Study Title: A comparison of clinical sample collection methods to detect Influenza virus and other respiratory pathogens using whole genome sequencing

Internal Reference Number / Short title: Influenza virus detection directly from clinical samples

Ethics Ref: Insert
IRAS Project ID: 269573
Date and Version No: 7 November 2019: v 1.0

Chief Investigator:

Professor Derrick Crook
Consultant in Infectious Diseases and Microbiology, Principal Investigator for Modernising Medical Microbiology Research Group, University of Oxford
Derrick.crook@ndcls.ox.ac.uk
07979 856473

Investigators:

Prof Philippa Matthews, Honorary Consultant Physician, Associate Professor in Infection and Immunity, University of Oxford

Prof Sarah Walker, Professor of Medical Statistics and Epidemiology
University of Oxford,

Dr Steven Pullan, Senior Laboratory Scientist
Public Health England

Dr Yifei Xu, Bioinformatician
University of Oxford

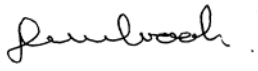
Dr Anna Zalevski
University of Oxford

Sponsor:

University of Oxford
Clinical Trials and Research Governance
1st floor, Boundary Brook House
Churchill Drive
Headington
Oxford OX3 7GB

Funder: Oxford Biomedical Research Centre, Oxford

Chief Investigator Signature:



The investigators have no potential conflicts of interest.

Confidentiality Statement

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

TABLE OF CONTENTS

1. KEY CONTACTS

Chief Investigator	Professor Derrick Crook Consultant in Infectious Diseases and Microbiology University of Oxford John Radcliffe Hospital, Oxford, OX3 9DU Derrick.crook@ndcls.ox.ac.uk Tel: 07979 856473 No fax
Sponsor	University of Oxford Clinical Trials and Research Governance Joint Research Office 1st floor, Boundary Brook House Churchill Drive, Headington, Oxford OX3 7GB ctrg@admin.ox.ac.uk Tel: 01865 616480
Funder(s)	Dr Vasiliki Kiparoglou, Chief Operating Officer Oxford Biomedical Research Centre Oxford vasiliki.kiparoglou@ouh.nhs.uk Tel: 01865 572308
Clinical Trials Unit	NA
Statistician	Prof Sarah Walker Professor of Medical Statistics and Epidemiology Microbiology, Level 7, University of Oxford John Radcliffe Hospital, Oxford, OX3 9DU Sarah.walker@ndm.ox.ac.uk Tel: 01865 220851
Committees	NA

2. LAY SUMMARY

Influenza virus infection causes an estimated 650,000 deaths globally each year with risk of severe infection being higher among certain groups, including older adults, infants and young children, pregnant women and those with underlying lung disease. Influenza virus infection is listed by the World Health Organisation (WHO) as one of the 'ten threats to global health' in 2019. Current diagnosis of influenza relies on a patient providing throat swab or nasopharyngeal aspirate that is collected in a special fluid to preserve the virus while it is transported to the laboratory. The hospital laboratory then tests this fluid by looking for small parts of the genetic material of the virus (a method called PCR). Diagnosis of other respiratory viruses requires different expensive tests. New technology allows all the genetic material in a sample to be tested using a method called whole genome sequencing, allowing us to identify both influenza and other viruses using the same test. The information that the test provides can also tell us other information that is important for providing clinical care, such as resistance to anti-viral drugs and detecting whether viruses are spreading between people.

We will recruit adults and children aged over one year who attend the Emergency Department at the John Radcliffe hospital who have respiratory symptoms suggesting they may have possible influenza infection. We are including adults and children because children have higher risks of getting influenza and we want to make sure that any new test works well in them, as well as in adults. In addition to the normal clinical sample that is taken, participants will be asked to provide for the study purposes an oral fluid sample and optionally an extra throat swab. We will also ask for information on flu vaccinations and duration of symptoms.

Our primary goal is to investigate whether a new sample, an oral fluid sample, is suitable to diagnose influenza virus by comparing the results from sequencing with the standard laboratory results. Our secondary goal is to determine whether the oral fluid sample is suitable to identify all respiratory viruses using one test, and to investigate the ways we can use this genetic information about the virus to identify resistance to anti-viral agents and to investigate how viruses are spreading in the community. This may also provide information that helps us to understand the extent to which flu infection is modified by the vaccination. In the long term, this work will contribute to better and quicker diagnosis, and therefore help the correct treatment choices to be made promptly. It will also provide insights into the spread of flu, and may help to inform improvements in immunisation so that we can protect more people.

3. SYNOPSIS

Study Title	A comparison of clinical sample collection methods to detect influenza virus and other respiratory pathogens using whole genome sequencing
Internal ref. no. / short title	Influenza virus detection directly from clinical samples
Study registration	Non-interventional study (information will be available on the HRA website)
Sponsor	University of Oxford Clinical Trials and Research Governance Joint Research Office 1st floor, Boundary Brook House Churchill Drive, Headington, Oxford OX3 7GB
Funder	Dr Vasiliki Kiparoglou, Chief Operating Officer Oxford Biomedical Research Centre Oxford vasiliki.kiparoglou@ouh.nhs.uk Tel: 01865 572308
Study Design	A feasibility study to evaluate different clinical samples for their potential for sequencing to diagnose influenza virus and other respiratory viruses
Study Participants	Symptomatic patients attending the Emergency Department at the John Radcliffe Hospital in Oxford, aged one year or older
Sample Size	Samples to be taken until a maximum of 200 influenza positives have been recruited (estimated to require total recruitment of around 1300 individuals, but this varies from season to season according to the proportion positive for influenza)
Planned Study Period	For each participant, their participation will be undertaken during the course of an emergency department visit. No follow up will be undertaken. The study is planned from 1 December 2019 to 30 November 2020.
Planned Recruitment period	Recruitment: 1 December 2019 – 30 April 2020

(ctd)

	Objectives	Outcome Measures	Timepoint(s)
Primary	Using sequencing of DNA/RNA extracted directly from oral fluids and standard clinical samples (throat swabs or nasopharyngeal aspirates (NPA)), to investigate whether it is possible to determine if *influenza* virus is present, i.e. causing influenza infection	The sensitivity and specificity of sequencing detection in oral fluids compared to existing clinical diagnosis from throat swabs or nasopharyngeal aspirates (NPA) for detection of *influenza* virus	Single cross-sectional study visit
Secondary	▪ Determine the presence of other respiratory viral infection agents in oral fluids compared to throat swabs/NPA ▪ Determine resistance to influenza antivirals detected from oral fluids or throat swabs/NPA ▪ Determine full genome sequences to provide insights into influenza diversity and spread	▪ The sensitivity and specificity of sequencing in oral fluids compared to existing clinical diagnosis (throat swabs or NPA) of viral respiratory pathogens and the turnaround time from starting sample processing to determination of influenza or other viral presence. ▪ The proportion of samples that are sequenced with adequate sequence depth and quality to identify resistance determinants ▪ Generation of phylogenetic trees to investigate relatedness between influenza (and other respiratory viruses).	Single cross-sectional study visit
Exploratory Objectives	▪ Optimization of laboratory methods for extraction of high-quality, long-fragment influenza virus directly from samples ▪ Development of novel bioinformatic algorithms for a) identifying influenza infection; b) identifying influenza virus resistance determinants; c)	▪ Increase in the number of Influenza or other respiratory viral genetic reads ▪ Improvements in bioinformatic processing time/accuracy and/or development of computational pipeline for high volume processing of genetic data. ▪ Physiological measurements, blood test results, and need for	

	identifying other viral pathogens ▪ Investigation of viral predictors of severity of presentation	antiviral/antibiotic medication (from routine electronic health records)	
Intervention(s)	NA		
Comparator	NA		

4. ABBREVIATIONS

CI	Chief Investigator
CTRG	Clinical Trials & Research Governance, University of Oxford
DNA	Deoxyribonucleic Acid
GCP	Good Clinical Practice
GP	General Practitioner
HRA	Health Research Authority
ICMJE	International Committee of Medical Journal Editors
NAI	Neuraminidase Inhibitors
NHS	National Health Service
NPA	Nasopharyngeal Aspirate
ONT	Oxford Nanopore Technologies
OUHFT	Oxford University Hospitals NHS Foundation Trust
PCR	Polymerase chain reaction
R&D	NHS Trust R&D Department
REC	Research Ethics Committee
RES	Research Ethics Service
RNA	Ribonucleic Acid
SISPA	Sequence-independent, single-primer amplification
WHO	World Health Organisation

5. BACKGROUND AND RATIONALE

Influenza virus infection is a substantial public health threat, causing an estimated 650,000 deaths globally each year[1]. Certain groups are particularly at risk of severe disease, including older adults, infants and young children, pregnant women, those with underlying lung disease, and the immunocompromised[2]. Influenza virus infection is listed by the World Health Organisation (WHO) as one of the 'ten threats to global health' in 2019[3].

The influenza virus can evolve by exchanging genetic material (called reassortment) from different hosts, such as birds, humans or pigs[4], creating new varieties of the virus that may cause more serious illness[5,6,7]. Influenza diagnostics and surveillance are fundamental to identify the emergence of these novel strains, to improve prediction of potential epidemics and pandemics[6,8]. Surveillance also informs vaccine strategy[9] by identifying whether any current vaccine protects people against the circulating influenza virus as well as potentially providing information for future vaccine composition. Diagnostic data also facilitates real-time surveillance, can underpin infection control protocols[10,11] and inform the prescription of Neuraminidase Inhibitors (NAI)[2].

The application of Oxford Nanopore Technologies (ONT; https://nanoporetech.com/) to generate full-length influenza sequences from clinical respiratory samples can address these challenges. ONT offers a 'third-generation', portable, real-time approach to generating long-read sequence data, with demonstrated success across a range of viruses[12-15]. Our group has undertaken a pilot laboratory study to improve the laboratory detection of influenza virus from anonymised throat swabs and naopharyngeal aspirates (NPA) that demonstrates the utility of ONT long-read sequencing in identifying the presence of influenza virus and simultaneously detecting other viruses[16].

Currently, the recommended viral respiratory clinical diagnosis pathways require either throat swab or NPA samples taken by medical personal and fast transport to the diagnostic laboratory to prevent samples from degrading. An improved sample for influenza viral testing requires the following; simple for non-professionals to administer, more acceptable to patients than the current tests and less time sensitive. One possibility is oral fluid samples, currently the standard for mumps testing by Public Health England; these are easy to obtain, acceptable to patients and may be posted to diagnostic laboratories. A small study found that oral fluid sampling could be used for detection of influenza A virus in a population that was tested for mumps[17]. However, this study did not compare whether oral fluid testing was as good as throat swabs in detecting influenza and other respiratory pathogens.

The primary outcome of this study is therefore to compare the use of oral fluid samples against the current sample standard of throat swabs or NPA for detecting influenza and other respiratory viral pathogens by whole genome sequencing.

6. OBJECTIVES AND OUTCOME MEASURES

Objectives	Outcome Measures	Timepoint(s) of evaluation of this outcome measure (if applicable)
▪ **Primary Objective** Using sequencing of DNA/RNA extracted directly from oral fluids and standard clinical samples (throat swabs or nasopharyngeal aspirates (NPA)), to investigate whether it is possible to determine if influenza virus is present	The sensitivity and specificity of sequencing detection in oral fluids compared to existing clinical diagnosis from throat swabs or nasopharyngeal aspirates (NPA) for detection of *influenza* virus	N/A – cross-sectional study at single timepoint
Secondary Objectives ▪ Determine the presence of other respiratory viral infection agents in oral fluids compared to throat swabs/NPA ▪ Determine resistance to influenza antivirals detected from oral fluids or throat swabs/NPA ▪ Determine full genome sequences to provide insights into influenza diversity and spread	▪ The sensitivity and specificity of sequencing in oral fluids compared to existing clinical diagnosis (throat swabs or NPA) of other viral respiratory pathogens* and the turnaround time from starting sample processing to determination of influenza or other viral presence ▪ The proportion of samples that are sequenced with adequate sequence depth and quality to identify resistance determinants ▪ Generation of phylogenetic trees to investigate relatedness between influenza (and other respiratory viruses).	N/A – cross-sectional study at single timepoint
Exploratory Objectives • Optimization of laboratory methods for extraction of high-quality, long-fragment influenza virus directly from samples ▪ Development of novel bioinformatic algorithms for a) identifying influenza infection; b) identifying influenza virus resistance determinants; c) identifying other viral pathogens • Investigation of viral predictors of severity of presentation	• Optimization of laboratory methods for extraction of high-quality, long-fragment influenza virus directly from samples. • Development of novel bioinformatic algorithms for a) identifying influenza infection; b) identifying influenza virus resistance determinants; c) identifying other viral pathogens* • Physiological measurements, blood test results, and need for antiviral/antibiotic medication (from routine electronic health records)	N/A – cross-sectional study at single timepoint

* specifically those tested by the Film Array Respiratory Panel 2: adenovirus, coronavirus 229E, coronavirus HKU1, coronavirus NL63, coronavirus OC43, human metapneumovirus, human rhinovirus/enterovirus, influenza virus A, influenza virus A H1, influenza virus A H1-2009, influenza virus A H3, influenza virus B, parainfluenza virus 1, parainfluenza virus 2, parainfluenza virus 3, parainfluenza virus 4, respiratory syncytial virus, *Bordetella pertussis*, *Chlamydia pneumoniae*, *Mycoplasma pneumonia*, Middle East respiratory syndrome coronavirus and *Bordetella parapertussis*.

7. STUDY DESIGN

This study is an observational, non-interventional, cross-sectional diagnostic study. The study is hospital based (John Radcliffe) at one study site where participants will be recruited. Laboratory processing and analysis of research samples will take place at the Nuffield Department of Medicine, University of Oxford (at laboratories within the John Radcliffe Hospital, OUH NHS Foundation Trust) and Porton Down, PHE. The expected duration of participant involvement is 30 minutes in total: that includes the consent process, a short questionnaire and taking the samples. There is no follow-up.

8. PARTICIPANT IDENTIFICATION

8.1. Study Participants

Participants with respiratory symptoms aged more than one year will be recruited at the Emergency Department at Oxford University Hospitals NHS Foundation Trust (OUHFT).

8.2. Inclusion Criteria

- Adults or children (over one year old) with symptomatic respiratory illness attending the Emergency Department in Oxford and providing a standard diagnostic specimen.
- Written informed consent by participant or parent/carer of children/adolescents.

8.3. Exclusion Criteria

There are no additional exclusion criteria.

9. PROTOCOL PROCEDURES

9.1. Recruitment

Symptomatic patients attending the Emergency department will be approached by a member of the clinical care staff to ask if they are prepared to participate in the study. Potential participants will be provided with a patient information sheet appropriate for their age and will have the opportunity to discuss the study with a study nurse, doctor or other health professional trained in Good Clinical Practice (GCP). A separate simplified version will be provided for parents and older children and adolescents under 16 years, with a separate pictoral version for children aged 6-10 years, which will be provided depending on illness state, capacity, and parental preference reflecting this. As children will be presenting acutely sick to the Emergency department, we will provide information only for parents for children aged 5 years and under.

9.2. Screening and Eligibility Assessment

A member of the medical team in the Emergency Department will screen patients as part of standard triage. Those patients from whom a standard diagnostic specimen is taken for the purposes of influenza diagnostics are potentially eligible and will be provided with the relevant patient information sheet, providing they are judged to be well enough to provide genuinely informed consent. There will be no exceptions made regarding eligibility, i.e., each participant must satisfy all the approved inclusion and exclusion criteria of the protocol.

9.3. Informed Consent

Written and verbal versions of the participant information and consent form will be presented to the participants detailing no less than: the exact nature of the study; what it will involve for the participant; the implications and constraints of the protocol; any risks involved in taking part. It will be clearly stated that the participant is free to withdraw from the study at any time for any reason without prejudice to future care, without affecting their legal rights, and with no obligation to give the reason for withdrawal. Simplified versions will be provided to children and adolescents aged 6 to 15 years old, as described above, depending on illness state, capacity and parental preference reflecting this).

Informed consent includes the ability to understand the information being presented, the ability to retain this information, and the means with which to process the information in order to make an individual decision.
Specific areas of consent sought include:

- Provision of study samples, namely oral fluid samples and extra throat swabs and use of remaining clinical samples.
- Provision of study questionnaire data
- Use of routine electronic health records relating to this hospital visit linked though their hospital number
- Publication of findings without any identifying information
- Public release of any sequence data generated, after this is fully anonymised and any human DNA sequence generated removed
- Storage of data and samples for defined periods

Participants will be assured that all study-associated records will be treated as confidential and kept in a secure location.

Written informed consent will be obtained by means of participant dated signature (adults 16 years or older, or parent/carer if <16 years at last birthday) and dated signature of the person who presented and obtained the informed consent. Older children and adolescents under 16 years will also be asked for written or verbal assent based on their capacity to understand and provided genuinely informed assent, which may also depend on their illness acuity. The person who obtained the consent (and assent as relevant) must be suitably qualified and experienced and have been authorised to do so by the Chief/Principal Investigator. A copy of the signed informed consent form (and assent form, as relevant) will be given to the participant. The original signed form(s) will be retained at the study site. These consent/assent forms will be held in a locked cabinet or room at the study site for 15 years after it ends. The consent form (and assent form as relevant) will be signed prior to the performance of any study-related activity.

9.4. Enrolment

Participants will be identified by healthcare workers in the Emergency Department at the John Radcliffe Hospital and approached to take part in the study as above. After allowing sufficient time for discussion and questions, participants or their parents/guardians will be asked to consent at this initial clinic visit, as this is when diagnostic samples are being obtained. The participant will be allowed as much time as wished to consider the information subject to the duration of the emergency department visit, and the opportunity to ask questions before they decide whether they will participate in the study.

There is no intervention in the study, and there is no randomisation.

9.5. Blinding and code-breaking

There is no blinding in the study, and/or no code breaking procedure.

9.6. Description of study intervention(s), comparators and study procedures (clinical)

There is no study intervention and/or comparator.

9.6.1. Description of study procedure(s)

Samples

Oral fluid sample

All participants will be asked to provide an oral fluid sample for research purposes in addition to the throat swab/NPA sample they provide for routine clinical testing:

- The sampling sponge head of the swab is rubbed over the teeth and gums for one to two minutes by the healthcare worker (based on the Oracol Oral fluid collection device).
- The swab is placed into the plastic swab holder tube by the healthcare worker.
- The sample tube will be transported to the clinical laboratory with the routine throat sample. Collection of the sample by research staff from the routine laboratory will take place Monday to Friday.

Throat swab - optional

Participants will also be asked to provide an extra throat swab in addition to any obtained for routine clinical testing in order to ensure that sufficient material is available for testing for this research study. Consent for this can be refused and the participant can still join the study:

- A flocked swab will be placed into the participant's throat and rotated for 10 seconds and withdrawn. The swab will be taken by healthcare staff, in exactly the same manner as the

standard clinical throat swab.
- The swab will be placed into 3ml of universal transport medium according to the manufacturer's instructions
- The sample tube will be transported to the clinical laboratory with the routine throat sample. Collection of the sample by research staff from the routine laboratory will take place Monday to Friday.

Questionnaire data

A written questionnaire (to be completed by the participant or the participant parent or guardian) will record:
- Participant study identifier (added by staff member)
- Date of sampling
- Age at last birthday: 1-4/5-12/13–17/18-24/ 25-34/35-44/45-54/55-64/65-74/75-84/85 years or older
- Gender
- How many days have you had flu-like symptoms for?
- Have you had ever an influenza/flu vaccine? If yes, did you receive vaccination this year/last year/Before last year (all that apply)
- The outward portion of postcode (not considered identifiable by NHS Digital) to be used to look for the potential for local transmission
- Have you been outside England in the last month? If so, what country/countries?

Note: the pseudonymised study identifier will be added by the clinical member of staff recruiting the patient. It will be linked on a separate CRF to hospital number (required to link to routine microbiology results and routinely collected hospital data about the visit to the hospital, see Section 12 below) and the laboratory identifier of the routine clinical sample (to be supplied by laboratory).

9.7. Study Visit

The study sampling and baseline data collection will occur while the participant is attending the Emergency Department. Medical care as deemed appropriate for the visit will occur. Participants will be approached to take part in the study, given the study information and time for discussion. If they consent, extra samples will be taken (outlined above) and a questionnaire given. The questionnaire can be completed by the participant or the parent or guardian. The total time for this will be 3-5 mins. There are no further visits required.

9.8. Subsequent Visits

Not applicable.

9.9. Sample Handling

9.9.1 Sample handling for study purposes

Oral fluid samples and the optional additional throat swab (research samples) will be returned to the local NHS microbiology laboratory together with the routine NHS diagnostic sample(s) (standard of care sample). On arrival in the routine microbiology laboratory, study samples will be stored at 4°C. Study samples will be retrieved daily on weekdays from the OUFT NHS microbiology laboratory by researchers and assayed in the University laboratory which is adjacent to the routine laboratory in the John Radcliffe hospital. Residual study samples will be

shipped to co-investigators at PHE Porton Down for any necessary work to further optimize laboratory methods for extraction of high-quality, long-fragment influenza virus directly from samples.

All study samples will be labelled with a pseudonymised participant study identifier. Hospital number will only be used to link to routine electronic health records to retrieve the results of the standard diagnostic tests and routinely collected information about the visit to hospital.

Firstly, DNA/RNA will be extracted from oral fluid and routine clinical samples using the method developed in an initial pilot study[16]. A spin step removes the majority of human DNA. Extracted material will be stored and then sequenced in batches using ONT sequencing. The sensitivity and specificity of bioinformatic detection of influenza infection and other viral infections from this sequencing will be compared with molecular detection of infection as determined by normal clinical processing, a molecular PCR test (GeneXpert, Cepheid) provided by the routine diagnostic microbiology laboratory. The results generated will not be released to clinicians, firstly given their experimental nature and secondly, because they are likely to be generated some time after relevant clinical decision-making has occurred. Individual participant results will therefore not form part of patient care or interfere with routine diagnostic testing.

The study involves metagenomic sequencing of all genetic material present in the study samples. This approach is what makes it possible to attempt to make a same-day diagnosis. As a consequence, some human DNA is likely to be sequenced from study samples. Output sequence data will be screened for human DNA reads before analysis. All human DNA will be permanently deleted before the samples are analysed. No human DNA sequence will be kept, analysed or made available as part of the study.

We will also investigate several approaches to maximize the amount of high-quality long fragment RNA available for sequencing. We aim to extract as much RNA as possible from each sample and where possible to selectively extract viral RNA. Initial work[16] demonstrates that our current method successfully detects influenza virus and other respiratory viruses with a limit of detection of 10^2-10^3 genome copies/ml. Coverage of virus when detected was adversely impacted by sample composition and the presence of bacterial and human DNA reads.

Strategies to reduce the extent of human DNA present will be tested, including comparing the current centrifugation processing to use of chemical lysis agents e.g. saponin. The protocol developed during the pilot study requires modification to reduce the time and complexity of sample processing. We will test modifications of this method (called Sequence-independent, single-primer amplification, SISPA) to reduce time taken in PCR tests as well as testing a Rapid sequencing kit (from Nanopore) that requires a shorter preparation time. Sequencing will be undertaken using a combination of the ONT MinION and GridION platforms.

To reconstruct whole viral genomes, we will use a mapping-based approach to compensate for the relatively high per-base error rate in ONT sequence data. Centrifuge will be used to screen out contaminating reads. The closest reference genome will be chosen a k-mer based similarity measure using distance measures based on the mash algorithm. Mapping will be undertaken using minimap2 followed by consensus and variant calling, specific algorithms for which will be developed during the project.

Where additional DNA/RNA is extracted that is not needed for immediate sequencing (nucleic acid extraction quantity can be variable depending on the sample), then consent to retain this,

within a secure freezer at the University of Oxford, for use in future research will be sought. This total extracted material will contain microbial DNA/RNA and could potentially also contain some human DNA. However, no analysis of any sequences of human DNA will be undertaken in any future research. All human DNA sequenced data will be permanently discarded at the earliest opportunity before any analysis is undertaken. Residual DNA/RNA extracts (not containing any cellular material and therefore outside of the Human Tissue Act) will be retained for up to 15 years and then securely destroyed. Swabs will be retained for 12 months after the end of the study and then securely destroyed. Therefore, any samples containing human cells will be stored for a maximum of 12 months after the end of the study whereas acellular extracted nucleic acids will be stored for up to 15 years.

9.9.2 Sample handling for standard of care

The standard of care throat swab will be processed in the routine microbiology laboratory using standard NHS accredited workflows. We will retrieve results from these routine tests from the laboratory information systems, using the hospital number and clinical laboratory identification identifiers provided by the routine laboratory staff matched to the hospital numbers associated with the diagnostic sample. Any residual material not required for routine testing will be used in the study with the research samples for collection and processing, as described above.

9.10. Early Discontinuation/Withdrawal of Participants

Participants may withdraw their consent, meaning that they wish to withdraw from the study completely. In this case samples will not be taken. If samples have already been taken, any remaining sample that has not already been tested will be destroyed. We will use any data provided (e.g. from the questionnaire) and/or generated (e.g. from sequencing) up to the time of withdrawal, and this is explained in the patient information sheet.

9.11. Definition of End of Study

The study will end at the completion of analysis, including analysis of sequences from study samples.

10. SAFETY REPORTING

Safely reporting is not required because there is no intervention (or comparator), and because the only additional samples being taken are either identical to those taken for routine clinical diagnostics (throat swab, NPA) or are less invasive (oral fluid sample). There is therefore minimal risk of harm to any patient from participating.

11. STATISTICS AND ANALYSIS

Bioinformatic algorithms developed in the pilot study[16] will be used to determine whether influenza virus and other viruses are present in the sequencing data obtained from the samples.

11.1. Analysis Plan

Analysis will firstly determine the sensitivity and specificity of ONT sequencing from oral fluid and throat swabs compared to existing molecular diagnostics (GeneXpert PCR) for detection of influenza virus. The same comparison will be undertaken for detection of other respiratory viruses. Estimates will be reported with 95% confidence intervals (CI).

The proportion of samples that are sequenced with adequate sequence depth and quality to identify resistance determinants in influenza virus will also be estimated, and we will determine the prevalence of key drug resistance mutations, at the level of consensus sequences and within quasispecies. Relatedness between influenza genomes will be estimated using standard phylogenetic models to assess close genetic clusters and to make inferences about the plausibility of transmission (indicated by very close genetic relationships).

11.2. Sample Size Determination

Research samples will be collected until a minimum of 200 positive influenza virus positive samples are obtained. 200 influenza positive samples provide a >0.95 probability that the lower limit of the 95% CI and the final estimate of sensitivity exceeds 95%, assuming the true sensitivity is 99% after method optimisation. Initial study sensitivity was 100% for samples with cycle threshold (CT) <30 (two-thirds of the samples) but only 46% for >30 (overall 83%).

If the overall sensitivity in this study is lower than expected after methods optimisation, say 85%, 200 influenza positives also provides a >0.95 probability that the lower limit of the 95% CI around the final estimate of sensitivity exceeds 75%.

It is not possible to know in advance of recruitment which patients will have which laboratory results. We estimate that 15% of symptomatic participants will have a diagnosis of influenza or other viral respiratory infection. Thus, we anticipate recruiting 1300 participants to get 200 influenza positive samples. Annually approximately 500 cases of molecular-confirmed influenza or other respiratory viral infection are identified in Oxford during the influenza season (November to May).

As this study is designed to test suitability of clinical samples for detection of influenza virus, the sample size is based on providing relatively precise estimates of sensitivity (as above), the number of participants that can feasibly be recruited in the fixed study duration and the size of the budget available for sequencing consumables.

11.3. Analysis populations

All samples provided by participants will be included in the analysis, providing there is a result from the standard diagnostic test and a result from sequencing (which may be that sequencing was attempted on a sample and failed).

11.4. Decision points

Not applicable.

11.5. Stopping rules

There are no stopping rules.

11.6. The Level of Statistical Significance

95% confidence intervals will be used.

11.7. Procedure for Accounting for Missing, Unused, and Spurious Data.

There are no specific methods for missing, unused or spurious data. Analysis compares results from different technologies on the same sample, and therefore by definition is restricted to participants with a result from the standard diagnostic test and who provided a sample in which sequencing could be attempted. Failure of sequencing will be counted as failure (i.e. incorrect result).

11.8. Procedures for Reporting any Deviation(s) from the Original Statistical Plan

Not applicable.

11.9. Health Economics Analysis

Not applicable.

12. DATA MANAGEMENT

The plans for the data management of the study are outlined below. There is not a separate Data Management document in use for the study.

The consent forms will be stored in a locked room or cabinet at the study site. The log linking hospital numbers to study identifiers will be held securely in a locked room or cabinet at the study site. The electronic version of this log will be stored within the NHS network, and will be destroyed 3 months after the end of the study. Questionnaire data will be identified only by a pseudonymised study identifier and will be entered into an Excel spreadsheet (since it is a very limited and simple questionnaire with a small number of fields). A limited number of people will have access to the spreadsheet, which will be locked/password protected and backed up at regular intervals (and only contain study numbers, not patient identifiable information).).

In order to reduce burden on participants and reduce duplication of effort, we will ask participants for consent to retrieve information about this hospital visit from their routine electronic records. Specifically, we will retrieve from underlying NHS data systems

1. Date and time of arrival at the Emergency Department and of any onward admission as an inpatient.
2. If admitted as an inpatient, details of the admission, including which wards the patient was admitted to together with information about specific bed locations (including side room or open bay) where available, date of discharge, consultant episodes including primary and secondary diagnoses from which co-morbidities will be assess (particularly immunosuppression, diabetes, chronic lung/renal/cardiovascular disease). This information will be used to investigate the possibility of nosocomial transmission of influenza and any other respiratory viruses identified.
3. All microbiological and virological test results, collection dates and times from this visit and any onward inpatient admission. This information will be used to investigate the ability of the new sequencing method to identify other pathogens.
4. Physiological measurements and dates/times of measurements from this visit and any onward inpatient admission, specifically temperature, respiratory rate, heart rate, oxygen saturations, blood pressure, weight, height where available (to calculate BMI). This information will be used to investigate whether virus characteristics are associated with more or less severe presentations with influenza, adjusting for known potential confounders (BMI), in exploratory analyses.
5. Laboratory test results and dates/times of specimens from this visit and any onward inpatient admission, specifically CRP, total white cell count and differential white cell count (neutrophils, lymphocytes, eosinophils etc), creatinine (to calculate estimated glomerular filtration rate), HbA1C (to assess diabetes). This information will be used to investigate whether virus characteristics are associated with more or less severe presentations with influenza, adjusting for known potential confounders (diabetes), in exploratory analyses.
6. Antiviral (e.g. tamiflu) and antibiotic prescription(s) and dates from this visit and any onward inpatient admission. This information will be used to investigate whether virus characteristics are associated with more or less severe presentations with influenza, as indicated by need for medication.

De-identified sequencing data and associated metadata from the participant questionnaires and electronic health records will be stored within the University of Oxford and retained for 15 years after the end of the study. De-identified data will be held securely on University network drives

or University encrypted laptops with password protected access restricted only to researchers directly involved in the project.

Metagenomic sequencing of oral and swab samples will unavoidably generate human DNA sequence data. The speed advantage of the technique is derived from sequencing all DNA/RNA present in the sample, and so sequencing of human DNA and other bacterial DNA alongside Influenza and other viral pathogens DNA/RNA cannot be avoided. All human DNA sequenced data will be permanently discarded at the earliest opportunity before further analysis is undertaken. No human DNA will be analysed as part of the project.

Sequence data generated from the project, after removal of human DNA, will be made publicly available via the European Bioinformatic Institute's nucleotide archive.

12.1. Source Data

Source documents are where data are first recorded, and from which participants' data are obtained. These include, but are not limited to, hospital records and laboratory records.

For this study, the electronic health record is the source data for age, gender and outward portion of postcode, and for the specific items extracted from it above. The questionnaire itself will be considered source data for the other questionnaire items because it is the site of the original recording (e.g. there is no other written or electronic record of data). All documents will be stored safely in confidential conditions. On all study-specific documents, other than the signed consent, the participant will be referred to by their study number, not by name.

12.2. Access to Data

Direct access will be granted to authorised representatives from the Sponsor and host institution for monitoring and/or audit of the study to ensure compliance with regulations.

12.3. Data Recording and Record Keeping

All questionnaire data will be entered onto an Excel spreadsheet on a password protected computer, located in locked offices at the Nuffield Department of Medicine, University of Oxford, John Radcliffe Hospital, Oxford.

The participants will be identified by a unique study specific number in any database. The hospital number (the only identifying details) will NOT be included in any data electronic files. See above for consent form storage.

The data will be kept for 15 years. The consent forms will be kept for 15 years as the stored acellular material may be used for further studies.

13. QUALITY ASSURANCE PROCEDURES

The study may be monitored, or audited in accordance with the current approved protocol, GCP, relevant regulations and standard operating procedures.

13.1. Risk assessment

No formal risk assessment is required.

The study involves recruiting symptomatic patients attending the Emergency department at the OUHFT. Participants will be asked to provide an oral fluid and optional extra throat sample at the clinic visit. The duration of participation in the study is limited to the clinic attendance.

The main burden of participating in the study is the time taken for the initial discussion of the study, the consent process and completing a brief study questionnaire. The time to obtain additional study samples will be limited.

There is minimal risk of harm to any patient from participating.

The study does not include any intervention and individual participant results will not form part of patient care or interfere with routine diagnostic testing. The results generated will not be released to clinicians, firstly given their experimental nature and secondly, they are likely to be generated some time after relevant clinical decision making has occurred.

The study involves metagenomic sequencing of all DNA/RNA present in the study samples. This approach is what makes it possible to attempt to make a same-day diagnosis. As a consequence of this approach human DNA is likely to be sequenced from study samples. Output sequence data will be screened for human DNA reads before analysis. All human DNA sequences will be permanently deleted before the samples are analysed. No human DNA will be kept, analysed or made available as part of the study.

13.2. Study monitoring

No GCP monitoring will be undertaken. As described above, there are minimal risks posed to patients by this observational and non-interventional study. The only data items are either retrieved directly from electronic hospital records (that is, are source documents in their own right against which no monitoring is possible) or are participant responses to a questionnaire. The only study procedures are completing the questionnaire and taking research samples – absence of sample by definition means that research procedures were not followed. One important goal of the study is to identify whether the oral sample is suitable for viral diagnostics: whilst staff will be trained in how to take these samples following recommended procedures, if in practice, it is not possible to gain sufficient material from how they are taken to sequence viruses within them, then this is an important study finding. The inclusion criteria are extremely simple – they will be recorded on the single study case record form with the study identifier, hospital number and laboratory number and will be confirmed when results are retrieved from the laboratory.

13.3. Study Committees

There are no oversight committees for this diagnostic study, as it poses minimal risks to patients.

14. PROTOCOL DEVIATIONS

A study related deviation is a departure from the ethically approved study protocol or other study document or process (e.g. consent process or administration of study intervention) or from Good Clinical Practice (GCP) or any applicable regulatory requirements. Any deviations from the protocol will be documented in a protocol deviation form and filed in the study master file.

15. SERIOUS BREACHES

A "serious breach" is a breach of the protocol or of the conditions or principles of Good Clinical Practice which is likely to affect to a significant degree

(a) the safety or physical or mental integrity of the trial subjects; or

(b) the scientific value of the research.

In the event that a serious breach is suspected the Sponsor must be contacted within 1 working day. In collaboration with the Chief Investigator, the serious breach will be reviewed by the Sponsor and, if appropriate, the Sponsor will report it to the approving REC committee and the relevant NHS host organisation within seven calendar days.

16. ETHICAL AND REGULATORY CONSIDERATIONS

16.1. Declaration of Helsinki

The Investigator will ensure that this study is conducted in accordance with the principles of the Declaration of Helsinki.

16.2. Guidelines for Good Clinical Practice

The Investigator will ensure that this study is conducted in accordance with relevant regulations and with Good Clinical Practice (GCP).

16.3. Approvals

Following Sponsor approval, the protocol, informed consent form and the participant information sheets will be submitted to an appropriate Research Ethics Committee (REC), and HRA (where required) and host institutions for written approval.

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

16.4. Other Ethical Considerations

For children and adolescents under 16 years, consent will be obtained from the parents/legal guardian. We propose to not record additional written assent from these child/adolescent participants, because they will be acutely unwell, and the study poses minimal additional burden to them. However, children and adolescents should provide verbal assent based on age and capacity, and the written consent from the parents/legal guardian will be assumed to reflect this.

16.5. Reporting

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

16.6 Transparency in Research

As the study is non-interventional it will not be registered on a clinical trials site. It will be registered on HRA Summaries (https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/) and (after approvals have been gained) it will also be described on the group's website (http://modmedmicro.nsms.ox.ac.uk/).

16.7 Participant Confidentiality

The study will comply with the General Data Protection Regulation (GDPR) and Data Protection Act 2018, which require data to be de-identified as soon as it is practical to do so. The processing of the personal data of participants will be minimised by making use of a unique participant study number only on all study documents and any electronic database(s). All documents will be stored securely and only accessible by study staff and authorised personnel. The study staff will safeguard the privacy of participants' personal data.

16.8 Expenses and Benefits

No expenses will be incurred by study participants. No payments will be made to study participants.

17. FINANCE AND INSURANCE

17.1. Funding

Funding is provided by the Oxford Biomedical Research Centre.

17.2. Insurance

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

17.3. Contractual arrangements

Appropriate contractual arrangements will be put in place with all third parties.

18. PUBLICATION POLICY

The Investigators will be involved in reviewing drafts of the manuscripts, abstracts, press releases and any other publications arising from the study. Authors will acknowledge that study funding by the Oxford Biomedical Research Centre. Authorship will be determined in accordance with

the International Committee of Medical Journal Editors (ICMJE) guidelines and other contributors will be acknowledged.

19. DEVELOPMENT OF A NEW PRODUCT/ PROCESS OR THE GENERATION OF INTELLECTUAL PROPERTY

Ownership of IP generated by employees of the University vests in the University. The University will ensure appropriate arrangements are in place as regards any new IP arising from the trial.

20. ARCHIVING

We will keep questionnaires identified only with participants' study identifiers in a card-protected research office for 15 years after the end of the study.

21. REFERENCES

1 Krammer F, Smith GJD, Fouchier RAM, Peiris M, Kedzierska K, Doherty PC, et al. Influenza. Nat Rev Dis Primers. 2018;4: 3.

2 Uyeki TM, Bernstein HH, Bradley JS, Englund JA, File TM Jr, Fry AM, et al. Clinical Practice Guidelines by the Infectious Diseases Society of America: 2018 Update on Diagnosis, Treatment, Chemoprophylaxis, and Institutional Outbreak Management of Seasonal Influenza. Clin Infect Dis. 2018; doi:10.1093/cid/ciy866 .

3 Ten health issues WHO will tackle this year [Internet]. [22 June 2019]. Available: https://www.who.int/emergencies/ten-threats-to-global-health-in-2019.

4 Joseph U, Su YCF, Vijaykrishna D, Smith GJD. The ecology and adaptive evolution of influenza A interspecies transmission. Influenza Other Respi Viruses. 2017;11: 74–84.
5 Humphreys M. The influenza of 1918: Evolutionary perspectives in a historical context. Evol Med Public Health. 2018;2018: 219–22.

6 Smith GJD, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, et al. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature. 2009;459: 1122–1125.

7 Sutton TC. The Pandemic Threat of Emerging H5 and H7 Avian Influenza Viruses. Viruses. 2018;10. doi:10.3390/v10090461.

8 WHO | Burden of disease. World Health Organization; 2018; Available: https://www.who.int/influenza/surveillance_monitoring/bod/en/ .

9 Morris DH, Gostic KM, Pompei S, Bedford T, Łuksza M, Neher RA, et al. Predictive Modeling of Influenza Shows the Promise of Applied Evolutionary Biology. Trends Microbiol. 2018;26: 102–118.

10 Blanco N, Eisenberg MC, Stillwell T, Foxman B. What Transmission Precautions Best Control Influenza Spread in a Hospital? Am J Epidemiol. 2016;183: 1045–1054.

11 Sansone M, Wiman Å, Karlberg ML, Brytting M, Bohlin L, Andersson L-M, et al. Molecular characterization of a nosocomial outbreak of influenza B virus in an acute care hospital setting. J Hosp Infect. 2019;101: 30–37. 11 Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 2014; **15**: 1.

12 Kafetzopoulou LE, Efthymiadis K, Lewandowski K, Crook A, Carter D, Osborne J, et al. Assessment of metagenomic Nanopore and Illumina sequencing for recovering whole genome sequences of chikungunya and dengue viruses directly from clinical samples. Euro Surveill. 2018;23. doi:10.2807/1560-7917.ES.2018.23.50.1800228 625.

13 Karamitros T, van Wilgenburg B, Wills M, Klenerman P, Magiorkinis G. Nanopore sequencing and full genome de novo assembly of human cytomegalovirus TB40/E reveals clonal diversity and structural variations. BMC Genomics. 2018;19: 577.

14 Greninger AL, Naccache SN, Federman S, Yu G, Mbala P, Bres V, et al. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. Genome Med. 2015;7: 99. 634.

15 McNaughton AL, Roberts HE, Bonsall D, de Cesare M, Mokaya J, Lumley SF, et al. Illumina and Nanopore methods for whole genome sequencing of hepatitis B virus (HBV) [Internet]. bioRxiv. 2018. p. 470633. doi:10.1101/470633.

16 Lewandowski K, Xu Y, Pullen ST, Lumley SF, Foster D, Sanderson N et al. Metagenomic Nanopore sequencing of influenza virus direct from clinical respiratory samples. Deposited bioRxiv 19 June 2019 (under review) https://www.biorxiv.org/content/10.1101/676155v1 .

17 Thompson CI, Ellis J, Galiano M, Ramsey M, Brown KE, Zambon M. Detection of influenza A(H3N2) virus in children with suspected mumps during winter 2014/15 in England. EuroSureill.2015;20(31).