A close up of a logo

Description automatically generated

**UTI diagnosis in pregnancy by VOC analysis**

**Point of care diagnosis of urinary tract infections (UTIs) in pregnancy by volatile organic compound (VOC) analysis**

**PROTOCOL**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **IRAS number** | | 282829 | | |
| **ClinicalTrials.gov number** | |  | | |
| **Sponsor** | | University of Warwick | | |
| **Sponsor reference** | |  | | |
| **Funder** | | Warwick-Wellcome Translational Partnership | | |
| **REC reference** | | **20/NW/0450** | | |
| **REC approval date** | | 30th March 2021 | | |
| **HRA approval date** | |  | | |
| **Version number** | | 1.0 | | |
| **Date** | |  | | |
| **Planned Start date** | | 1.11.20 | | |
| **Planned End date** | | 31.12.2023 | | |
| Protocol Amendments | | | | |
| **Amendment Number** | Protocol Version | | Date of Amendment | Date of Approval |
|  |  | |  |  |

**Confidentiality statement**

All information contained within this document is regarded as, and must be kept, confidential. No part of this document may be disclosed to any Third Party without the written permission of the Chief Investigator and/or Sponsor.

# SIGNATURE PAGE

The undersigned confirm that the following protocol has been agreed and accepted and that the Chief Investigator agrees to conduct the trial in compliance with the approved protocol and will adhere to the principles outlined in the UK Policy Framework for Health and Social Care Research, the ICH Good Clinical Practice guidelines and the Sponsor’s SOPs.

I agree to ensure that the confidential information contained in this document will not be used for any other purpose other than the evaluation or conduct of the clinical investigation without the prior written consent of the Sponsor.

I also confirm that I will make the findings of the study publically available through publication or other dissemination tools without any unnecessary delay and that an honest accurate and transparent account of the study will be given; and that any discrepancies from the study as planned in this protocol will be explained.

|  |  |  |
| --- | --- | --- |
| **For and on behalf of the Study Sponsor:** | | |
| Signature:  .............................................................................. |  | Date:  ......../......../........ |
| Name (please print):  .............................................................................. |  |  |
| Position:  .............................................................................. |  |  |
| **Chief Investigator:** | | |
| Signature:  .............................................................................. |  | Date:  ......../......../........ |
| Name: (please print):  ..............................................................................  Position:  .............................................................................. |  |  |

# KEY TRIAL CONTACTS

|  |  |
| --- | --- |
| **Chief Investigator** | ***Lauren Lacey***  *NIHR Academic Clinical Lecturer*  *UHCW NHS Trust*  *Clifford Bridge Road*  *CV2 2DX*  *l.lacey.1@warwick.ac.uk* |
| **Co-investigators** | ***Siobhan Quenby***  *Professor of Obstetrics*  *UHCW NHS Trust*  *Clifford Bridge Road*  *CV2 2DX*  *S.quenby@warwick.ac.uk*  ***James Covington***  *Professor of Engineering*  *Department of Engineering*  *University of Warwick*  *CV4 7AL*  *j.a.covington@warwick.ac.uk*  ***Ian Henderson***  *NIHR Academic Clinical Fellow*  *UHCW NHS Trust*  *Clifford Bridge Road*  *CV2 2DX*  *Ian.henderson@uhcw.nhs.uk* |
| **Sponsor** | *UHCW NHS Trust*  *Clifford Bridge Road*  *CV2 2DX*  *Contact details* |
| **Funder** | *Warwick Wellcome Translational Partnership* |
| **Trial Co-ordinator / Co-ordination Centre** | *Lauren Lacey*  *C/O CSRL*  *UHCW NHS Trust*  *Clifford Bridge Road*  *CV2 2DX* |
| **Trial Steering Committee / Trial Management Group** |  |

**STUDY SUMMARY**

|  |  |  |
| --- | --- | --- |
| **Full study title** | Point of care diagnosis of urinary tract infections (UTIs) in pregnancy by volatile organic compound (VOC) analysis | |
| **Short study title** | UTI diagnosis in pregnancy by VOC analysis | |
| **Study aim** | We aim to verify that VOCs measured from urine samples of pregnant women without UTIs will be different to those who are infected. | |
| **Study design** | Single centre cohort | |
| **Study participants** | Pregnant patients | |
| **Study arms** | N/a | |
| **Sample size** | Training set plus test set for screening (test set n= 360) and test set for diagnosis (n=4700) | |
| **Planned study period** | 36 months (or until study complete) | |
| **Planned recruitment start date** | 1.11.2020 | |
| **Planned recruitment end date** | 30.9.2023 | |
| **Planned study end date** | 31.12.2023 | |
|  | **Objectives** | **Outcome Measures** |
| **Primary** | To optimise VOC detection technology for UTIs and then be able to screen for and diagnose culture positive UTIs in pregnancy including asymptomatic bacteriuria, symptomatic cystitis and pyelonephritis. | Identify optimal VOC detection technology and classifier to accurately diagnose UTIs from urine samples of pregnant women and to translate this into a clinical bedside test. |
| **Secondary** | To undertake VOC detection using this technology to distinguish between different pathogens that cause UTIs. This will allow targeted antibiotic treatment as the pathogen will be known. | To identify VOC patterns associated with specific bacteria causing urinary tract infections to allow prescription of specific antibiotic treatment regimes |

Key Words: **VOCs, volatile organic compounds, diagnosis, urinary tract infections, UTIs, pregnancy**

**STUDY FLOW CHART**

*Figure 1: Flow of participants through the study*

Contents

[SIGNATURE PAGE 2](#_Toc502674539)

[KEY TRIAL CONTACTS 3](#_Toc502674540)

[STUDY SUMMARY 5](#_Toc502674541)

[STUDY FLOW CHART 7](#_Toc502674542)

[SCHEDULE OF EVENTS 8](#_Toc502674543)

[LIST OF ABBREVIATIONS 10](#_Toc502674544)

[1. INTRODUCTION 11](#_Toc502674545)

[2. RATIONALE 16](#_Toc502674546)

[3. OBJECTIVES AND OUTCOME MEASURES/ENDPOINTS 16](#_Toc502674547)

[4. STUDY DESIGN 19](#_Toc502674548)

[5. STUDY SETTING 19](#_Toc502674549)

[6. ELIGIBILITY CRITERIA 19](#_Toc502674550)

[7. TRIAL PROCEDURES 20](#_Toc502674551)

[8. STATISTICS AND DATA ANALYSIS 26](#_Toc502674552)

[9. DATA MANAGEMENT 29](#_Toc502674553)

[10. TRIAL OVERSIGHT 30](#_Toc502674554)

[11. MONITORING, AUDIT & INSPECTION 32](#_Toc502674555)

[12. ETHICAL AND REGULATORY CONSIDERATIONS 33](#_Toc502674556)

[13. DISSEMINATION POLICY 34](#_Toc502674557)

[14. REFERENCES 35](#_Toc502674558)

# LIST OF ABBREVIATIONS

|  |  |
| --- | --- |
| AUC | Area under curve |
| CI | Chief investigator |
| CRF | Case report form |
| GCP | Good Clinical Practice |
| ISF | Investigator site file |
| M C & S | Microscopy culture and sensitivity |
| MSU | Midstream urine |
| NHS R&D | National Health Service Research & Development |
| NPV | Negative predictive value |
| PIS | Participant information sheet |
| PPV | Positive predictive value |
| REC | Research ethics committee |
| ROC | Receiver operator characteristics |
| SOP | Standard operating procedure |
| TMF | Trial master file |
| UHCW | University Hospitals Coventry & Warwickshire |
| UTI | Urinary tract infection |
| VOC | Volatile organic compound |

**STUDY PROTOCOL**

Point of care diagnosis of urinary tract infections (UTIs) in pregnancy by volatile organic compound (VOC) analysis

# INTRODUCTION

UTIs in pregnancy affect 5-10% of women, complicating 60000 pregnancies annually in the UK (1). This incidence is due to the physiological/anatomical changes that occur during pregnancy. UTIs are caused by a variety of pathogens, most commonly *Escherichia coli*. Many women are initially asymptomatic which poses a challenge for prompt diagnosis and treatment. Screening and accurately diagnosing women is important as untreated, asymptomatic bacteriuria progresses within days to pyelonephritis in 40-50% of women which can rapidly lead to life-threatening maternal sepsis and preterm delivery, the major cause of neonatal death/disability (2).

The accurate diagnosis/treatment of UTIs requires urinary culture and sensitivity testing in the microbiology laboratory at present. This takes at least 24-hours to obtain results, leading to delays in definitive treatment. A point of care test is therefore vital to reduce the risk of complications. The rapid dipstick chemical test is commonly used in clinical practice, but this has a low sensitivity (3). Many women are treated empirically and a urine sample not sent for culture (behaviour contributing to increasing antibiotic resistance). There is therefore a need for an accurate point of care test for the diagnosis of UTIs in pregnancy to help to reduce serious maternal/neonatal morbidity.

Volatile organic compounds (VOCs) are organic chemicals which have a high vapour pressure at room temperature as a result of their low boiling point. VOCs are produced as a result of physiological/metabolic body processes, of pathogens, commensals, and the host response to microbial infections. Engineering and bioinformatic advances in the University of Warwick have enabled the detection and analysis of very low quantities of thousands of VOCs, which have increasingly been utilised in medical diagnostics (4-10).

We will develop a test which can detect UTIs (including the pathogen) with high sensitivity and specificity, rapidly in pregnancy, building upon our previous work and the previously studies in the non-pregnant population (11-13) to outperform the urine dipstick test. We will then transfer this technology to the hospital for clinical validation.

* 1. Background

Urinary tract infection (UTI) is common in pregnancy (1). Pathogenic colonisation of the urinary tract is increased during pregnancy due to both physiological and anatomical changes that occur (2). UTIs in pregnancy can be both symptomatic and asymptomatic which can make diagnosis a challenge, nevertheless accurate diagnosis is imperative as it can lead to complications which can affect both the mother and fetus, including life threatening sepsis and preterm birth (2).

UTIs have three clinical presentations in pregnancy, asymptomatic bacteriuria, symptomatic cystitis and pyelonephritis. Asymptomatic bacteriuria is defined a significant colonisation by bacteria without clinical symptoms. It affects 2-10% of pregnancies (14, 15). Symptomatic cystitis is defined as dysuria, urinary frequency, haematuria, nocturia or suprapubic pain but the absence of pyrexia or systemic symptoms, its incidence is difficult to establish as many women are treated empirically and a urine sample not sent for culture. Pyelonephritis is the presence of significant bacteriuria with systemic symptoms, it is estimated to complicate 0.5-2% of pregnancies (16-18) and recurs in almost one quarter of affected patients (17). Asymptomatic bacteriuria progresses to cystitis and pyelonephritis in up to 30% and 50% of patients respectively if untreated (2).

Normal urine is sterile but when voided it can become contaminated by the distal urethra, a further diagnostic challenge is therefore distinguishing contamination from infection. To reduce the risk of contamination, a midstream urine sample should be collected. To help to distinguish between contamination and significant bacteriuria suggesting a UTI, quantitative urine culture in the microbiology laboratory is currently necessary. This is required as the performance of rapid screening tests for UTI in pregnancy is poor (15, 19). The most widely used rapid test in clinical practice outside of pregnancy is the dipstick chemical test. This test is based upon strips that have reagent pads for semiquantitative assessment of nitrites (a product of common urinary pathogens), leukocyte esterase (a byproduct of leucoctyes), protein, and blood (as a sign of inflammation). A meta-analysis of urine dipstick testing in pregnancy demonstrated the low sensitivity of nitrites and leucocyte esterase for UTI detection (0.46 and 0.68 respectively) (3). More recent studies have also highlighed that urine dipstick testing has a poor negative predictive value which makes it unsuitable to exclude the presence of urine infection, especially in pregnancy (20). A further diagnostic tool is the use of flow cytometry analysers (FCA). This test is based upon detection and quantification of both leukocytes and bacteria and is currently being evaluated. This analysis would reduce the number of samples cultured with a sharp decrease in workload, time, and costs (21). Additionally, negative results could be informed considerably earlier, which would reduce unnecessary empirical antibiotic prescriptions. Some units in the UK utilise this method as a form of screening but this is not recommended for pregnancy.

Despite urine culture being the “Gold Standard” the true bacterial concentration which is diagnostic for infection is debated. The urine culture accounts for logarithmic bacterial proliferation rates. A concentration of >105 bacteria/ml is highly suggestive of infection with <1% chance of this level being secondary to contamination. At 104-105 bacteria/ml, there is a higher risk of contamination, microbiologists diagnose a UTI if a single strain of a uropathogen is isolated. At 103-104 bacteria/ml, there is a 50% chance of contamination and therefore microbiology laboratories request a repeat urine sample for culture, if the same organism is isolated on the second sample, this is more indicative of a significant bacteriuria (2), this again adds a further delay to the time of diagnosis.

The current UK guidance advocates that as part of routine antenatal care all women have a midstream urine (MSU) sample is sent to the laboratory for culture at the booking appointment for their pregnancy (22). The rational for this is due to the association of UTIs with an increased risk of preterm birth (16, 23, 24). Urine culture is the “Gold standard” for diagnosis of UTI but creates a significant workload and is time-consuming with a high percentage yielding no growth. However, as discussed, the other diagnostic tools discussed have a low dianostic accuracy especially in pregnancy.

There is a need to develop an accurate point of care test for UTI in pregnancy which can be taken at the bedside and produce results within a few minutes, to reduce the burden of UTIs in pregnancy and its’ complications. Additionally, it would be beneficial due to the large scale of women who need to be tested if this cost of this test could be minimised.

The detection of specific patterns of volatile organic compounds (VOCs) in urine, breath, sweat and faeces is a novel tool that has been developing over recents years for the detection of various diseases (5, 6, 9, 10, 25-31). Gas phase analytical instruments including gas chromatography and mass spectrometry (GC-MS), selective ion flow mass spectrometry (SIFT) and the electronic nose (enose) have been used to detect VOCs. The pattern or “finger print” of VOCs reflect changes in the pathogens causing the disease (fermentone). There are several examples demonstrating that VOC analysis has been able to recognise various microorganisims including, *Clostridium difficile* in stool samples (17), Group B Streptococcus from vaginal swabs (8) and bacterial respiratory tract infection from breath (19). Utilising instruments that detect VOC mean that it is possible to obtain results within minutes of samples being taken.

* 1. Proposed study

The objective of this study is to determine the ability of this VOC detection technology to detect culture positive maternal UTI including asymptomatic bacteriuria, symptomatic cystitis and pyelonephritis.

We will examine gas phase VOCs emanating from MSU samples taken from women in pregnancy (index test) and compare this to the results of the dipstick chemical test and urine culture method (reference standard). The ultimate aim is to establish a point of care test for diagnosis that can be undertaken as part of a pregnancy pathway at the bedside, as the rapid dipstick test is used at present but with improved diagnostic accuracy.

As discussed, asymptomatic bacteriuria progresses to cystitis and pyelonephritis in up to 30% and 40-50% of patients respectively if untreated (1, 2). This new test will identify this and prevent disease progression by acute targeted administration of appropriate antibiotics to those who need it, reducing complications for both the mother and the child. From a global health presective, appropriate administration will reduce antibiotic resistance and this test will reduce the cost to screen women for these infections with a novel, rapid and increasingly available tool.

* 1. Study population

Pregnant women presenting to antenatal clinic, ultrasound department, labour ward triage or the early pregnancy assessment unit who would routinely be asked for a MSU specimen are identified by the clinical team.

* 1. Intervention

Pregnant women will be asked to produce a midstream urine sample, specific instructions about how to do this will be given to reduce the risks of contamination of the sample. This will be as follows:

To collect a clean urine sample:

* You must produce the sample at the hospital
* Wash your hands
* Part your labia
* Start to urinate, but don’t collect the first part of urine that comes out
* Collect a sample of urine "mid-stream" in a sterile container
* Wash your hands thoroughly

The reference standard will be sent immediately to the microbiology laboratory for processing in a boric acid container which helps to maintain the microbiological quality of the specimen, by preventing cell degradation and overgrowth of organisms. A rapid chemical dipstick test will also be performed on the sample

* 1. Clinical data

Several previous studies have demonstrated the disease detecting potential of this technology for UTIs. From a total of 680 urine samples from patients attending hospitals with suspected UTIs, they report sensitivity and specificity of various VOC detection technologies ranging from 83-100% and 88-100% respectively (11-13), a much higher sensitivity and equivalent specificity to the currently used rapid dipstick test. Our own teams study from UTI subjects recruited from UHCW (with a range of other conditions) have provided similar diagnostic performance (sensitivity/specificitiy >90%, unpublished). Forty midstream urine samples were transferred to the University of Warwick from UHCW NHS Trust. These samples were healthy volunteers and subjects with urinary tract infections (20 of each group). Samples contained multiple samples from the same patient in 2 ml vials. Samples were tested with an Owlstone Lonestar instrument (serial no. 135), Atlas sampling system and Split flow box. The instrument was located in a Category 2 bio-hazard laboratory. The unit was supplied with compressed air which was filtered using a water and VOC trap (details of filters available on request).

Samples testing procedure:

* Samples were placed in -20 oC freezer on arrival to Warwick University;
* On day of test, samples were placed in a laboratory fridge (3 oC) for 12 hours before testing;
* 5 ml of urine were aliquoted into a new 20 ml vial;
* These were placed in the Atlas sampling system (set to 40 oC) for 10 minutes to warm up;
* Flow over the sample = 200 ml/min; make-up air 1800 ml/min; total flow 2 L/min;
* Instrument was set to scan from 0-100% d.f. in 51 steps and from -6 to +6V c.v. in 512 steps;
* Each sample was tested sequentially four times;
* Instrument was left to clean for 25 mins in between samples.

Figure 2 demonstrates an output from a culture positive urine sample

*Figure 2: Example outputs from a culture positive urine sample is shown below (positive ions only)*

A close up of a logo

Description automatically generated:

Statistical analysis was done in-line with current methods used at the University of Warwick. These in brief are:

* 2D wavelet applied to each sample (second “run” used throughout);
* Feature selection by Wilcoxon test;
* 10 fold-cross validation;
  + Classification model training (90% of data) with three classifiers;
  + Test model prediction (10% of data)
  + ROC curve
* Performance metrics (AUC, sensitivity, specificity)

The diagnostic accuracy of the common machine learning models applied are illustrated in Table 1. The ROC curve for the sparse logistic regression model is illustrated in Figure 3.

*Table 1: Analysis Results*

|  |  |  |  |
| --- | --- | --- | --- |
| **Metric/Classifier** | **Sparse Logistic Regression (CI 95%)** | **Random Forrest (CI 95%)** | **Support Vector machine (CI 95%)** |
| AUC | 0.98 (0.93 - 1) | 0.95 (0.86 - 1) | 0.93 (0.82 - 1) |
| Sensitivity | 0.9 (0.55 - 1) | 0.8 (0.44 - 0.97) | 0.8 (0.44 - 0.97) |
| Specificity | 0.9 (0.55 - 1) | 0.8 (0.44 - 0.97) | 0.8 (0.44 - 0.97) |

*Figure 3: ROC Curve (Sparse Logistic Regression) for the diagnosis of culture positive urinary tract infections in the non-pregnant population (unpublished data)*



# RATIONALE

* 1. Aims and hypothesis

We aim to verify that VOCs measured from urine samples of pregnant women without UTIs will be different to those who are infected and translate this into a clincial test which can be used in a ward based setting. This test will more accurately diagnose urinary tract infections in pregannt women. This will allow rapid and appropriate antibiotic administration to those who need it only. This will also reduce inappropriate antibiotic administration.

* 1. Justification

The current rapid clinical urine dipstick test has a poor sensitivity and therefore women with urinary tract infections are being missed and are therefore at risk of the sequaelae of UTIs which can impact both the health of the mother and the fetus. The current “Gold standard” urine culture test is time-consuming, delaying diagnosis and expensive. VOC technology has the potential to rapidly and accurately diagnose UTIs in pregnant women at the bedside. This work is needed to validate the previous work into VOCs for the diagnosis of urinary tract infections specific to the pregnant population and then translate this into a clinical test.

* 1. Assessment and management of risk

Women are asked for a urine sample at every contact with healthcare professionals in pregnancy. This is normal practice and therefore there is minimal risk associated with this.

# OBJECTIVES AND OUTCOME MEASURES/ENDPOINTS

* 1. Primary objective

Our primary objectives is:

* To optimise VOC detection technology for UTIs to specifically identify culture positive UTIs in pregnancy including asymptomatic bacteriuria, symptomatic cystitis and pyelonephritis in pregnant women and to translate this into a clinically useful bedside test.

P – Pregnant women attending the hospital for any clincial review

I – VOC analysis of a midstream urine sample

C – Microscopy culture and sensitivity testing of midstream urine sample (and chemical dipstick

test results)

O – Diagnostic accuracy of VOC analysis of midstream urine sample when compared to the

“Gold Standard” of microscopy, culture and sensitivity testing

* 1. Secondary objectives

To undertake VOC detection using this technology to distinguish between different pathogens that cause UTIs. This will allow targeted antibiotic treatment as the pathogen will be known.

* 1. Primary endpoint/outcome

This project will consist of two phases.

Phase 1: Lab based testing

Collect urine samples from pregnant women until thirty confirmed UTI cases have been attained. At the same time, we will collect negative urine cases as controls, from suspected women. All samples will be cultured and used as the gold standard.

All samples will be frozen at point of collection in liquid nitrogen and then stored at -80°C (in line with our existing protocol). From previous experiments (4, 5, 26), we know samples will be store for up to nine months in these conditions. Samples will be tested in a single batch on multiple platforms, with sample aliquoted for multi-instrument testing. Specifically:

1. GC-TOF-MS: For chemical indentification
2. GC-IMS: For chemical indentification and potential ward use case
3. Electronic Nose: As a potential technology for ward use.

For I and II, chemical identification will be undertaken by NIST, followed by chemical confirmation. The School of Engineering have developed dedicated statistical pipelines for processing data from these platforms. These provide traditional outputs (including AUC, sensitivity, specificity, PPV and NPV), but also biomarkers/potential biomarkers. These will be compared with our previous UTI findings in non-pregnant individuals and the literature. Any differences will be investigated and protected as needed. This phase will also create classification/diagnostic models that will be applied in Phase 2.

Phase 2: Ward based testing

Depending on diagnostic performance, one of the VOC analysers used within Phase 1 (instruments ii and iii) will be installed at UHCW. These instruments do not require additonal services beyond power and are simple to use. Training will be provided to the ward staff and patients recruited as before. In this case, samples will be tested as soon as possible (or briefly stored in a fridge until analysis). The VOC detection equipment will be supported by Profesor James Covington or one of his team from his laboratory from the department of engineering. The equipment will be mantained and calibrated as per the manufacturers instructions. If we have any concerns or problems with the equipment we will contact Professor Covington directly. As before, culture will be undertaken as the gold standard and a chemical dipstick test performed.

* 1. Secondary endpoints/outcomes

Further analysis of the VOC patterns will be performed to identify specific patterns associated with specific organisms.

# STUDY DESIGN

Prospective cohort study

# STUDY SETTING

* Single centre
* Recruitment from antenatal clinic, obstetric ultrasound scan department, labour ward triage and the early pregnancy assessment unit

# ELIGIBILITY CRITERIA

* All adult pregnant women presenting to the hospital for clinical care will be eligible (aged =>18 years)
  1. Inclusion criteria

Pregnant women, either confirmed by urinary pregnancy test in first trimester or with other clinical signs of pregnancy

* 1. Exclusion criteria

Women who are not able to give informed consent

# TRIAL PROCEDURES

* 1. Recruitment

Patients will be recruited by the research team after being identified by their clinical team when they attend for clinical care during their pregnancy

* + 1. Patient identification

Patients will be identfied by the clinical team as being eligible to take part in the study and then will be approched by any member of the research team who has been trained to do so. This will include all team members who have completed their GCP training and are on the delegation log for the study. Women will be asked to consent on the day that they are approached about the study. This is clinically important as many women are treated based upon their urine chemical disptick tests and symptoms as is discussed in the background to the project. These women therefore are commenced on antibiotics and so it is not possible to wait for 24 hours to take patient consent and obtain a urine sample the following day.

* + 1. Screening

If patients are recruited from the EPAU and do not have an ultrasound scan to confirm pregnancy status, a urinary pregnancy test will be performed by the clinical team before they are approached about the study. All other participants will be attedning pregancy services provided as part of routine low risk and high risk antenatal care.

* + 1. Payment

Not applicable

* 1. Consent

Written informed consent will be obtained from patients after discussion with the study team and after they have received a patient information sheet about the study. The patient will be made aware that if they do not wish to particpate in the study their clinical care will not be affected in any way. The participant will remain free to withdraw at any time from the study without giving reasons and without prejudicing her further treatment and will be provided with a contact point where she can obtain further information about the study. Data and samples collected up to the point of withdrawal will only be used after withdrawal if the participant consents for this. Any intention to utilise such data will be outlined in the consent literature. Where a participant is required to re-consent or new information is required to be provided to a participant it will be the responsibility of the PI to ensure this is done in a timely manner. If a potential participant does not speak English, a transalator will be used to obtain written consent.

Participants must be capable of giving consent for themselves. A capable person will:

* understand the purpose and nature of the research
* understand what the research involves, its benefits (or lack of benefits), risks and burdens
* understand the alternatives to taking part
* be able to retain the information long enough to make an effective decision
* be able to make a free choice
* be capable of making this particular decision at the time it needs to be made (though their capacity may fluctuate, and they may be capable of making some decisions but not others depending on their complexity)
  1. Blinding

During the lab based testing phase, the study team will be aware of the results of the standard test as this will be needed to develop the index test. After this phase, the index test result will be rapidly available prior to the standard test. The results of the standard test will be chased by the research team and following any positive results patients will be contacted and advised to attend for clinical review and antibiotics. In view of the poor diagnostic accuracy of the chemical dipstick test, the results of this will not be used in patient care for study patients.

* 1. Baseline data

Baseline data will be collected on a case report form (CRF). These data will include:

* Date of sample
* Age at booking
* BMI at booking
* Gestational age
* Urine chemical analysis dipstick test results
* Ethnicity
* Smoking status
* Parity
* Symptoms of UTi at time of sample
* Medications taking at time of sample
* Allergies
* Time of day sample obtained
* M C & S results

This will be stored by the Tommy’s Biobank team

* 1. End of study definition

The study will be closed after the sample size has been achieved. We are aiming to complete by the end of 2023

# STATISTICS AND DATA ANALYSIS

* 1. Sample size calculation

Phase I

Samples will be collected until 30 culture positive samples have been obtained

Phase II

The sensitivity of the chemical analysis dipstick test for UTI in pregnancy is mediocre at approximately 0.55. To determine the suitability of VOC analysis of urine as a replacement screening test, we used a minimum acceptable lower confidence limit for sensitivity of 0.60 (3), an expected sensitivity of 0.90 based on our pilot data, and a UTI prevalence of 5%. We used the one-sided method outlined in Flahault et al. 2005 (32),which identified a sample size requirement of 18 cases, and therefore 342 controls. We have funding for the project up until this stage from the Warwick Wellcome Translational Partnership. Once the data has been analysed, we will use this to apply for larger scale funding for the second part of Phase II below for the use of VOC detection technology as a diagnostic tool for urinary tract infections.

To consider the suitability of VOC analysis of urine as a diagnostic test, we used a sensitivity of 0.90 and a minimum acceptable lower confidence limit of 0.80, and identified a sample size requirement of 235 cases and 4,465 controls, where the probability that the 95% lower confidence interval is above the minimum lower confidence limit is 95%.

* 1. Planned recruitment rate

UHCW has approximately 6000 deliveries per year. Women attend multiple times throughout their pregnancy. We aim to recruit and collected 200 samples per month once the project enters Phase II.

* 1. Statistical analysis plan
     1. Summary of baseline data and flow of patients

Baseline characteristics of the cohort will be reported. Parametric data will be reported with mean and standard deviation and non-parametric data reported as median and interquartile range

* + 1. Primary outcome analysis

The data will be analysed using the statistical pipeline successfully used in (9, 33, 34). In summary, the VOC output data will be extracted using the L.A.V. software (v2.2.1, G.A.S, Germany), which converts the data from its native file format to a text file. This will be followed by a pre-processing step to reduce the dimentionality of the data, making the statistical analysis less computationaly intensive. A typical output file (of a single sample) contains typically 11 million data points. Though the number of data points is high, the information content is sparse, with the all of the values containing non-background information being located around the centre of the dataset. Thus, we are able to crop the central section of the data and then apply a threshold to make the background values all be zero. These values are selected by visual inspection of the data using the LAV software and results in around a 500 fold reduction in the number of non-zero data points. Once completed, the data will analysed using a 10-fold cross validation approach. In each fold, the data was split into a 90% training set and a 10% test set. Features with discriminary power will identified from the training set using a rank-sum test and 50 features with the lowest p-value were taken forward for classification. Here, five different classifiers will be used, specifically sparse logistic regression, random forest, Gaussian process classifier, support vector machine and neural network (this set is commonly used within our pipeline). Once the training models have been created in Phase 1, they will be applied to the same features in the test set (Phase 2). This process will be repeated ten times until all the data has a test result. This process will provide test probabilities for each sample and from this, statistical values, including sensitivity and specificity will be calculated.

* + 1. Secondary outcome analysis

Samples positive for specific micro-organisms will have their VOC analysis interrogated to identify if a specific fingerprint or specific VOCs are associated with each pathogen. Once identified in future work this pattern will be reviewed in further samples to investigate diagnostic accuracy.

# DATA MANAGEMENT

* 1. Data collection tools and source document identification

Baseline data will be collected from the patient on the day of recruitment. This will use a standardised case report form (CRF). This original document will be stored by the Tommy’s Biobank team in a locked cupboard in the Biomedical Research Unit at UHCW. This document will contain patient identifiable data. These data will be transferred onto an electronic database and stored by the Tommy’s Biobank Team only. This electronic database will not have patient identifiable data and each sample will have a study number. This number will also be recorded on the CRF. Patient contact details will be kept by the research team to ensure the results of the reference standard (midstream urine sent to the microbiology laboratory for microscopy, culture and sensitivity testing) are chased and acted upon if needed. With patient permission we will collect each patients telephone number and call them with the results of the reference standard if they need further discussion or treatment.

A copy of the consent form will be kept in a file in the Biomedical Research Unit and each patient will be given a copy of the consent form.

* 1. Archiving

Following the resolution of queries and confirmation of study close-out by the Chief Investigator, all essential documentation will transferred to a third party archiving service, which provides suitable fire and water-resistant facilities.  The data will be stored in accordance with the sponsor recommendations. The standard data retention period for University of Warwick is 10 years.

# TRIAL OVERSIGHT

* 1. Role and responsibilities of the Sponsor

University of Warwick has agreed to act as sponsor for this study and will undertake the responsibilities of sponsor as defined by the UK Policy Framework for Health and Social Care Research and ICH Good Clinical Practice. An authorised representative of the Sponsor has approved the final version of this protocol with respect to the study design, conduct, data analysis and interpretation and plans for publication and dissemination of results. As sponsor, University of Warwick provides indemnity for this study and, as such, will be responsible for claims for any non-negligent harm suffered by anyone as a result of participating in this trial. The indemnity is renewed on an annual basis and will continue for the duration of this trial.

# MONITORING, AUDIT & INSPECTION

The study may be monitored by the Univerosty of Warwick as representatives of the Sponsor or by the Research & Development Department at UHCW, to ensure that the study is being conducted as per protocol, adhering to Research Governance and GCP. The approach to, and extent of, monitoring may be specified in a trial monitoring plan determined by the risk assessment undertaken prior to the start of the study.

# ETHICAL AND REGULATORY CONSIDERATIONS

* 1. Ethical approval and research governance

The study will be conducted in compliance the principles of the ICH GCP guidelines and in accordance with all applicable regulatory guidance, including, but not limited to, the UK policy framework for health and social care research. Ethical approval for this study will be sought from the Research Ethics Committee combined with Health Research Authority (HRA) approval. No study activities will commence until favourable ethical opinion and HRA approval has been obtained. Progress reports and a final report at the conclusion of the trial will be submitted to the approving REC within the timelines defined by the committee. Confirmation of capacity and capability will be obtained from the R&D department prior to commencement of the study at all participating sites.”

* 1. Peer review

This study was peer reviewed by the Warwick-Wellcome Translational Partnership team prior to funding being offered for the project

* 1. Data protection and patient confidentiality

The study will comply with the current Data Protection regulations and regular checks and monitoring will be undertaken to ensure compliance. Participants will be assigned a unique identifier upon enrolment into the study to allow pseudonymisation of patient-identifiable data. Access to patient identifiable data will be restricted to members of the study co-ordination team who require it for the performance of their role. Electronic data will be stored on password protected encrypted drives and hard copies of study documents will be stored in locked filing cabinets in secure entry-card protected sites.

# DISSEMINATION POLICY

The study findings will be reported to the Warwick-Wellcome Translational Partnership and we aim to publish the results in high impact journals. We will apply for patent with Warwick Ventures.

# REFERENCES

1. Mikhail MS, Anyaegbunam A. Lower urinary tract dysfunction in pregnancy: a review. Obstet Gynecol Surv 1995;50:675-83.

2. McCormick T, Ashe RG, Kearney PM. Urinary tract infection in pregnancy. The Obstetrician & Gynaecologist 2008;10:156-62.

3. Devillé WLJM, Yzermans JC, van Duijn NP, Bezemer PD, van der Windt DAWM, Bouter LM. The urine dipstick test useful to rule out infections. A meta-analysis of the accuracy. BMC Urology 2004;4:4.

4. Arasaradnam RP, Ouaret N, Thomas MG, Quraishi N, Heatherington E, Nwokolo CU *et al.* A novel tool for noninvasive diagnosis and tracking of patients with inflammatory bowel disease. Inflamm Bowel Dis 2013;19:999-1003.

5. Arasaradnam RP, McFarlane MJ, Ryan-Fisher C, Westenbrink E, Hodges P, Thomas MG *et al.* Detection of colorectal cancer (CRC) by urinary volatile organic compound analysis. PLoS One 2014;9:e108750.

6. Arasaradnam RP, McFarlane M, Ling K, Wurie S, O'Connell N, Nwokolo CU *et al.* Breathomics--exhaled volatile organic compound analysis to detect hepatic encephalopathy: a pilot study. J Breath Res 2016;10:016012.

7. Esfahani S, Wicaksono A, Mozdiak E, Arasaradnam PR, Covington AJ. Non-Invasive Diagnosis of Diabetes by Volatile Organic Compounds in Urine Using FAIMS and Fox4000 Electronic Nose. Biosensors 2018;8.

8. Lacey L, Daulton E, Wicaksono A, Covington JA, Quenby S. Detection of Group B Streptococcus in pregnancy by vaginal volatile organic compound analysis: a prospective exploratory study. Transl Res 2020;216:23-9.

9. Lewis JM, Savage RS, Beeching NJ, Beadsworth MBJ, Feasey N, Covington JA. Identifying volatile metabolite signatures for the diagnosis of bacterial respiratory tract infection using electronic nose technology: A pilot study. PLoS One 2017;12:e0188879.

10. van Keulen BJ, Benninga MA, Berkhout DJC, Schouten A, Klaver E, Bessem JR *et al.* Late-onset Sepsis in Preterm Infants Can Be Detected Preclinically by Fecal Volatile Organic Compound Analysis: A Prospective, Multicenter Cohort Study. Clinical Infectious Diseases 2018;68:70-7.

11. Aathithan S, Plant JC, Chaudry AN, French GL. Diagnosis of bacteriuria by detection of volatile organic compounds in urine using an automated headspace analyzer with multiple conducting polymer sensors. Journal of clinical microbiology 2001;39:2590-3.

12. Kodogiannis VS, Lygouras JN, Tarczynski A, Chowdrey HS. Artificial Odor Discrimination System Using Electronic Nose and Neural Networks for the Identification of Urinary Tract Infection. IEEE Transactions on Information Technology in Biomedicine 2008;12:707-13.

13. Roine A, Saviauk T, Kumpulainen P, Karjalainen M, Tuokko A, Aittoniemi J *et al.* Rapid and accurate detection of urinary pathogens by mobile IMS-based electronic nose: a proof-of-principle study. PLoS One 2014;9:e114279.

14. Foley ME, Farquharson R, Stronge JM. Is screening for bacteriuria in pregnancy worthwhile? Br Med J (Clin Res Ed) 1987;295:270.

15. Smaill FM, Vazquez JC. Antibiotics for asymptomatic bacteriuria in pregnancy. Cochrane Database of Systematic Reviews 2015.

16. Wing DA, Fassett MJ, Getahun D. Acute pyelonephritis in pregnancy: an 18-year retrospective analysis. Am J Obstet Gynecol 2014;210:219.e1-6.

17. Gilstrap LC, Ramin SM. Urinary tract infections during pregnancy. Obstet Gynecol Clin North Am 2001;28:581-91.

18. Hill JB, Sheffield JS, McIntire DD, Wendel GD. Acute pyelonephritis in pregnancy. Obstet Gynecol 2005;105:18-23.

19. Mignini L, Carroli G, Abalos E, Widmer M, Amigot S, Nardin JM *et al.* Accuracy of diagnostic tests to detect asymptomatic bacteriuria during pregnancy. Obstet Gynecol 2009;113:346-52.

20. Eigbefoh JO, Isabu P, Okpere E, Abebe J. The diagnostic accuracy of the rapid dipstick test to predict asymptomatic urinary tract infection of pregnancy. Journal of Obstetrics and Gynaecology 2008;28:490-5.

21. Grosso S, Bruschetta G, De Rosa R, Avolio M, Camporese A. Improving the efficiency and efficacy of pre-analytical and analytical work-flow of urine cultures with urinary flow cytometry. New Microbiol 2008;31:501-5.

22. National Institute for Health and Care Excellence. Antenatal care for uncomplicated pregnancies CG62. In: National Institute for Health and Care Excellence,, 2008.

23. Agger WA, Siddiqui D, Lovrich SD, Callister SM, Borgert AJ, Merkitch KW *et al.* Epidemiologic factors and urogenital infections associated with preterm birth in a midwestern U.S. population. Obstet Gynecol 2014;124:969-77.

24. KASS EH. Bacteriuria and pyelonephritis of pregnancy. Arch Intern Med 1960;105:194-8.

25. Arasaradnam RP, Westenbrink E, McFarlane MJ, Harbord R, Chambers S, O'Connell N *et al.* Differentiating coeliac disease from irritable bowel syndrome by urinary volatile organic compound analysis--a pilot study. PLoS One 2014;9:e107312.

26. Arasaradnam RP, McFarlane M, Daulton E, Westenbrink E, O'Connell N, Wurie S *et al.* Non-invasive distinction of non-alcoholic fatty liver disease using urinary volatile organic compound analysis: early results. J Gastrointestin Liver Dis 2015;24:197-201.

27. Arasaradnam RP, McFarlane M, Daulton E, Skinner J, O'Connell N, Wurie S *et al.* Non-invasive exhaled volatile organic biomarker analysis to detect inflammatory bowel disease (IBD). Dig Liver Dis 2016;48:148-53.

28. Arasaradnam RP, Wicaksono A, O'Brien H, Kocher HM, Covington JA, Crnogorac-Jurcevic T. Noninvasive Diagnosis of Pancreatic Cancer Through Detection of Volatile Organic Compounds in Urine. Gastroenterology 2018;154:485-7.e1.

29. Bomers MK, Menke FP, Savage RS, Vandenbroucke-Grauls CM, van Agtmael MA, Covington JA *et al.* Rapid, accurate, and on-site detection of C. difficile in stool samples. Am J Gastroenterol 2015;110:588-94.

30. Covington JA, Westenbrink EW, Ouaret N, Harbord R, Bailey C, O'Connell N *et al.* Application of a novel tool for diagnosing bile acid diarrhoea. Sensors (Basel) 2013;13:11899-912.

31. Plat VD, van Gaal N, Covington JA, Neal M, de Meij TGJ, van der Peet DL *et al.* Non-Invasive Detection of Anastomotic Leakage Following Esophageal and Pancreatic Surgery by Urinary Analysis. Digestive Surgery 2018.

32. Flahault A, Cadilhac M, Thomas G. Sample size calculation should be performed for design accuracy in diagnostic test studies. Journal of Clinical Epidemiology 2005;58:859-62.

33. Rouvroye DM, Wicaksono A, Bosch S, Savelkoul E, Covington AJ, Beaumont H *et al.* Faecal Scent as a Novel Non-Invasive Biomarker to Discriminate between Coeliac Disease and Refractory Coeliac Disease: A Proof of Principle Study. Biosensors 2019;9.

34. Tiele A, Wicaksono A, Kansara J, Arasaradnam PR, Covington AJ. Breath Analysis Using eNose and Ion Mobility Technology to Diagnose Inflammatory Bowel Disease—A Pilot Study. Biosensors 2019;9.