Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Molecular diagnosis of multiple bacterial zoonoses in febrile outpatients in north-eastern Kenya

**Martin Wainaina** (markimwa@gmail.com)
German Federal Institute for Risk Assessment
**Johanna F. Lindahl**
International Livestock Research Institute
**Anne Mayer-Scholl**
German Federal Institute for Risk Assessment
**Christoph-Martin Ufermann**
German Federal Institute for Risk Assessment
**Jean-Baka Domelevo Entfellner**
International Livestock Research Institute
**Uwe Roesler**
Freie Universität Berlin
**Kristina Roesel**
International Livestock Research Institute
**Delia Grace**
International Livestock Research Institute
**Bernard Bett**
International Livestock Research Institute
**Sascha Al Dahouk**
German Federal Institute for Risk Assessment

**Article**

**Keywords:**

**Posted Date:** September 14th, 2023

**DOI:** https://doi.org/10.21203/rs.3.rs-3336661/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

# Abstract

Bacterial zoonoses are important causes of non-malarial fevers in Kenya, yet their epidemiology remains unclear. We investigated brucellosis, Q-fever and leptospirosis in 216 malaria-negative febrile patients recruited in two health centres in Garissa County in north-eastern Kenya.

We determined exposure to the three zoonoses using serological and molecular testing and identified risk factors for exposure. We also used metagenomic sequencing on nine selected patients to assess the presence of other possible bacterial causes of non-malarial fevers.

High PCR positivity was found for *Brucella*and *Leptospira* spp., and high endpoint titres were observed against leptospiral serovar Grippotyphosa from the serological testing. Patients aged 5-17 years old had 4.02 (95% CI: 1.18-13.70, p-value = 0.03) and 2.42 (95% CI: 1.09- 5.34, p-value = 0.03) times higher odds of infection with *Brucella*spp. and *C. burnetii* than those of ages 35-80. The source of water was also a significant predictor for exposure to *C. burnetii*. Other bacterial infections including those caused by *Streptococcus*spp. were determined using metagenomic sequencing.

The substantial evidence for non-malarial fevers in healthcare centres in Garissa County calls for their inclusion in routine diagnosis to lower the human disease burden and ensure better health outcomes, especially in younger populations.


# Introduction

Kenya has had remarkable success in the reduction of overall morbidity and mortality, and there has been a marked increase in healthy life expectancy from 54.1 years in 1990 to 58.6 years in 2016. This is largely due to the introduction, and scaling-up from the early 2000s, of many health interventions such as the treatment and control of HIV/AIDS and malaria, child immunisations, and free maternal healthcare [1]. Despite this, communicable diseases are responsible for a large proportion of disability-adjusted life years (DALYs) in the country (DALYs are an aggregate measure of overall health loss which incorporates morbidity and mortality) [1]. Due to the large number of endemic agents that can cause these illnesses, metagenomics can play a role in detecting pathogens in a non-targeted manner in clinical samples when the clinical presentation is ambiguous [2]. Unlike traditional microbiological methods, metagenomic sequencing can detect multiple pathogens present in a clinical sample, enables the detection of unculturable agents, and can be used even after patients have started antibiotic treatment [2]. Consequently, clinical metagenomics has proven useful in several contexts [3,4]. Metagenomics allows for the profiling of the microbial community in a sample by sequencing the DNA using next-generation sequencing methods. When compared to 16S rRNA metagenomics, shotgun metagenomics can determine the bacterial profile of samples even when their 16S rRNA sequences are not well described, thereby enabling the discovery of novel species. The approach can also provide functional insights by illustrating functional genes, including those that confer antimicrobial resistance (AMR), is more sensitive and can offer resolution beyond bacterial genera and species such as strain, and identifies non-bacterial

microorganisms such as viruses and protists. This untargeted approach can be quite useful, especially in East Africa where non-malarial fevers are caused by a wide array of bacterial pathogens [5,6].

There are still huge regional disparities in health outcomes in Kenya. Due to historical inequitable allocation of resources, poverty and insecurity, Kenya's expansive arid and semi-arid (ASAL) areas have poor coverage of treatment and control programs for important infections. For instance, ownership of insecticide-treated bed nets is lowest in these counties, thereby posing serious challenges in malaria control. In addition, the high prevalence of non-malarial fevers in the region complicates malaria control because most of these illnesses have similar clinical presentations. Brucellosis, Q-fever and leptospirosis are important but often under-diagnosed non-malarial febrile illnesses in ASAL Kenya [7–9]. We therefore undertook a study in Garissa County in ASAL Kenya to estimate the prevalence of brucellosis, leptospirosis and Q-fever in febrile patients (otherwise termed proportional morbidity rates [PMr] when determined in a sick population) and to identify risk factors for exposure to these pathogens. We also used metagenomic sequencing for a non-targeted detection of other important bacterial causes of non-malarial fevers in selected patients. These findings will lead to improved management of febrile illnesses for patients with a malaria-negative test result.

## Results

# Epidemiological results

We recruited 216 febrile patients, 118 from Sangailu and 98 from Ijara health centre. Most were female (n = 150), and the median age of patients was 25 years (interquartile range [IQR]: 18–35). Most patients never attended school (n = 148), had no source of income (n = 167) and drew water from unprotected wells (n = 93). The median distance to the healthcare facility for the patients was 7km (IQR: 2–13). The highest PMr estimates observed from the serological testing were those of *C. burnetii* (45.8%, 95% confidence intervals [CI]: 39.1–52.7), and fewer patients were exposed to *Brucella* spp. (13.9%, 95% CI: 9.6–19.2) and *Leptospira* spp. (3.7%, 95% CI: 1.6–7.2). The distribution of the PMr estimates with the various demographic characteristics and putative risk factors is given in supplementary table 1.

# Risk factors for seropositivity

Results from the univariable logistic regression model for *Brucella* spp. exposure showed income source, water source, age categories, camel herding, disposal of carcasses, sleeping outside with herds, assisting with birthing livestock, disposal of aborted foetuses, and taking care of sick animals as significant predictors (p-value < 0.2). However, only age remained a significant predictor at the multivariable level (Table 1). The final logistic regression model showed that patients aged 5–17 had 4.02 (95% CI: 1.18–13.70, p-value 0.03) times higher odds of exposure than those aged 35–80 years. Patients who drew water from dams and springs had 1.09 times higher odds of exposure to *Brucella* spp. than those who did so from unprotected wells, although this observation was not statistically significant (p-value = 0.07).

Table 1
Results of the final multivariable logistic regression models to determine risk factors for exposure to the selected bacterial zoonoses. Due to the few positives observed, we excluded leptospires from these analyses.

| Pathogen | Variable | Categories | Total | OR (95% CI) | SE | Z-value | P-value |
|---|---|---|---|---|---|---|---|
| *Brucella* spp. | Age categories | 35 to 80 | 49 | Ref. | | | |
| | | 25 to 34 | 45 | 2.47 (0.71–8.62) | 0.64 | 1.42 | 0.16 |
| | | 18 to 24 | 63 | 2.22 (0.58–8.48) | 0.68 | 1.17 | 0.24 |
| | | 5 to 17 | 59 | 4.02 (1.18–13.70) | 0.63 | 2.22 | **0.03** |
| | Water source | Dams and springs | 61 | Ref. | | | |
| | | Unprotected wells | 93 | 0.43 (0.17–1.08) | 0.47 | -1.79 | 0.07 |
| | | Others * | 62 | 0.47 (0.17–1.30) | 0.52 | -1.45 | 0.15 |
| *Coxiella burnetii* | Age categories | 35 to 80 | 49 | Ref. | | | |
| | | 25 to 34 | 45 | 1.11 (0.53–2.31) | 0.38 | 0.27 | 0.79 |
| | | 18 to 24 | 63 | 1.12 (0.50–2.50) | 0.41 | 0.28 | 0.78 |
| | | 5 to 17 | 59 | 2.42 (1.09–5.34) | 0.40 | 2.18 | **0.03** |
| | Water source | Unprotected wells | 93 | Ref. | | | |
| | | Dams and springs | 61 | 2.39 (1.22–4.68) | 0.34 | 2.53 | **0.01** |
| | | Others * | 62 | 2.24 (1.15–4.36) | 0.34 | 2.36 | **0.02** |

OR: Odds ratio, CI: confidence interval, SE: standard error, Ref.: Reference category

* Category comprised protected wells, boreholes, bottled water, and water pans

The univariable logistic regression models for *C. burnetii* showed several variables significant at p-value < 0.2, but the final model demonstrated that patients aged 5–17 had 2.42 (95% CI: 1.09–5.34) times higher odds of exposure to *C. burnetii* than those aged 35–80 (p-value = 0.03). In addition, patients who sourced

water from dams/springs, and other sources (protected wells, boreholes, bottled water, and water pans), had 2.39 (95% CI: 1.22–4.68, p-value = 0.01) and 2.24 (1.15–4.35, p-value = 0.02) times higher odds of exposure than those who used unprotected wells. Results from the goodness of fit are given in the supplementary information.

# PCR results and leptospiral serovars

PCR positivity was highest for *Brucella* spp., with 19.4% (95% CI: 14.2–25.5) patients testing positive. This was followed by *Leptospira* spp., with 1.7% (95% CI: 0.4–4.9), and *C. burnetii* (0.0%, 95% CI: 0.0-1.5%). All patients testing positive for *Brucella* were also confirmed to be positive for *B. melitensis* using the species-specific assay.

According to the MAT test results, the highest agglutinating titres were observed in serovar Grippotyphosa (n = 4), with endpoint titres of 1:800, 1:800, 1:1600, and 1:12800. In addition, serovars Bataviae and Tarassovi showed agglutination in one patient and both had endpoint titres of 1:200. The remaining three patients showed agglutination for serovars Icterohaemorrhagiae (1:100), Pyrogenes (1:200), and Sejroe (1:400).

# Clinical signs and symptoms

The clinical signs and symptoms of patients positive for *Brucella* by serology and PCR were varied (Supplementary Fig. 3). However, most patients exhibited minor splenomegaly and had marginally higher median diastolic pressure and respiratory rate and lower median systolic pressure.

*Coxiella*-seropositive patients had even more diverse clinical manifestations, with minor splenomegaly, moderate headaches, severe seizures, moderate rashes, and moderate presence of blood in stool or vomit, no headaches, and higher median diastolic and marginally lower systolic pressures (Supplementary Fig. 3).

There were no consistent symptoms in *Leptospira*-positive patients, except for slight elevations of the median diastolic and systolic pressures, and respiratory rates (Supplementary Fig. 3).

# Metagenomic results

We obtained a median of 783,410 (IQR: 732,804–947,650) pair-end reads using the metagenomic sequencing after quality trimming (Supplementary table 2). The peak insert size was 149 bases, indicating possible DNA sample degradation which is likely during sample transportation from the remote study sites, and long storage of the archived samples, leading to shorter average fragment lengths. Various bacteria were detected, most from the phyla Firmicutes, Proteobacteria and Tenericutes (Fig. 1). The NEC had the lowest alpha diversity scores (observed, Chao1 and Shannon), and samples showed varying alpha diversities, with patient 142 having the lowest and patient 3 having the highest observed score (Fig. 2). There were also reads from the top twenty commonest causes of non-malarial fevers identified, with the most abundant from the metagenomic analyses being *Staphylococcus* and *Klebsiella* spp (Fig. 3). We confirmed 500 random reads taxonomically assigned to the top 20 agents using BLASTn

to increase diagnostic specificity. Agents with fewer than 500 had their entire taxonomically assigned reads blasted for confirmation. The most confirmed reads were those of *Streptococcus* spp., with no reads from these 20 agents being detected in the negative control (Fig. 4). AMR genes were found in the following patients; patient 110 (*lmo0919_fam*, coverage = 38.3, depth = 0.9), patient 203 (*qacC*, coverage = 38.3, depth = 1.6), patient 2 (*fosX*, coverage = 43.5, depth = 1.71), patient 3 (*fosX*, coverage = 52.2, depth = 1.0), and patient 205 (*fosX*, coverage = 40.1, depth = 1.6). There were no virulence factors detected.

# Discussion

We investigated bacterial zoonoses (*Brucella*, *C. burnetii*, and *Leptospira*) in 216 febrile patients who visited two health centres in Garissa County using both targeted and untargeted metagenomic detection methods. High PCR positivity was observed for *Brucella* and *Leptospira* spp. High endpoint titres were observed against leptospiral serovar Grippotyphosa. Young patients had higher odds of seropositivity to *Brucella* spp. and *C. burnetii* than their older counterparts. Lastly, AMR genes and other bacteria such as *Streptococcus* were found using metagenomic sequencing. Results of this survey reveal multiple causes of non-malarial fevers in Garissa County. This calls for their inclusion in routine diagnosis of non-malarial fevers to lower disease burden.

Brucellosis is an important cause of non-malarial fevers in Kenya and prevalence at the community level is high. The PMr in our study of 13.9% was comparable to other studies of febrile patients in the region [10]. The prevalence of animal brucellosis is considerably lower than that of humans in the area [11], and households with one seropositive animal have been shown to have higher odds of having a seropositive human [12]. The monitoring of animal brucellosis through One Health surveillance and control strategies is therefore necessary for lowering the human brucellosis prevalence because exposure is usually from livestock and animal products.

A previous study on fever patients in the region showed 19.1% exposure by IgG antibodies against *C. burnetii*, an estimate much lower than that in our study [13]. However, cross-reactivity can occur when patients have been exposed to *Bartonella* spp., *Legionella* spp., or *Chlamydia* spp. [14,15], thereby overestimating the seroprevalence figures.

Higher leptospirosis PMr has been observed in febrile patients in Kenya [16], and 3/12 (25%) patients were *Leptospira* IgM positive by ELISA in a 2005 outbreak of acute febrile illness in the region [7]. The exposure of 3.7% of fever patients therefore demonstrates that leptospirosis should be considered in the differential diagnosis of acute non-malarial fevers in the region.

Our study also determined that younger patients (5–17 years) had higher seroprevalence estimates and were at higher odds of being exposed to *Brucella* and *C. burnetii* than their older counterparts between 35 and 80 years. This may be because younger people come into contact with infected animal hosts more regularly than those aged 35–80, through any activities in the production, slaughter, processing, and retail of animal source foods. These professional activities and consumption habits also have gender-specific considerations which may play a role in this risk factor [17]. Healthcare-seeking behaviour is also higher in

younger people with febrile illnesses in Kenya than in older patients [18], and our observation could therefore have arisen from selection bias. However, community exposure in the country is higher in older patients [12].

Patients that sourced water from unprotected wells had lower odds of exposure to *C. burnetii*, and tended towards significantly lower odds of exposure to brucellae than those getting water from dams, springs, and other sources. Even though exposure to *C. burnetii* through contaminated environmental water has been documented, it is an uncommon transmission pathway and in need of further elucidation [19]. Contamination of water sources with animal waste can lead to the spread of *C. burnetii*[20]. Brucellae can contaminate water sources when animal waste pollutes water sources, and the bacteria can survive for long periods in water, thereby also posing a risk of human exposure [21,22]. Therefore, the protection of water sources from possible contamination is vital in reducing human exposure to zoonoses.

Our study also showed a seroprevalence of 3.7% for leptospires in fever patients. This is lower than estimates observed in fever patients in neighbouring Tanzania [23]. Serovars Pyrogenes, Sejroe, Bataviae, Tarassovi and Icterohaemorrhagiae were found in the patients, and Grippotyphosa showed particularly high endpoint titres. Even though high titres against serovars other than the infective one (paradoxical reactions) can happen in acute leptospirosis, these are likely important serovars responsible for human leptospirosis in Garissa and MAT panels should include them.

Reads for *Streptococcus* were confirmed by BLASTn analyses in the metagenomic sequencing and these were also not observed in the negative control used, implying this could be a cause of non-malarial fevers in some patients. The metagenomic datasets generated have the potential to reveal AMR genes which are clinically relevant and cannot be obtained with 16S rRNA amplicon sequencing. Results from shotgun metagenomic experiments can be confounded by external contaminants at various stages of the workflows. These can originate from extraction and library prep reagents [24] and various bacteria have been identified as common contaminants [25,26]. The effect of these contaminants is more pronounced in low biomass samples such as serum, as the sample is inundated by contaminating nucleic acid due to the low amount of starting material from the patient, consequently generating misrepresentative results [25]. In addition, index hopping and cross-contamination by carry-over of amplicons from previous runs are sources of internal contamination [24] and can bias metagenomics studies. Even though some reads were found in our NEC, they could not be assigned to the top 20 leading fever agents, validating our results as free from internal contamination. We also extracted circulating cell-free DNA which according to the manufacturer may not be detected with spectrophotometric methods such as the Qubit™ assay. There is also growing consensus that a blood biome exists for healthy individuals and may be constituted largely by bacteria speculated to be transient and from sources such as the gastrointestinal tract, skin and oral microbiomes. The putative blood biome is largely constituted by *Proteobacteria*, *Actinobacteria*, *Firmicutes*, and *Bacteroidetes*[27] which were also the largest bacterial phyla found in this study. Therefore, the identification of these bacteria does not offer definitive proof that they are the causative agents of non-malarial fevers. Lastly, a few important AMR genes were determined in some patients. The gene

coverage and coverage depth for most was not high, and these results could have originated from low abundance [28]. We therefore cannot establish whether these genes were clinically relevant. However, studies utilising higher sequencing depth could add value in monitoring AMR in fever patients using metagenomic sequencing.

Clinical signs and symptoms for the different pathogens were not distinctive, with a selected few being associated with positive patients using both serology and PCR tests. Future studies making use of the gold standard diagnostic tests can therefore determine the positive predictive values of these clinical signs and symptoms to aid physicians in the region in the diagnosis of these non-malarial fevers. Most of the *Brucella* isolated from humans in Kenya has been *B. melitensis*[9], as is the case in other countries [29], a likely result of *B. melitensis* being more prevalent than *B. abortus*. *B. melitensis* is often considered more pathogenic than *B. abortus* due to its higher association with human brucellosis. However, human brucellosis resulting from *B. abortus* can be equally severe [30], and preventative measures that lower exposure from cattle, sheep and goats should be adopted in the country.

Our study had several limitations. We utilised shallow metagenomic sequencing due to limited resources which may lead to decreased sensitivity in detecting low concentrations of circulating DNA in clinical samples. Our study also had a relatively small sample size of inpatients and this can lead to failure to detect significant effects of putative risk factors in causing disease exposure.

In conclusion, Garissa County is an ASAL area with a high poverty index. As all patients recruited in this study were malaria-negative, investments in the diagnosis and treatment of zoonotic diseases are key in lowering disease burden, improving health outcomes and increasing the productivity of human capital. Prevention of exposure through occupational risks or consumption of contaminated foods, especially to younger individuals, should be considered to lower the high exposures observed. The presence of multiple zoonotic pathogens in febrile patients presenting in local health centres shows the urgent need for surveillance and control programs in the county. Using non-targeted sequencing approaches can add value in detecting uncommon causes of febrile illnesses, especially when recently collected samples are obtained and appropriate contamination controls are included in the sequencing process. Lowering the burden of these diseases in animal hosts is cost-effective. Investments in the local health centres for routine screening of these zoonoses, and sensitisation of medical professionals on the importance of non-malarial fevers in the region should also be carried out to ensure proper use of anti-malarial and antibiotic therapies. Protection of water sources from animal waste should also be encouraged to minimise zoonotic disease exposure.

## Methods

# Study patients and inclusion criteria

Patients seeking treatment for febrile illnesses were recruited and blood samples were obtained according to a protocol approved by the Ethics and Scientific Review Committee of the Africa Medical

and Research Foundation (Approval number REF: AMREF-ESRC P65/2013). Only participants above five years old were recruited and informed consent was sought from all participants (or their guardians for minors) and all methods in this study done according to the approved ethical guidelines and regulations. We collected blood samples from 216 febrile patients from Ijara and Sangailu health centres in Garissa County of Kenya (Supplementary Fig. 1) using systematic random sampling from March to September 2014. Patients were sampled from each health centre using sampling fractions estimated from the relative number of patients served at each centre on an average day. Within a health centre, the first patient was selected randomly, followed by sampling the $n^{th}$ individual in the queue depending on the average number of fever cases a facility handles in a day. Pretested questionnaires uploaded to smartphones using the open data kit tool were administered to the participants in Somali and Swahili by clinical officers as they awaited services. The questionnaire data collected included: location, gender, occupation, age, and level of education. The potential risk factors for exposure to important and common non-malarial febrile illnesses were derived from the literature and covered topics such as the source of water used, livestock ownership, contact with livestock, and ease of access to public health services. We considered only malaria-negative patients to avoid confounding results with malarial infections which have a similar clinical presentation.

## Laboratory analyses

## Serological and PCR testing

Blood in plain vacutainers was centrifuged at 1500g for 20 minutes to obtain serum. Testing for malaria was done using the CareStart™ Malaria Pf (HRP2) Ag rapid diagnostic test (Access Bio, Inc.) according to the manufacturer's instructions. Serological testing for the three bacterial infections was performed as follows: the Rose Bengal test (Pourquier® Rose Bengale Ag, IDEXX) for *Brucella* spp., the *Coxiella burnetii* Phase 2 IgG and IgM ELISA (Institut Virion\Serion GmbH) for *C. burnetii*, and lastly, the microscopic agglutination test (MAT) for *Leptospira* spp. We included the following 12 leptospiral reference strains which are relevant in Kenya: Ballico, Akiyami A, Mus 127, Swart, Hond Utrecht IV, Moskva V, RGA, Hebdomadis, Pomona, Salinem, M 84, and Perepelitsin. Samples with MAT titres > 1:100 were regarded as positive and were further tested to 1:12800 to determine endpoint titres.

DNA was isolated from heparinised blood and tested using PCR as previously done in a related study [11]. This involved testing for *Brucella* spp. by targeting both the *bcsp31* and IS*711* loci, and further testing positive samples for *B. melitensis* and *B. abortus* with the species-specific IS*711* assays. PCRs for pathogenic leptospires and *C. burnetii* targeted *lipL32* and IS*1111*, respectively.

## Next-generation sequencing

We randomly selected a subset of nine patient samples for these analyses. Circulating cell-free DNA was isolated from serum samples using the QIAamp minElute ccfDNA mini kit (Qiagen), and included nuclease-free water during DNA extraction as a negative extraction control (NEC). Sequence libraries for

the DNA were prepared using the Nextera DNA flex library prep kit (Illumina) and sequencing was performed on an Illumina NextSeq 500 platform with a 150-cycle paired-end configuration.

# Bioinformatic analyses

Trimming and quality checks of sequence reads were done with *fastp* version 0.20.1, and a mean phred score of 30 or above (Q30) was used to filter low-quality reads. Human reads were further filtered out from trimmed reads using *bowtie2* version 2.3.5.1. Taxonomic identification and abundance estimation was performed with *kraken2* version 2.1.2, and later re-estimated with *bracken* version 2.5, using the RefSeq release 95 database. Pathogenic species were filtered as per the guidelines of the American Biological Safety Association (www.absa.org), and a subset not exceeding 500 reads taxonomically assigned to these pathogenic species was extracted and confirmed using BLASTn as detailed by Gruetzke et al. [28]. We reported any agents comprising the twenty leading causes of non-malarial fevers in eastern Africa as determined by Wainaina et al. [6]. We determined the presence of AMR genes and virulence factors by SRST2 version 0.2.0 (minimum coverage = 30 % and minimum depth = 1) using the Virulence Factor Database (VFDB set A_nt downloaded on March 22, 2018) and National Center for Biotechnology Information (NCBI) AMRfinder database version 2019-10-30.1, respectively [28]. Plots were created in R statistical environment using *phyloseq* version 1.38.0, *patchwork* version 1.1.2, *pheatmap* version 1.0.12 and *ggplot2* version 3.4.0. Raw dehumanised reads generated by next-generation sequencing have been uploaded to the European Nucleotide Archive (ENA) under project accession PRJEB65963 (https://www.ebi.ac.uk/ena/browser/view/PRJEB65963).

# Epidemiological analyses

Data analysis was done on R. We initially determined the distribution of PMr estimates for *Brucella* spp., *C. burnetii* and *Leptospira* spp. with the patient metadata. Univariable logistic regression models were fitted with healthcare facilities as random effects. The variables with $p < 0.2$ were included in the multivariable models and variables were removed using backward elimination until there was no evidence of confounding. Significant association was deemed when $p < 0.05$. The generalized linear mixed models were fitted using *lme4* version 1.1–31. Leptospires were excluded from these analyses due to the few positives observed. We also determined the distribution of clinical signs and symptoms of patients that were sero- and PCR-positive. Further details on these analyses are given in the supplementary information.

# Declarations

## Acknowledgements

We thank Anne Stephan and Enno Luge for their excellent laboratory assistance. We also thank Matthias Flor for his assistance in formatting the questionnaire data. We lastly thank Josephine Gruetzke for her invaluable input in the metagenomic analyses and Bernd-Alois Tenhagen for going through the manuscript.

### Author contributions

Conceptualization (D.G., B.B., S.A.D.). Methodology (M.W., J.F.L., A.M.S., C.M.U., J-B.D.E., B.B., S.A.D.). Validation (M.W., J-B.D.E.). Formal Analysis (M.W.). Investigation (M.W.). Resources (D.G., B.B). Writing-Original Draft (M.W.). Writing- Review & Editing (all authors). Visualization (M.W.). Supervision (S.A.D.). Project Administration (S.A.D.). Funding Acquisition (M.W., J.F.L., K.R., U.R., B.B., S.A.D.).

### Data availability statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

### Competing interests

The authors declare no conflicts of interest.

### Funding source

We acknowledge funding from the CGIAR Research Program on Agriculture for Nutrition and Health (A4NH), led by the International Food Policy Research Institute (IFPRI) and the German Federal Institute for Risk Assessment (Project No.: 1322-709). This work is part of the CGIAR initiative on One Health supported by contributors to the consortium for International Agricultural Research (CGIAR) Trust Fund (https://www.cgiar.org/funders/). The German Academic Exchange Service (DAAD) supported M.W.'s research fellowship (Ref. No.: 91693043).

### Ethical Approval statement

Ethical approval was given by the Ethics and Scientific Review Committee of the Africa Medical and Research Foundation (Approval number REF: AMREF-ESRC P65/2013). Only participants above five years old were recruited and informed consent was sought from all participants (or their guardians for minors).

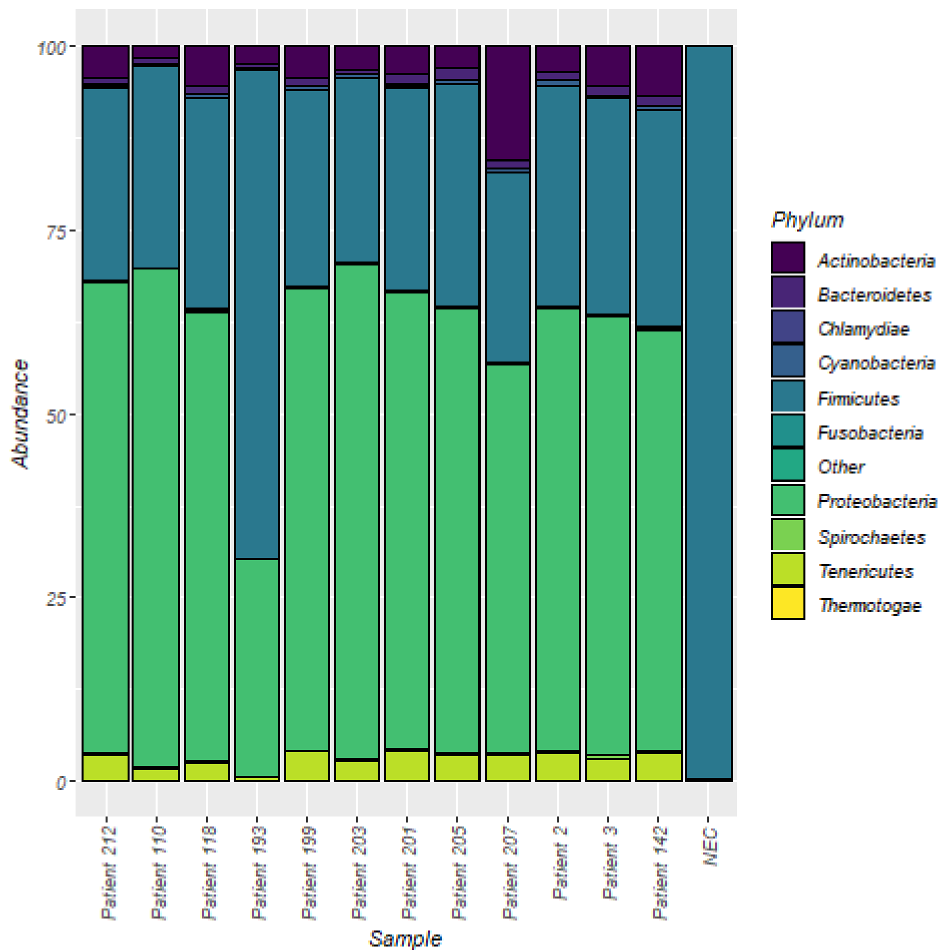
## References

1. Achoki, T. *et al.* Health disparities across the counties of Kenya and implications for policy makers, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Global Health* **7**, e81-e95, doi:10.1016/S2214-109X(18)30472-8 (2019).
2. Chiu, C. Y. & Miller, S. A. Clinical metagenomics. *Nat. Rev. Genet.* **20**, 341-355, doi:10.1038/s41576-019-0113-7 (2019).
3. Somasekar, S. *et al.* Viral surveillance in serum samples from patients with acute liver failure by metagenomic next-generation sequencing. *Clin. Infect. Dis.* **65**, 1477-1485, doi:10.1093/cid/cix596 (2017).
4. Wilson, M. R. *et al.* Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *N. Engl. J. Med.* **370**, 2408-2417, doi:10.1056/NEJMoa1401268 (2014).

5. Aguiar-Pulido, V. *et al.* Metagenomics, metatranscriptomics, and metabolomics approaches for microbiome analysis. *Evol. Bioinform. Online* **12**, 5-16, doi:10.4137/ebo.S36436 (2016).
6. Wainaina, M. *et al.* A systematic review and meta-analysis of the aetiological agents of non-malarial febrile illnesses in Africa. *PLoS Negl. Trop. Dis.* **16**, e0010144, doi:10.1371/journal.pntd.0010144 (2022).
7. Ari, M. D. *et al.* Challenges of establishing the correct diagnosis of outbreaks of acute febrile illnesses in Africa: the case of a likely *Brucella* outbreak among nomadic pastoralists, northeast Kenya, March-July 2005. *Am. J. Trop. Med. Hyg.* **85**, 909-912, doi:10.4269/ajtmh.2011.11-0030 (2011).
8. Njeru, J., Henning, K., Pletz, M. W., Heller, R. & Neubauer, H. Q fever is an old and neglected zoonotic disease in Kenya: a systematic review. *BMC Public Health* **16**, 297, doi:10.1186/s12889-016-2929-9 (2016).
9. Njeru, J. *et al.* Systematic review of brucellosis in Kenya: disease frequency in humans and animals and risk factors for human infection. *BMC Public Health* **16**, 853, doi:10.1186/s12889-016-3532-9 (2016).
10. Njeru, J. *et al.* Human brucellosis in febrile patients seeking treatment at remote hospitals, northeastern Kenya, 2014–2015. *Emerg. Infect. Dis.* **22**, 2160 (2016).
11. Wainaina, M. *et al.* Longitudinal study of selected bacterial zoonoses in small ruminants in Tana River County, Kenya. *Microorganisms* **10**, 1546 (2022).
12. Kairu-Wanyoike, S. *et al.* Positive association between *Brucella* spp. seroprevalences in livestock and humans from a cross-sectional study in Garissa and Tana River Counties, Kenya. *PLoS Negl. Trop. Dis.* **13**, e0007506 (2019).
13. Njeru, J. *et al.* Febrile patients admitted to remote hospitals in Northeastern Kenya: seroprevalence, risk factors and a clinical prediction tool for Q-Fever. *BMC Infect. Dis.* **16**, 1-15 (2016).
14. Musso, D. & Raoult, D. Serological cross-reactions between *Coxiella burnetii* and *Legionella micdadei*. *Clin. Diagn. Lab. Immunol.* **4**, 208-212, doi:10.1128/cdli.4.2.208-212.1997 (1997).
15. Maurin, M., Eb, F., Etienne, J. & Raoult, D. Serological cross-reactions between *Bartonella* and *Chlamydia* species: implications for diagnosis. *J. Clin. Microbiol.* **35**, 2283-2287 (1997).
16. de Vries, S. G. *et al.* Leptospirosis in Sub-Saharan Africa: a systematic review. *Int. J. Infect. Dis.* **28**, 47-64, doi:https://doi.org/10.1016/j.ijid.2014.06.013 (2014).
17. Grace, D., Roesel, K., Kang'ethe, E., Bonfoh, B. & Theis, S. *Gender roles and food safety in 20 informal livestock and fish value chains*. Vol. 1489 (Intl Food Policy Res Inst, 2015).
18. Emukule, G. O. *et al.* Healthcare-seeking behavior for respiratory illnesses in Kenya: implications for burden of disease estimation. *BMC Public Health* **23**, 1-13 (2023).
19. D'Ugo, E. *et al.* Detection of Coxiella burnetii in urban river water. *Vector Borne Zoonotic Dis.* **17**, 514-516, doi:10.1089/vbz.2017.2107 (2017).
20. Shabbir, M. Z. *et al.* Evidence of *Coxiella burnetii* in Punjab province, Pakistan. *Acta Trop.* **163**, 61-69, doi:https://doi.org/10.1016/j.actatropica.2016.07.017 (2016).

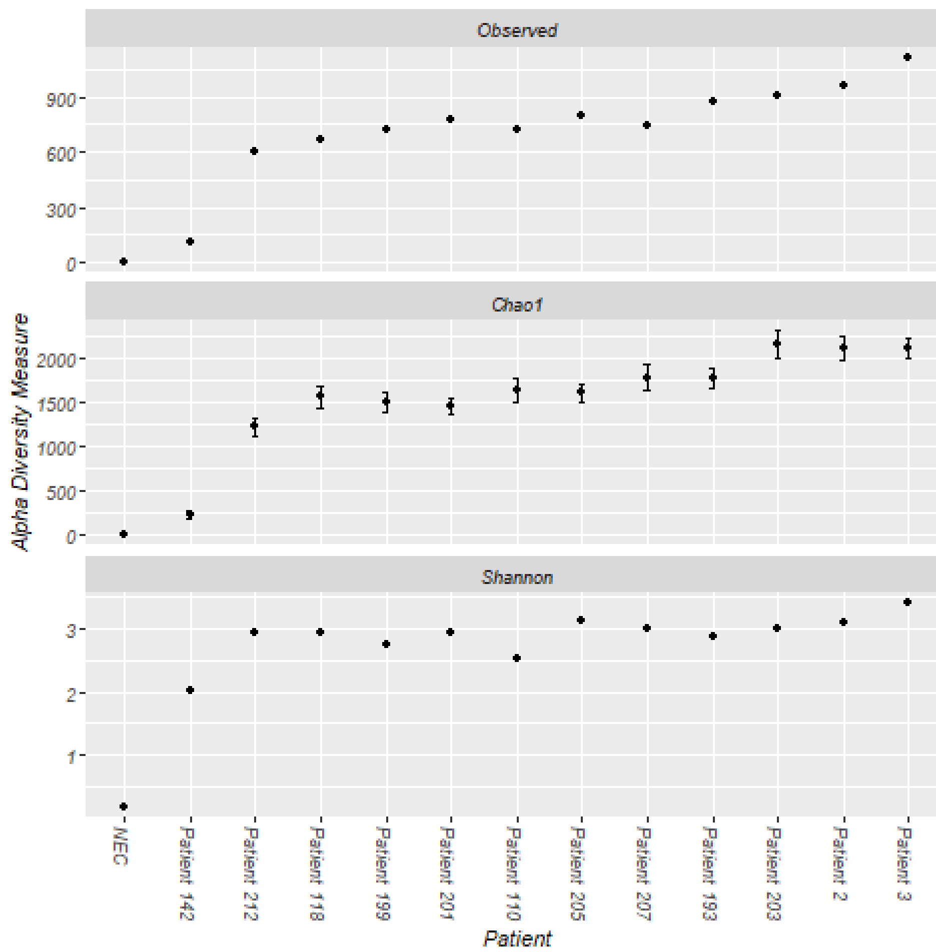
21. Kaden, R. *et al. Brucella abortus*: determination of survival times and evaluation of methods for detection in several matrices. *BMC Infect. Dis.* **18**, 259, doi:10.1186/s12879-018-3134-5 (2018).
22. El-Tras, W. F., Tayel, A. A., Eltholth, M. M. & Guitian, J. *Brucella* infection in fresh water fish: Evidence for natural infection of Nile catfish, *Clarias gariepinus*, with *Brucella melitensis*. *Vet. Microbiol.* **141**, 321-325, doi:https://doi.org/10.1016/j.vetmic.2009.09.017 (2010).
23. Motto, S. K., Shirima, G. M., de Clare Bronsvoort, B. M. & Cook, E. A. J. Epidemiology of leptospirosis in Tanzania: A review of the current status, serogroup diversity and reservoirs. *PLoS Negl. Trop. Dis.* **15**, e0009918, doi:10.1371/journal.pntd.0009918 (2021).
24. Jurasz, H., Pawłowski, T. & Perlejewski, K. Contamination issue in viral metagenomics: Problems, solutions, and clinical perspectives. *Front. Microbiol.* **12**, 745076, doi:10.3389/fmicb.2021.745076 (2021).
25. Salter, S. J. *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 87, doi:10.1186/s12915-014-0087-z (2014).
26. Glassing, A., Dowd, S. E., Galandiuk, S., Davis, B. & Chiodini, R. J. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathog.* **8**, 24, doi:10.1186/s13099-016-0103-7 (2016).
27. Castillo, D. J., Rifkin, R. F., Cowan, D. A. & Potgieter, M. The healthy human blood microbiome: fact or fiction? *Frontiers in Cellular and Infection Microbiology* **9**, doi:10.3389/fcimb.2019.00148 (2019).
28. Grützke, J. *et al.* Fishing in the soup - Pathogen detection in food safety using metabarcoding and metagenomic sequencing. *Front. Microbiol.* **10**, 1805, doi:10.3389/fmicb.2019.01805 (2019).
29. Jiang, W. *et al.* Epidemiological characteristics, clinical manifestations and laboratory findings in 850 patients with brucellosis in Heilongjiang Province, China. *BMC Infect. Dis.* **19**, 1-6 (2019).
30. Dokuzoğuz, B. *et al.* Characteristics of *B. melitensis* versus *B. abortus* bacteraemias. *J. Infect.* **50**, 41-45 (2005).

# Figures




## Figure 1

Comparison of bacterial community compositions in patient samples based on metagenomic sequencing. The relative abundance of each bacterial phylum is shown, as identified using kraken2 and further re-estimated using a Bayesian approach with bracken. The data are presented as percentages of the total bacterial community in each sample and “Other” comprised bacteria that had less than 5% relative abundance.




## Figure 2

Alpha diversity scores of the metagenomic profiles, calculated using the observed, Chao1, and Shannon indices. These indices represent a measure of species richness and evenness in the datasets, which were obtained through shotgun metagenomic sequencing of patient sera.

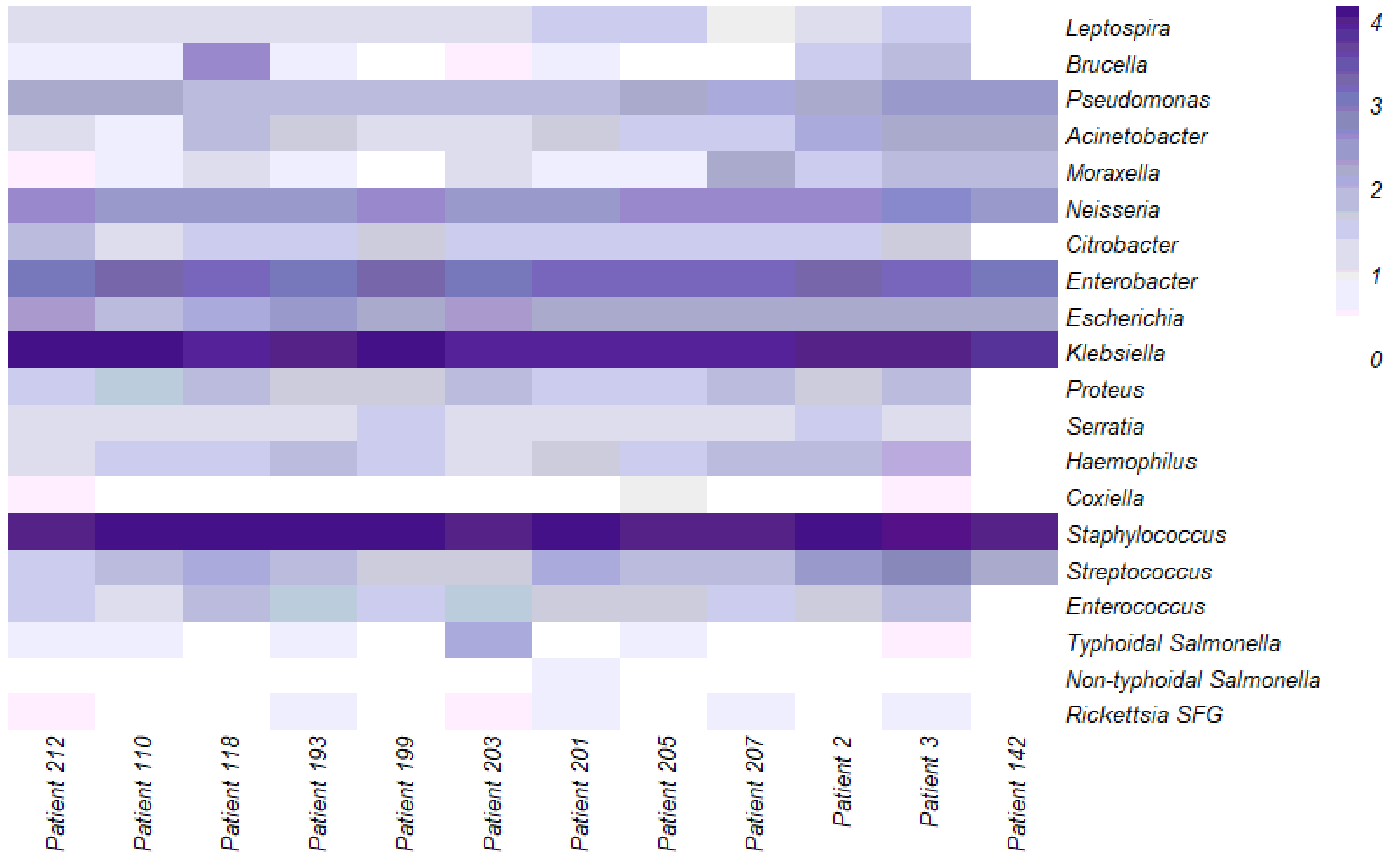



**Figure 3**

Classified genus- and species-specific reads per million total reads in the metagenomic analyses. To enhance the visualisation and avoid computing $\log_{10}$ of zero, a pseudo count of 1 was added to the original data and the transformed values were expressed as $\log_{10}$. Therefore, 4 on the heat map scale should be interpreted as 3 in the original data for instance (or 1,000 reads per 1,000,000 total reads).

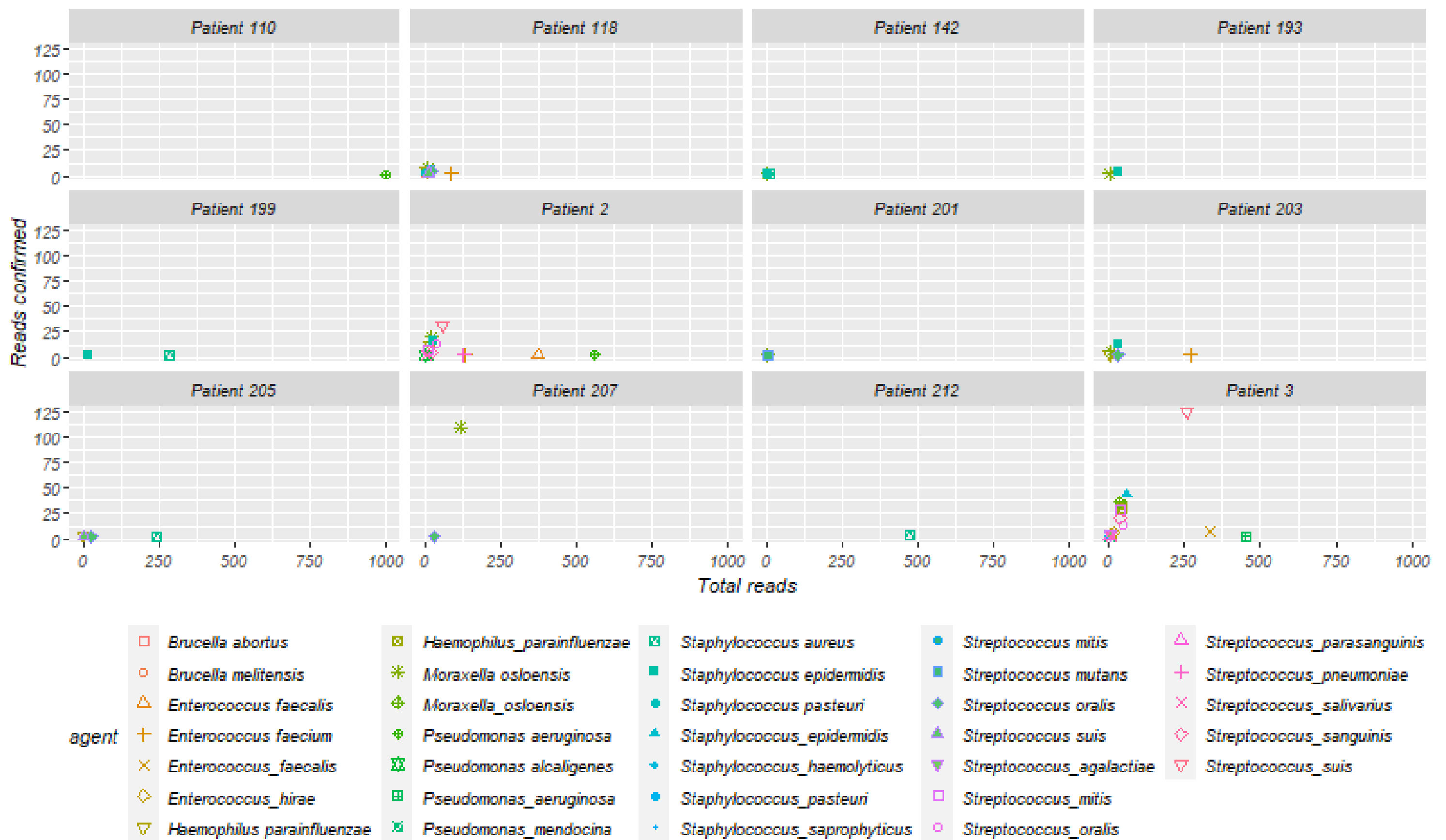



**Figure 4**

A scatter plot summarizing the number of BLASTn-confirmed reads for the top twenty fever-causing agents in the metagenomic datasets. Each dot represents one agent and its coordinates are the number of reads confirmed against the number of total reads BLAST'ed. The x-axis is capped at 1000 reads to show only the agents that reached the cap, while the y-axis shows the number of confirmed reads. The plot suggests that many agents were confirmed by a low number of reads, and highlights the agents that were confirmed by the most reads.

# Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryinformationscientificreports.pdf