BMJ Open

Characterizing clinical, epidemiological, and etiological aspects in relation to global disease burden assessment of leptospirosis in Sri Lanka: Study Protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-027850
Article Type:	Protocol
Date Submitted by the Author:	20-Nov-2018
Complete List of Authors:	Agampodi, Suneth; Rajarata University, Department of Community Medicine Warnasekara, Janith; Rajarata University of Sri Lanka Faculty of Medicine and Allied Sciences, Department of Community Medicine Jayasundara, Dinesha; Rajarata University of Sri Lanka Faculty of Medicine and Allied Sciences, Department of Microbiology Senawirathna, Indika; Rajarata University of Sri Lanka Faculty of Medicine and Allied Sciences, Department of Biochemistry Gamage, Chandika; Faculty of Medicine. University of Peradeniya, Department of Microbiology Kularatne, Senanayake; Peradeniya University,SriLanka, Medicine Siribaddana, Sisira; Rajarata University of Sri Lanka Faculty of Medicine and Allied Sciences, Department of Medicine Maththias, Michael; Yale University School of Medicine Vinetz, Joseph; Yale University School of Medicine
Keywords:	Leptospirosis, Sri Lanka, Leptospira, Isolates, Whole genome sequencing, Clinical disease

Characterizing clinical, epidemiological, and etiological aspects in relation to global disease burden assessment of leptospirosis in Sri Lanka: Study Protocol

Agampodi SB[1*], Warnasekara YPJN[1], Jayasundara JMDD[2], Senavirathna RMISK[3], Gamage CD[4], Kularatna SAM[5], Siribaddana SH[6], Maththias M[7], Vinetz JV[7]

[1]Department of Community Medicine, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka

[2]Department of Microbiology, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka

[3]Department of Biochemistry, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka

[4]Department of Microbiology, Faculty of Medicine, University of Peradeniya

[5] Department of Medicine, Faculty of Medicine, University of Peradeniya

[6] Department of Medicine, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka

[7] Yale School of Medicine, Section of Infectious Diseases, Department of Internal Medicine New Haven, CT 06520

ASB: sunethagampodi@yahoo.com,

WYPJN: jwarnasekara@yahoo.com

JJMDD: dineshajayasundara@yahoo.com

SRMISK: irathnabahu@gmail.com

GCD: chandika@pdn.ac.lk

KSAM: samkul@sltnet.lk

SSH: sisira.siribaddana@gmail.com

MM: mikematthias34@icloud.com

VJV: joseph.vinetz@yale.edu

*Corresponding author

Abstract

Introduction

Sri Lanka has one of the highest incidence of leptospirosis worldwide. We hypothesized that different geographical locations and patient context will have a distinct molecular epidemiology of leptospirosis, based on; microgeographic characteristics related to regionally-specific *Leptospira* predominance. Our objective is to characterize clinical, epidemiological, and molecular aspects of leptospirosis in Sri Lanka to understand the disease progression, risk factors and also to have isolates of *Leptospira* from well characterized population to enhance the knowledge on molecular epidemiology of leptospirosis in Sri Lanka.

Methods and Analysis

We designed a multi-center study in Sri Lanka to recruit undifferentiated febrile patients and to follow them up during hospital stay. Patients will be recruited by visiting out patient department and all medical wards. Two main sites (Anuradhapura and Peradeniya) and several additional sites (Awissawella, Ratnapura and Polonnaruwa) will be selected for this study. Blood and urine will be collected from patients on the day of admission to the ward or presentation to the outpatient department. Bedside inoculation of 2 and 4 drops of venous blood will be done into EMJH semisolid media with added antibiotics. Disease confirmation will be done using a regionally optimized MAT (microscopic agglutination test), culture, and qPCR-evidence of the presence of *Leptospira* in blood. For newly isolated *Leptospira*, whole genome sequencing will be done for all isolates. Multi Locus Sequence Typing (MLST) will be used for the genotyping of the new isolates. Sri Lankan isolates will be identified using three published MLST schemes for *Leptospira.*

Ethics and Dissemination

Ethical clearance for the study was obtained from Ethics Review Committees of Rajarata University of Sri Lanka and University of Peradeniya. All data generated through this project will be available for researchers and policy makers on a reasonable request.

Strengths and limitations of this study

- This study will be the largest multi-centered prospective study on leptospirosis in Sri Lanka
- This study will provide robust data on clinical disease progression and predictors of clinical disease, which will be helpful in clinical managment of leptospirosis.
- Culture isolation of *Leptospira* is not carried out for 30 years in Sri Lanka at the time of writing of this protocol and this study will fill that research gap.
- While having one of the highest incidence, Sri Lankan Leptospira isolates has not been undergone whole genome sequencing. This study will enrich the global databases on *Leptospira* genome.

Article Summary

- A prospective recruitment of undifferentiated febrile patients over 3 years
- Diagnosis of leptospirosis using MAT, qPCR and rapid diagnostic tests
- Building disease progression models to predict adverse outcomes
- Culture isolation and typing of *Leptospira spp* to understand the infecting serovar diversity.
- Whole genome sequencing to expand the knowledge on *Leptospira* genome

Introduction

Leptospirosis is a globally widespread, neglected and emerging zoonotic disease[1] with estimated 1.03 million cases and 58,900 annual deaths attribute to the disease.[2] An average of 2.9 million daily adjusted life years (DALYs) are estimated to be lost every year due to leptospirosis.[3] Emerging leptospirosis mostly affects vulnerable communities living in resource poor settings. Often, the disease is either not suspected or is under or misdiagnosed due to the need for laboratory resources to confirm leptospirosis; typically, such resources are neither accessible nor affordable. From the clinical perspective, better understanding of clinical disease associated with diverse *Leptospira* species is required and better diagnostics are needed to prevent severe complications and death. From the public health perspective, the lack of reliable and efficient diagnostics tests makes assessing the burden of disease--whether regionally or globally--difficult.

Sri Lanka has emerged as a country with high incidence of leptospirosis since 2007[4] and in 2008, the total number of clinically suspected cases reported to the surveillance system was 7,406 with 204 deaths.[5] In 2009, 4980 cases and 145 deaths were reported[6] and the outbreak persisted until 2013 with more than 4000 reported cases each year.[7] The probable case incidence during this time period was more than 20 per 100,000 population, indicating that Sri Lanka has one of the highest reported incidence of leptospirosis worldwide.[8] Our extensive work on 2008 outbreak suggested that clinical diagnosis of leptospirosis is highly inaccurate and we could be missing as much as 40% of the leptospirosis patients in hospitals.[9] However, all Sri Lankan studies including our previous study have the same drawback, as observed in most of the published reports in global literature: hospital admission bias, where large number of patients with leptospirosis might have treated as out-patients, which severely underestimate the actual disease burden. Precise disease burden estimates needs seroprevalence studies coupled with population based incidence studies.

Our studies on 2011 outbreak of leptospirosis in Anuradhapura showed that the molecular epidemiology, disease transmission and the clinical manifestations of this outbreak were different from what we observed in Central Province (wet zone) in 2008.[10] One of the main concern about these marked variations in *Leptospira* strain and the clinical disease is whether this is due to microgeographical variations, environmental conditions or reservoir mammal infection. It has been shown for other disease like malaria[11, 12] and schistosomiasis[13-15] that microgeography may have a major influence on disease epidemiology. Geochemistry is well described as a major contributory factor in human health.[16] Despite the extensive literature available on leptospirosis, studies on the microgeographic variation of *Leptospira* is scarce.

The systematic review done on Sri Lankan isolates showed that more than 40 strains from 20 serovars of *Leptospira* have been reported from Sri Lanka.[17] Our previous studies on 2008 and 2011 outbreaks of leptospirosis in Sri Lanka expanded the knowledge of circulating *Leptospira* and was consistent with a diversity of *Leptospira* infection in Sri Lanka. We showed that 2008 outbreak was predominantly with *L. interrogans*, compared to post flood outbreak in 2011, in which *L.krischneri* was the predominant species. However, the molecular epidemiology of *Leptospira* species in endemic setting is still unknown. In addition to the STNPCR method, we, for the first time showed that previously published MLST schemes could be used with modifications for direct patients samples to study the genetic diversity of infecting leptospiral species in resource-poor settings, where culture and isolation facilities are minimal.[18]

One of the most important conclusions we draw from our published work is that qPCR is promising as a clinical diagnostic tool in the diagnosis of leptospirosis during the acute phase with a wider window of positivity than previously noted.[19] The editorial by Katz accompanying this paper[20] endorsed the qPCR approach but also emphasized the need for a prospective evaluation of this method. However, most countries with high leptospirosis burden still find it prohibitively expensive in terms of cost and logistics to have molecular based methods for rapid diagnosis.

Disease burden assessment, public health interventions, and clinical management of leptospirosis are challenges due to lack of diagnostic facilities. Any population-based attempt to estimate the disease burden is often challenged by unavailability of diagnostic tests or poor performance of available diagnostic tests. Development of diagnostic tests; global and country specific, requires knowledge on circulating serovars. Further, well-characterized samples are required with species and subspecies level identification of infecting *Leptospira* backed by culture isolation. In Sri Lanka, culture isolation of *Leptospira* has not been reported since 1970s. Even though we showed the microgeographical changes of leptospirosis in 2014[10], no prospective studies have been conducted in leptospirosis literature to primarily identify the disease diversity with specific research design. We hypothesized that different study site and patient context—inpatient vs. outpatient—will have a distinct molecular epidemiology of leptospirosis, based on; Microgeographic characteristics related to regionally-specific *Leptospira* predominance, clinical characteristics related to pathogenesis/virulence potential of specific *Leptospira* species, strains, serovars and the acquisition of infection related to occupational and activities of daily living, reflecting different environmental exposure contexts. Based on these background analyses and hypothesis, the objectives of the present study included;

1. To determine the prevalence of leptospirosis among undifferentiated febrile patients (inpatient and outpatient) in contrasting geographical settings in Sri Lanka,

2. To describe the full clinical spectrum of leptospirosis in patients infected with *Leptospira*,

3. To determine the predictors of leptospirosis progression using a nested case control approach among febrile patients, both inpatient and ambulatory,

4. To determine the *Leptospira* species and type diversity in different microgeographical settings in Sri Lanka,

5.To characterize Sri Lankan *Leptospira* using MLST and whole genome sequencing methods and

6. To have a well characterized collection of serum, whole blood and urine sample for future development and validation of leptospirosis diagnostics

Methods and Analysis

Study sites

This study will have four primary subject enrolment/data collection centers in two main contrasting sites; Kandy and Anuradhapura. In addition, small component of the study will be carried out in Sabaragamuwa, Western, North Central and Southern provinces for species diversity (Figure 1). Kandy and Anuradhapura was selected based on our previous observation of *Leptospira* diversity in these two settings, supported by historically distant identification of diverse *Leptospira* in Sri Lanka (1960s)[21], without any more recent work; hence this project will yield modern information at the molecular and genomically-determined taxonomical levels in addition to conventional methods of identifying *Leptospira*). The first focus is on Anuradhapura, located in North Central province of Sri Lanka, in the dry zone of the country. The annual rainfall is 1,200-1600mm and the mean annual temperature is ~30°C. In Anuradhapura area, paddy farming is carried out by traditional, full-time farmers and provides the main mode of income. Cattle, water buffalo, certain rodent species—well known sources of *Leptospira* transmission—are in and round the paddy fields. Paddy fields are large and well maintained. There are no wetlands or marshy lands in these areas, except the paddy fields during working seasons. Paddy field work depends on irrigation systems, so between farming seasons, paddy fields become completely dry. There are probably other risk factors for acquiring leptospirosis from activities of daily living in economically poor conditions, for example living in proximity to dogs and rodents in houses and neighborhoods and shoe-ware is often no more than sandals.

The soil structure, water quality/hardness, and ecological systems in Anuradhapura contrast with those of the wet zone of Sri Lanka. Kandy (another historic site in Sri Lanka/Ceylon,

the cultural capital of the country since ancient times) is situated in the wet zone with more than 3000mm rainfall and located in the middle, hilly parts of the country. Paddy farming in this area is based on rainwater and the paddy fields remains wet throughout the years. These are small fields adjacent to wildlife and the soil and ecological structure is different from Anuradhapura. Nonetheless, economic conditions of residents in this region are poor and rural, so that rodent- and dog-transmitted *Leptospira* in the context of activities of daily living is likely common but completely unexplored.

The purpose of including additional sites is primarily to identify differences in leptospirosis outbreaks especially during floods and unsuspected outbreaks. Only selected components of the study will be done in those areas if there are outbreaks.

Study population

Study population will include all undifferentiated fever cases either presented to the outpatient department or admitted to medical wards/ICUs in TH Anuradhapura and TH Peradeniya (main sites) and selected hospitals in Western, Sabaragamuwa and Southern provinces.

Inclusion criteria:

- age ≥12yrs
- temperature≥ 38^0C
- self-reported fever ≥ 2 days

Exclusion criteria:

- Influenza-like illness with runny nose and upper respiratory tract symptoms
- physician-diagnosed probable or definite meningitis or lower respiratory tract infection (e.g. consolidated lobar pneumonia)

- traumatic or post-operative fever per physician discretion,
- fever due to nosocomial infections
- any patient with confirmed diagnosis as a cause for the fever

Patient recruitment

MBBS graduates awaiting their internship appointment and established registrars will be located in all four centers, and directly supervised on a daily basis. They will screen all febrile patients to identify and perform directed physician exams in outpatient department as well as inpatients admitted to medical wards. Once the diagnostic facilities are established, we will let all clinicians in the selected districts know that these facilities are available. If they are interested we will provide diagnostic facilities including cultures, specifically to look at the species diversity.

Procedure:

All possible undifferentiated fever cases will be interviewed and examined by the onsite pre-intern medical officers and suspected leptospirosis cases (fulfilling inclusion and exclusion criteria) will be referred to the study project. In the OPD, once the OPD medical officers complete the history taking and examination, data collector will collect the data. For inward patients, pre-intern medical officer will screen all fever patients on admission to ward, to select eligible patients. In Anuradhapura, four medical units are there and the study will be carried out in all four medical wards. In Peradeniya, a single medical ward is available. Data collectors will visit all wards and ICUs daily for this purpose. Once the routine ward procedures are completed, data collection will be done. All eligible patients will be given an explanatory statement of the study and written informed consent will be obtained before enrolment. Data collection will be done using a fully structured, interviewer administered questionnaire. An additional clinical data check list will be used to assess the clinical features. Daily follow up will be done for all hospitalized patients to observe the clinical progress of the disease. This will be done by using a clinical

checklist. The questionnaires and the checklists will be adopted from previously published time tested protocol.[22]

All patients will be given an appointment for follow up after two weeks. We will maximize the follow up through reminders, telephone calls, and the transport fee will be paid. Two ml of blood and a urine sample will be obtained from the follow up patients.

Sample size

Assuming that around 70% of patients are infected with same infective *Leptospira* strain in specific setting and to detect at least 20% difference of in OPD and hospitalized patients, at least 62 leptospirosis patients with molecular identification are needed from each site. This calculation was done to have a power of 80% with an alpha value of 0.05. Based on our previous studies, interpretable sequencing data are available for around 50% of confirmed cases of leptospirosis[23] and leptospirosis is accounted for 15% of acute undifferentiated fever[24]. Based on these assumptions, the total number of febrile patients that is required from one sample collection site for this study is 827. This number of patients is conservatively estimated and feasible to obtain within the project period.

Sample collection, procession, and storage procedure

Blood and urine will be collected from patients on the day of admission to the ward or presentation to the outpatient department. Total of 7ml of blood will be collected and bed side inoculation of 2 and 4 drops (100µl-500µl) will be done into two tubes containing 9ml of EMJH semisolid media with added antibiotics (5-Fluorouracil and Neomycin). 2.5ml of venous blood will be collected to an EDTA tube and the rest into plain tubes for serum separation. For serum separation, blood in the plain tubes will be centrifuged at 1300rcf for 10 minutes. All blood and serum samples will be stored at -80^{0}C for further analysis. Blood collected into EMJH media for culture will be incubated at 30^{0}C incubator.

Ten to fifteen milliliters of clean catch mid stream urine will be collected into a sterile wide mouth container and processed within 2 hours of collection. In a situation where urine can't be processed during the specified time period, pH of urine will be measured with strips and adjusted to around 7.4 using 1N NaOH. Initially urine will be centrifuged at 3000rpm for five minutes to sediment white cells, epithelial cells and other crystals. Then the supernatant will be separated and centrifuged at 15000 rpm for 10 minutes. The resultant sediment will be taken for PCR studies, labeled and stored at -20^{0}C. [25-27] A follow up blood sample (2ml) will be collected into plain tubes for serology and 10-15ml of urine will be collected for PCR studies after 3 weeks from the onset of fever.

Disease confirmation

Disease confirmation will be done using a regionally optimized MAT (microscopic agglutination test), culture, and qPCR-evidence of the presence of *Leptospira* in serum.

MAT

For this purpose, first we will use the U.S. Centers for Disease Control (CDC) MAT panel with already-available five local strains obtained from Royal Tropical Institute, Amsterdam (KIT). Once Sri Lankan isolates are available, validation of best panel suitable for Sri Lanka will be determined as part of this study. MAT will be done in two main steps. At the first step, serum samples will be screened at a titre of 1/50 by using CDC recommended panel of 24 strains which includes five local strains. At the next step positive sera of screening step will be subjected to the run out test which uses a serial dilution from 1/50 to 1/3200. MAT evidence for seropositivity will be defined as one or more of the following: seroconversion from negative to positive; 4-fold rise in titer between acute-phase and convalescent-phase (follow-up) samples; or a single titer of ≥800.

Culture isolation of Leptospira

EMJH semisolid media will be prepared by adding 2.3g of EMJH base, 1.5g of bacteriological agar and 100mg of sodium pyruvate into 785ml of distilled water and adjusting the pH to 7.4. The media will be autoclaved and once it cooled to around 50^0C, 100ml of Leptospira enrichment media, and 100ml of foetal bovine serum will be added.*[28-31]* To suppress the growth of possible contaminants, 5 Fluorouracil and Neomycin will be added in 100µg/ml and 25µg/ml concentrations respectively. Inoculated media with blood will be inspected under dark field microscopy for the presence of mortile leptospires initially after 3 weeks and then on monthly basis. When a positive growth is detected, sub cultures will be made into liquid and semisolid media and an aliquot will be fixed with 5% DMSO (Dimethyl Sulfoxide) and stored in -80^0C. Isolates will be sub cultured in liquid media 2 weekly and on semi solid media 3 monthly. Serotyping of newly isolated *Leptospira* strains will be done in the Pasteur Institute, France

Quantitative PCR

Diagnosis

For the diagnosis, DNA will be extracted from whole blood, serum and urine using QIAamp DNA Blood Mini Kit (Qiagen, USA) according to the manufacture's instruction. Extracted DNA will be quantified using Invitrogen Qubit 4 Flurometer. Two previously published protocols will be used for detection of pathogenic *Leptospira* DNA. A quantitative PCR based on SYBR Green chemistry targeting 16s ribosomal RNA and Lipl32 genes will be used for the diagnosis (Table 1). [32]

Table 1: Primers for detection of *Leptospira* DNA in clinical samples.

Primer/Probe Sequence	Annealing temperature	Complementary target species
Forward: 5' AAG CAT TAC CGC TTG TGG TG3'	60°C	lipL32 gene. of pathogenic

Reverse: 5'GAA CTC CCA TTT CAG CGA TT3'		Leptospira spp
Forward: 5'TAA AGG CTC ACC AAG GCG AC3' Reverse: 5'TTA GCC GGT GCT TTA GGC AG3'	60 °C	16s gene of pathogenic Leptospira spp

Speciation

For the determination of *Leptospira* species in clinical samples, a previously published quantitative PCR protocol based on SYBR Green chemistry (Table 2) will be used. For early identification species in cultures (before sequencing) we will used the same protocol. This will low identification of L*. interrogans, L. borgpetersenii, L. kirschneri and L. noguchii.*

Table 2. Primer pairs for detection of species.

Primer/Probe Sequence	**Annealing temperature**	**Complementary target species**
Forward: 5'CTT GAG CCT GCG CGT TAY 3' Reverse: 5'CCG ATA ATT CCA GCG AAG ATC3'	63°C	secY gene of *L. interrogans*
Forward: 5'GAT TCG GGT TAC AAT TAG ACC3' Reverse: 5'TTG ATC TAA CCG GAC CAT AGT3'	65°C	ompL1 gene of *L. borgpetersenii*
Forward: 5' CTG GCT TAA TCA ATG CTT CTG 3' Reverse: 5'CTC TTT CGG TGA TCT GTT CC3'	60°C	secY gene of *L. kirschneri*
Forward: 5'TCA GGG TGT AAG AAA GGT TC 3' Forward: 5'CAA AAT TAA AGA AGA AGC AAA GA3'	63°C	secY gene of *L. noguchii*

Molecular studies

For newly isolated *Leptospira*, whole genome sequencing will be done using two technologies; MinION Nanopore sequencing in the Leptospirosis research lab, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka and Pacbio sequencing at the Institute for Genomic Medicine, Universiy of California, San Diego.

BAS and BAX files contain all information on DNA sequences will be converted to fasta file. Prior to the assembly the quality (coverage) of the raw data will be assessed. Using PacBioEDA software the quality summary of the data will be obtained. CanuPacBiO assembly

program will be used for genome assembly. The raw data of Bas/Bax files will be converted to Bam files by using the bax2bam software. These bam files of reads will be aligned to our references using Pbalingn software. The software Quiver will be used for the polishing of data. Assembled data will be annotated by Prokka[33] and Roary[34] software and annotations will be inspected and curated using Artemis. The functional annotation will be done using eggNOG and eggNOG MAPPER online tools.[35] Genomes of the Sri Lankan isolates will be compared using MAUVE software. The organism will be roughly identified from the direct raw data using KRAKEN.[36]

Multi Locus Sequence Typing (MLST) will be used for the genotyping of the new isolates. Sri Lankan isolates will be identified using three published MLST schemes for *Leptospira*.[37-39] Genetic similarity between isolates will be determined using SPLITSTREE software.

Taxonogenomics, pan-genome and phylogenetic analyses

In order to establish how this strain was related to other *Leptospira* , we will use several genomic analysis. First, 16S rRNA sequences will be used to construct a phylogeny. High sequence identities (≥97%) between newly sequenced strain and other recognized *Leptospira* strains will be used preliminarily as suggestive of phylogenetic relationships. Accordingly, 16S rRNA gene sequences of *Leptospira* spp. deposited in GenBank database will be acquired. Multiple sequence alignment based on 16S rRNA will be performed using a ClustalX v1.81. These result will be then used to construct the maximum likelihood tree implementing MEGA v5.05. To evaluate clade support, we will construct a bootstrap analysis with 1,000 replicates. Relationship of each sequenced genome to previously described or novel leptospiral species will be determined by Overall Genetic Relatedness Indices (OGRIs). Average Nucleotide Identity (ANI) and the Average Amino Acid Identity (AAI). Both indices will be automatically calculated using two-way BLAST + blastnand blastp. For ANI calculation, default

parameters: sequence identity cut-off, 30%; alignment cut-off, 70%; and query length, 1,020 bp will be applied.

Data management and analysis

Patient interviews and patient health records will be used as data sources. On site data entry will be done in real time for all collected data. EpiInfo database will be prepared using cross checks, field validation and other inbuilt techniques available for data quality improvement. Investigators will frequently visit the hospitals to cross validate collected data. All databases will be kept under password-protected computers. Only the investigators and data entry assistants will have access to the database. Databases will not include the personal identification data and those information in the hardcopies will be kept in separately.

Proportions and 5% confidence intervals will be used to describe the confirmed cases, diversity of infecting *Leptospira* and sequelae. Chi-square test for trend will be used to test significant differences of *Leptospira* diversity at different microgeographical and clinical settings. Predictors of leptospirosis and severe disease will be analyzed using a logistic regression model. Model building will be done using a conceptual hierarchy of determinants, which we proposed for leptospirosis earlier. Online database for *Leptospira* MLST will be used for typing analysis.

Patient and Public Involvement

The study was developed using the patients' feedback we received on leptospirosis disease burden, it's impact on their lives, and especially the problems they are facing during the illness. During the previous studies, we specifically looked at patients as well as general public views on leptospirosis disease and prevention, which was used for designing this study. Results of the study will be disseminated in a continues basis for patients. We have already conducted several public awareness as well as health worker education workshops based on our findings.

Ethics and dissemination

Undifferentiated fever patients eligible for the study will be given an explanatory statement of the study and written informed consent will be obtained before enrolment. Rapid diagnostic kits will be used and the results will be available for treating physicians. qPCR diagnostic facilities will be provided for all patients whenever the urgent tests are requested by treating physicians. All test results will be informed to the patients and the wards. Ethical clearance for the study was obtained from the Ethic Review Committees of Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka (ERC/2015/18).

All genomic data will be deposited in Gene Bank. Characterized serum samples will be stored in -80C in the Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka. Data pertaining to this study will be available for researchers and policy makers on a reasonable request address to the corresponding author. After the completion of work, data will be deposited in a repository.

References

1. WER: **Leptospirosis worldwide, 1999**. *Wkly Epidemiol Rec* 1999, **74**(29):237-242.
2. Costa F, Hagan JE, Calcagno J, Kane M, Torgerson P, Martinez-Silveira MS, Stein C, Abela-Ridder B, Ko AI: **Global Morbidity and Mortality of Leptospirosis: A Systematic Review**. *PLoS Negl Trop Dis* 2015, **9**(9):e0003898.
3. Taylor AJ, Paris DH, Newton PN: **A Systematic Review of the Mortality from Untreated Leptospirosis**. *PLoS Negl Trop Dis* 2015, **9**(6):e0003866.
4. Agampodi S, Peacock SJ, Thevanesam V: **The potential emergence of leptospirosis in Sri Lanka**. *Lancet Infect Dis* 2009, **9**(9):524-526.
5. Epidemiological Unit: **Selected notifiable diseases reported by Medical Officers of Health**. *Weekly Epidemiological Report* 2009, **36**(1):4.
6. Epidemiological Unit: **Selected notifiable diseases reported by Medical Officers of Health**. *Weekly Epidemiological Report* 2010, **37**(1):4.
7. Warnasekara JN, Agampodi S: **Leptospirosis in Sri Lanka**. *Sri Lankan Journal of Infectious Diseases* 2017, **7**(2).
8. Pappas G, Papadimitriou P, Siozopoulou V, Christou L, Akritidis N: **The globalization of leptospirosis: worldwide incidence trends**. *Int J Infect Dis* 2008, **12**(4):351-357.
9. Agampodi SB, Peacock SJ, Thevanesam V, Nugegoda DB, Smythe L, Thaipadungpanit J, Craig SB, Burns MA, Dohnt M, Boonsilp S *et al*: **Leptospirosis outbreak in Sri Lanka in 2008: lessons for assessing the global burden of disease**. *Am J Trop Med Hyg* 2011, **85**(3):471-478.

10. Agampodi SB, Dahanayaka NJ, Bandaranayaka AK, Perera M, Priyankara S, Weerawansa P, Matthias MA, Vinetz JM: **Regional differences of leptospirosis in Sri Lanka: observations from a flood-associated outbreak in 2011**. *PLoS Negl Trop Dis* 2014, **8**(1):e2626.
11. Myers WP, Myers AP, Cox-Singh J, Lau HC, Mokuai B, Malley R: **Micro-geographic risk factors for malarial infection**. *Malaria journal* 2009, **8**:27.
12. Schellenberg JA, Newell JN, Snow RW, Mung'ala V, Marsh K, Smith PG, Hayes RJ: **An analysis of the geographical distribution of severe malaria in children in Kilifi District, Kenya**. *Int J Epidemiol* 1998, **27**(2):323-329.
13. Hassan AO, Amoo AOJ, Deji-Agboola AM, Adeleke MA, Muhibi MA, Hassan RO: **Microgeographical patterns of schistosomiasis and anthropometric indices of children in eko-ende south west, nigeria**. *Int J Biol Med Res* 2911, **2**(3):656-660.
14. Kloos H, Gazzinelli A, Van Zuyle P: **Microgeographical patterns of schistosomiasis and water contact behavior; examples from Africa and Brazil**. *Mem Inst Oswaldo Cruz* 1998, **93 Suppl 1**:37-50.
15. Baliraine FN, Afrane YA, Amenya DA, Bonizzoni M, Menge DM, Zhou G, Zhong D, Vardo-Zalik AM, Githeko AK, Yan G: **High prevalence of asymptomatic plasmodium falciparum infections in a highland area of western Kenya: a cohort study**. *J Infect Dis* 2009, **200**(1):66-74.
16. Dissanayake CB, Chandrajith R: **Medical geochemistry of tropical environments**. *Earth-Science Reviews* 1999, **47**(3–4):219-258.
17. Naotunna C, Agampodi SB, Agampodi TC: **Etiological agents causing leptospirosis in Sri Lanka: A review**. *Asian Pac J Trop Med* 2016, **9**(4):390-394.
18. Agampodi SB, Moreno AC, Vinetz JM, Matthias MA: **Utility and limitations of direct multi-locus sequence typing on qPCR-positive blood to determine infecting Leptospira strain**. *Am J Trop Med Hyg* 2013, **88**(1):184-185.
19. Agampodi SB, Miatthias MA, Moreno AC, Vinetz JM: **Utility of Quantitative PCR in Leptospirosis Diagnosis: Association of Level of Leptospiremia and Clinical Manifestations in Sri Lanka**. *Clinical Infectious Disease* 2012, **In Press**.
20. Katz AR: **Quantitative polymerase chain reaction: filling the gap for early leptospirosis diagnosis**. *Clin Infect Dis* 2012, **54**(9):1256-1258.
21. Nityananda K, Harvey T: **Leptospirosis in Ceylon-Epidemiological and Laboratory Studies**. *Ceylon J Med Sci* 1971, **20**(1):5-14.
22. Agampodi SB, Nugegoda DB, Thevanesam V: **Determinants of leptospirosis in Sri Lanka: Study Protocol**. *BMC Infect Dis* 2010, **10**(1):332.
23. Agampodi SB, Matthias MA, Moreno AC, Vinetz JM: **Utility of quantitative polymerase chain reaction in leptospirosis diagnosis: association of level of leptospiremia and clinical manifestations in Sri Lanka**. *Clin Infect Dis* 2012, **54**(9):1249-1255.
24. Reller ME, Bodinayake C, Nagahawatte A, Devasiri V, Kodikara-Arachichi W, Strouse JJ, Flom JE, Dumler JS, Woods CW: **Leptospirosis as frequent cause of acute febrile illness in southern Sri Lanka**. *Emerg Infect Dis* 2011, **17**(9):1678-1684.
25. Lucchesi PM, Arroyo GH, Etcheverria AI, Parma AE, Seijo AC: **Recommendations for the detection of Leptospira in urine by PCR**. *Rev Soc Bras Med Trop* 2004, **37**(2):131-134.
26. Bal AE, Gravekamp C, Hartskeerl RA, De Meza-Brewster J, Korver H, Terpstra WJ: **Detection of leptospires in urine by PCR for early diagnosis of leptospirosis**. *J Clin Microbiol* 1994, **32**(8):1894-1898.
27. Van Eys GJ, Gravekamp C, Gerritsen MJ, Quint W, Cornelissen MT, Schegget JT, Terpstra WJ: **Detection of leptospires in urine by polymerase chain reaction**. *J Clin Microbiol* 1989, **27**(10):2258-2262.

28. Wuthiekanun V, Amornchai P, Langla S, Oyuchua M, Day NP, Limmathurotsakul D: **Maintenance of leptospira species in leptospira Vanaporn Wuthiekanun agar**. *J Clin Microbiol* 2014, **52**(12):4350-4352.
29. Wuthiekanun V, Amornchai P, Langla S, White NJ, Day NP, Limmathurotsakul D, Peacock SJ: **Antimicrobial Disk Susceptibility Testing of Leptospira spp. Using Leptospira Vanaporn Wuthiekanun (LVW) Agar**. *Am J Trop Med Hyg* 2015, **93**(2):241-243.
30. Wuthiekanun V, White NJ, Amornchai P, Day NP, Limmathurotsakul D: **Quality controls for antimicrobial disk diffusion testing on Leptospira Vanaporn Wuthiekanun agar**. *Trans R Soc Trop Med Hyg* 2016, **110**(11):673-675.
31. Picardeau M, Bulach DM, Bouchier C, Zuerner RL, Zidane N, Wilson PJ, Creno S, Kuczek ES, Bommezzadri S, Davis JC *et al*: **Genome sequence of the saprophyte Leptospira biflexa provides insights into the evolution of Leptospira and the pathogenesis of leptospirosis**. *PLoS ONE* 2008, **3**(2):e1607.
32. Backstedt BT, Buyuktanir O, Lindow J, Wunder EA, Reis MG, Usmani-Brown S, Ledizet M, Ko A, Pal U: **Efficient Detection of Pathogenic Leptospires Using 16S Ribosomal RNA**. *PLOS ONE* 2015, **10**(6):e0128913-e0128913.
33. Seemann T: **Prokka: rapid prokaryotic genome annotation**. *Bioinformatics* 2014, **30**(14):2068-2069.
34. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J: **Roary: rapid large-scale prokaryote pan genome analysis**. *Bioinformatics* 2015, **31**(22):3691-3693.
35. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, Bork P: **Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper**. *Mol Biol Evol* 2017, **34**(8):2115-2122.
36. Wood DE, Salzberg SL: **Kraken: ultrafast metagenomic sequence classification using exact alignments**. *Genome Biol* 2014, **15**(3):R46.
37. Varni V, Ruybal P, Lauthier JJ, Tomasini N, Brihuega B, Koval A, Caimi K: **Reassessment of MLST schemes for Leptospira spp. typing worldwide**. *Infect Genet Evol* 2014, **22**:216-222.
38. Ahmed N, Devi SM, Valverde Mde L, Vijayachari P, Machang'u RS, Ellis WA, Hartskeerl RA: **Multilocus sequence typing method for identification and genotypic classification of pathogenic Leptospira species**. *Ann Clin Microbiol Antimicrob* 2006, **5**:28.
39. Boonsilp S, Thaipadungpanit J, Amornchai P, Wuthiekanun V, Bailey MS, Holden MT, Zhang C, Jiang X, Koizumi N, Taylor K *et al*: **A single multilocus sequence typing (MLST) scheme for seven pathogenic Leptospira species**. *PLoS Negl Trop Dis* 2013, **7**(1):e1954.

Author statement

Authors' contributions

JMV and SBA drafted the initial protocol. YPJNW, JMDDJ and RMISKS are responsible for the detailed design, field/ laboratory work analysis and interpretation of objectives 1-3, 4 and 5

respectively. CDG, SAMK and SHS helped in design, field, hospital and laboratory work. MM participated in scientific design and guided all laboratory components. All authors have read and approved the manuscript.

Funding

This project is the Sri Lankan component of a multi country study funded by the NIAID of the National Institutes of Health under award number U19AI115658. Funders had no role in the design, execution, analyses, data interpretation, or decision to submit results.

Conflicts of interests

The authors declare that they have no competing interests.

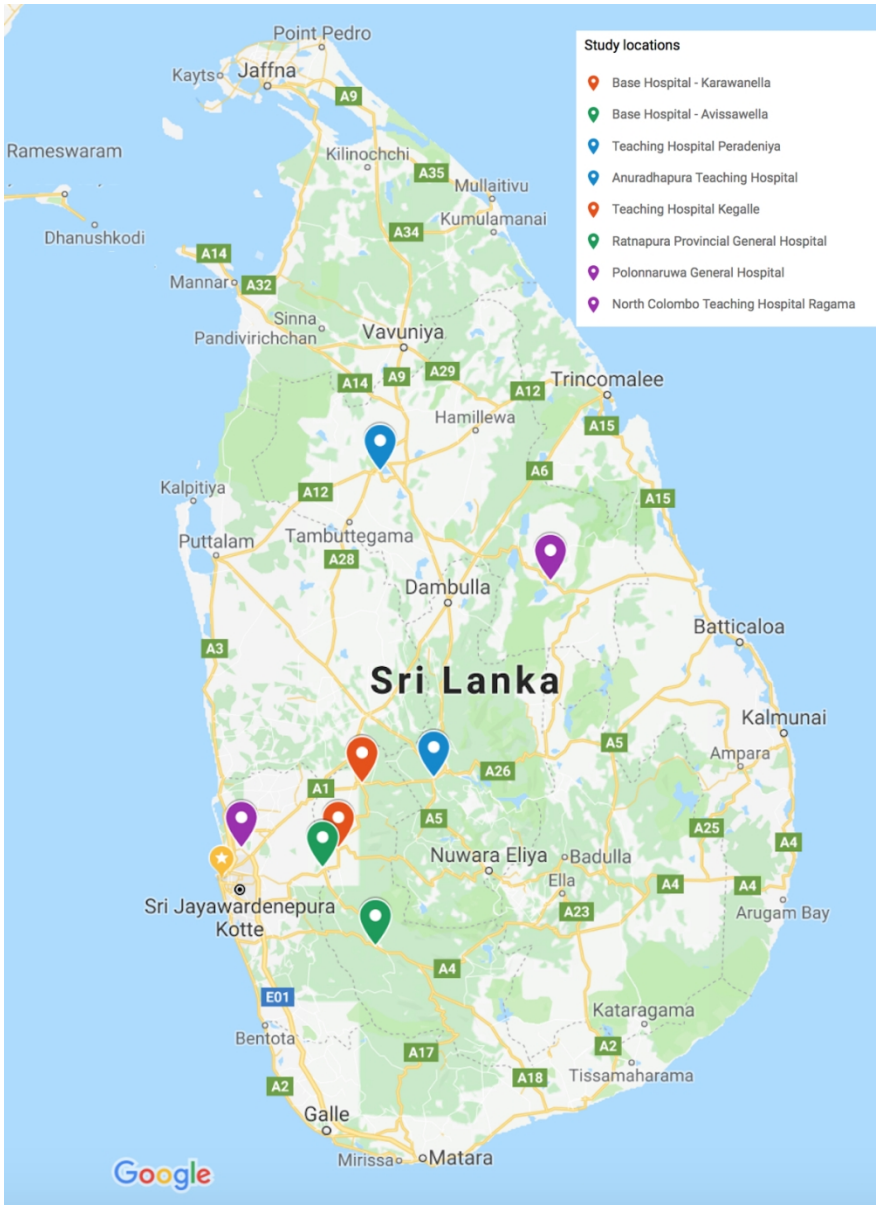
Word count

3595

Figure titles

Figure 1: Locations of study sites.

Figure was prepared by the first author using Google maps (https://www.google.com/maps).



Locations of study sites

BMJ Open

Study protocol: Characterizing the clinical, epidemiological, and etiological aspects of leptospirosis in Sri Lanka

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-027850.R1
Article Type:	Protocol
Date Submitted by the Author:	29-May-2019
Complete List of Authors:	Agampodi, Suneth; Rajarata University, Department of Community Medicine Warnasekara, Janith; Rajarata University of Sri Lanka Faculty of Medicine and Allied Sciences, Department of Community Medicine Jayasundara, Dinesha; Rajarata University of Sri Lanka Faculty of Medicine and Allied Sciences, Department of Microbiology Senawirathna, Indika; Rajarata University of Sri Lanka Faculty of Medicine and Allied Sciences, Biochemistry Gamage, Chandika; University of Peradeniya Faculty of Medicine, Microbiology Kularatne, Senanayake; University of Peradeniya Faculty of Medicine, Medicine Siribaddana, Sisira; Rajarata University of Sri Lanka Faculty of Medicine and Allied Sciences, Department of Medicine Maththias, Michael; Yale University School of Medicine Vinetz, Joseph; Yale University School of Medicine
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Epidemiology, Global health, Public health, Infectious diseases
Keywords:	Leptospirosis, Sri Lanka, Leptospira, Isolates, Whole genome sequencing, Clinical disease

Study protocol: Characterizing the clinical, epidemiological, and etiological aspects of leptospirosis in Sri Lanka

Agampodi SB[1*], Warnasekara YPJN[1], Jayasundara JMDD[2], Senawirathna RMISK[3], Gamage CD[4], Kularatne SAM[5], Siribaddana SH[6], Maththias M[7], Vinetz JV[7]

[1] Department of Community Medicine, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka

[2] Department of Microbiology, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka

[3] Department of Biochemistry, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka

[4] Department of Microbiology, Faculty of Medicine, University of Peradeniya

[5] Department of Medicine, Faculty of Medicine, University of Peradeniya

[6] Department of Medicine, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka

[7] Yale School of Medicine, Section of Infectious Diseases, Department of Internal Medicine New Haven, CT 06520

ASB: sunethagampodi@yahoo.com,
WYPJN: jwarnasekara@yahoo.com
JJMDD: dineshajayasundara@yahoo.com
SRMISK: irathnabahu@gmail.com
GCD: chandika@pdn.ac.lk
KSAM: samkul@sltnet.lk
SSH: sisira.siribaddana@gmail.com
MM: mikematthias34@icloud.com
VJV: joseph.vinetz@yale.edu

*Corresponding author.

ABSTRACT

Introduction: Sri Lanka has one of the highest incidences of leptospirosis worldwide. We hypothesized that different geographical locations and patient context will have a distinct molecular epidemiology of leptospirosis, based on microgeographic characteristics related to regiona-specific *Leptospira* predominance. Our objective is to characterize the clinical, epidemiological, and molecular aspects of leptospirosis in Sri Lanka to understand disease progression, risk factors, and obtain isolates of *Leptospira* from well-characterized populations to enhance the understanding of molecular epidemiology of leptospirosis in Sri Lanka. **Methods and Analysis:** We designed a multi-center prospective study in Sri Lanka to recruit undifferentiated febrile patients and conduct follow-ups during hospital stays. Patients will be recruited from outpatient departments and medical wards. This study will be conducted at two main sites (Anuradhapura and Peradeniya) and several additional sites (Awissawella, Ratnapura and Polonnaruwa). Blood and urine will be collected from patients on the day of admission to the ward or presentation to the outpatient department. Bedside inoculation of 2 and 4 drops of venous blood will be performed with EMJH semisolid media supplemented with antibiotics. Regionally optimized microscopic agglutination test, culture, and qPCR-evidence will be performed to confirm the presence of *Leptospira* in blood, which in turn will confirm the presence of disease. Whole genome sequencing will be carried out for all isolates recovered from patients. Multi locus sequence typing (MLST) will be used for the genotyping of new isolates. Sri Lankan isolates will be identified using three published MLST schemes for *Leptospira*. **Ethics and Dissemination:** Ethical clearance for the study was obtained from Ethics Review Committees of Rajarata University of Sri Lanka and University of Peradeniya. All genomic data generated through this project will be available at GenBank. For the ethics approval, anonymized data will be deposited at the Ethics Review Committee of Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka and is available for researchers and policy makers.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- This is a multi-center study involving wet and dry zones, low and highlands and high endemic and low endemic areas of the country covering different geographical and climate zones to provide a better understanding of leptospirosis in Sri Lanka.
- We will use a prospective study design focusing on culture isolation as a main aim with specific protocols for the particular objective to describe the strain diversity of *Leptospira spp*. in Sri Lanka.
- Clinical and epidemiological data will be collected as primary data to ensure high quality data using clinically trained health professionals.
- Since *Leptospira* are fastidious organisms, the culture yield will be low and linking the clinical disease with isolated *Leptospira* may be partly confounded by the growth potentials of different *Leptospira*.
- Since PCR and culture-based methods are typically valid only during the first 7-10 days of the illness and late presentations may lead to reduce sensitivity of these tests, MAT is best interpreted with paired samples, and a lack of convalescent samples will impair the interpretation of MAT results.

INTRODUCTION

Leptospirosis is a globally widespread, neglected, and emerging zoonotic disease[1] with estimated 1.03 million cases and 58,900 annual deaths attributed to the disease.[2] An average of 2.9 million daily adjusted life years (DALYs) are estimated to be lost every year due to leptospirosis.[3] Emerging leptospirosis mostly affects vulnerable communities living in resource-poor settings. Often, the disease is either not suspected or is under or misdiagnosed due to the need for laboratory resources to confirm leptospirosis; typically, such resources are neither accessible nor affordable. From the clinical perspective, a better understanding of the clinical disease associated with diverse *Leptospira* species is required and superior diagnostics are needed to prevent severe complications and death. From the public health perspective, the lack of reliable and efficient diagnostic tests makes assessing the burden of disease—whether regionally or globally—difficult.

Sri Lanka has emerged as a country with a high incidence of leptospirosis since 2007[4] and in 2008, the total number of clinically suspected cases reported to the surveillance system was 7,406 with 204 deaths.[5] In 2009, 4980 cases and 145 deaths were reported[6], and the outbreak persisted until 2013 with more than 4000 reported cases each year.[7] The probable case incidence during this time period was more than 20 per 100,000 population, indicating that Sri Lanka has one of the highest reported incidences of leptospirosis worldwide.[8] Our extensive work on the 2008 outbreak suggested that clinical diagnosis of leptospirosis is highly inaccurate, since as much as 40% of the leptospirosis patients are misdiagnosed in hospitals.[9] However, all Sri Lankan studies, including our previous study, have the same drawback, as observed in most of the published reports in global literature: hospital admission bias, where a large number of patients with leptospirosis may have been treated as outpatients, which leads to a severe underestimate of the actual disease burden. A precise disease burden estimate requires seroprevalence studies coupled with population-based incidence studies.

Our studies on the 2011 outbreak of leptospirosis in Anuradhapura showed that the molecular epidemiology, disease transmission, and clinical manifestations of this outbreak were different

from those observed in Central Province (wet zone) in 2008.[10] One of the main concerns regarding these marked variations in the *Leptospira* strain and the clinical disease was whether this was due to microgeographical variations, environmental conditions, or reservoir mammal infection. It has been shown for other diseases such as malaria[11, 12] and schistosomiasis[13-15], that microgeography may have a major influence on disease epidemiology. Geochemistry is a well-known and major contributory factor in human health.[16] Despite the extensive literature available on leptospirosis, studies on the microgeographic variation of *Leptospira* are scarce.

A systematic review performed on Sri Lankan isolates showed that more than 40 strains from 20 serovars of *Leptospira* have been reported from Sri Lanka.[17] Our previous studies on the 2008 and 2011 outbreaks of leptospirosis in Sri Lanka expanded the knowledge of circulating *Leptospira* and were consistent with a diversity of *Leptospira* infection in Sri Lanka. We showed that in the 2008 outbreak, *L. interrogans* was the predominant species, compared to the post-flood outbreak in 2011, in which *L. krischneri* was the predominant species. However, the molecular epidemiology of *Leptospira* species in endemic settings remains unknown in Sri Lanka. In addition to the STNPCR method, for the first time, we showed that previously published multi locus sequence typing (MLST) schemes could be used with modifications for direct patients samples to study the genetic diversity of infecting *Leptospira* species in resource-poor settings, where the facilities needed to perform cultures and isolations are minimal.[18]

One of the most important conclusions we draw from our published work is that qPCR is promising as a clinical diagnostic tool in the diagnosis of leptospirosis during the acute phase, with a wider window of positivity than previously noted.[19] The editorial by Katz[20] endorsed the qPCR approach but also emphasized the need for a prospective evaluation of this method. However, most countries with high leptospirosis burdens still find it prohibitively expensive to have molecular based methods for rapid diagnosis, in terms of both the costs and logistics.

Disease burden assessment, public health interventions, and clinical management of leptospirosis are challenges due to the lack of diagnostic facilities. Any population-based attempt

to estimate the disease burden is often limited by the unavailability of diagnostic tests or the poor performance of the available diagnostic tests. The development of diagnostic tests that are both global and country-specific requires knowledge on circulating serovars. Furthermore, well-characterized samples are required with species and subspecies level identification of infecting *Leptospira* backed by culture isolations. In Sri Lanka, a culture isolation of *Leptospira* has not been reported since the 1970s. Even though we showed the microgeographical changes of leptospirosis in 2014,[10] no prospective studies have been conducted in a Sri Lankan cohort to primarily identify the disease diversity with specific research design. We hypothesized that a different study site and patient context—inpatient vs. outpatient—will have a distinct molecular epidemiology of leptospirosis, based on microgeographic characteristics related to region-specific *Leptospira* predominance, clinical characteristics related to pathogenesis/virulence potential of specific *Leptospira* species, strains, serovars, and the acquisition of infection related to occupational and activities of daily living, reflecting different environmental exposure contexts. Based on these background analyses and hypotheses, the objectives of the present study were as follows:

1. To determine the prevalence of leptospirosis among undifferentiated febrile patients (inpatient and outpatient) in contrasting geographical settings in Sri Lanka.

2. To describe the full clinical spectrum of leptospirosis in patients infected with *Leptospira*.

3. To determine the predictors of leptospirosis progression using a nested case control approach among febrile patients, both inpatient and ambulatory.

4. To determine the *Leptospira* species and type diversity in different geographical settings in Sri Lanka.

5. To characterize Sri Lankan *Leptospira* using MLST and whole genome sequencing methods.

6. To have a well-characterized collection of serum, whole blood, and urine samples for the future development and validation of leptospirosis diagnostics.

METHODS AND ANALYSIS

Study sites

This study will be performed in four primary subject enrolment/data collection centers in two main contrasting sites: Kandy and Anuradhapura. In addition, a small component of the study will be carried out in Sabaragamuwa, Western, North Central, and Southern provinces for species diversity (Figure 1). Kandy and Anuradhapura was selected based on our previous observations of *Leptospira* diversity in these two settings, supported by historically distant identification of diverse *Leptospira* in Sri Lanka (1960s)[21]. Since no recent studies have been conducted, this project will yield up-to-date information at the molecular and genomically-determined taxonomical levels in addition to conventional methods of identifying *Leptospira*. The initial focus is on Anuradhapura, located in North Central province of Sri Lanka, in the dry zone of the country. The annual rainfall is 1,200-1,600 mm, with a mean annual temperature of ~30°C. In Anuradhapura, paddy farming is carried out by traditional, full-time farmers and constitutes the main mode of income. Cattle, water buffalo, certain rodent species—well-known sources of *Leptospira* transmission—are found in and around the paddy fields. Paddy fields are large and well maintained. There are no wetlands or marshy lands in these areas, except the paddy fields during working seasons. Paddy field work depends on irrigation systems, such that, between farming seasons, the paddy fields become completely dry. Other risk factors for acquiring leptospirosis include daily activities involved in living in economically poor conditions, for example living in proximity to dogs and rodents in houses and neighborhoods and the wearing of shoes that are often no more than sandals.

The soil structure, water quality/hardness, and ecological systems in Anuradhapura contrast with those of the wet zone of Sri Lanka. Kandy (another historic site in Sri Lanka/Ceylon, the cultural capital of the country since ancient times) is situated in the wet zone, with over 3,000

mm of rainfall, and is located in the middle, hilly parts of the country. Paddy farming in this area is based on rainwater and the paddy fields remain wet throughout the year. These are small fields adjacent to wildlife and the soil and ecological structure is different from that in Anuradhapura. Nonetheless, the economic conditions of residents in this region are poor and rural, such that rodent- and dog-transmitted *Leptospira* in the context of activities of daily living is likely common but completely unexplored.

The purpose of including additional sites is primarily to identify differences in leptospirosis outbreaks, especially during floods and unsuspected outbreaks. Only selected components of the study will be performed in those areas if there are outbreaks.

Study population

The study population will include all undifferentiated fever cases either presented to the outpatient department or admitted to medical wards/ICUs in TH Anuradhapura and TH Peradeniya (main sites) and selected hospitals in Western, Sabaragamuwa, and Southern provinces.

Inclusion criteria

- Age ≥ 12 years.
- Temperature ≥ 38°C.
- Self-reported fever ≥ 2 days.

Exclusion criteria

- Fever with running nose without any other symptom or sign.
- Any patient with confirmed diagnosis/foci of infection as a cause for the fever.
- Traumatic or post-operative fever per physician discretion.
- Fever due to nosocomial infections.

Patient recruitment

Graduates with a Bachelor of Medicine and Bachelor of Surgery (MBBS) awaiting their internship appointment and established registrars will be assigned to all four centers and directly supervised on a daily basis. They will screen all febrile patients and perform directed physician exams in the outpatient department, as well as on inpatients admitted to the medical wards. Once the diagnostic facilities are established, we will inform all clinicians in the selected districts about the available facilities. If they express interest, diagnostic facilities, including cultures to determine species diversity will be provided.

Procedure

All possible undifferentiated fever cases will be interviewed and examined by the on-site pre-intern medical officers and suspected leptospirosis cases (fulfilling inclusion and exclusion criteria) will be referred to the study project. In the outpatient department (OPD), once the medical officers complete the history taking and examination, the data collector will collect the data. For inward patients, the pre-intern medical officer will screen all fever patients upon admission to the ward, for the selection of eligible patients. In Anuradhapura, four medical units are available and the study will be carried out in all four medical wards. In Peradeniya, a single medical ward is available. Data collectors will visit all wards and ICUs daily for the purpose of data collection. Once the routine ward procedures are completed, data will be collected. All eligible patients will be provided an explanatory statement of the study. The data collector will also explain the benefits and data confidentiality, as well as the right to withdraw from the study at any given time. Once a patient is well informed, written consent will be obtained before enrolment. Data will be collected in the ward, at the patient bed side to ensure the patient's privacy and confidentiality, as a routine clinical procedure. In the outpatient department, a separate location will be used to collect the data. Data will be collected using a fully structured,

interviewer-administered questionnaire. An additional clinical data check list will be used to assess the clinical features. A daily follow-up will be performed for all hospitalized patients to observe the clinical progress of the disease using a clinical checklist. The questionnaires and the checklists will be adopted from previously published time-tested protocols.[22]

All patients will be given an appointment for follow up 3 weeks after the initial assessment. To maximize patient follow-up, reminders and telephone calls will be used to contact participants before their appointment, and the fee for transportation will be paid. Both blood (2 mL) and urine (10-15 mL) samples will be collected from patients during the follow-up.

Sample size

Assuming that at least 70% of patients are infected with same infective *Leptospira* strain in a specific setting, and to detect at least a 20% difference between the OPD and hospitalized patients, at least 62 leptospirosis patients with molecular identification are needed from each site. This calculation was performed with a power of 80% and an alpha value of 0.05. Based on our previous studies, the interpretable sequencing data are available for around 50% of confirmed cases of leptospirosis[23] and leptospirosis accounts for 15% of acute undifferentiated fever.[24] Based on these assumptions, the total number of febrile patients that is required from one sample collection site for this study is 827. This number of patients is conservatively estimated and feasible to obtain within the project period.

Variables and epidemiological data

We will be looking at socio-demographic, exposure and clinical data for the clinic-epidemiological components of the study. The socio-demographic data will include age, sex, ethnicity, religion, occupation, income and residence. These variables are usually the distance determinants of leptospirosis and considered as important in determining the social determinants of the disease. Exposure variables will be looking at several types of exposure. First set of

variables will include the residential and working environment and hosts. Next set will include actual individual exposure during last three weeks. Clinical data will be obtained initially to describe the presenting complains. All symptoms will be recorded as described by the patients. We will be using a set of probing questions to clarify the presenting complains. Signs will be documented after the physical examination by the clinically qualified data collector. If a sign is documented after the admission by a treating physician it will also be included, even if it is not demonstrable at the time of examination.

Sample collection, procession, and storage procedure

Blood and urine samples will be collected from patients on the day of admission to the ward or upon presentation to the outpatient department. A total of 7 ml of blood will be collected and a bed side inoculation of 2 to 4 drops (100-500 µl) will be performed using two tubes containing 9 ml of EMJH semisolid media with added antibiotics (5-Fluorouracil and Neomycin). Collected blood added to EMJH media for culture will be incubated at 30ºC in an incubator.

Venous blood (2.5 mL) will be collected in an EDTA tube and the remainder will be placed into plain tubes for serum separation. For serum separation, the blood in the plain tubes will be centrifuged at 1300 rpm for 10 minutes. The blood in the EDTA tubes will be used for PCR assay and the serum samples for serological assays. Aliquots (500 µl) will be prepared from the whole blood and serum samples. These will then be frozen at –20ºC or –80ºC for short-term or long-term storage, respectively. This procedure will be completed within 2 hours.

Clean catch mid-stream (10-15 mL) urine will be collected into a sterile wide mouth container and processed within 2 hours of collection. Initially, the urine will be centrifuged at 3000 rpm for five minutes to sediment the white cells, epithelial cells, and other crystals. Next, the supernatant will be separated and further centrifuged at 15000 rpm for 10 minutes. The sediment obtained will be used for PCR studies and stored at –20ºC until further use.[25-27] If the urine cannot be processed during the specified time period, the pH of the urine will be adjusted to

approximately 7.4 using 1 M NaOH solution. Three weeks after the onset of fever, a follow-up blood sample (2 mL) will be collected in plain tubes for paired sera and 10-15 ml of urine will be collected for PCR assay.

Disease confirmation

Disease will be confirmed using a regionally optimized microscopic agglutination test (MAT), culture, and qPCR-evidence of the presence of *Leptospira* in the serum. The criteria for confirmed and probable cases of leptospirosis is presented in Table 1.[28]

Table 1. Criteria for the diagnosis of confirmed and probable cases of leptospirosis

Case	Definition
Confirmed	**Clinical signs and symptoms consistent with leptospirosis with any one of the following criteria:** 1. Four-fold increase in MAT titer in acute and convalescent serum samples. 2. MAT titer ≥ 1:400 in single or paired serum samples. 3. Isolation of pathogenic *Leptospira* species from normally sterile sites. 4. Pathogenic Leptospira species DNA detected by PCR.
Probable case	**Clinical signs and symptoms consistent with leptospirosis with any one of the following criteria:** 1. Presence of IgM antibodies by enzyme-linked immunosorbent assay (ELISA) or dipstick. 2. MAT titer ≥ 1:100 in single acute-phase serum sample (with no convalescent serum).

MAT

For this purpose, we will use the U.S. Center for Disease Control's (CDC) MAT panel with five already available local strains obtained from the Royal Tropical Institute, Amsterdam (KIT). Once the Sri Lankan isolates are obtained, a validation of best panel suitable for Sri Lanka will be determined. MAT will be carried out in two main steps. First, the serum samples will be screened at a titer of 1/50 using the MAT panel of 24 strains, which includes the five local strains. Positive sera from the screening test will be subjected to a run out test with the serial dilution of sera from 1/50 to 1/3200. MAT evidence for seropositivity will be defined as one or more of the following: seroconversion (1/100) from negative to positive; 4-fold increase in titer between acute-phase and convalescent-phase (follow-up) samples; or a single titer of ≥400.

Culture isolation of *Leptospira*

For the culture isolation, we adopted the previously validated protocol published by Wuthiekanan et al. [29]. EMJH semisolid media will be prepared by adding 2.3 g of EMJH base, 1.5 g of bacteriological agar, and 100 mg of sodium pyruvate into 785 ml of distilled water, then adjusting the pH to 7.4. The media will be autoclaved and cooled to around 50ºC, Next, 100 ml of *Leptospira* enrichment media and 100 ml of fetal bovine serum will be added to the media.[30-33] To suppress the growth of possible contaminants, 5-Fluorouracil and Neomycin will be added, at concentrations of 100 µg/ml and 25 µg/ml, respectively. The media inoculated with blood will be inspected using dark field microscopy for the presence of motile *Leptospira*, initially after 3 weeks and then on a monthly basis. When a positive growth is detected, sub cultures will be transferred into liquid and semisolid media. An aliquot will then be obtained and fixed with 5% DMSO (Dimethyl Sulfoxide) and stored in -80ºC. To maintain live cultures, the isolates will be subcultured in liquid media on a bi-weekly basis and on semisolid media every 3 months. Serotyping of the newly isolated *Leptospira* strains will be carried out at the Pasteur Institute, France.

Quantitative PCR

Diagnosis

For the diagnosis, DNA will be extracted from whole blood, serum, and urine samples using the QIAamp DNA Blood Mini Kit (Qiagen, USA), according to the manufacturer's instructions. The extracted DNA will be quantified using an Invitrogen Qubit 4 Flurometer. Two previously published quantitative PCR protocols targeting 16s ribosomal RNA and Lipl32 genes will be used for the detection of pathogenic *Leptospira* DNA based on SYBR Green Chemistry (Table 2).[34]

Table 2. Primers for the detection of *Leptospira* DNA in clinical samples

Primer/probe sequence	**Annealing temperature**	**Complementary target species**
Forward: 5'-AAG CAT TAC CGC TTG TGG TG-3' Reverse: 5'-GAA CTC CCA TTT CAG CGA TT-3'	60ºC	lipL32 gene of pathogenic *Leptospira* spp.
Forward: 5'-TAA AGG CTC ACC AAG GCG AC-3' Reverse: 5'-TTA GCC GGT GCT TTA GGC AG-3'	60ºC	16s gene of pathogenic *Leptospira* spp.

Speciation

To determine the *Leptospira* species in the positive cultures, a previously published quantitative PCR protocol based on SYBR Green Chemistry (Table 3) will be used (before sequencing).[35] This will facilitate the identification of L. *interrogans, L. borgpetersenii, L. kirschneri*, and *L. noguchii*.

Table 3. Primer pairs for detection of species

Primer/Probe sequence	Annealing temperature	Complementary target species
Forward: 5'-CTT GAG CCT GCG CGT TAY-3' Reverse: 5'-CCG ATA ATT CCA GCG AAG ATC-3'	63ºC	secY gene of *L. interrogans*
Forward: 5'-GAT TCG GGT TAC AAT TAG ACC-3' Reverse: 5'-TTG ATC TAA CCG GAC CAT AGT-3'	65ºC	ompL1 gene of *L. borgpetersenii*
Forward: 5'-CTG GCT TAA TCA ATG CTT CTG-3' Reverse: 5'-CTC TTT CGG TGA TCT GTT CC-3'	60ºC	secY gene of *L. kirschneri*
Forward: 5'-TCA GGG TGT AAG AAA GGT TC-3' Forward: 5'-CAA AAT TAA AGA AGC AAA GA-3'	63ºC	secY gene of *L. noguchii*

Molecular studies

For newly isolated *Leptospira*, whole genome sequencing will be performed using the following two technologies: MinION Nanopore sequencing in the Leptospirosis Research Laboratory (Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka, India) and PacBio third generation sequencing at the of University) Medicine Genomic for Institute California, San Diego, USA).

BAS and BAX files containing all the DNA sequence information will be converted into FASTA files. Prior to assembly, the quality (coverage) of the raw data will be assessed. Using the PacBioEDA software, the quality summary of the data will be obtained. The CanuPacBiO assembly program will be used for genome assembly. The raw data of Bas/Bax files will be converted into bam files using the bax2bam software. These bam files will be aligned to our references using the Pbalingn software. The software Quiver will be used to polish the data. The assembled data will be annotated using Prokka[36] and Roary[37], and annotations will be

inspected and curated using Artemis. The functional annotation will be performed using the eggNOG and eggNOG MAPPER online tools.[38] The genomes of the Sri Lankan isolates will be compared using MAUVE software. The organism will be roughly identified from the direct raw data using KRAKEN.[39]

MLST will be used for the genotyping of the new isolates. Sri Lankan isolates will be identified using three published MLST schemes for *Leptospira*.[40-42] Genetic similarity between the isolates will be determined using SPLITSTREE.

Taxonogenomics, pan-genome, and phylogenetic analyses

To establish how this strain was related to other *Leptospira* strains, we will use several genomic analysis tools. First, the 16S rRNA sequences will be used to construct a phylogeny. High sequence identities (≥97%) between the newly sequenced strain and previously identified *Leptospira* strains will be preliminarily used as suggestive of phylogenetic relationships. Accordingly, the 16S rRNA gene sequences of *Leptospira* spp. deposited in the GenBank database will be acquired. Multiple sequence alignment based on 16S rRNA will be performed using ClustalX v1.81. These results will be then used to construct the maximum likelihood tree using MEGA v5.05. To evaluate clade support, we will perform a bootstrap analysis with 1,000 replicates. The relationship of each sequenced genome to the previously described or novel *Leptospira* species will be determined according to the Overall Genetic Relatedness Indices (OGRIs), Average Nucleotide Identity (ANI), and the Average Amino Acid Identity (AAAI). The indices will be automatically calculated using two-way BLA blastn and blastp. For the ANI calculation, the default parameters shall be as follows: sequence identity cut-off, 30%; alignment cut-off, 70%; query length, 1,020 bp.

Data management and analysis

Patient interviews and patient health records will be used as data sources. On-site data entry will be performed in real-time for all the collected data. An EpiInfo database will be prepared using cross-checks, field validation, and other in-built techniques available for data quality improvement. The investigators will frequently visit the hospitals to cross-validate the data collected. All databases will be stored on password-protected computers. The main database will be stored at the Leptospirosis Research Laboratory of the Faculty of Medicine and Allied Science, Rajarata University of Sri Lanka. Personal identification data will be removed from all the databases, such that only aliquots and predesigned patient identifiers are available. Hardcopies of the questionnaires will be archived using the standard protocols for archiving hard copies of questionnaires. Only the investigators will have access to the database. The database will be continuously updated with laboratory data using a linking patient identifier (Key variable). The lab results will be generated for all samples daily, once a fully functioning lab is set up, and the treating physicians will be informed about this over the phone. Separate laboratory registers will be maintained for the receipt of samples and the delivery of the results.

For the missing data, case deletions and pairwise deletions will be carried out, based on the amount of missing data. Proportions and 5% confidence intervals will be used to describe the confirmed cases, diversity of infecting *Leptospira*, and sequelae. A Chi-square test will be used to identify any trends and test significant differences in the *Leptospira* diversity in the different microgeographical and clinical settings. Predictors of severe disease will be analyzed using a logistic regression model. Model building will be performed using a conceptual hierarchy of determinants, which we have previously proposed for leptospirosis. An online database for *Leptospira* MLST will be used for typing analysis.

Patient and public involvement

The study was developed based on the patients' feedback regarding leptospirosis disease burden, its impact on their lives, and in particular, the problems faced during illness. In previous

studies, we specifically explored patients as well as general public views on leptospirosis disease and prevention, and used this information for the design of this study. Moreover, the results of this study will be disseminated on a continuous basis for the patients. We have already conducted several public awareness and health worker education workshops based on our findings.

ETHICS AND DISSEMINATION

Undifferentiated fever patients eligible for the study will be given an explanatory statement of the study and written informed consent will be obtained before enrolment. Rapid diagnostic kits will be used on site to assist treating physicians. qPCR diagnostic facilities will be provided for all patients whenever the urgent tests are requested by treating physicians. All test results will be informed to the patients and the wards. Ethical clearance for the study was obtained from the Ethic Review Committees of Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka (ERC/2015/18).

All genomic data will be deposited in Gene Bank. Characterized serum samples will be stored in -80C in the Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka. Data pertaining to this study will be available for researchers and policy makers on a reasonable request address to the corresponding author. After the completion of work, data will be deposited in a repository.

REFERENCES

1. WER: **Leptospirosis worldwide, 1999**. *Wkly Epidemiol Rec* 1999, **74**(29):237-242.
2. Costa F, Hagan JE, Calcagno J, Kane M, Torgerson P, Martinez-Silveira MS, Stein C, Abela-Ridder B, Ko AI: **Global Morbidity and Mortality of Leptospirosis: A Systematic Review**. *PLoS Negl Trop Dis* 2015, **9**(9):e0003898.
3. Taylor AJ, Paris DH, Newton PN: **A Systematic Review of the Mortality from Untreated Leptospirosis**. *PLoS Negl Trop Dis* 2015, **9**(6):e0003866.
4. Agampodi S, Peacock SJ, Thevanesam V: **The potential emergence of leptospirosis in Sri Lanka**. *Lancet Infect Dis* 2009, **9**(9):524-526.
5. Epidemiological Unit: **Selected notifiable diseases reported by Medical Officers of Health**. *Weekly Epidemiological Report* 2009, **36**(1):4.
6. Epidemiological Unit: **Selected notifiable diseases reported by Medical Officers of Health**. *Weekly Epidemiological Report* 2010, **37**(1):4.
7. Warnasekara JN, Agampodi S: **Leptospirosis in Sri Lanka**. *Sri Lankan Journal of Infectious Diseases* 2017, **7**(2).

8. Pappas G, Papadimitriou P, Siozopoulou V, Christou L, Akritidis N: **The globalization of leptospirosis: worldwide incidence trends**. *Int J Infect Dis* 2008, **12**(4):351-357.
9. Agampodi SB, Peacock SJ, Thevanesam V, Nugegoda DB, Smythe L, Thaipadungpanit J, Craig SB, Burns MA, Dohnt M, Boonsilp S *et al*: **Leptospirosis outbreak in Sri Lanka in 2008: lessons for assessing the global burden of disease**. *Am J Trop Med Hyg* 2011, **85**(3):471-478.
10. Agampodi SB, Dahanayaka NJ, Bandaranayaka AK, Perera M, Priyankara S, Weerawansa P, Matthias MA, Vinetz JM: **Regional differences of leptospirosis in Sri Lanka: observations from a flood-associated outbreak in 2011**. *PLoS Negl Trop Dis* 2014, **8**(1):e2626.
11. Myers WP, Myers AP, Cox-Singh J, Lau HC, Mokuai B, Malley R: **Micro-geographic risk factors for malarial infection**. *Malaria journal* 2009, **8**:27.
12. Schellenberg JA, Newell JN, Snow RW, Mung'ala V, Marsh K, Smith PG, Hayes RJ: **An analysis of the geographical distribution of severe malaria in children in Kilifi District, Kenya**. *Int J Epidemiol* 1998, **27**(2):323-329.
13. Hassan AO, Amoo AOJ, Deji-Agboola AM, Adeleke MA, Muhibi MA, Hassan RO: **Microgeographical patterns of schistosomiasis and anthropometric indices of children in eko-ende south west, nigeria**. *Int J Biol Med Res* 2911, **2**(3):656-660.
14. Kloos H, Gazzinelli A, Van Zuyle P: **Microgeographical patterns of schistosomiasis and water contact behavior; examples from Africa and Brazil**. *Mem Inst Oswaldo Cruz* 1998, **93 Suppl 1**:37-50.
15. Baliraine FN, Afrane YA, Amenya DA, Bonizzoni M, Menge DM, Zhou G, Zhong D, Vardo-Zalik AM, Githeko AK, Yan G: **High prevalence of asymptomatic plasmodium falciparum infections in a highland area of western Kenya: a cohort study**. *J Infect Dis* 2009, **200**(1):66-74.
16. Dissanayake CB, Chandrajith R: **Medical geochemistry of tropical environments**. *Earth-Science Reviews* 1999, **47**(3–4):219-258.
17. Naotunna C, Agampodi SB, Agampodi TC: **Etiological agents causing leptospirosis in Sri Lanka: A review**. *Asian Pac J Trop Med* 2016, **9**(4):390-394.
18. Agampodi SB, Moreno AC, Vinetz JM, Matthias MA: **Utility and limitations of direct multi-locus sequence typing on qPCR-positive blood to determine infecting Leptospira strain**. *Am J Trop Med Hyg* 2013, **88**(1):184-185.
19. Agampodi SB, Miatthias MA, Moreno AC, Vinetz JM: **Utility of Quantitative PCR in Leptospirosis Diagnosis: Association of Level of Leptospiremia and Clinical Manifestations in Sri Lanka**. *Clinical Infectious Disease* 2012, **In Press**.
20. Katz AR: **Quantitative polymerase chain reaction: filling the gap for early leptospirosis diagnosis**. *Clin Infect Dis* 2012, **54**(9):1256-1258.
21. Nityananda K, Harvey T: **Leptospirosis in Ceylon-Epidemiological and Laboratory Studies**. *Ceylon J Med Sci* 1971, **20**(1):5-14.
22. Agampodi SB, Nugegoda DB, Thevanesam V: **Determinants of leptospirosis in Sri Lanka: Study Protocol**. *BMC Infect Dis* 2010, **10**(1):332.
23. Agampodi SB, Matthias MA, Moreno AC, Vinetz JM: **Utility of quantitative polymerase chain reaction in leptospirosis diagnosis: association of level of leptospiremia and clinical manifestations in Sri Lanka**. *Clin Infect Dis* 2012, **54**(9):1249-1255.
24. Reller ME, Bodinayake C, Nagahawatte A, Devasiri V, Kodikara-Arachichi W, Strouse JJ, Flom JE, Dumler JS, Woods CW: **Leptospirosis as frequent cause of acute febrile illness in southern Sri Lanka**. *Emerg Infect Dis* 2011, **17**(9):1678-1684.
25. Lucchesi PM, Arroyo GH, Etcheverria AI, Parma AE, Seijo AC: **Recommendations for the detection of Leptospira in urine by PCR**. *Rev Soc Bras Med Trop* 2004, **37**(2):131-134.

26. Bal AE, Gravekamp C, Hartskeerl RA, De Meza-Brewster J, Korver H, Terpstra WJ: **Detection of leptospires in urine by PCR for early diagnosis of leptospirosis**. *J Clin Microbiol* 1994, **32**(8):1894-1898.
27. Van Eys GJ, Gravekamp C, Gerritsen MJ, Quint W, Cornelissen MT, Schegget JT, Terpstra WJ: **Detection of leptospires in urine by polymerase chain reaction**. *J Clin Microbiol* 1989, **27**(10):2258-2262.
28. World Health Organization: **Report of the second meeting of leptospirosis burden epidemiology reference group**. In. Geneva: World Health Organization 2011: 1-34.
29. Wuthiekanun V, Chierakul W, Limmathurotsakul D, Smythe LD, Symonds ML, Dohnt MF, Slack AT, Limpaiboon R, Suputtamongkol Y, White NJ *et al*: **Optimization of culture of Leptospira from humans with leptospirosis**. *J Clin Microbiol* 2007, **45**(4):1363-1365.
30. Wuthiekanun V, Amornchai P, Langla S, Oyuchua M, Day NP, Limmathurotsakul D: **Maintenance of leptospira species in leptospira Vanaporn Wuthiekanun agar**. *J Clin Microbiol* 2014, **52**(12):4350-4352.
31. Wuthiekanun V, Amornchai P, Langla S, White NJ, Day NP, Limmathurotsakul D, Peacock SJ: **Antimicrobial Disk Susceptibility Testing of Leptospira spp. Using Leptospira Vanaporn Wuthiekanun (LVW) Agar**. *Am J Trop Med Hyg* 2015, **93**(2):241-243.
32. Wuthiekanun V, White NJ, Amornchai P, Day NP, Limmathurotsakul D: **Quality controls for antimicrobial disk diffusion testing on Leptospira Vanaporn Wuthiekanun agar**. *Trans R Soc Trop Med Hyg* 2016, **110**(11):673-675.
33. Picardeau M, Bulach DM, Bouchier C, Zuerner RL, Zidane N, Wilson PJ, Creno S, Kuczek ES, Bommezzadri S, Davis JC *et al*: **Genome sequence of the saprophyte Leptospira biflexa provides insights into the evolution of Leptospira and the pathogenesis of leptospirosis**. *PLoS ONE* 2008, **3**(2):e1607.
34. Backstedt BT, Buyuktanir O, Lindow J, Wunder EA, Reis MG, Usmani-Brown S, Ledizet M, Ko A, Pal U: **Efficient Detection of Pathogenic Leptospires Using 16S Ribosomal RNA**. *PLOS ONE* 2015, **10**(6):e0128913-e0128913.
35. Ferreira AS, Costa P, Rocha T, Amaro A, Vieira ML, Ahmed A, Thompson G, Hartskeerl RA, Inacio J: **Direct detection and differentiation of pathogenic Leptospira species using a multi-gene targeted real time PCR approach**. *PLoS One* 2014, **9**(11):e112312.
36. Seemann T: **Prokka: rapid prokaryotic genome annotation**. *Bioinformatics* 2014, **30**(14):2068-2069.
37. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J: **Roary: rapid large-scale prokaryote pan genome analysis**. *Bioinformatics* 2015, **31**(22):3691-3693.
38. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, Bork P: **Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper**. *Mol Biol Evol* 2017, **34**(8):2115-2122.
39. Wood DE, Salzberg SL: **Kraken: ultrafast metagenomic sequence classification using exact alignments**. *Genome Biol* 2014, **15**(3):R46.
40. Varni V, Ruybal P, Lauthier JJ, Tomasini N, Brihuega B, Koval A, Caimi K: **Reassessment of MLST schemes for Leptospira spp. typing worldwide**. *Infect Genet Evol* 2014, **22**:216-222.
41. Ahmed N, Devi SM, Valverde Mde L, Vijayachari P, Machang'u RS, Ellis WA, Hartskeerl RA: **Multilocus sequence typing method for identification and genotypic classification of pathogenic Leptospira species**. *Ann Clin Microbiol Antimicrob* 2006, **5**:28.
42. Boonsilp S, Thaipadungpanit J, Amornchai P, Wuthiekanun V, Bailey MS, Holden MT, Zhang C, Jiang X, Koizumi N, Taylor K *et al*: **A single multilocus sequence typing (MLST) scheme for seven pathogenic Leptospira species**. *PLoS Negl Trop Dis* 2013, **7**(1):e1954.

AUTHORS' CONTRIBUTIONS

JMV and SBA drafted the initial protocol. YPJNW, JMDDJ and RMISKS are responsible for the detailed design, field/ laboratory work analysis and interpretation of objectives 1-3, 4 and 5 respectively. CDG, SAMK and SHS helped in design, field, hospital and laboratory work. MM participated in scientific design and guided all laboratory components. All authors have read and approved the manuscript.

FUNDING

This project is the Sri Lankan component of a multi country study funded by the NIAID of the National Institutes of Health under award number U19AI115658. Funders had no role in the design, execution, analyses, data interpretation, or decision to submit results.

CONFLICTS OF INTERESTS

The authors declare that they have no competing interests.

WORD COUNT

3791

FIGURES

Figure 1: Locations of study sites.

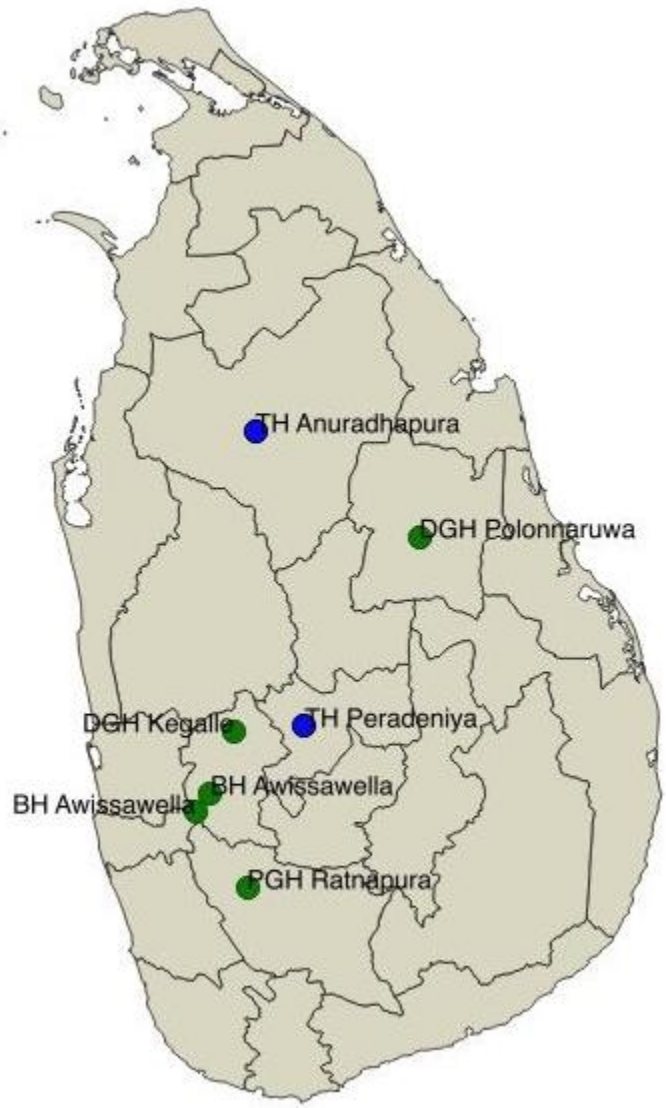


Figure 1. Study sites

147x201mm (72 x 72 DPI)

BMJ Open

Study protocol: Characterizing the clinical, epidemiological, and etiological aspects of leptospirosis in Sri Lanka

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-027850.R2
Article Type:	Protocol
Date Submitted by the Author:	11-Jul-2019
Complete List of Authors:	Agampodi, Suneth; Rajarata University, Department of Community Medicine Warnasekara, Janith; Rajarata University of Sri Lanka Faculty of Medicine and Allied Sciences, Department of Community Medicine Jayasundara, Dinesha; Rajarata University of Sri Lanka Faculty of Medicine and Allied Sciences, Department of Microbiology Senawirathna, Indika; Rajarata University of Sri Lanka Faculty of Medicine and Allied Sciences, Biochemistry Gamage, Chandika; University of Peradeniya Faculty of Medicine, Microbiology Kularatne, Senanayake; University of Peradeniya Faculty of Medicine, Medicine Siribaddana, Sisira; Rajarata University of Sri Lanka Faculty of Medicine and Allied Sciences, Department of Medicine Maththias, Michael; Yale University School of Medicine Vinetz, Joseph; Yale University School of Medicine
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Epidemiology, Global health, Public health, Infectious diseases
Keywords:	Leptospirosis, Sri Lanka, Leptospira, Isolates, Whole genome sequencing, Clinical disease

Study protocol: Characterizing the clinical, epidemiological, and etiological aspects of leptospirosis in Sri Lanka

Agampodi SB[1*], Warnasekara YPJN[1], Jayasundara JMDD[2], Senawirathna RMISK[3], Gamage CD[4], Kularatne SAM[5], Siribaddana SH[6], Maththias M[7], Vinetz JV[7]

[1] Department of Community Medicine, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka

[2] Department of Microbiology, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka

[3] Department of Biochemistry, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka

[4] Department of Microbiology, Faculty of Medicine, University of Peradeniya

[5] Department of Medicine, Faculty of Medicine, University of Peradeniya

[6] Department of Medicine, Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka

[7] Yale School of Medicine, Section of Infectious Diseases, Department of Internal Medicine New Haven, CT 06520

ASB: sunethagampodi@yahoo.com,
WYPJN: jwarnasekara@yahoo.com
JJMDD: dineshajayasundara@yahoo.com
SRMISK: irathnabahu@gmail.com
GCD: chandika@pdn.ac.lk
KSAM: samkul@sltnet.lk
SSH: sisira.siribaddana@gmail.com
MM: mikematthias34@icloud.com
VJV: joseph.vinetz@yale.edu

*Corresponding author.

ABSTRACT

Introduction: Sri Lanka has one of the highest incidences of leptospirosis worldwide. We hypothesized that different geographical locations and patient context will have a distinct molecular epidemiology of leptospirosis, based on microgeographic characteristics related to regiona-specific *Leptospira* predominance. Our objective is to characterize the clinical, epidemiological, and molecular aspects of leptospirosis in Sri Lanka to understand disease progression, risk factors, and obtain isolates of *Leptospira*. **Methods and Analysis:** We designed a multi-center prospective study in Sri Lanka to recruit undifferentiated febrile patients and conduct follow-ups during hospital stays. Patients will be recruited from outpatient departments and medical wards. This study will be conducted at two main sites (Anuradhapura and Peradeniya) and several additional sites (Awissawella, Ratnapura and Polonnaruwa). Blood and urine will be collected from patients on the day of admission to the ward or presentation to the outpatient department. Bedside inoculation of 2 and 4 drops of venous blood will be performed with EMJH semisolid media supplemented with antibiotics. Regionally optimized microscopic agglutination test, culture, and qPCR-evidence will be performed to confirm the presence of *Leptospira* in blood, which in turn will confirm the presence of disease. Whole genome sequencing will be carried out for all isolates recovered from patients. Multi locus sequence typing (MLST) will be used for the genotyping of new isolates. Sri Lankan isolates will be identified using three published MLST schemes for *Leptospira.* **Ethics and Dissemination:** Ethical clearance for the study was obtained from Ethics Review Committees (ERC), Medicine and Allied Sciences (FMAS), Rajarata University of Sri Lanka (RUSL) and University of Peradeniya. All genomic data generated through this project will be available at GenBank. Anonymized data will be deposited at the ERC, FMAS, RUSL.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- This is a multi-center study involving wet and dry zones, low and highlands and high endemic and low endemic areas of the country covering different geographical and climate zones to provide a better understanding of leptospirosis in Sri Lanka.
- We will use a prospective study design focusing on culture isolation as a main aim with specific protocols for the particular objective to describe the strain diversity of *Leptospira spp*. in Sri Lanka.
- Clinical and epidemiological data will be collected as primary data to ensure high quality data using clinically trained health professionals.
- Since *Leptospira* are fastidious organisms, the culture yield will be low and linking the clinical disease with isolated *Leptospira* may be partly confounded by the growth potentials of different *Leptospira*.
- Since PCR and culture-based methods are typically valid only during the first 7-10 days of the illness and late presentations may lead to reduce sensitivity of these tests, MAT is best interpreted with paired samples, and a lack of convalescent samples will impair the interpretation of MAT results.

INTRODUCTION

Leptospirosis is a globally widespread, neglected, and emerging zoonotic disease[1] with estimated 1.03 million cases and 58,900 annual deaths attributed to the disease.[2] An average of 2.9 million daily adjusted life years (DALYs) are estimated to be lost every year due to leptospirosis.[3] Emerging leptospirosis mostly affects vulnerable communities living in resource-poor settings. Often, the disease is either not suspected or is under or misdiagnosed due to the need for laboratory resources to confirm leptospirosis; typically, such resources are neither accessible nor affordable. From the clinical perspective, a better understanding of the clinical disease associated with diverse *Leptospira* species is required and superior diagnostics are needed to prevent severe complications and death. From the public health perspective, the lack of reliable and efficient diagnostic tests makes assessing the burden of disease—whether regionally or globally—difficult.

Sri Lanka has emerged as a country with a high incidence of leptospirosis since 2007[4] and in 2008, the total number of clinically suspected cases reported to the surveillance system was 7,406 with 204 deaths.[5] In 2009, 4980 cases and 145 deaths were reported[6], and the outbreak persisted until 2013 with more than 4000 reported cases each year.[7] The probable case incidence during this time period was more than 20 per 100,000 population, indicating that Sri Lanka has one of the highest reported incidences of leptospirosis worldwide.[8] Our extensive work on the 2008 outbreak suggested that clinical diagnosis of leptospirosis is highly inaccurate, since as much as 40% of the leptospirosis patients are misdiagnosed in hospitals.[9] However, all Sri Lankan studies, including our previous study, have the same drawback, as observed in most of the published reports in global literature: hospital admission bias, where a large number of patients with leptospirosis may have been treated as outpatients, which leads to a severe underestimate of the actual disease burden. A precise disease burden estimate requires seroprevalence studies coupled with population-based incidence studies.

Our studies on the 2011 outbreak of leptospirosis in Anuradhapura showed that the molecular epidemiology, disease transmission, and clinical manifestations of this outbreak were different from

those observed in Central Province (wet zone) in 2008.[10] One of the main concerns regarding these marked variations in the *Leptospira* strain and the clinical disease was whether this was due to microgeographical variations, environmental conditions, or reservoir mammal infection. It has been shown for other diseases such as malaria[11, 12] and schistosomiasis[13-15], that microgeography may have a major influence on disease epidemiology. Geochemistry is a well-known and major contributory factor in human health.[16] Despite the extensive literature available on leptospirosis, studies on the microgeographic variation of *Leptospira* are scarce.

A systematic review performed on Sri Lankan isolates showed that more than 40 strains from 20 serovars of *Leptospira* have been reported from Sri Lanka.[17] Our previous studies on the 2008 and 2011 outbreaks of leptospirosis in Sri Lanka expanded the knowledge of circulating *Leptospira* and were consistent with a diversity of *Leptospira* infection in Sri Lanka. We showed that in the 2008 outbreak, *L. interrogans* was the predominant species, compared to the post-flood outbreak in 2011, in which *L. krischneri* was the predominant species. However, the molecular epidemiology of *Leptospira* species in endemic settings remains unknown in Sri Lanka. In addition to the STNPCR method, for the first time, we showed that previously published multi locus sequence typing (MLST) schemes could be used with modifications for direct patients samples to study the genetic diversity of infecting *Leptospira* species in resource-poor settings, where the facilities needed to perform cultures and isolations are minimal.[18]

One of the most important conclusions we draw from our published work is that qPCR is promising as a clinical diagnostic tool in the diagnosis of leptospirosis during the acute phase, with a wider window of positivity than previously noted.[19] The editorial by Katz[20] endorsed the qPCR approach but also emphasized the need for a prospective evaluation of this method. However, most countries with high leptospirosis burdens still find it prohibitively expensive to have molecular based methods for rapid diagnosis, in terms of both the costs and logistics.

Disease burden assessment, public health interventions, and clinical management of leptospirosis are challenges due to the lack of diagnostic facilities. Any population-based attempt

to estimate the disease burden is often limited by the unavailability of diagnostic tests or the poor performance of the available diagnostic tests. The development of diagnostic tests that are both global and country-specific requires knowledge on circulating serovars. Furthermore, well-characterized samples are required with species and subspecies level identification of infecting *Leptospira* backed by culture isolations. In Sri Lanka, a culture isolation of *Leptospira* has not been reported since the 1970s. Even though we showed the microgeographical changes of leptospirosis in 2014,[10] no prospective studies have been conducted in a Sri Lankan cohort to primarily identify the disease diversity with specific research design. We hypothesized that a different study site and patient context—inpatient vs. outpatient—will have a distinct molecular epidemiology of leptospirosis, based on microgeographic characteristics related to region-specific *Leptospira* predominance, clinical characteristics related to pathogenesis/virulence potential of specific *Leptospira* species, strains, serovars, and the acquisition of infection related to occupational and activities of daily living, reflecting different environmental exposure contexts. Based on these background analyses and hypotheses, the objectives of the present study were as follows:

1. To determine the prevalence of leptospirosis among undifferentiated febrile patients (inpatient and outpatient) in contrasting geographical settings in Sri Lanka.

2. To describe the full clinical spectrum of leptospirosis in patients infected with *Leptospira*.

3. To determine the predictors of leptospirosis progression using a nested case control approach among febrile patients, both inpatient and ambulatory.

4. To determine the *Leptospira* species and type diversity in different geographical settings in Sri Lanka.

5. To characterize Sri Lankan *Leptospira* using MLST and whole genome sequencing methods.

6. To have a well-characterized collection of serum, whole blood, and urine samples for the future development and validation of leptospirosis diagnostics.

METHODS AND ANALYSIS

Study sites

This study will be performed in four primary subject enrolment/data collection centers in two main contrasting sites: Kandy and Anuradhapura. In addition, a small component of the study will be carried out in Sabaragamuwa, Western, North Central, and Southern provinces for species diversity (Figure 1). Kandy and Anuradhapura was selected based on our previous observations of *Leptospira* diversity in these two settings, supported by historically distant identification of diverse *Leptospira* in Sri Lanka (1960s)[21]. Since no recent studies have been conducted, this project will yield up-to-date information at the molecular and genomically-determined taxonomical levels in addition to conventional methods of identifying *Leptospira*. The initial focus is on Anuradhapura, located in North Central province of Sri Lanka, in the dry zone of the country. The annual rainfall is 1,200-1,600 mm, with a mean annual temperature of ~30°C. In Anuradhapura, paddy farming is carried out by traditional, full-time farmers and constitutes the main mode of income. Cattle, water buffalo, certain rodent species—well-known sources of *Leptospira* transmission—are found in and around the paddy fields. Paddy fields are large and well maintained. There are no wetlands or marshy lands in these areas, except the paddy fields during working seasons. Paddy field work depends on irrigation systems, such that, between farming seasons, the paddy fields become completely dry. Other risk factors for acquiring leptospirosis include daily activities involved in living in economically poor conditions, for example living in proximity to dogs and rodents in houses and neighborhoods and the wearing of shoes that are often no more than sandals.

The soil structure, water quality/hardness, and ecological systems in Anuradhapura contrast with those of the wet zone of Sri Lanka. Kandy (another historic site in Sri Lanka/Ceylon, the cultural capital of the country since ancient times) is situated in the wet zone, with over 3,000 mm of rainfall, and is located in the middle, hilly parts of the country. Paddy farming in this area is based on rainwater and the paddy fields remain wet throughout the year. These are small fields adjacent to wildlife and the soil and ecological structure is different from that in Anuradhapura. Nonetheless, the economic conditions of residents in this region are poor and rural, such that rodent-

and dog-transmitted *Leptospira* in the context of activities of daily living is likely common but completely unexplored.

The purpose of including additional sites is primarily to identify differences in leptospirosis outbreaks, especially during floods and unsuspected outbreaks. Only selected components of the study will be performed in those areas if there are outbreaks.

Study population

The study population will include all undifferentiated fever cases either presented to the outpatient department or admitted to medical wards/ICUs in TH Anuradhapura and TH Peradeniya (main sites) and selected hospitals in Western, Sabaragamuwa, and Southern provinces.

Inclusion criteria

- Age ≥ 12 years.
- Temperature ≥ 38°C.
- Self-reported fever ≥ 2 days.

Exclusion criteria

- Fever with running nose without any other symptom or sign.
- Any patient with confirmed diagnosis/foci of infection as a cause for the fever.
- Traumatic or post-operative fever per physician discretion.
- Fever due to nosocomial infections.

Patient recruitment

Graduates with a Bachelor of Medicine and Bachelor of Surgery (MBBS) awaiting their internship appointment and established registrars will be assigned to all four centers and directly supervised on a daily basis. They will screen all febrile patients and perform directed physician

exams in the outpatient department, as well as on inpatients admitted to the medical wards. Once the diagnostic facilities are established, we will inform all clinicians in the selected districts about the available facilities. If they express interest, diagnostic facilities, including cultures to determine species diversity will be provided.

Procedure

All possible undifferentiated fever cases will be interviewed and examined by the on-site pre-intern medical officers and suspected leptospirosis cases (fulfilling inclusion and exclusion criteria) will be referred to the study project. In the outpatient department (OPD), once the medical officers complete the history taking and examination, the data collector will collect the data. For inward patients, the pre-intern medical officer will screen all fever patients upon admission to the ward, for the selection of eligible patients. In Anuradhapura, four medical units are available and the study will be carried out in all four medical wards. In Peradeniya, a single medical ward is available. Data collectors will visit all wards and ICUs daily for the purpose of data collection. Once the routine ward procedures are completed, data will be collected. All eligible patients will be provided an explanatory statement of the study. The data collector will also explain the benefits and data confidentiality, as well as the right to withdraw from the study at any given time. Once a patient is well informed, written consent will be obtained before enrolment. Data will be collected in the ward, at the patient bed side to ensure the patient's privacy and confidentiality, as a routine clinical procedure. In the outpatient department, a separate location will be used to collect the data. Data will be collected using a fully structured, interviewer-administered questionnaire. Once the initial data collation is completed, all biological sample will be collected (see the section on sample collection, procession, and storage procedure). An additional clinical data check list will be used to assess the clinical features. A daily follow-up will be performed for all hospitalized patients to observe the clinical progress of the disease using a clinical checklist. The questionnaires and the checklists will be adopted from previously published time-tested protocols.[22]

All patients will be given an appointment for follow-up 3 weeks after the initial assessment. To maximize patient follow-up, reminders and telephone calls will be used to contact participants before their appointment, and the fee for transportation will be paid. The main purpose of the follow-up would be to obtain a sample for disease confirmation.

Sample size

Assuming that at least 70% of patients are infected with same infective *Leptospira* strain in a specific setting, at least 24 leptospirosis patients with molecular identification are needed from each site. This calculation was performed with a power of 80% and an alpha value of 0.05. Based on our previous studies, the interpretable sequencing data are available for around 50% of confirmed cases of leptospirosis [23] and leptospirosis accounts for 15% of acute undifferentiated fever.[24] Based on these assumptions, the total number of febrile patients that is required from one sample collection site for this study is 320. This number of patients is conservatively estimated and feasible to obtain within the project period.

Variables and epidemiological data

We will be looking at socio-demographic, exposure and clinical data for the clinic-epidemiological components of the study. The socio-demographic data will include age, sex, ethnicity, religion, occupation, income and residence. These variables are usually the distance determinants of leptospirosis and considered as important in determining the social determinants of the disease. Exposure variables will be looking at several types of exposure. First set of variables will include the residential and working environment and hosts. Next set will include actual individual exposure during last three weeks. Clinical data will be obtained initially to describe the presenting complains. All symptoms will be recorded as described by the patients. We will be using a set of probing questions to clarify the presenting complains. Signs will be documented after the physical examination by the clinically qualified data collector. If a sign is

documented after the admission by a treating physician it will also be included, even if it is not demonstrable at the time of examination.

Sample collection, procession, and storage procedure

Blood and urine samples will be collected from patients on the day of admission to the ward or upon presentation to the outpatient department. A total of 7 ml of blood will be collected and a bed side inoculation of 2 to 4 drops (100-500 µl) will be performed using two tubes containing 9 ml of EMJH semisolid media with added antibiotics (5-Fluorouracil and Neomycin). Collected blood added to EMJH media for culture will be incubated at 30ºC in an incubator.

Venous blood (2.5 mL) will be collected in an EDTA tube and the remainder will be placed into plain tubes for serum separation. For serum separation, the blood in the plain tubes will be centrifuged at 1300 rpm for 10 minutes. The blood in the EDTA tubes will be used for PCR assay and the serum samples for serological assays. Aliquots (500 µl) will be prepared from the whole blood and serum samples. These will then be frozen at –20ºC or –80ºC for short-term or long-term storage, respectively. This procedure will be completed within 2 hours.

Clean catch mid-stream (10-15 mL) urine will be collected into a sterile wide mouth container and processed within 2 hours of collection. Initially, the urine will be centrifuged at 3000 rpm for five minutes to sediment the white cells, epithelial cells, and other crystals. Next, the supernatant will be separated and further centrifuged at 15000 rpm for 10 minutes. The sediment obtained will be used for PCR studies and stored at –20ºC until further use.[25-27] The available protocols for qPCR based urine analysis will be tested and optimized. If the urine cannot be processed during the specified time period, the pH of the urine will be adjusted to approximately 7.4 using 1 M NaOH solution. Three weeks after the onset of fever, a follow-up blood sample (2 mL) will be collected in plain tubes for paired sera and 10-15 ml of urine will be collected for PCR assay.

Disease confirmation

Disease will be confirmed using a regionally optimized microscopic agglutination test (MAT), culture, and qPCR-evidence of the presence of *Leptospira* in the serum. The criteria for confirmed and probable cases of leptospirosis is presented in Table 1.[28]

Table 1. Criteria for the diagnosis of confirmed and probable cases of leptospirosis

Case	Definition
Confirmed	**Clinical signs and symptoms consistent with leptospirosis with any one of the following criteria:** 1. Four-fold increase in MAT titer in acute and convalescent serum samples. 2. MAT titer ≥ 1:400 in single or paired serum samples. 3. Isolation of pathogenic *Leptospira* species from normally sterile sites. 4. Pathogenic Leptospira species DNA detected by PCR.
Probable case	**Clinical signs and symptoms consistent with leptospirosis with any one of the following criteria:** 1. Presence of IgM antibodies by enzyme-linked immunosorbent assay (ELISA) or dipstick. 2. MAT titer ≥ 1:100 in single acute-phase serum sample (with no convalescent serum).

MAT

For this purpose, we will use the U.S. Center for Disease Control's (CDC) MAT panel with five already available local strains obtained from the Royal Tropical Institute, Amsterdam (KIT). Once the Sri Lankan isolates are obtained, a validation of best panel suitable for Sri Lanka will be determined. MAT will be carried out in two main steps. First, the serum samples will be screened

at a titer of 1/50 using the MAT panel of 24 strains, which includes the five local strains. Positive sera from the screening test will be subjected to a run out test with the serial dilution of sera from 1/50 to 1/3200. MAT evidence for seropositivity will be defined as one or more of the following: seroconversion from negative to positive (<1/50- >1/50); 4-fold increase in titer between acute-phase and convalescent-phase (follow-up) samples; or a single titer of ≥400.

Culture isolation of *Leptospira*

For the culture isolation, we adopted the previously validated protocol published by Wuthiekanan et al. [29]. EMJH semisolid media will be prepared by adding 2.3 g of EMJH base, 1.5 g of bacteriological agar, and 100 mg of sodium pyruvate into 785 ml of distilled water, then adjusting the pH to 7.4. The media will be autoclaved and cooled to around 50ºC, Next, 100 ml of *Leptospira* enrichment media and 100 ml of fetal bovine serum will be added to the media.[30-33] To suppress the growth of possible contaminants, 5-Fluorouracil and Neomycin will be added, at concentrations of 100 µg/ml and 25 µg/ml, respectively. The media inoculated with blood will be inspected using dark field microscopy for the presence of motile *Leptospira*, initially after 3 weeks and then on a monthly basis. When a positive growth is detected, sub cultures will be transferred into liquid and semisolid media. An aliquot will then be obtained and fixed with 5% DMSO (Dimethyl Sulfoxide) and stored in -80ºC. To maintain live cultures, the isolates will be subcultured in liquid media on a bi-weekly basis and on semisolid media every 3 months. Serotyping of the newly isolated *Leptospira* strains will be carried out at the Pasteur Institute, France.

Quantitative PCR

Diagnosis

For the diagnosis, DNA will be extracted from whole blood, serum, and urine samples using the QIAamp DNA Blood Mini Kit (Qiagen, USA), according to the manufacturer's instructions. The extracted DNA will be quantified using an Invitrogen Qubit 4 Flurometer. Two previously

published quantitative PCR protocols targeting 16s ribosomal RNA and Lipl32 genes will be used for the detection of pathogenic *Leptospira* DNA based on SYBR Green Chemistry (Table 2).[34]

Table 2. Primers for the detection of *Leptospira* DNA in clinical samples

Primer/probe sequence	**Annealing temperature**	**Complementary target species**
Forward: 5'-AAG CAT TAC CGC TTG TGG TG-3' Reverse: 5'-GAA CTC CCA TTT CAG CGA TT-3'	60ºC	lipL32 gene of pathogenic *Leptospira* spp.
Forward: 5'-TAA AGG CTC ACC AAG GCG AC-3' Reverse: 5'-TTA GCC GGT GCT TTA GGC AG-3'	60ºC	16s gene of pathogenic *Leptospira* spp.

Speciation

To determine the *Leptospira* species in the positive cultures, a previously published quantitative PCR protocol based on SYBR Green Chemistry (Table 3) will be used (before sequencing).[35] This will facilitate the identification of L. *interrogans, L. borgpetersenii, L. kirschneri*, and *L. noguchii*.

Table 3. Primer pairs for detection of species

Primer/Probe sequence	**Annealing temperature**	**Complementary target species**

Forward: 5'-CTT GAG CCT GCG CGT TAY-3' Reverse: 5'-CCG ATA ATT CCA GCG AAG ATC-3'	63ºC	secY gene of *L. interrogans*
Forward: 5'-GAT TCG GGT TAC AAT TAG ACC-3' Reverse: 5'-TTG ATC TAA CCG GAC CAT AGT-3'	65ºC	ompL1 gene of *L. borgpetersenii*
Forward: 5'-CTG GCT TAA TCA ATG CTT CTG-3' Reverse: 5'-CTC TTT CGG TGA TCT GTT CC-3'	60ºC	secY gene of *L. kirschneri*
Forward: 5'-TCA GGG TGT AAG AAA GGT TC-3' Forward: 5'-CAA AAT TAA AGA AGC AAA GA-3'	63ºC	secY gene of *L. noguchii*

Molecular studies

For newly isolated *Leptospira*, whole genome sequencing will be performed using the following two technologies: MinION Nanopore sequencing in the Leptospirosis Research Laboratory (Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka, India) and PacBio third generation sequencing at the of Universiy) Medicine Genomic for Institute California, San Diego, USA).

BAS and BAX files containing all the DNA sequence information will be converted into FASTA files. Prior to assembly, the quality (coverage) of the raw data will be assessed. Using the PacBioEDA software, the quality summary of the data will be obtained. The CanuPacBiO assembly program will be used for genome assembly. The raw data of Bas/Bax files will be converted into bam files using the bax2bam software. These bam files will be aligned to our references using the Pbalingn software. The software Quiver will be used to polish the data. The assembled data will be annotated using Prokka[36] and Roary[37], and annotations will be inspected and curated using Artemis. The functional annotation will be performed using the eggNOG and eggNOG MAPPER online tools.[38] The genomes of the Sri Lankan isolates will be compared using MAUVE software. The organism will be roughly identified from the direct raw data using KRAKEN.[39]

MLST will be used for the genotyping of the new isolates. Sri Lankan isolates will be identified using three published MLST schemes for *Leptospira*.[40-42] Genetic similarity between the isolates will be determined using SPLITSTREE.

Taxonogenomics, pan-genome, and phylogenetic analyses

To establish how this strain was related to other *Leptospira* strains, we will use several genomic analysis tools. First, the 16S rRNA sequences will be used to construct a phylogeny. High sequence identities (≥97%) between the newly sequenced strain and previously identified *Leptospira* strains will be preliminarily used as suggestive of phylogenetic relationships. Accordingly, the 16S rRNA gene sequences of *Leptospira* spp. deposited in the GenBank database will be acquired. Multiple sequence alignment based on 16S rRNA will be performed using ClustalX v1.81. These results will be then used to construct the maximum likelihood tree using MEGA v5.05. To evaluate clade support, we will perform a bootstrap analysis with 1,000 replicates. The relationship of each sequenced genome to the previously described or novel *Leptospira* species will be determined according to the Overall Genetic Relatedness Indices (OGRIs), Average Nucleotide Identity (ANI), and the Average Amino Acid Identity (AAAI). The indices will be automatically calculated using two-way BLA blastn and blastp. For the ANI calculation, the default parameters shall be as follows: sequence identity cut-off, 30%; alignment cut-off, 70%; query length, 1,020 bp.

Data management and analysis

Patient interviews and patient health records will be used as data sources. On-site data entry will be performed in real-time for all the collected data. An EpiInfo database will be prepared using cross-checks, field validation, and other in-built techniques available for data quality improvement. The investigators will frequently visit the hospitals to cross-validate the data collected. All databases will be stored on password-protected computers. The main database will be stored at the

Leptospirosis Research Laboratory of the Faculty of Medicine and Allied Science, Rajarata University of Sri Lanka. Personal identification data will be removed from all the databases, such that only aliquots and predesigned patient identifiers are available. Hardcopies of the questionnaires will be archived using the standard protocols for archiving hard copies of questionnaires. Only the investigators will have access to the database. The database will be continuously updated with laboratory data using a linking patient identifier (Key variable). The lab results will be generated for all samples daily, once a fully functioning lab is set up, and the treating physicians will be informed about this over the phone. Separate laboratory registers will be maintained for the receipt of samples and the delivery of the results.

For the missing data, case deletions and pairwise deletions will be carried out, based on the amount of missing data. Proportions and 5% confidence intervals will be used to describe the confirmed cases, diversity of infecting *Leptospira*, and sequelae. A Chi-square test will be used to identify any trends and test significant differences in the *Leptospira* diversity in the different microgeographical and clinical settings. Predictors of severe disease will be analyzed using a logistic regression model. Model building will be performed using a conceptual hierarchy of determinants, which we have previously proposed for leptospirosis. An online database for *Leptospira* MLST will be used for typing analysis.

Patient and public involvement

The study was developed based on the patients' feedback regarding leptospirosis disease burden, its impact on their lives, and in particular, the problems faced during illness. In previous studies, we specifically explored patients as well as general public views on leptospirosis disease and prevention, and used this information for the design of this study. Moreover, the results of this study will be disseminated on a continuous basis for the patients. We have already conducted several public awareness and health worker education workshops based on our findings.

ETHICS AND DISSEMINATION

Undifferentiated fever patients eligible for the study will be given an explanatory statement of the study and written informed consent will be obtained before enrolment. Rapid diagnostic kits will be used on site to assist treating physicians. qPCR diagnostic facilities will be provided for all patients whenever the urgent tests are requested by treating physicians. All test results will be informed to the patients and the wards. Ethical clearance for the study was obtained from the Ethic Review Committees of Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka (ERC/2015/18).

All genomic data will be deposited in Gene Bank. Characterized serum samples will be stored in -80C in the Faculty of Medicine and Allied Sciences, Rajarata University of Sri Lanka. Data pertaining to this study will be available for researchers and policy makers on a reasonable request address to the corresponding author. After the completion of work, data will be deposited in a repository.

REFERENCES

1. WER: **Leptospirosis worldwide, 1999**. *Wkly Epidemiol Rec* 1999, **74**(29):237-242.
2. Costa F, Hagan JE, Calcagno J, Kane M, Torgerson P, Martinez-Silveira MS, Stein C, Abela-Ridder B, Ko AI: **Global Morbidity and Mortality of Leptospirosis: A Systematic Review**. *PLoS Negl Trop Dis* 2015, **9**(9):e0003898.
3. Taylor AJ, Paris DH, Newton PN: **A Systematic Review of the Mortality from Untreated Leptospirosis**. *PLoS Negl Trop Dis* 2015, **9**(6):e0003866.
4. Agampodi S, Peacock SJ, Thevanesam V: **The potential emergence of leptospirosis in Sri Lanka**. *Lancet Infect Dis* 2009, **9**(9):524-526.
5. Epidemiological Unit: **Selected notifiable diseases reported by Medical Officers of Health**. *Weekly Epidemiological Report* 2009, **36**(1):4.
6. Epidemiological Unit: **Selected notifiable diseases reported by Medical Officers of Health**. *Weekly Epidemiological Report* 2010, **37**(1):4.
7. Warnasekara JN, Agampodi S: **Leptospirosis in Sri Lanka**. *Sri Lankan Journal of Infectious Diseases* 2017, **7**(2).
8. Pappas G, Papadimitriou P, Siozopoulou V, Christou L, Akritidis N: **The globalization of leptospirosis: worldwide incidence trends**. *Int J Infect Dis* 2008, **12**(4):351-357.
9. Agampodi SB, Peacock SJ, Thevanesam V, Nugegoda DB, Smythe L, Thaipadungpanit J, Craig SB, Burns MA, Dohnt M, Boonsilp S *et al*: **Leptospirosis outbreak in Sri Lanka in 2008: lessons for assessing the global burden of disease**. *Am J Trop Med Hyg* 2011, **85**(3):471-478.
10. Agampodi SB, Dahanayaka NJ, Bandaranayaka AK, Perera M, Priyankara S, Weerawansa P, Matthias MA, Vinetz JM: **Regional differences of leptospirosis in Sri Lanka:**

observations from a flood-associated outbreak in 2011. *PLoS Negl Trop Dis* 2014, **8**(1):e2626.

11. Myers WP, Myers AP, Cox-Singh J, Lau HC, Mokuai B, Malley R: **Micro-geographic risk factors for malarial infection**. *Malaria journal* 2009, **8**:27.
12. Schellenberg JA, Newell JN, Snow RW, Mung'ala V, Marsh K, Smith PG, Hayes RJ: **An analysis of the geographical distribution of severe malaria in children in Kilifi District, Kenya**. *Int J Epidemiol* 1998, **27**(2):323-329.
13. Hassan AO, Amoo AOJ, Deji-Agboola AM, Adeleke MA, Muhibi MA, Hassan RO: **Microgeographical patterns of schistosomiasis and anthropometric indices of children in eko-ende south west, nigeria**. *Int J Biol Med Res* 2911, **2**(3):656-660.
14. Kloos H, Gazzinelli A, Van Zuyle P: **Microgeographical patterns of schistosomiasis and water contact behavior; examples from Africa and Brazil**. *Mem Inst Oswaldo Cruz* 1998, **93 Suppl 1**:37-50.
15. Baliraine FN, Afrane YA, Amenya DA, Bonizzoni M, Menge DM, Zhou G, Zhong D, Vardo-Zalik AM, Githeko AK, Yan G: **High prevalence of asymptomatic plasmodium falciparum infections in a highland area of western Kenya: a cohort study**. *J Infect Dis* 2009, **200**(1):66-74.
16. Dissanayake CB, Chandrajith R: **Medical geochemistry of tropical environments**. *Earth-Science Reviews* 1999, **47**(3–4):219-258.
17. Naotunna C, Agampodi SB, Agampodi TC: **Etiological agents causing leptospirosis in Sri Lanka: A review**. *Asian Pac J Trop Med* 2016, **9**(4):390-394.
18. Agampodi SB, Moreno AC, Vinetz JM, Matthias MA: **Utility and limitations of direct multi-locus sequence typing on qPCR-positive blood to determine infecting Leptospira strain**. *Am J Trop Med Hyg* 2013, **88**(1):184-185.
19. Agampodi SB, Miatthias MA, Moreno AC, Vinetz JM: **Utility of Quantitative PCR in Leptospirosis Diagnosis: Association of Level of Leptospiremia and Clinical Manifestations in Sri Lanka**. *Clinical Infectious Disease* 2012, **In Press**.
20. Katz AR: **Quantitative polymerase chain reaction: filling the gap for early leptospirosis diagnosis**. *Clin Infect Dis* 2012, **54**(9):1256-1258.
21. Nityananda K, Harvey T: **Leptospirosis in Ceylon-Epidemiological and Laboratory Studies**. *Ceylon J Med Sci* 1971, **20**(1):5-14.
22. Agampodi SB, Nugegoda DB, Thevanesam V: **Determinants of leptospirosis in Sri Lanka: Study Protocol**. *BMC Infect Dis* 2010, **10**(1):332.
23. Agampodi SB, Matthias MA, Moreno AC, Vinetz JM: **Utility of quantitative polymerase chain reaction in leptospirosis diagnosis: association of level of leptospiremia and clinical manifestations in Sri Lanka**. *Clin Infect Dis* 2012, **54**(9):1249-1255.
24. Reller ME, Bodinayake C, Nagahawatte A, Devasiri V, Kodikara-Arachichi W, Strouse JJ, Flom JE, Dumler JS, Woods CW: **Leptospirosis as frequent cause of acute febrile illness in southern Sri Lanka**. *Emerg Infect Dis* 2011, **17**(9):1678-1684.
25. Lucchesi PM, Arroyo GH, Etcheverria AI, Parma AE, Seijo AC: **Recommendations for the detection of Leptospira in urine by PCR**. *Rev Soc Bras Med Trop* 2004, **37**(2):131-134.
26. Bal AE, Gravekamp C, Hartskeerl RA, De Meza-Brewster J, Korver H, Terpstra WJ: **Detection of leptospires in urine by PCR for early diagnosis of leptospirosis**. *J Clin Microbiol* 1994, **32**(8):1894-1898.
27. Van Eys GJ, Gravekamp C, Gerritsen MJ, Quint W, Cornelissen MT, Schegget JT, Terpstra WJ: **Detection of leptospires in urine by polymerase chain reaction**. *J Clin Microbiol* 1989, **27**(10):2258-2262.
28. World Health Organization: **Report of the second meeting of leptospirosis burden epidemiology reference group**. In. Geneva: World Health Organization 2011: 1-34.
29. Wuthiekanun V, Chierakul W, Limmathurotsakul D, Smythe LD, Symonds ML, Dohnt MF, Slack AT, Limpaiboon R, Suputtamongkol Y, White NJ *et al*: **Optimization of**

culture of Leptospira from humans with leptospirosis. *J Clin Microbiol* 2007, **45**(4):1363-1365.
30. Wuthiekanun V, Amornchai P, Langla S, Oyuchua M, Day NP, Limmathurotsakul D: **Maintenance of leptospira species in leptospira Vanaporn Wuthiekanun agar**. *J Clin Microbiol* 2014, **52**(12):4350-4352.
31. Wuthiekanun V, Amornchai P, Langla S, White NJ, Day NP, Limmathurotsakul D, Peacock SJ: **Antimicrobial Disk Susceptibility Testing of Leptospira spp. Using Leptospira Vanaporn Wuthiekanun (LVW) Agar**. *Am J Trop Med Hyg* 2015, **93**(2):241-243.
32. Wuthiekanun V, White NJ, Amornchai P, Day NP, Limmathurotsakul D: **Quality controls for antimicrobial disk diffusion testing on Leptospira Vanaporn Wuthiekanun agar**. *Trans R Soc Trop Med Hyg* 2016, **110**(11):673-675.
33. Picardeau M, Bulach DM, Bouchier C, Zuerner RL, Zidane N, Wilson PJ, Creno S, Kuczek ES, Bommezzadri S, Davis JC *et al*: **Genome sequence of the saprophyte Leptospira biflexa provides insights into the evolution of Leptospira and the pathogenesis of leptospirosis**. *PLoS ONE* 2008, **3**(2):e1607.
34. Backstedt BT, Buyuktanir O, Lindow J, Wunder EA, Reis MG, Usmani-Brown S, Ledizet M, Ko A, Pal U: **Efficient Detection of Pathogenic Leptospires Using 16S Ribosomal RNA**. *PLOS ONE* 2015, **10**(6):e0128913-e0128913.
35. Ferreira AS, Costa P, Rocha T, Amaro A, Vieira ML, Ahmed A, Thompson G, Hartskeerl RA, Inacio J: **Direct detection and differentiation of pathogenic Leptospira species using a multi-gene targeted real time PCR approach**. *PLoS One* 2014, **9**(11):e112312.
36. Seemann T: **Prokka: rapid prokaryotic genome annotation**. *Bioinformatics* 2014, **30**(14):2068-2069.
37. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J: **Roary: rapid large-scale prokaryote pan genome analysis**. *Bioinformatics* 2015, **31**(22):3691-3693.
38. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, Bork P: **Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper**. *Mol Biol Evol* 2017, **34**(8):2115-2122.
39. Wood DE, Salzberg SL: **Kraken: ultrafast metagenomic sequence classification using exact alignments**. *Genome Biol* 2014, **15**(3):R46.
40. Varni V, Ruybal P, Lauthier JJ, Tomasini N, Brihuega B, Koval A, Caimi K: **Reassessment of MLST schemes for Leptospira spp. typing worldwide**. *Infect Genet Evol* 2014, **22**:216-222.
41. Ahmed N, Devi SM, Valverde Mde L, Vijayachari P, Machang'u RS, Ellis WA, Hartskeerl RA: **Multilocus sequence typing method for identification and genotypic classification of pathogenic Leptospira species**. *Ann Clin Microbiol Antimicrob* 2006, **5**:28.
42. Boonsilp S, Thaipadungpanit J, Amornchai P, Wuthiekanun V, Bailey MS, Holden MT, Zhang C, Jiang X, Koizumi N, Taylor K *et al*: **A single multilocus sequence typing (MLST) scheme for seven pathogenic Leptospira species**. *PLoS Negl Trop Dis* 2013, **7**(1):e1954.

AUTHORS' CONTRIBUTIONS

JVV and SBA drafted the initial protocol. YPJNW, JMDDJ and RMISKS are responsible for the detailed design, field/ laboratory work analysis and interpretation of objectives 1-3, 4 and 5 respectively. CDG, SAMK and SHS helped in design, field, hospital and laboratory work. MM

participated in scientific design and guided all laboratory components. All authors have read and approved the manuscript.

FUNDING

This project is the Sri Lankan component of a multi country study funded by the NIAID of the National Institutes of Health under award number U19AI115658. Funders had no role in the design, execution, analyses, data interpretation, or decision to submit results.

CONFLICTS OF INTERESTS

The authors declare that they have no competing interests.

WORD COUNT

3791

FIGURES

Figure 1: Locations of study sites.

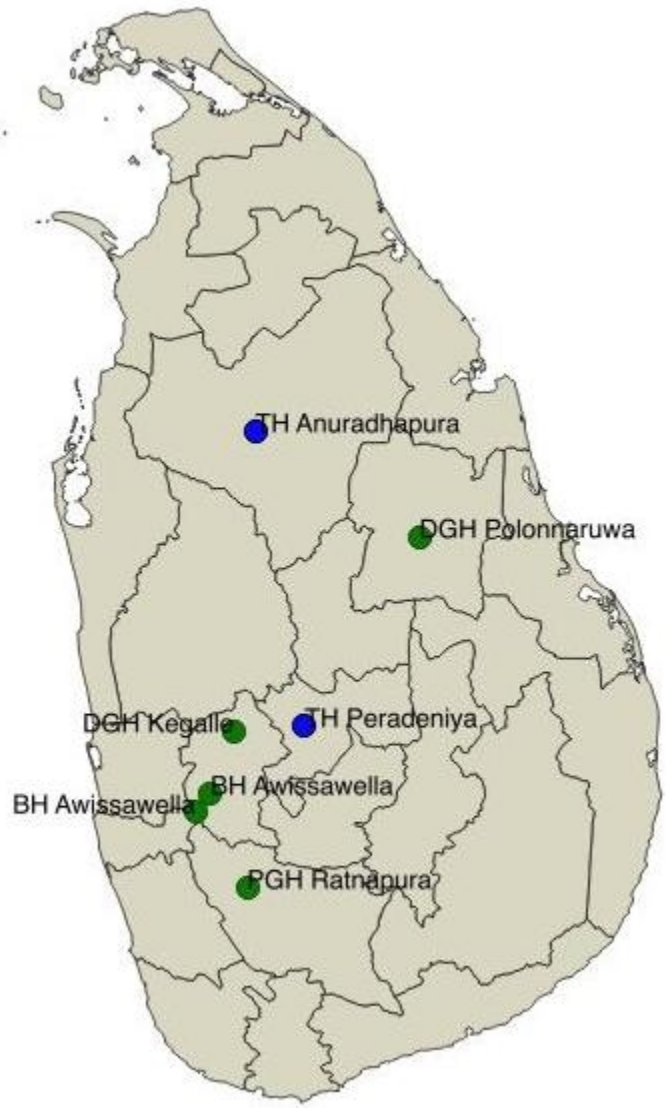


Figure 1. Study sites

147x201mm (72 x 72 DPI)