). Provide in the abstract an informative and balanced summary of what was done and what was found	ROSES-I 1.1: The term "seroepidemiologic," "seroepidemiology," "seroprevalence," or "seroincidence" should be applied to the study in the title or abstract, and the medical subject heading "Seroepidemiologic Studies" be used when the report is of a population-based serological survey.	Pages 3, 4
Introduction	2	Explain the scientific background and rationale for the investigation being reported	ROSES-I 2.1: State what is known about the kinetics of antibody rise, decay, and persistence following infection for the particular virus being studied and the justification for threshold antibody titers or changes in titers used to define evidence of infection ROSES-I 2.2: State what is known about the sensitivity and specificity of the antibody detection assay being used	Pages 5-11
	3	State specific objectives, including any prespecified hypotheses	ROSES-I 3.1: State the specific measure of occurrence that is being estimated, for example, point seroprevalence, cumulative incidence of infection, secondary infection risk	Page 6, 11
EPIDEMIOLOGIC METHODS				
Study design	4	Present key elements of study design early in the paper	ROSES-I 4.1: State which specific seroepidemiologic study design was chosen and why (see Table 1)	Cross- Sectional seroprevalance Page 7

	Item No	STROBE Items	ROSES- I items	Reported on Page
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, methods for sampling, and data collection	ROSES-I 5.1: Describe the timing of the biological sampling in relation to the disease epidemiology in the study population (the beginning, peak, and end of virus transmission) ROSES-I 5.2: Where known, describe the timing of biological sampling in individuals in relation to disease onset and to exposures of interest ROSES-I 5.3: State the interval between sequential biological samples (serial cross-sectional or longitudinal studies), or specify whether only a single sample was collected (cross-sectional study)	NA Methods described on pages 7-11 NA Single Samples Taken
Participants	6	(a). Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case–control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants (b). Cohort study—For matched studies, give matching criteria and the number of exposed and unexposed Case–control study—For matched studies, give matching criteria and the number of controls per case	ROSES-I 6.1: For case-ascertained transmission studies, describe the method of case ascertainment and criteria for defining a “case” ROSES-I 6.2: For household-or institution-based transmission studies, describe the definition of a household or the institution ROSES-I 6.3: For outbreak investigations involving serologic sampling, describe the setting in which the cases were identified, for example, village/ residential setting, occupational workplace ROSES-I 6.4: To aid the interpretation of seroepidemiologic studies of novel influenza A virus subtypes, the results from exposed populations should be compared with the results from unexposed populations. Efforts to validate the assay in virologically confirmed cases should be reported	NA

	Item No	STROBE Items	ROSES- I items	Reported on Page
Variables	7	Clearly define all outcomes, exposures, predictors, potential risk factors, and effect modifiers. Give diagnostic criteria, if applicable	ROSES-I 7.1 The median age and range for each exposure group should be reported ROSES-I 7.2: Describe the potential for immunization (specify vaccine and timing of vaccination in relationship to collection of serum), if applicable, to affect the outcome measures ROSES-I 7.3: Describe any known or potential immunological cross-reactivity that may bias the outcome measures ROSES-I 7.4: Describe illness definitions and methods for ascertaining the presence or absence of clinical illness in subjects	Page 7 NA Page 5 NA
Data sources/ measurement biases	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	ROSES-I 8.1: If relevant, describe measures taken to identify and record immunization history	Page 9-10
Bias	9	Describe any efforts to address the potential sources of bias	ROSES-I 9.1: If relevant, describe efforts to control for the potential effect of immunization on estimates of outcomes	Page 9-10
Study Size	10	Explain how the study size was arrived at	ROSES-I 10.1: Describe the baseline estimated seroprevalence at given antibody titers or incidence of infection and cite published literature to support these estimates	Pages 7-11

	Item No	STROBE Items	ROSES- I items	Reported on Page
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	ROSES-I 11.1: Describe the serological assay's limit of detection and how this limit is defined or calculated. Describe how samples with a result below or on the borderline of the limit were handled in the analysis ROSES-I 11.2: Describe and justify the titer or other result used to define "seropositivity," or the antibody titer change or change in other assay result used to define "seroconversion." Avoid the term "seroconversion" unless referring to change from undetectable to detectable antibody level. Otherwise report the fold-rise in titer. Avoid the term "infection" but report "seroprevalence at a titer of …." ROSES-I 11.3: If statements or inferences are made about protection from infection, describe what is known about the correlation between the assay results and protection from infection and illness	No Results available yet as this study is ongoing
Statistical methods	12	(a). Describe all statistical methods, including those used to control for confounding (b). Describe any methods used to examine subgroups and interactions (c). Explain how missing data were addressed (d). Cohort study—If applicable, explain how loss to follow up was addressed Case–control study—If applicable, explain how matching of cases and controls was addressed Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy (e). Describe any sensitivity analyses	ROSES-I 12.1: if relevant, state how the non-independence of data was managed ROSES-I 12.2: if relevant, report methods used to account for the probability of seropositivity or seroconversion if infected, and to account for decay in antibody titers over time	No Results available yet as this study is ongoing

	Item No	STROBE Items	ROSES- I items	Reported on Page
Laboratory methods	12a		ROSES-I 12a.1: Describe the sample type—serum or plasma. If plasma is used, specify the anticoagulant used (heparin, sodium citrate, EDTA, etc.) ROSES-I 12a.2: Describe the specimen storage conditions (4°C, −20 °C, −80 °C). If frozen prior to the analysis, describe the time to freezing and the number of freeze/thaw cycles prior to testing Serological assays ROSES-I 12a.3: Specify the assay type (e.g., hemagglutination inhibition; virus neutralization/microneutralization; ELISA; other) and methods used to determine the endpoint titer ROSES-I 12a.4: Reference a previously published, CONSISE consensus serologic assay or WHO protocol if used, and any modifications of the protocol. If a previously published protocol is not used, provide full details in supplementary materials ROSES-I 12a.5: State what is known about the determinants of the variability of the antibody detection assay being used ROSES-I 12a.6: Specify the antigen(s) used in the assay, including virus strain name, subtype, lineage or clade, with standardized nomenclature and reference; specify whether live virus or inactivated virus was used (where applicable) ROSES-I 12a.7: Report if antigen(s) from potentially cross-reactive pathogens/strains were used in order to identify cross-reactivity, and specify which antigen was used, including virus name, subtype, strain, lineage and clade, with standardized nomenclature and reference ROSES-I12a.8: If red blood cells were used for a hemagglutinin inhibition assay, specify the animal species from which they were obtained and concentration (v/v) used ROSES-I 12a.9: Describe positive and negative controls used ROSES-I 12a.10: Describe starting and end dilutions ROSES-I 12a.11: Specify laboratory biosafety conditions ROSES-I 12a.12: Specify whether replication was performed, and if so, the acceptable replication parameters ROSES-I 12a.13: Specify whether a confirmatory assay was performed and all specifics of this assay, at the same level of detail ROSES-I 12a.14: Specify international standards used, if appropriate	Page 8-9
RESULTS				
Participants	13	(a). Report the numbers of individuals at each stage of the study—the numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed (b). Give reasons for non-participation at each stage (c). Consider use of a flow diagram	See STROBE item	No Results available yet as this study is Ongoing

	Item No	STROBE Items	ROSES- I items	Reported on Page
Descriptive data	14	a). Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential risk factors (b). Indicate the number of participants with missing data for each variable of interest (c). Cohort study—summarize follow-up time (e.g., average and total amount)	See STROBE item	Page 7
Outcome data	15*	Cohort study—report the numbers of outcome events or summary measures over time Case–control study—report the numbers in each exposure category, or summary measures of exposure Cross-sectional study—report the numbers of outcome events or summary measures	See STROBE item	NA
Main results	16	(a). Give unadjusted estimates and, if applicable, risk factor-adjusted estimates andtheir precision (e.g., 95% confidence interval). Make clear which risk factors were adjusted for and why they were included (b). Report category boundaries when continuous variables were categorized (c). If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	ROSES-I 16.1: Report unadjusted estimates of distribution of titers by age Group ROSES-I 16.2: Report methods to standardize the results from the study sample to the target population	No results available yet as this study is ongoing
Other analyses	17	Report other analyses performed—analyses of subgroups and interactions, and sensitivity analyses	See STROBE item	NA
DISCUSSION				
Key results	18	Summarize key results with reference to study objectives	See STROBE item	NA
Interpretation	19	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	ROSES-I 20.1: Discuss the interpretation of the results in the context of known or potential cross-reactivity	NA

	Item No	STROBE Items	ROSES- I items	Reported on Page
Generalizability	21	Discuss the generalizability (external validity) of the study results	See STROBE item	NA
Other information **Funding**	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	ROSES-I 22.1: Specify if institutional review board approval was received; if not, specify reason (e.g., public health outbreak response/non-research designation)	Page 18

BMJ Open

Serological surveillance of influenza in an English sentinel network: pilot study protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-024285.R2
Article Type:	Protocol
Date Submitted by the Author:	11-Jan-2019
Complete List of Authors:	de Lusignan, Simon; University of Surrey, Department of Clinical and Experimental Medicine; St. Georges University of London, Division of Population Health Sciences and Education Borrow, Ray; Public Health England, Vaccine Evaluation Unit, Manchester Royal Infirmary Tripathy, Manasa ; University of Surrey, Department of Clinical and Experimental Medicine Linley, Ezra; Public Health England, Vaccine Evaluation Unit, Manchester Royal Infirmary Zambon, Maria; Public Health England, Hoschler, Katja; Public Health England Ferreira, Filipa; University of Surrey, Department of Clinical and Experimental Medicine Andrews, Nick; Public Health England, Modelling and Economics Department Yonova, Ivelina; University of Surrey, Department of Clinical and Experimental Medicine; Royal College of General Practitioners, CIRC Hriskova, Mariya ; University of Surrey, Department of Clinical and Experimental Medicine; Royal College of General Practitioners, CIRC Rafi, Imran; Royal College of General Practitioners, Pebody, Richard; Public Health England, Centre for Infectious Disease Surveillance and Control
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Public health, Global health, General practice / Family practice, Infectious diseases
Keywords:	Medical Records System, Computerized, Serology, Data Collection, Population Surveillance, Records as topic, Influenza, Human

Serological surveillance of influenza in an English sentinel network: pilot study protocol

Simon de Lusignan[1,2]
Professor of Primary Care and Clinical Informatics
Head of Department of Clinical and Experimental Medicine
s.lusignan@surrey.ac.uk
SdeL

Ray Borrow[4]
Head of the Vaccine Evaluation Unit
Ray.Borrow@phe.gov.uk
RB

Manasa Tripathy[1]
Practice Liaison Officer
m.tripathy@surrey.ac.uk
MT

Ezra Linley[4]
Deputy Head of Vaccine Evaluation Unit
Ezra.Linley@phe.gov.uk
EL

Maria Zambon[3]
Head of Reference Microbiology
Maria.Zambon@phe.gov.uk
MZ

Katja Hoschler[3]
Senior Scientist, Respiratory Virus Unit
Katja.hoschler@phe.gov.uk
KH

Filipa Ferreira[1]
Project Manager
f.ferreira@surrey.ac.uk
FF

Nick Andrews[3]
Senior statistician
nick.andrews@phe.gov.uk
NA

Ivelina Yonova[1,2]
Practice Liaison Officer
i.yonova@surrey.ac.uk
IY

Mariya Hriskova[1,2]
Practice Liaison Officer
m.hriskova@surrey.ac.uk
MH

Imran Rafi[2]
Chair of Clinical Innovation and Research Centre
Irafi@rcgp.org.uk
IR

Richard Pebody[3]
Head of Flu Surveillance
richard.pebody@phe.gov.uk
RP

[1] Department of Clinical and Experimental Medicine, University of Surrey, GUILDFORD, GU2 7XH, UK
[2] Royal College of General Practitioners Research and Surveillance Centre, Euston Square, LONDON,
[3] Public Health England, London, UK
[4] Vaccine Evaluation Unit, Public Health England, Manchester Royal Infirmary, Manchester, UK

Correspondence to:
Professor Simon de Lusignan
s.lusignan@surrey.ac.uk
T: 01483 683089
Section of Clinical Medicine & Ageing
Department of Clinical and Experimental Medicine
University of Surrey
GUILDFORD, GU2 7XH, UK

Word Count: 5847

Serological surveillance of influenza in an English sentinel network: pilot study protocol

ABSTRACT

Introduction

Background:
Rapidly undertaken age-stratified serology studies can produce valuable data about a new emerging infection including background population immunity and seroincidence during an influenza pandemic. Traditionally seroepidemiology studies have used surplus laboratory sera with little or no clinical information or have been expensive detailed population based studies. We propose collecting population based sera from the Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC), a sentinel network with extensive clinical data.

Aim:
To pilot a mechanism to undertake population based surveys that collect serological specimens and associated patient data to measure sero-positivity and seroincidence due to seasonal influenza, and create a population based serology bank.

Methods and Analysis:

Setting and Participants: We will recruit 6 RCGP RSC practices already taking nasopharyngeal virology swabs. Patients who attend for a scheduled blood test will be consented to donate additional blood samples. Approximately 100-150 blood samples will be collected from each of the following age-bands - 18-29, 30-39, 40-49, 50-59, 60-69, 70+ years.

Methods: We will send the samples to the Public Health England (PHE) Seroepidemiology Unit (SEU) for processing and storage. These samples will be tested for influenza antibodies, using haemagglutination inhibition assays (HAI). Serology results will be pseudonymised, sent to the RCGP RSC and combined using existing processes at the RCGP RSC secure hub. The influenza seroprevalence results from the RCGP cohort will be compared against those from the annual PHE influenza residual serosurvey.

Ethics and dissemination

Ethical approval was granted by the Proportionate Review Sub- Committee of the London – Camden & Kings Cross on 6 February 2018. This study received approval from Health Research Authority on 7 February 2018. On completion the results will be made available via peer- reviewed journals.

Strengths and Limitations of this study

Strengths:

- The Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) is one of the oldest sentinel networks in Europe, it has completed its 51st season of surveillance.
- Practices in RCGP take microbiological samples, including influenza virology specimens which are linked at individual level.
- We have the potential through this network to link vaccine exposure and serology.

Limitations:

- This is a pilot study to demonstrate we can collect samples across pre-defined adult age-groups - we will not collect samples from children, who may be exposed to live attenuated intranasal vaccine (LAIV) in recent years.
- Pharmacist and other non-GP vaccination can lead to missing information on the computerised medical record system at General Practice.

Keywords:

Medical records systems, computerized
Population Surveillance
Serology
Influenza, human
Data collection
Records as topic
Seroepidemiologic Studies
Primary Health Care
Pandemics

INTRODUCTION

Rationale for the pilot study and background

There have been calls for a World Serology Bank as serology could tell us much about the susceptibility of the population to infectious disease.[1] This pilot study explores the potential to establish a serology bank, based on a sentinel network, and focussed on influenza.

Serological data potentially allow the assessment of the severity of a new influenza strain by providing the capability to detect asymptomatic and mild infections and thus determine the symptomatic proportion.[2] The number of infections can be determined if the age-specific prevalence of immunity prior to and then during and after the pandemic are known. Thus, the number of people infected (and therefore no longer susceptible) can be calculated. If these data are available early (particularly the background population immunity and the symptomatic proportion), they can be used to adjust planning assumptions and to help predict the impact of the pandemic on health care services and optimal intervention strategies.

Serum archives would help assess the severity of a novel influenza virus and allow modification of local, national and international pandemic plans.[3] This requirement was a lesson from the 2009 A(H1N1) pandemic and its importance is recognised by the World Health Organization (WHO), the European Centre for Disease Control (ECDC), the UK Department of Health (DH), and the Joint Committee on Vaccination and Immunisation (JCVI).

Despite its apparent merit, the use of serology for seasonal influenza surveillance has several limitations. Seasonal influenza infections occur regularly and individuals are re-infected throughout life with related strains. The prior (lifetime) exposure of individuals to circulating influenza viruses in previous years and a high degree of cross-reactivity between antibodies to different seasonal influenza virus strains can lead to results which are difficult to interpret. Distinguishing between recent infection and recent vaccination also presents further difficulties. Consequently, seroepidemiological studies have not often been used for the investigation of seasonal influenza at a population level.

The UK undertook a series of influenza serosurveys during the 2009 pandemic based on residual blood samples from the Public Health England (PHE) National Seroepidemiology Programme.[4] These are samples submitted to PHE and National Health Service (NHS) laboratories for routine diagnostic purposes. Although this work delivered critical information on background population seroprevalence and seroincidence, several issues were raised in post-pandemic reviews. These reviews highlighted that, although this information was gathered and published earlier than almost any other country, even earlier availability of this intelligence would have been critical to inform important national policy decisions. The following key recommendations have been made in relation to influenza seroepidemiology:

- The Science and Technology Committee (3rd report 2010-12) stated that seroepidemiological data need to be available earlier in the time course of a future pandemic to help with risk assessment (i.e. the likely spread of the disease across the population);[5]
- The Chief Medical Officer (CMO)-Statistics Legacy Group (SLG) determined that serosurveillance is critical to determine population immunity and community infection rates. These data cannot be obtained from other sources and are vital to making modelling predictions of the pandemic.[6]
- The 2011 UK Influenza Pandemic Preparedness Strategy includes seroepidemiology as a key surveillance initiative that will be required at the start of any pandemic, and states that work should be underway to enhance capability to respond, based on the H1N1 (2009) influenza pandemic.[7]
- The Scientific Pandemic Influenza Advisory Committee (SPI), Subgroup on Modelling (SPI-M) group recommended that the PHE strengthen population based influenza seroepidemiology, including collection of key epidemiological information on vaccination status and underlying risk status.[8]

- Finally ECDC has highlighted the importance of influenza seroepidemiology, as has WHO as part of the Fineberg report into the pandemic response and the lack of preparedness including the need for a proper assessment of severity at national and subnational levels early in a pandemic.[9]

The RCGP RSC and PHE, and its predecessor organisations have an over 50 year history of collaboration in influenza and respiratory disease surveillance and vaccine effectiveness studies.[10] This study builds on this long term collaboration.

The RCGP RSC provides a suitable venue for collecting influenza serology data. This sentinel network is one of the longest established and has a nationally representative network of practices.[11] The network has a long history of feedback about data quality, particularly in the areas of influenza like illness (ILI) and other respiratory infections.[12] The RCGP RSC also collects data about vaccine exposure. Combining information about ILI, virology results and vaccine exposure enables the estimation of influenza vaccine effectiveness.[13] Its dashboard capability provides a method for near real time feedback to practices about sample collection.[14]

This pilot study tests whether collecting serology data from a sentinel network and linking it to clinical record data at the individual level would provide a low cost way of creating a high quality sero-epidemiological resource.[15]

Aim:

To pilot a mechanism to undertake population based surveys that collect serological specimens that will be linked with key epidemiological information at strategic time points after each influenza season as a resource to be deployed in a future pandemic, for seasonal influenza and potentially other infections of public health importance.

Objectives:

- Establish a system that allows volunteer patients from practices to provide a serology sample during a routine blood test. The result of that sample will be linked to that patient's pseudonymised record which will provide high quality data about vaccine exposure, and any medically reported influenza like illnesses (ILI) or other condition of scientific interest, included within an approved scientific protocol, or meeting public health needs in a pandemic.
- Measure the feasibility of collecting 100-150 samples from each of the following age-bands - 18-29, 30-39, 40-49, 50-59, 60-69, 70+ years.
- Pilot laboratory programme for the processing and storage of these specimens
- Pilot linkage of these biological specimens to the Public Health England Respiratory Virus Unit (RVU) analysis programme for the detection of influenza antibodies
- Link the results to epidemiological data in particular vaccination history, age and underlying clinical risk factor status and any swab results. The serology results can be linked to vaccine brand.
- Estimate the costs of providing a national seroepidemiology service based on samples from a subset of RCGP RSC practices.

METHODS AND ANALYSIS

Study Design

Study setting and population:

The project will consist of four interlinked work-packages:
1) Population sampling and collection of biological specimens

2) Laboratory analysis
3) Data management
4) Statistical and modelling analysis.

1) Population sampling and collection of biological specimens

The RCGP RSC practices will collect the biological specimens:
The RCGP RSC will be used to collect the biological specimens. Practices who participated in the virology swabbing scheme will be invited.

The project proposes to pilot a population-based seroprevalence survey, involving 100-150 individuals across the following age range: 18-29, 30-39, 40-49, 50-59, 60-69, 70+ years, following the 2017/18 influenza season. Patients, who attend their pilot sentinel network practice for routine blood test during the study period will be asked to also provide an additional blood sample for serology. This will provide information on seroprevalence to set in context other measures of impact of an influenza season/epidemic in a population, and provide the most accurate measures of population exposure. This new approach to serology banking represents a compromise between using residual serum samples from laboratories where less is known about the patient's medical and immunisation history and formal surveys that can collect such data, but can have non- response bias.

Analysis will be carried out using the blood collected in RCGP RSC practices and sent to the Seroepidemiology Unit (SEU) archive. The SEU archive is a collection of anonymised residual serum samples from routine microbiological testing, submitted voluntarily each year from laboratories throughout England. SEU archive sera are stored at the PHE North West regional laboratory in Manchester and are anonymised and permanently unlinked from any patient identifying information, with only age, gender, date of collection (if available) and contributing laboratory retained.

	All practices (5-6 practices) during the study period	
Age Band	100 sample threshold crossed after the following number of weeks	150 sample threshold crossed after the following number of weeks
18-29	3	5
30-39	3	5
40-49	2	2
50-59	1	2
60-69	1	2
70+	1	1

Table 1: Blood test plan for the pilot study. This table represents the week in which sample collection would be complete, if all patients consented. The pilot study will take place as soon as ethical approval is achieved and would involve the collection of a total of up to 150 samples from approximately six practices over the study period.

Serology samples will be analysed at the Respiratory Virus Unit (RVU) at the PHE Colindale using HAI with representative vaccine strains. Each sample will be tested once and not in duplicate, the result will have strong identifiers pseudonymised (in accordance with current best practice) and be returned to the RCGP RSC hub's encrypted server, decrypted and linked to individual patients pseudonymised data.

The RCGP RSC will recruit six practices (depending on practice list size) to ascertain the feasibility and establish the approach. We will aim to collect an evenly distributed set of sera across all age bands that are set out in Table 1. Older people have more chronic disease and hence have higher number of blood tests performed

than younger people; so the target number of samples would be achieved sooner. There will be no attempt to select patients for serology on the basis of whether they have had an influenza immunisation, or not or had a virology specimen taken for influenza. The number of samples is based on a combined list size of 50,000 across the participating practices.

The proposed method would involve feeding back to practices as each age band reaches the target number of samples. (minimum 100, maximum 150 per age band). University of Surrey has developed methods for RCGP RSC to give practice specific feedback about vaccine exposure and data quality - these methods could be used to provide feedback to practices participating in serum collection.

Patient selection:

Inclusion criteria: All patients of age 18 and over who visit their practice for a routine blood test and provide another sample for serology are eligible for inclusion in the analysis. The main inclusion criteria for practices is that practices are within our influenza swabbing practices' list with quota sampling according to the table above.

Exclusion criteria: Patients who have explicitly opted out of data sharing will be excluded from the analysis. We will identify these patients using the opt-out codes within GP information systems where the patients have made an explicit choice to opt out; patients will be informed of their option to opt-out via posters in the practices and information sheets.

2) Laboratory analysis

Samples will be submitted to the PHE Manchester laboratory in practice batches of clotted vacutainer bottles, and will be accompanied by the standard request form. This form will be generated by existing ICE pathology request software used by the GP practices.

General Practices use a pathology request software, such as the ICE system to send biological specimens to their local laboratory for testing. When a patient requires a blood test, the GP can print a test request form via their ICE system. To send samples to PHE's SEU/VEU:

- The practice will need to print out an additional test request form for routine blood sampling
- The University of Surrey will provide the practices with detailed guidance on sample collection and postage to PHE's SEU/VEU

Upon receiving consent, practices will be able to send specimens to PHE's SEU/VEU lab via pre-paid envelopes. A Material Transfer Agreement (MTA) will be put in place for the transport of blood samples from GP practices to Manchester SEU/VEU.

We suggest a single Read code[16] is allocated to mark that the serology specimen has been consented to and sent. This will facilitate one-to-one matching of specimens with the consenting patient/subjects clinical history.

Whole blood samples will be received, processed to obtain sera, catalogued and stored in -80° C archive freezers within the Serum Archives section of the Vaccine Evaluation Unit (VEU) at Manchester according to existing SOPs, modified where necessary. It is envisioned that sera collected will be permanently transferred to the SEU archive once this study is complete.

Sera will be processed to undergo analysis at the Respiratory Virus Unit (RVU) of the National Infection Service of PHE (London, UK) using haemagglutination inhibition (HAI) assays and other measures of functional antibody status according to established protocols to detect antibody levels against relevant circulating and/or vaccine influenza strains.[17] Antigens will be grown in-house in egg and cell-culture and Influenza B antigens

will be diethyl ether extracted as previously described.[18] Briefly, for the HAI assay, sera will be treated to remove non-specific inhibitors using receptor destroying enzyme (RDEII) and then twofold serially diluted starting at a 1:10 dilution, followed by mixing with an equal volume (25µl) of PBS containing 4HA units of each of the strains. Turkey red blood cells (RBC) will be used for the influenza A(H1N1)pdm09 and influenza B components, and Guinea pig RBC for H3N2. HAI titres will be expressed as the reciprocal of the last serum dilution that results in complete inhibition of agglutination. Where sample volumes after completion of the HAI permit, additional analysis using influenza neutralisation and/or Neuraminidase Antibody Inhibition assays in the format of the Enzyme-linked Lectin Assay [16] may be considered to investigate the functionality of the antibody further.

Following analysis, the results will be reported in form of an Excel table (containing titres against each antigen/influenza virus for every sample) to the data manager of the project. The data (with pseudonymised strong identifiers) will be linked to RCGP RSC data to allow analysis of relevant data including vaccine exposure and previous influenza.

Using the UK laboratory bounded code list

UK laboratories are currently obligated to use the Pathology Bounded Code List (PBCL), a subset of NHS Digital Read codes, when electronically reporting pathology results to GPs, but are freely allowed to choose which PBCL codes are used for each test. A PBCL code assigned to the test (Table 2) will allow one week in areas for the RCGP RSC team to follow the collection of the samples by age-band.

V2_READ_CODE	V2_TERM
4JDb.	Influenza (A&B) serology
43L..	Sample serology

Table 2: Read codes to flag that a specimen has been taken

3) Data extraction and data management

Data collection from volunteer RCGP RSC practice

Data will be extracted from RCGP RSC databases which store pseudonymised data received from participating RCGP practices.[10] A UK general practice is a registration based system where all citizens can register with a single General Practice (GP) of their choice. Practices are computerised, and data entered into computerised medical record systems either as coded data, or free text.

Practices will code 43L.. (Sample serology) into patient record when a patient consents to providing a sample for Serology. We will extract the coded data, and our results will be based on this element of the record. We will extract all coded data, pseudonymising as close to sources as possible. Where patients have a range of codes inserted in their record suggesting they opt out of record sharing we will not analyse their data.

The RCGP RSC will only extract coded data, e.g. where the GP or other health professional codes a disease or symptom into their computerised medical record system, and will process these data as required for this project.[16,19] Coded routine data from UK primary care has been widely used in research.[20]

Data collection methods will follow best practice, including pseudonymisation

These data extractions will be conducted in accordance with best practice, using the Clinical Informatics and Health Outcomes Research Group's standard operating procedures for data extraction, pseudonymisation, and transfer. The method and governance procedure has been developed by the University of Surrey, using an approved provider.

Pseudonymisation is the standard approach for protecting patient's privacy. It is a process that involves the removal of all personal identifiers from data – such as name, date of birth, etc. However, there is a risk that if

data are linked to other data a person might be identified.[21,22] Therefore although all identifiers are removed we keep data encrypted during transfer and on a secure network that meets NHS Information Governance standards to minimise the risk of re-identification. A legally binding definition of pseudonymisation has been introduced into European law on the recommendation of the European Data Protection Supervisor (EDPS).[19]

We "pseudonymise" strong identifiers (in this study NHS number) so that we can link further data to the same individual's record. For this study we need, for example, to be able to link whether there have been any immune changes to the individual that had been vaccinated (and with the specific brand and batch number). Pseudonymisation allows us to do this without knowing any of the strong personal identifiers of that individual.

All data processing and analysis in the present proposed study will be conducted within the secure IT environment of the Clinical Informatics and Health Outcomes Research Group, at the University of Surrey. The information security policies and procedures of the Research Group have been approved by the NHS Digital as meeting the Information Governance Toolkit (IGT) standards.[23]

The following routinely collected patient data will be collected for the study:

- Demographic information: age, gender, ethnicity, registered date.
- Lower Super Output Area (LSOA): full postcodes will be automatically and immediately transformed into Lower Super Output Area (LSOA) which can be used for calculating deprivation scores, using the Index of Multiple Deprivation (IMD), within GP computer systems upon extraction. This would provide information about any inequities in access according to level of social deprivation using geographical information system (GIS) methods.
- Influenza vaccination – including date of vaccination and brand/lot
- Primary care consultations following vaccination, any other markers of health care utilisation, and referral to further care.
- Reactogenicity outcomes of seasonal influenza vaccination as listed in the research literature and any contemporary EU guidance.
- Life-style/risk factors – e.g. Body Mass Index (BMI), smoking status.
- Records of other diseases and long term conditions – e.g. chronic respiratory disease, chronic heart disease, chronic kidney disease, chronic liver disease, chronic neurological disease, diabetes, immunosuppression, pneumonia, etc.
- Pregnancy.
- PHE results of influenza HAI testing will be linked to this data-set at the University of Surrey according to established information governance (IG) procedures

Practices will receive weekly feedback via our dashboard system about progress to target within each age-band (Table 1).[14]

Data tables for analysis:

RCGP RSC will prepare tables for analysis as set out in the section below on statistical modelling and analysis. A schedule for reports would be created, as is currently carried out for reporting influenza vaccine effectiveness studies, and blank tables prepared in advance as outlined in this protocol.

4) Statistical and modelling analysis.

PHE will lead on this analysis and follow the outline plan set out below.

The proof-of-concept study will consist of a total of up to 100-150 participants per age-band from whom blood samples have been taken during the survey period.

The evaluation will include the following elements:

- A comparison of representativeness of the recruited study population will be undertaken in relation to national surveys.
- A comparison of the influenza seroprevalence results from the RCGP cohort against those from the annual PHE influenza residual serosurvey
- Demographic, epidemiological and lab data will be checked for completeness, errors and inconsistencies
- Samples with titres >40 by HAI will be considered seropositive to report prevalence of influenza antibody detected by HAI.[4] In addition, geometric mean titres and reverse cumulative distribution (RCD) curves will be calculated for some preliminary, exploratory comparisons of groups such as those vaccinated versus unvaccinated – although conclusions will be limited by the samples available.
- Acceptability to participate amongst practices and patients ascertained through measures such as response rate
- Feasibility of extending the pilot approach to a wider number of practices

University of Surrey will report the extent to which we have samples from common households – this might provide information about shared infections/levels of immunity within households.

We will produce a weekly report on progress to target by age-band; and provide feedback to practices about the utility of the data provided. This will be a simple table – based on Table 1, explaining the anticipated number of weeks to complete the sampling by age-band.

5) Patient and Public Involvement

No patients or public were involved in the development of the research question or design of this study.

Project Management

The pilot project will be a collaborative project lead by Prof S de Lusignan at University of Surrey, with RCGP and PHE as collaborators. The RCGP-PHE scientific committee will oversee the project in collaboration with the principal investigator.

This research and information governance framework for RCGP RSC sits within the University of Surrey's formal frameworks for information and research governance. In addition, all externally funded projects and collaborative projects with external partners are supported and guided by the University's Research and Enterprise Support (RES) service. RES ensures that university-supported projects are financially viable, and that legal issues of knowledge transfer and intellectual properties are addressed. The project team is supported by IT services dedicated to the Faculty and to the Department of Clinical and Experimental Medicine. Our secure analysis servers are optimised for routine healthcare data processing, to provide faster deliveries for our projects.

ETHICS AND DISSEMINATION

Ethical Approval

PHE has ethical approval (05/Q0505/45) for the collection and use of unlinked and anonymised residual serum samples in cross-sectional antibody prevalence studies for the surveillance of population immunity to vaccine preventable diseases of public health importance and the collection has been extensively used for this purpose.

We will seek to collect serum samples from a cohort of patients that are registered with one of the RCGP RSC practices. Potential participants will be attending a pre-scheduled blood appointment, where the healthcare professional treating them will inform them about the study and seek their consent about whether they would be interested in donating an additional blood sample as part of the study.

Ethical approval was granted by the Proportionate Review Sub- Committee of the London – Camden & Kings Cross on 6 February 2018. This study received approval from Health Research Authority on 7 February 2018.

Blood serum is acellular and not considered a material subject to the Human Tissue Act 2004.[24] However, practices will all need to put in place a Material Transfer Agreement with Public Health England prior to starting their surveillance.[25]

Information Governance

The Clinical Informatics and Health Outcomes Research Group at the University of Surrey has worked with routinely collected healthcare data in a number of research and evaluation projects for over 20 years.[26] The Research Group works within the research and Information Governance frameworks for health and social care in the United Kingdom, and is compliant with the University's best practice standards. The University of Surrey is registered with the Information Commissioner's Office Data Protection Register, and is compliant with the Data Protection Act, and other legislations.

In addition, the Research Group reviewed its departmental information governance policies and procedures, against the requirements of the NHS Information Governance Toolkit (IGT) for Hosted Secondary Use Team/ Project, Version 14.1.[20]

Dissemination and Public Register Disclosure

The outputs from the research will be disseminated primarily through peer review papers in high impact journals within the domains of primary care, surveillance, vaccines, and infectious diseases.[27,28] We will present findings at relevant seminars and conferences. The University of Surrey, in accordance with PHE policy, will post a summary of the study protocol and results within 12 months of study completion.

DISCUSSION

Strengths:

The strengths of this application are that it uses an established sentinel network and builds on PHE expertise in serological analysis to establish a serology bank.

The RCGP RSC is a sentinel network that collects and monitors data from primary care, particularly influenza and other respiratory illnesses, with some practices in the network that have been providing data for decades.[11] These practices have data quality that is as good as it gets in primary care and the practices are used to taking specimens.

PHE has expertise in serology, but many of serosurveys used residual serum samples from diagnostic blood tests[4] and hence the clinical information available with these is very limited. Detailed serology surveys are also conducted, but these are extremely expensive.

In this pilot the serology sampling has been designed to provide almost real time results which can be linked to pseudonymised patient data extracted by RCGP RSC. Conducting the sampling alongside routine blood tests reduces overall time for practices and costs overall.

Collection from the same household may results in selection bias as it is likely that they will have had similar exposures, however it is also a possible strength of the network in trying to understand more about transmission within households or communal establishments, such as old peoples' homes.

We will create a database with individual level serology, virology specimens (if any) confirming influenza or other respiratory disease diagnoses, details of past medical history and vaccine exposure data.

Limitations:

This pilot is limited by its design as a pilot within the available resource envelope and data quality. The limitations of this study are: its small size; limitation to adult specimens only; no specific targeting of those who have had previous virology or at specific points within the annual cycle of vaccination or of influenza infection; and data quality particularly of out-of-practice vaccine exposure.

We set an arbitrary collection strategy across adult age-bands. However, this distribution is not representative of our practices populations' age distribution. Any full scale study would set out to represent the age-sex profile of the population and its geographical distribution.

The study is not powered to detect differences between groups, instead to demonstrate our ability to collect samples across all age groups. The ability to collect across all age-groups have been questioned – both by the original authors of the call for a World Serology Bank,[29] and set as a principal challenge by potential funders if the pilot is a success. Hence the focus of the pilot is on sample collection across age-groups.

We have not included children or young people under 18 years in this pilot. We feel this is appropriate for the first pilot of this type. In a substantive study this group may be important because children are important vectors of disease[30] and the UK is one of the few countries to systematically immunise the population using live attenuated intranasal vaccination (LAIV).[31]

This pilot is not targeting those who have had virology specimens, nor to be aligned with a particular point in the vaccination cycle. Generally vaccination takes place in the early autumn in the UK, with seasonal influenza starting to circulate around the year end. In a major study it may be possible to look at immunity in the population, exposed and unexposed to vaccine, and at residual immunity to the circulating strain of influenza.

Patients vaccinated against influenza outside of General Practice may not have information coded into their computerised medical record system at General Practice.[32] Vaccination data are sometimes missing, or sometimes incomplete as the standard reporting form from pharmacist to practice only indicate that the person has been vaccinated against flu, not which brand of vaccine or batch.

CONCLUSION

This pilot tests our ability to collect samples across all adult age-bands within a sentinel network. If successful we will pursue resources to expand this into a larger study.

ACKNOWLEDGEMENTS

Patients for allowing their data to be used for surveillance and research. Practices who have agreed to be part of the RCGP RSC and have allowed us to extract and use health data for surveillance and research. Other members of the Clinical Informatics and Health Outcomes Research Group at University of Surrey. Apollo Medical Systems for data extraction. Collaboration with EMIS, TPP, In-Practice and Micro-test CMR supplier for facilitating data extraction. Patient advisers at practices and colleagues at Public Health England.

LIST OF ABBREVIATIONS

BMI	Body Mass Index
CHPR	Centre for Health Protection Research
ECDC	European Centre for Disease Control
GIS	Geographical information system
GP	General Practitioner – A family physician providing NHS care to a registered list of patients
HAI	Haemagglutination inhibition assays
IG	Information Governance
IGT	Information Governance Toolkit – the NHS standard required for securely holding individual patient level daata
IMD	Index of Multiple Deprivation – a measure of socioeconomic status
JCVI	Joint Committee on Vaccination and Immunisation
LAIV	Live attenuated intranasal vaccination
LSOA	Lower Super Output Area
NHS	National Health Service
NIHR	National Institute for Health Research
PBCL	Pathology Bounded Code List
PHE	Public Health England
RCD	Reverse Cumulative Distribution
RCGP	Royal College of General Practitioners
RSC	Research and Surveillance Centre (within RCGP)
RVU	Respiratory Virus Unit
SEU	Seroepidemiology Unit
WHO	World Health Organization

REFERENCES

[1] Metcalf CJ, Farrar J, Cutts FT, Basta NE, Graham AL, Lessler J, Ferguson NM, Burke DS, Grenfell BT. Use of serological surveys to generate key insights into the changing global landscape of infectious disease. *Lancet*. 2016 Aug 13;388(10045):728-30. doi: 10.1016/S0140-6736(16)30164-7.

[2] Li ZN, Lin SC, Carney PJ, Li J, Liu F, Lu X, Liu M, Stevens J, Levine M, Katz JM, Hancock K. IgM, IgG, and IgA antibody responses to influenza A(H1N1)pdm09 hemagglutinin in infected persons during the first wave of the 2009 pandemic in the United States. *Clin Vaccine Immunol*. 2014 Aug;21(8):1054-60. doi: 10.1128/CVI.00129-14.

[3] Laurie KL, Huston P, Riley S, Katz JM, Willison DJ, Tam JS, Mounts AW, Hoschler K, Miller E, Vandemaele K, Broberg E, Van Kerkhove MD, Nicoll A. Influenza serological studies to inform public health action: best practices to optimise timing, quality and reporting. *Influenza Other Respir Viruses*. 2013;7(2):211-24. doi: 10.1111/j.1750-2659.2012.0370a.x.

[4] Hoschler K, Thompson C, Andrews N, Galiano M, Pebody R, Ellis J, Stanford E, Baguelin M, Miller E, Zambon M. Seroprevalence of influenza A(H1N1)pdm09 virus antibody, England, 2010 and 2011. *Emerg Infect Dis*. 2012 Nov;18(11):1894-7. doi: 10.3201/eid1811.120720.

[5] House of Commons Science and Technology Committee. Scientific advice and evidence in emergencies. Third Report of Session 2010–11. London, House of Commons, 2011. URL: http://www.publications.parliament.uk/pa/cm201011/cmselect/cmsctech/498/498.pdf (accessed 16 May 2018).

[6] Chief Medical Officer's Statistical Legacy Group (CMO-SLG) pandemic influenza preparedness programme. London; Department of Health 2010. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/215871/dh_122754.pdf (accessed 16 May 2018).

[7] DH Pandemic Influenza Preparedness Team . UK Influenza Pandemic Preparedness Strategy. London; Department of Health 2011. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213717/dh_131040.pdf (accessed 16 May 2018).

[8] Scientific Pandemic Influenza Advisory Committee (SPI-M) Modelling Summary. The Scientific Pandemic Influenza Advisory Committee, London, Department of Health, 2013. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/624155/SPI-M_influenza_modelling.pdf (accessed 16 May 2018).

[9] Fineberg HV. Pandemic preparedness and response--lessons from the H1N1 influenza of 2009. *N Engl J Med.* 2014 Apr 3;370(14):1335-42. doi:10.1056/NEJMra1208802.

[10] de Lusignan S, Correa A, Smith GE, Yonova I, Pebody R, Ferreira F, Elliot AJ, Fleming D. RCGP Research and Surveillance Centre: 50 years' surveillance of influenza, infections, and respiratory conditions. *Br J Gen Pract.* 2017 Oct;67(663):440-441. doi: 10.3399/bjgp17X692645.

[11] Correa A, Hinton W, McGovern A, van Vlymen J, Yonova I, Jones S, de Lusignan S. Royal College of General Practitioners Research and Surveillance Centre (RCGP RSC) sentinel network: a cohort profile. *BMJ Open*. 2016 Apr 20;6(4):e011092. doi:10.1136/bmjopen-2016-011092.

[12] de Lusignan S, Correa A, Pathirannehelage S, Byford R, Yonova I, Elliot AJ, Lamagni T, Amirthalingam G, Pebody R, Smith G, Jones S, Rafi I. RCGP Research and Surveillance Centre Annual Report 2014-2015: disparities in presentations to primary care. *Br J Gen Pract.* 2017 Jan;67(654):e29-e40.doi: 10.3399/bjgp16X688573.

[13] Pebody R, Warburton F, Ellis J, Andrews N, Potts A, Cottrell S, Johnston J, Reynolds A, Gunson R, Thompson C, Galiano M, Robertson C, Byford R, Gallagher N, Sinnathamby M, Yonova I, Pathirannehelage S, Donati M, Moore C, de Lusignan S, McMenamin J, Zambon M. Effectiveness of seasonal influenza vaccine for adults and children in preventing laboratory-confirmed influenza in primary care in the United Kingdom: 2015/16 end-of-season results. *Euro Surveill*. 2016 Sep 22;21(38). doi: 10.2807/1560-7917.ES.2016.21.38.30348.

[14] Pathirannehelage S, Kumarapeli P, Byford R, Yonova I, Ferreira F, de Lusignan S. Uptake of a Dashboard Designed to Give Realtime Feedback to a Sentinel Network About Key Data Required for Influenza Vaccine Effectiveness Studies. *Stud Health Technol Inform.* 2018;247:161-165.

[15] de Lusignan S, Correa A. Opportunities and challenges of a World Serum Bank. *Lancet*. 2017 Jan 21;389(10066):250-251. doi: 10.1016/S0140-6736(17)30046-6.

[16] de Lusignan S. Codes, classifications, terminologies and nomenclatures: definition, development and application in practice. *Inform Prim Care.* 2005;13(1):65-70.

[17] Hoschler K, Andrews NJ, Faust SN, Finn A, Pollard AJ, Snape MD, et al. Administration of AS03B-adjuvanted A(H1N1)pdm09 vaccine in children aged <3 years enhances antibody response to H3 and B viruses following a single dose of trivalent vaccine one year later. *Clin Infect Dis.* 2014 Jan;58(2):181-7. doi: 10.1093/cid/cit692.

[18] Höschler K, Southern J, Thompson C, Warburton F, Andrews NJ, Miller E, Zambon M. Responses to live attenuated influenza vaccine in children vaccinated previously with Pandemrix (ASO3(B) adjuvanted pandemic A/H1N1pdm09). *Vaccine*. 2018 May 17;36(21):3034-3040. doi: 10.1016/j.vaccine.2018.04.017.

[19] de Lusignan S, Metsemakers JF, Houwink P, Gunnarsdottir V, van der Lei J. Routinely collected general practice data: goldmines for research? A report of the European Federation for Medical Informatics Primary Care Informatics Working Group (EFMI PCIWG) from MIE2006, Maastricht, The Netherlands. *Inform Prim Care* 2006;14(3):203-9.

[20] Kousoulis AA, Rafi I, de Lusignan S. The CPRD and the RCGP: building on research success by enhancing benefits for patients and practices. *Br J Gen Pract*. 2015 Feb;65(631):54-5. doi: 10.3399/bjgp15X683353.

[21] De Meyer F, De Moor G, Reed-Fourquet L. Privacy Protection through pseudonymisation in eHealth. *Stud Health Technol Inform*. 2008;141:111-8.

[22] de Lusignan S. Effective pseudonymisation and explicit statements of public interest to ensure the benefits of sharing health data for research, quality improvement and health service management outweigh the risks. *Inform Prim Care.* 2014;21(2):61-3. doi: 10.14236/jhi.v21i2.68.

[23] NHS Digital. Information Governance Toolkit (IGT). URL: https://www.igt.hscic.gov.uk/ (accessed 16 May 2018).

[24] List of materials considered to be 'relevant material' under the Human Tissue Act 2004. Human Tissue Authority (HTA). URL: https://www.hta.gov.uk/policies/list-materials-considered-be-%E2%80%98relevant-material%E2%80%99-under-human-tissue-act-2004 (accessed 16 May 2018).

[25] Mascalzoni D, Dove ES, Rubinstein Y, Dawkins HJ, Kole A, McCormack P, Woods S, Riess O, Schaefer F, Lochmüller H, Knoppers BM, Hansson M. International Charter of principles for sharing bio-specimens and data. *Eur J Hum Genet.* 201 Jun;23(6):721-8. doi: 10.1038/ejhg.2014.197. Epub 2014 Sep 24. Erratum in: *Eur J Hum Genet.* 2016 Jul;24(7):1096.

[26] Clinical Informatics and Health Outcomes Research Group. URL: www.clininf.eu (accessed 16 May 2018).

[27] MRC Policy and Guidance on Sharing of Research Data from Population and Patient Studies (Sept 2011) URL: https://mrc.ukri.org/publications/browse/mrc-policy-and-guidance-on-sharing-of-research-data-from-population-and-patient-studies/ (accessed 16 May 2018).

[28] Economic Co-operation and Development (OECD). "Promoting Access to Public Research Data for Scientific, Economic and Social Development" 2007 URL: http://www.oecd.org/sti/sci-tech/38500813.pdf (accessed 16 May 2018).

[29] Metcalf CJ, Mina MJ, Winter AK, Grenfell BT. Opportunities and challenges of a World Serum Bank - Authors' reply. *Lancet.* 2017 Jan 21;389(10066):252. doi:10.1016/S0140-6736(17)30054-5.

[30] Gerlier L, Lamotte M, Dos Santos Mendes S, Damm O, Schwehm M, Eichner M. Estimates of the Public Health Impact of a Pediatric Vaccination Program Using an Intranasal Tetravalent Live-Attenuated Influenza Vaccine in Belgium. *Paediatr Drugs*. 2016 Aug;18(4):303-18. doi: 10.1007/s40272-016-0180-6.

[31] Pebody RG, Green HK, Andrews N, Boddington NL, Zhao H, Yonova I, Ellis J, Steinberger S, Donati M, Elliot AJ, Hughes HE, Pathirannehelage S, Mullett D, Smith GE, de Lusignan S, Zambon M. Uptake and impact of vaccinating school age children against influenza during a season with circulation of drifted influenza A and B strains, England, 2014/15. *Euro Surveill*. 2015;20(39). doi:10.2807/1560-7917.ES.2015.20.39.30029.

[32] de Lusignan S, Hoghton M, Rafi I. Flu vaccination by pharmacists leads to suboptimal medical records. *BMJ*. 2017 Nov 9;359:j5084. doi: 10.1136/bmj.j5084.

AUTHOR'S CONTRIBUTIONS

SdeL
Simon de Lusignan is the lead author who developed the study design and wrote the first draft, edited and contributed to subsequent versions of this protocol.

MT
Manasa Tripathy liaised with the participating practices, edited and contributed to the writing.

RB, EL, KH
Ray Borrow, Ezra Linley and Katja Hoschler helped develop methods, edited and contributed to subsequent versions of the protocol.

FF
Filipa Ferreira manages this study and extensively reviewed the protocol, and contributed to the writing.

MZ:
Maria Zambon reviewed the protocol and contributed to the writing

NA:
Nick Andrews reviewed the protocol and contributed to the writing

IY
Ivelina Yonova helped develop the practice recruitment methods, edited and contributed to the writing.

MH
Mariya Hriskova helped develop the practice recruitment methods, edited and contributed to the writing.

IR
Imran Rafi reviewed the protocol and contributed to the writing

RP
Richard Pebody helped develop methods and contributed to the subsequent versions of the protocol.

FUNDING STATEMENT

This work was supported by the University of Surrey. The funding was obtained as part of the Higher Education Funding Council for England's (HEFCE) Industrial Strategy funding allocated to the University.

COMPETING INTERESTS STATEMENT

Simon de Lusignan has received grant funding through University of Surrey from GSK to report vaccine adverse events, and attended advisory boards for Sanofi and Seqirus.

COMPONENTS OF THE ROSES-I STATEMENT FOR STANDARDIZATION OF THE REPORTING OF SEROEPIDEMIOLOGIC STUDIES

	Item No	STROBE Items	ROSES- I items	Reported on Page
Title and abstract	1	(a). Indicate the study's design with a commonly used term in the title or the abstract (b). Provide in the abstract an informative and balanced summary of what was done and what was found	ROSES-I 1.1: The term "seroepidemiologic," "seroepidemiology," "seroprevalence," or "seroincidence" should be applied to the study in the title or abstract, and the medical subject heading "Seroepidemiologic Studies" be used when the report is of a population-based serological survey.	Pages 3, 4
Introduction	2	Explain the scientific background and rationale for the investigation being reported	ROSES-I 2.1: State what is known about the kinetics of antibody rise, decay, and persistence following infection for the particular virus being studied and the justification for threshold antibody titers or changes in titers used to define evidence of infection ROSES-I 2.2: State what is known about the sensitivity and specificity of the antibody detection assay being used	Pages 5-11
	3	State specific objectives, including any prespecified hypotheses	ROSES-I 3.1: State the specific measure of occurrence that is being estimated, for example, point seroprevalence, cumulative incidence of infection, secondary infection risk	Page 6, 11
EPIDEMIOLOGIC METHODS				
Study design	4	Present key elements of study design early in the paper	ROSES-I 4.1: State which specific seroepidemiologic study design was chosen and why (see Table 1)	Cross- Sectional seroprevalance Page 7

	Item No	STROBE Items	ROSES- I items	Reported on Page
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, methods for sampling, and data collection	ROSES-I 5.1: Describe the timing of the biological sampling in relation to the disease epidemiology in the study population (the beginning, peak, and end of virus transmission) ROSES-I 5.2: Where known, describe the timing of biological sampling in individuals in relation to disease onset and to exposures of interest ROSES-I 5.3: State the interval between sequential biological samples (serial cross-sectional or longitudinal studies), or specify whether only a single sample was collected (cross-sectional study)	NA Methods described on pages 7-11 NA Single Samples Taken
Participants	6	(a). Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case–control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants (b). Cohort study—For matched studies, give matching criteria and the number of exposed and unexposed Case–control study—For matched studies, give matching criteria and the number of controls per case	ROSES-I 6.1: For case-ascertained transmission studies, describe the method of case ascertainment and criteria for defining a “case” ROSES-I 6.2: For household-or institution-based transmission studies, describe the definition of a household or the institution ROSES-I 6.3: For outbreak investigations involving serologic sampling, describe the setting in which the cases were identified, for example, village/ residential setting, occupational workplace ROSES-I 6.4: To aid the interpretation of seroepidemiologic studies of novel influenza A virus subtypes, the results from exposed populations should be compared with the results from unexposed populations. Efforts to validate the assay in virologically confirmed cases should be reported	NA

	Item No	STROBE Items	ROSES- I items	Reported on Page
Variables	7	Clearly define all outcomes, exposures, predictors, potential risk factors, and effect modifiers. Give diagnostic criteria, if applicable	ROSES-I 7.1 The median age and range for each exposure group should be reported ROSES-I 7.2: Describe the potential for immunization (specify vaccine and timing of vaccination in relationship to collection of serum), if applicable, to affect the outcome measures ROSES-I 7.3: Describe any known or potential immunological cross-reactivity that may bias the outcome measures ROSES-I 7.4: Describe illness definitions and methods for ascertaining the presence or absence of clinical illness in subjects	Page 7 NA Page 5 NA
Data sources/ measurement biases	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	ROSES-I 8.1: If relevant, describe measures taken to identify and record immunization history	Page 9-10
Bias	9	Describe any efforts to address the potential sources of bias	ROSES-I 9.1: If relevant, describe efforts to control for the potential effect of immunization on estimates of outcomes	Page 9-10
Study Size	10	Explain how the study size was arrived at	ROSES-I 10.1: Describe the baseline estimated seroprevalence at given antibody titers or incidence of infection and cite published literature to support these estimates	Pages 7-11

	Item No	STROBE Items	ROSES- I items	Reported on Page
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	ROSES-I 11.1: Describe the serological assay's limit of detection and how this limit is defined or calculated. Describe how samples with a result below or on the borderline of the limit were handled in the analysis ROSES-I 11.2: Describe and justify the titer or other result used to define "seropositivity," or the antibody titer change or change in other assay result used to define "seroconversion." Avoid the term "seroconversion" unless referring to change from undetectable to detectable antibody level. Otherwise report the fold-rise in titer. Avoid the term "infection" but report "seroprevalence at a titer of …." ROSES-I 11.3: If statements or inferences are made about protection from infection, describe what is known about the correlation between the assay results and protection from infection and illness	No Results available yet as this study is ongoing
Statistical methods	12	(a). Describe all statistical methods, including those used to control for confounding (b). Describe any methods used to examine subgroups and interactions (c). Explain how missing data were addressed (d). Cohort study—If applicable, explain how loss to follow up was addressed Case–control study—If applicable, explain how matching of cases and controls was addressed Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy (e). Describe any sensitivity analyses	ROSES-I 12.1: if relevant, state how the non-independence of data was managed ROSES-I 12.2: if relevant, report methods used to account for the probability of seropositivity or seroconversion if infected, and to account for decay in antibody titers over time	No Results available yet as this study is ongoing

	Item No	STROBE Items	ROSES- I items	Reported on Page
Laboratory methods	12a		ROSES-I 12a.1: Describe the sample type—serum or plasma. If plasma is used, specify the anticoagulant used (heparin, sodium citrate, EDTA, etc.) ROSES-I 12a.2: Describe the specimen storage conditions (4°C, −20 °C, −80 °C). If frozen prior to the analysis, describe the time to freezing and the number of freeze/thaw cycles prior to testing Serological assays ROSES-I 12a.3: Specify the assay type (e.g., hemagglutination inhibition; virus neutralization/microneutralization; ELISA; other) and methods used to determine the endpoint titer ROSES-I 12a.4: Reference a previously published, CONSISE consensus serologic assay or WHO protocol if used, and any modifications of the protocol. If a previously published protocol is not used, provide full details in supplementary materials ROSES-I 12a.5: State what is known about the determinants of the variability of the antibody detection assay being used ROSES-I 12a.6: Specify the antigen(s) used in the assay, including virus strain name, subtype, lineage or clade, with standardized nomenclature and reference; specify whether live virus or inactivated virus was used (where applicable) ROSES-I 12a.7: Report if antigen(s) from potentially cross-reactive pathogens/strains were used in order to identify cross-reactivity, and specify which antigen was used, including virus name, subtype, strain, lineage and clade, with standardized nomenclature and reference ROSES-I12a.8: If red blood cells were used for a hemagglutinin inhibition assay, specify the animal species from which they were obtained and concentration (v/v) used ROSES-I 12a.9: Describe positive and negative controls used ROSES-I 12a.10: Describe starting and end dilutions ROSES-I 12a.11: Specify laboratory biosafety conditions ROSES-I 12a.12: Specify whether replication was performed, and if so, the acceptable replication parameters ROSES-I 12a.13: Specify whether a confirmatory assay was performed and all specifics of this assay, at the same level of detail ROSES-I 12a.14: Specify international standards used, if appropriate	Page 8-9
RESULTS				
Participants	13	(a). Report the numbers of individuals at each stage of the study—the numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed (b). Give reasons for non-participation at each stage (c). Consider use of a flow diagram	See STROBE item	No Results available yet as this study is Ongoing

	Item No	STROBE Items	ROSES- I items	Reported on Page
Descriptive data	14	a). Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential risk factors (b). Indicate the number of participants with missing data for each variable of interest (c). Cohort study—summarize follow-up time (e.g., average and total amount)	See STROBE item	Page 7
Outcome data	15*	Cohort study—report the numbers of outcome events or summary measures over time Case–control study—report the numbers in each exposure category, or summary measures of exposure Cross-sectional study—report the numbers of outcome events or summary measures	See STROBE item	NA
Main results	16	(a). Give unadjusted estimates and, if applicable, risk factor-adjusted estimates andtheir precision (e.g., 95% confidence interval). Make clear which risk factors were adjusted for and why they were included (b). Report category boundaries when continuous variables were categorized (c). If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	ROSES-I 16.1: Report unadjusted estimates of distribution of titers by age Group ROSES-I 16.2: Report methods to standardize the results from the study sample to the target population	No results available yet as this study is ongoing
Other analyses	17	Report other analyses performed—analyses of subgroups and interactions, and sensitivity analyses	See STROBE item	NA
DISCUSSION				
Key results	18	Summarize key results with reference to study objectives	See STROBE item	NA
Interpretation	19	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	ROSES-I 20.1: Discuss the interpretation of the results in the context of known or potential cross-reactivity	NA

	Item No	STROBE Items	ROSES- I items	Reported on Page
Generalizability	21	Discuss the generalizability (external validity) of the study results	See STROBE item	NA
Other information **Funding**	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	ROSES-I 22.1: Specify if institutional review board approval was received; if not, specify reason (e.g., public health outbreak response/non-research designation)	Page 18