DAVIS EXPEDITION FUND

REPORT ON EXPEDITION / PROJECT

Expedition/Project Title: **"Old Books, New Data: Fever Rhythms in Historical Human Case Studies"**

Travel Dates: N/A

Location: N/A desk-based

Group Members: N/A

Aims: This project will exploit a rich source of unexamined historical case study data held by the Wellcome Trust Library, to generate and interrogate ecological hypotheses about the causes and consequences of biological rhythms in malaria infections. This project has three stepwise, and specific, outcomes:

1) Extraction: Digitally curate data from hundreds of hand-drawn fever plots in historical texts, using image processing software (ImageJ) developed by the National Institutes of Health.
2) Manipulation: Generate batch and individual data to useable formats via a bespoke coding pipeline, and extract circadian parameters (e.g., the rhythmic characteristics of fever) using circadian processing packages implemented in the statistical programming software R.
3) Statistical Analysis: Perform robust statistical analysis of rhythmicity and metadata correlates from which we will test existing hypotheses and derive new hypotheses to explain the ecological and evolutionary causes and consequences of individual variability in rhythmic infection parameters.

Photography consent form attached: (*please refer to your award letter)* ☒ Yes ☐ No

Outcome (a minimum of 500 words):-

During the period of this work, I was aware that I would need to work independently, apply myself to new methods, and be willing to re-evaluate my aims as I explored these uncharacterised resources.

Using a bespoke pipeline I developed as part of my Undergraduate Thesis, I was able to transform hundreds of hand-drawn clinical charts from over a century ago into digital and testable data. This work is more than quantitative, or code-based: Every historical resource in any archive has its own natural history. Many texts, especially those published by health commissions and surveys, like this exist worldwide. Clinical details from these texts contain valuable data which can go unused when it cannot be digitised, or when it is assumed too "old" for current investigations. I would like to challenge this notion. The biological patterns seen in some "old" records are not influenced by new drug therapies, by vaccination or prophylaxis – which can be barriers to current research in humans.

Introduction

An emerging topic in infection biology concerns the interactions between biological rhythms, disease severity, and pathogen activities (Westwood *et al.*, 2019). Humans, like all mammals, have circadian rhythms (from the Latin *circa* meaning *about; dies* meaning *day*) that schedule their behaviours and physiologies across 24-hr periods using molecular "clocks" (Buhr and Takahashi, 2013). These rhythms have been known to create opportunities and pressures for pathogens - for example, circadian clocks that control the production of surface proteins dictate the time-of-day that SARS-CoV-2 can invade and replicate (Borrmann, McKeating and Zhuang, 2021; Sengupta *et al.*, 2021). And there is increasing evidence that disruption of these rhythms could be a risk factor for, or result of, disease – such as findings that Hepatitis B virus proteins can disrupt normal, 24-hour rhythms in the expression of some genes, a process linked to the development of liver cancer (Yang *et al.*, 2014).

Temperature and pulse are two of the most tightly controlled rhythms in the human body, and both vary over 24-hr periods (Refinetti and Menaker, 1992). Fever, a rise of the body's temperature "set-point" and the loss of circadian regulation, is a common symptom of parasitic, bacterial, and viral infections, and is well-studied and easily recorded even in low-resource settings (Mackowiak, 1998; Ogoina, 2011; Sajadi *et al.*, 2012). Disruptions of circadian pulse changes are similarly associated with infection and correlated with disease severity (Ahmad *et al.*, 2009; Ogoina, 2011). These perturbations of the internal environment can be indicative of infection dynamics, pathology, and prognosis, and are modulated by host (e.g. age, sex) and pathogen (e.g. species, replication) factors (Elliot, Blanford and Thomas, 2002; Plaza *et al.*, 2016).

Yellow fever (YF) is a vector-borne disease caused by a flavivirus; Although there is a licensed vaccine, ~ 200,000 cases occur each year, approximately 90% of which are in Africa (Monath and Barrett, 2003; Gardner and Ryman, 2010). There is no specific treatment for YF, and 30,000 deaths are reported each year, but many argue that the burden is underestimated and that annual fatalities may number ~51,000 (Monath, 2008; Gardner and Ryman, 2010).

Most cases, especially in those previously exposed, are asymptomatic, but infections can progress to fatal disease - which manifests as Viral Haemorrhagic Fever (VHF), a syndrome also caused by Ebola and Marburg (Feldmann and Geisbert, 2011; Abdallah El Bouzedi, 2015). Beyond fever, patients with severe YF can experience abnormal variations in heart rate, renal failure, internal haemorrhage, and coma (Barnett, 2007). Although YF has long been a "model" for the study of VHF, advances in YF detection and supportive therapy preclude the study of rhythms associated with the syndrome, and the introduction of a vaccine may have dampened interests.

However, some historical records of YF circumvent these barriers.

YF became a disease of interest in the West when the colonisation of endemic regions by European and American citizens, and the slave trade, resulted in YF epidemics in naïve populations (McCarthy, 2001; Güereña-Burgueño, 2002). The Yellow Fever Commission was established by the US government in 1900 to report on YF in South America and (West) Africa (Güereña-Burgueño, 2002). Many of the reports exist today and contain hundreds of clinical datasets and patient records, with some being held in archives at the Wellcome Trust Library (WTL).

Figure 1: An image taken at the West African Fever Commission laboratory, 1928. Source: Rockefeller

These resources provide a rare opportunity to ask questions about the drivers of disease severity and explain variability between patients. It should be noted that this project is an exercise: the texts are yet uncharacterised and will present their own challenges. I will also provide a key proof of principle: that these resources, and others like them, can be transformed into informative data yielding new insights, using a protocol we will develop.

Aims

Preliminary Aims – Incorporated within "Methods"

- Evaluate the quality and contents of texts on YF held by the WTL.
- Characterise the natural history of these texts.
- Modify an existing pipeline developed during my thesis for application to the texts.

Central Aims

1) Extraction: Digitally curate data of temperature and pulse from hundreds of hand-drawn fever plots in historical texts, using image processing software (ImageJ; developed by the National Institutes of Health). Gather metadata from each case.

2) Manipulation: Transform batch and individual data to useable formats via a bespoke coding pipeline (see above), and extract circadian parameters using rhythm processing packages in the statistical programming software R.

3) Statistical Analysis: Perform robust statistical analysis of rhythmicity and metadata correlates, from which we could explore existing hypotheses or derive new hypotheses about the interacting rhythms in host-pathogen relationships

Materials

Several reports from the Yellow Fever Commission (West African Division) are owned by the Wellcome Trust Library, in conjunction with the London School of Hygiene and Tropical Medicine. Three volumes, containing reports gathered 1912 - 1916, were digitised on request ; All 147 charts were scanned as high-resolution JPEG files (Figure 2). Records of the texts are available online, and physical copies at the WTL Medical Collection ◊. The following is an initial brief of the texts, with all further details being a part of the project's outcomes in their own right.

As the viral agent was unknown in this era authors presented probable cases, cases of interest, and febrile cases of several kinds (McCarthy, 2001). The commission operated in several locales, and most reports were written by Western Physicians or entomologists. Most clinical cases are native patients, and the commission's focus was on mapping the landscape of febrile illnesses in colonial regions and major ports used for Western trade.

It must be acknowledged that the data contained by the reports is a product of Western colonialism. The informed consent of any patients cannot be confirmed is almost certainly not properly obtained (Güereña-Burgueño, 2002). Any patient names have been anonymised.

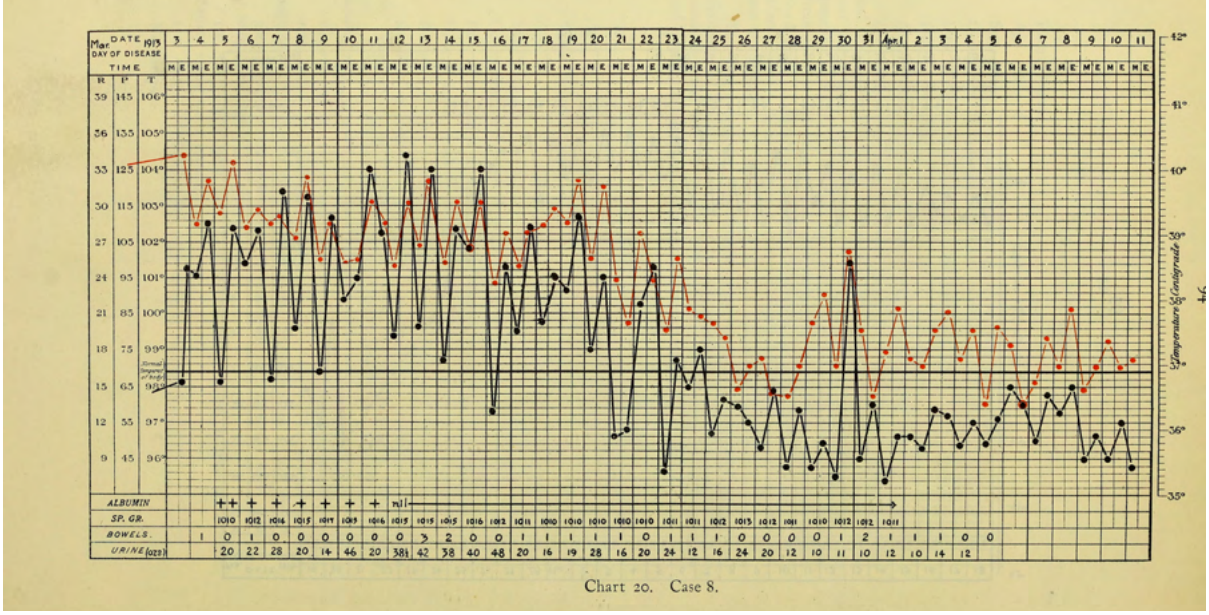


Figure 2: A representative plot from a sub-section of one volume. This hand-drawn chart displays observations of patient temperature (black) and pulse (red) in degrees Fahrenheit (°F) and beats per minute (BPM). This chart's format is a template used across most of the cases or has been adapted by some authors.

◊

Volume I:
Wyler, E. J., Russell Leonard, T. M. and O'Brien, J. M. (no date) 'Yellow Fever Commission (West Africa); Reports on Questions Connected with the Investigation of Non-malarial Fevers in West Africa; Volume 1.', in Fowler, J. K. et al. (eds) *Yellow Fever Bureau Bulletin Supplement*. Liverpool: The University Press of Liverpool, Liverpool School of Tropical Medicine.

Volume II:
Statham, J. C. B. *et al.* (no date) 'Yellow Fever Commission (West Africa); Reports on Questions Connected with the Investigation of Non-malarial Fevers in West Africa; Volume 2.', in Fowler, J. K. et al. (eds) *Yellow Fever Bureau Bulletin Supplement*. Liverpool: The University Press of Liverpool, Liverpool School of Tropical Medicine.

Volume III:
Bacot, A. W. *et al.* (1916) 'Yellow Fever Commission (West Africa); Reports on Questions Connected with the Investigation of Non-malarial Fevers in West Africa; Volume 3.', in Fowler, J. K. et al. (eds) *Yellow Fever Bureau Bulletin Supplement*. London: J. & A. Churchill, 7, Great Marlborough Street, London.

Methods

Evaluation of contents

The text contained time-series data collected from 147 individuals. Not every individual had both temperature and pulse recorded. I explored the data by dividing the charts into their respective sub-sections, which varied in several ways that might affect findings, and assessed each subsection for its

quality. For each subsection, a randomised sample of its charts, relative to the proportion of the total charts it contained, were inspected, and scored for their contents using set criteria (Table 1).

Table 1: Criteria used to evaluate the contents of each report section and subsection. Variables are selected for their relevance to data quantity, coverage, rhythm variability (drugs), and additional unusual parameters of interest (leukocyte counts).

Component	Description	Scoring system
Duration	The number of days spanned by the total chart	Raw; the number of days
Sampling	How frequently were vitals taken over a 24-hr period?	# per day (e.g., 2 = twice daily), taken from the maximum frequency.
Pulse	Has the author recorded patient pulse?	Binary: yes = 10 no = 0
Coverage – temperature	How much temperature data is recorded?	Scale: Low ($^1/_4$ *to* $^1/_3$) = 1 Medium ($^1/_2$) = 2 High ($^2/_3$ *to* $^3/_4$) = 3 All (every sample) = 4
Coverage – pulse	How much pulse data is recorded? * N.B. if absent, default is 0	As above
Drugs	Is there any record of drug treatment in chart or prose?	Binary: yes = 10 no = 0
Drug start time	Has the author recorded a clear drug start time?	As above
Leukocyte counts	Has the author recorded patient leukocyte counts?	As above

Through this assessment framework I characterised the type of data contained in the books (which was unknown), its resolution, and how much this differed across subsections. I was able to show what kinds of parameters could be found from data, and what should be expected to be available for extraction. I used these criteria on the randomly selected samples and gave each section a quality score, to ask which section(s) were suitable for further use, and which (if any) could be prioritised.

Data extraction

Using ImageJ, with the "Figure Calibration" plug-in, I extracted the temperature and, where present, pulse data from all charts containing more than two measurements of each variable. In several cases, the author supplemented data contained in the charts with time-stamped written records. I manually recorded and merged these data.

Metadata was recorded in Microsoft Excel. A "maximal" approach was taken, to record all plausible variables, making the information available in a digital format should it be of future use, and then "condense" to focus on more relevant factors: in the 'condensed' metadata set any symptoms described in less than 10% of the cases were removed.

Data manipulation

I created a coding pipeline that imported temperature and pulse data in two batches, assigned unique IDs to represent each case, and then merged pulse and temperature data. The pipeline then converts the data into a format recognised by R as "time", both in seconds and in a 24-hr manner. The pipeline yielded over 1353 days' worth of time-series data. I produced plots of case vitals and visually examined these against originals. Following the results of metadata extraction, I imported this data in R and derived additional variables that helped to characterise patient rhythms and factors. The outcomes of this process are presented in Results.

Circadian analyses

Cases with sufficient data were explored through the Rethomics framework (Geissmann *et al.*, 2019), and period estimates were produced alongside Lomb-Scargle periodograms (the Lomb-Scargle approach is better suited for data with missing measures or irregular intervals (VanderPlas, 2018). The power and significance thresholds were also calculated. The estimates of all cases exceeded their threshold values, so there was no need to omit estimates. Periodograms are shown in the Appendix.

Statistical analyses

For all statistical analyses, carried out in R, linear models were fitted and AICc were used to compare permutations of models to avoid overfitting. The exact approaches used are described below, as the models were the result of iterative evaluation of the texts and their contexts, and continual assessment of the data's suitability as it was discovered.

Results and Discussion

Evaluation of contents

All subsections containing charts held common variables. Although there was variation in "quality" between the subsections (Table 2), I decided to extract data from all, both to test the pipeline's function in all, and to maximise the material garnered.

Table 2: The "quality" scores of cases randomly selected from each subsection of all three volumes. The number of cases randomly selected from each subsection is proportional to the number of total cases it contains. Scores are calculated according to Table 1. * 1_1_4_3_73 is highlighted as being only two days in duration but having a large number of observations at uneven sampling intervals. This may affect the subsection total (also marked *).

Chart	Duration (days)	Sampling	Pulse	Pulse coverage	Temperature coverage	Drugs	Leukocyte Count	Total
1_1_4_1_8	4	2	-	-	4	10	-	**19**
1_1_4_1_9	13	2	10	4	4	-	-	**33**
1_1_4_1_12	23	2	10	4	4	10	-	**53**
Subsection total: 105								
1_1_4_2_12	23	2	10	4	4	10	-	**53**
1_1_4_2_15	8	2	10	1	4	10	-	**35**
1_1_4_2_17	10	2	10	2	4	-	-	**28**
1_1_4_2_38	5	2	10	4	4	-	-	**24**
1_1_4_2_49	4	2	10	4	4	-	-	**24**
1_1_4_2_65	5	2	10	3	4	-	-	**24**
1_1_4_2_18	24	2	10	4	4	-	-	**44**
Subsection total: 135								
1_1_4_3_68	6	6	-	-	2	10	-	**24**
1_1_4_3_73*	2	8	10	3	4	10	-	**37***
1_1_4_3_76	11	6	10	3	3	10	-	**53**
1_1_4_3_77	23	2	10	3	4	10	-	**53**
Subsection total: 114*								
1_2_8	4	2	10	4	4	-	10	**34**
1_2_12	21	2	10	1	4	-	10	**48**
1_2_19	16	2	10	3	4	-	-	**35**
1_2_25	14	2	10	4	4	-	-	**34**
1_2_35	15	6	10	3	3	-	-	**37**
Subsection total: 188								
1_3_7	5	2	10	3	4	-	10	**34**
1_3_10	11	2	10	4	4	-	10	**41**
1_3_1	13	2	10	3	4	-	10	**42**
Subsection total: 75								

That the quality of data varied across sub-sections and across charts suggested that some variation in the magnitude or reliability of period estimates may differ across subsections. What is not able to be recorded with certainty is the approach used by authors; temperature readings vary depending on the site taken, as well as the equipment used (Ogoina, 2011). Both pulse and temperature are also open to observer bias and to influences of environment, interactions and behaviours of patients and physicians (Elliot, Blanford and Thomas, 2002; Lambin *et al.*, 2010; Ogoina, 2011). Further, there is little known about the conditions of each healthcare facility of clinic featuring in each subsection, which may have large effects on the quantity or quality of data.

Although initially a variable of interest (Levi and Schibler, 2007; Adam, 2019), drug treatment was not described in many cases. The duration of the pilot study charts also varied greatly. It was of note that sampling was generally only twice per 24-hour period. This raised initial concerns that the resolution of the data may be low with respect to estimating period of temperature and pulse, but the decision was taken to proceed and to see what was retrievable.

To further test "quality" and suitability, I used the highest scoring chart from the sub-sample of each section (n = 5, see Table 2) to undertake a pilot of the data extraction, manipulation, and

parameterisation. I used this process to refine the pipeline and workflow and created a 'dummy' metadata table. All five charts successfully progressed through ImageJ processing and the data pipeline in R. Using the dummy metadata and Rethomics, I found that this pilot data could be used to calculate circadian parameters and return period estimates, which confirmed that the methods could generate useable data.

Data extraction

The extraction process produced 146 csv files of temperature data, and 130 of pulse data, which were compiled in pre-processing folders. In one case, I found that data from two separate admissions of the same patient was present, and data was extracted only from the second, longer admission. Another chart was omitted as this was found upon comprehensive reading to be a chart of a previous admission of one patient, with no reliable accompanying information or direct measures. As mentioned in brief, most authors used template charts that allowed for measurements to be recorded in 12-hour windows, e.g., morning and evening (Figure 2), but there were rarely specific times assigned. I found that the most consistent approach was to plot in ImageJ directly onto the hand-plotted point, whether this appeared on a solid line or within the "box".

This process suggested that ImageJ was a suitable tool for extracting the time-series data as csv files, but that difficulties arise from the variation in technique between authors. Further, when historical texts are less focused on quantitative recording and more interested in qualitative descriptions of disease or disease patterns, it is true that patterns may not be adequately reflected by quantitative measures (Mackowiak, 1998).

Exploration of the full texts and case records (with quality scores), data collection, and curation produced a "curated" metadata sheet of 30 variables. The value of each was determined using criteria that also resulted from full study of the texts. Both the curated metadata and these criteria are included in the Appendix. The resolution of metadata varied across sub-sections.

It is important to note that exploration led me to recorded variables describing symptoms as binary, to describe their presence or absence in clinical description. This was done to avoid interpreting or extrapolating meaning from the records, but also means there are limitations to what we can know: Absence of mention does not necessarily mean the symptom or syndrome did not occur, but that it was not highlighted by the author. For example, in cases where blood slides were not examined for *Plasmodium* parasites, the default value for "presence of *Plasmodium*" is necessarily zero.

These complexities demonstrate that metadata, as well as quantitative data, is affected by observer bias: authors may make conscious or unconscious choices about what to note or may fail to detect some symptoms. For example, rashes associated with VHF often go unnoticed in people of colour, especially when physicians are white and frequencies of this symptom may be underestimated (Nijhawan and Alexis, 2011). Additionally, haemorrhage in women with YF is often mistaken for regular menstruation, and the symptom is underreported (Wiwanitkit and Wiwanitkit, 2013). However, as there were only 7 female cases, this is unlikely to be a factor relevant to study of these volumes.

Data manipulation and exploration

The coding pipeline was successful in importing and merging temperature and pulse data and assigning unique IDs to represent each case. The pipeline converted the data into a format recognised by R as "time", both in seconds and in a 24-hr manner. The process yielded 1342 days' worth of patient observations. I produced time series plots for case vitals and visually examined these against originals. Plots illustrating these results are included in the Appendix.

These results lead me to characterise the data and understand more about the types of parameters that can be derived from it once digitised. Through several R packages, I compiled details that revealed more about the reports, individual patients, and the spread of quantitative and meta- data. The patient age and ethnicity and the locality, which is also an identifier of sub-section are shown in Figure 3. Further information is available via the metadata (Appendix) or on request.
I found that in the metadata, symptoms of severe YF were common, and calculated the case fatality rates

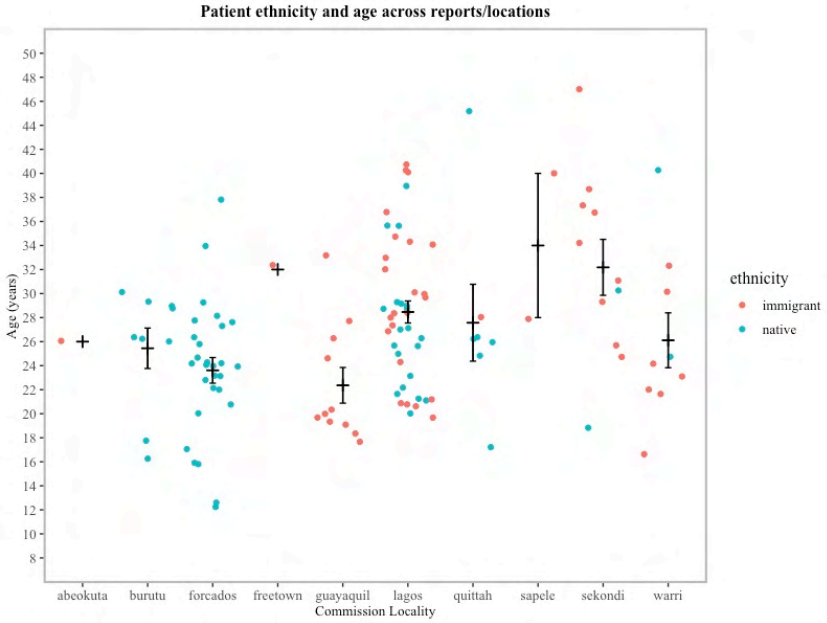


Figure 3: The ages and ethnicities of patients in the texts, by their locality. Different localities also represent different sub-sections of the texts. The ethnicities of patients should be interpreted with caution as authors recorded little detail and the distinction between natives (those born in West Africa) and "settled" peoples (white Westerners) was sometimes unclear. It is of note that racial slurs were often used to identify patients, a further sign of inaccuracy but also a vital ethical consideration important in the use of all historical texts. Whether ethnicity should be taken as a proxy of previous exposure to YF (which confers some protection) is thus also unclear.

(CFRs) associated with these symptoms (not exclusive of other symptoms) (Table 4).

The overall CFR for all cases was 15%, higher than recent estimates of 7.5% CFR – but this figure is with treatment, which has of course advanced since the texts were published > 100 years ago. The higher CFR suggested in patients with haemorrhage is in alignment with current literature (approximately 50% CFR in patients with severe YF and haemorrhage; (Monath and Barrett, 2003; Monath, 2008)). The frequency of albuminuria and jaundice is high considering contemporary estimates: only 20% of patients with YF develop severe YF, which is associated with these symptoms, that are not common features of the non-severe form (Gardner and Ryman, 2010). This may reflect the inclusion of specific cases by authors: reports are not random samples, but cases are likely selected for the appearance of symptoms or interesting features and are not representative of YF in the general population. Further, symptomatic, or severe cases may be more likely to attend hospital or clinics than people with mild YF. Asymptomatic YF infections may go completely undetected, especially in native patients who have a level of immunity to YF (Monath and Barrett, 2003). These factors may explain the relatively high occurrence of severe symptoms. The lower CFR is more difficult to understand, but the course of disease may be a source of bias: patients with symptomatic YF generally enter a period of apparent "recovery" after ~ 6 to 10 days of disease; bit remission and rapid escalation of disease to severe YF occur in 20% of these patients after around 24 – 48-hours (Barnett, 2007; Staples *et al.*, 2020). In some of the historical cases, patients were discharged or not followed-up shortly after a decrease in fever. It is possible that some of these patients may have undergone remission to severe YF, in which case the overall CFR (15%) may be an underestimate.

Table 4: Symptoms of severe YF seen across cases and their frequencies, details on their types or indications, and the case fatality rate (CFR) associated with their presence in a case.

Symptom	% of patients	Details	CFR
Albuminuria	89.73	Indicative of renal failure	16%
Jaundice	45.21	Indicative of renal failure	19.7%
Bleeding	10.96	Indicative of internal haemorrhage	25%
Dark vomit/stools	21.23	Indicative of internal haemorrhage	45.2%
Haemorrhage (one or both)	26.7	Bleeding and/or dark excrement and/or rash	38.5%

The presence of symptoms associated with renal failure and with haemorrhage is in accordance with the known disruption of hepatic and metabolic rhythms during YF (Quaresma *et al.*, 2006; Quaresma, Duarte and Vasconcelos, 2006; Song and Carneiro D'Albuquerque, 2019)These rhythms are "peripheral rhythms", meaning that they operate alongside the centrally controlled rhythms such as temperature and heart rate (Yamamoto *et al.*, 2020; Manella *et al.*, 2021). Central rhythms are set and maintained by a "master" clock, the supra-chiasmatic nucleus in the hypothalamus, which is entrained to dependable circadian stimuli, with the principle one being light (Mohawk, Green and Takahashi, 2012). But rhythms do not only operate at a central level with a single control: distinct rhythms are present even within organs, or single cells (Buhr and Takahashi, 2013; Schibler *et al.*, 2016; Kelly *et al.*, 2018). Although rhythms in tissues of the liver and kidneys are generally synchronised to central circadian rhythms such as temperature and heart rate, and therefore under some control of the SCN, they also can be entrained by other stimuli; For example, liver rhythms can be modulated by the time-of-day of feeding, but also by the levels of different nutrients (Manella *et al.*, 2021), and some studies have demonstrated that liver rhythms can even become uncoupled from central rhythms (Schibler *et al.*, 2016). Disruption of liver rhythms, in association with, or without, disruption of central rhythms, is an important piece of the puzzle that is severe YF, especially the build-up of toxic by-products and haemorrhage.

Circadian analyses

Estimating period for pulse and temperature was the first objective. Of the 146 with extracted temperature data, period could be estimated for 142 cases. The mean period estimate was 40.2 hours (Figure 3). Of the 130 cases for which pulse data could be extracted, period estimates could be calculated for 126, with the mean period being 43.6 hours (Figure 4).

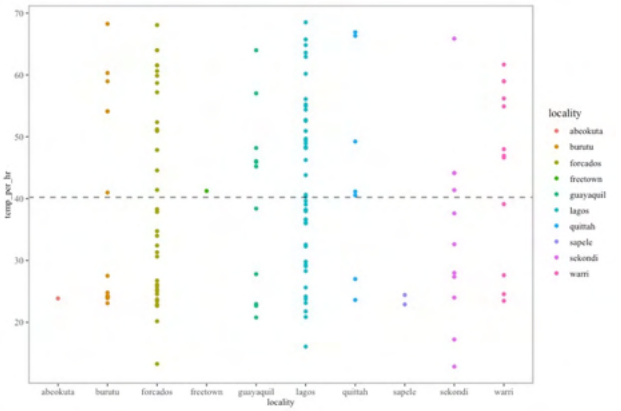


Figure 3: The estimated period of **temperature** in hours, by locality, which is also a proxy of sub-section. Dashed line represents the mean period estimate length.

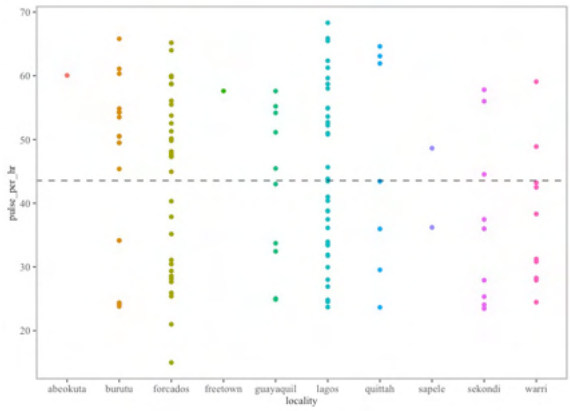


Figure 4: The estimated period of **pulse** in hours, by locality, which is also a proxy of sub-section. Dashed line represents the mean period estimate length.

There were no significant trends in period associated with the locality of the cases, which is also indicative of subsection (subsections have different regional settings).

The expected period for both temperature and pulse in "healthy" individuals would be roughly 24 hours, as pulse tends to peak during an active phase (day), and dip during resting phase (night), with a circadian rhythm (Mackowiak, 1998; Ahmad *et al.*, 2009; Plaza *et al.*, 2016). The same patterns are seen in body temperature although these peaks and troughs may not be at the same time-of-day (Refinetti and Menaker, 1992). If the data showed clustering of period estimates at a specific value below or above 24-hours, it might be possible to infer disruption of this rhythm; however, estimates of period length for both temperature and pulse were widely ranging and showed no specific patterns (Figures 3 and 4).

The above demonstrates that the data in these reports were unsuitable for circadian analysis using Rethomics, likely due to the broad sampling intervals, and possibly lack of precision in the plotting of

time-of-day of samples, due to the methods used by authors in charts. Therefore, the data was not amenable to asking focused questions about the effects of infection and patient characteristics on host rhythms.

Further Analyses

As a **result** of these findings, further exploration of the data was carried out. Of particular interest was the relationship between two host rhythms, known to be a feature of YF. As mentioned in the Introduction, abnormal heart rates are a signature of YF. Specifically, heart rates which are at odds with body temperature. In "healthy" individuals, heart rate is expected to rise by 10 beats per minute (BPM) with every 1.8°F increase in body temperature **over the fever threshold, of 99.5°F** (Ogoina, 2011)**.** Greater or smaller changes are known as temperature-pulse disassociation, or Relative Bradycardia (RB; (Ye *et al.*, 2018). RB is a hallmark symptom of YF as well as other viral infections and is associated with pathology and the disruption of organ and other bodily functions (Monath and Barrett, 2003; Quaresma, Duarte and Vasconcelos, 2006; Song and Carneiro D'Albuquerque, 2019). It has potential as a prognostic and/or diagnostic marker but is poorly understood, especially when affected by chemotherapy (Adam, 2019).

The identification of RB was also a process that did not require the production of periodograms or period estimates, and which could be suggested with relatively little data. For this reason, the following was done to evaluate RB within the historical cases.

The data I mined permitted me to calculate a new variable describing the total observation period of each case (its *max_time*); To visualise the range in recorded temperatures for each patient I also calculated the maximum and minimum temperature readings (*max_temp, min_temp)*, and the resulting range (*temp_range*). I calculated equivalent parameters for pulse data. I found that the data was suitable for the "binned" function within the Rethomics framework - I was able to obtain daily parameters of temperature and pulse across the observation window. From this the change in a patient's temperature or pulse across each 24-hr period was successfully derived. I defined a fever threshold of 99.5°F - as suggested by large-scale studies of average body temperatures of healthy individuals, and as according to the above details on RB (Ogoina, 2011).

To detect cases suggestive of RB:

- Days on which patients were febrile were defined as days where temperature readings exceeded 99.5°F. The temperature ranges on these days were calculated as relative to this threshold, giving a measure of fever height (below).
- For all fever days, the temperature-pulse relationship (TPR) was thus calculated using the following formula:
    - Daily fever height, *DFH* = Total 24-hour temperature change in °F – 99.5 °F.
    - $\frac{IP}{DFH/1.8}$ where *IP* = increase in pulse in BPM, *DFH* = as above, 1.8 corresponding to the relative of °F to pulse.
- Where the result was lower than 10 (BPM), this a day on which a patient was suggested to experience RB.
- An example is shown below (Figure 5).
- The total number of days each patient was suggested to have RB was calculated, followed by the proportion of the total number of their observation days this represented.

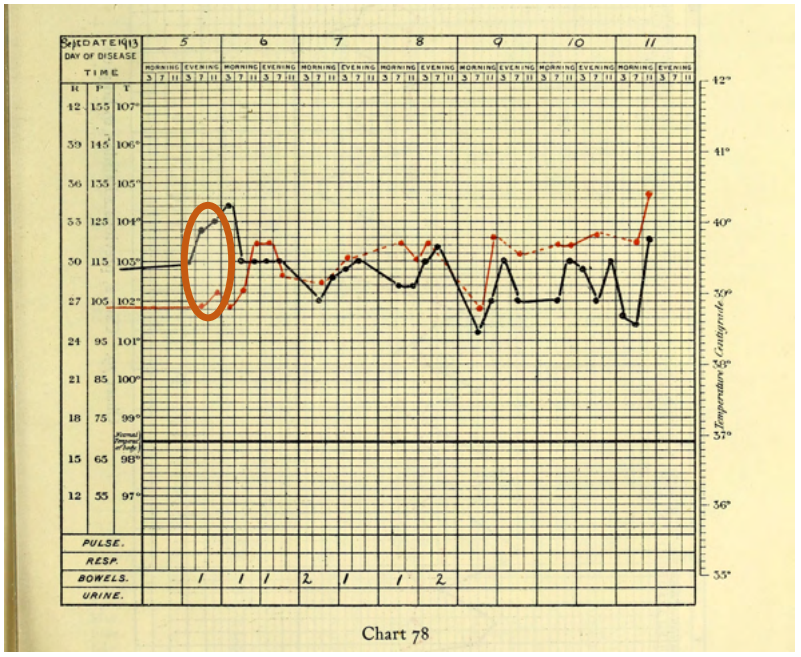


Figure 5: An example of how TPR was calculated and how RB was detected.

In the chart on the left, the *DFH* within the highlighted area is 1°F. The change in pulse is 4 BPM. Therefore:

$$\text{TPR} = \frac{4\ BPM}{1\,℉\ /1.8\,℉}$$

TPR = 7; suggestive of RB.

All patients had febrile days; on average, patients had fever for ½ of the days they were observed. RB was suggested in 74.65% of cases (109/146), across age and ethnicities.

Statistical Analyses

The results of the above processes which were of interest were symptoms of YF and their relationship with case outcome, and RB and its relationship with

First, I fit a linear model that asked which symptoms (if any) of severe YF might be the best indicators of outcome (Table 5).

Table 5: The output of an AICc used to compare the suitability of possible models, in which it was asked if outcome was most likely to be associated with several symptoms of severe YF. Maximal model: **Outcome ~ albuminuria + bleeding + dark excrement + jaundice.** A + in a column indicates this factor is included in that permutation. A delta of less than 2 indicates a competitive model, compared with 0, which is the most probable model of those used.

(Intercept)	albuminuria	bleeding	dark excrement	jaundice	family	df	logLik	AICc	delta	weight
1	NA	NA	+	NA	gaussian(identity)	4	-41.99	92.26	0	0.23
0.88	+	NA	+	NA	gaussian(identity)	5	-41.12	92.67	0.41	0.19
0.93	NA	NA	+	+	gaussian(identity)	5	-41.38	93.19	0.93	0.15
1	NA	+	+	NA	gaussian(identity)	5	-41.64	93.72	1.46	0.11
0.84	+	NA	+	+	gaussian(identity)	6	-40.73	94.08	1.81	0.09
0.88	+	+	+	NA	gaussian(identity)	6	-40.78	94.17	1.91	0.09
0.88	+	+	+	NA	gaussian(identity)	6	-40.78	94.17	1.91	0.09
0.85	+	+	+	+	gaussian(identity)	7	-40.5	95.82	3.56	0.04

The results of this analysis suggest that dark excrement, a sign of internal haemorrhage, may be the most important prognostic indicator (Table 5, delta = 0). Crucially, all permutations outperformed a null model. I chose not to model "haemorrhage" more generally as it subsumes two other factors and would distort results. Dark excrement is generally indicative of renal failure and liver damage (Abdallah El Bouzedi, 2015; Table 4). This failure leads to the build-up of toxic by-products within tissues but pathology underlying haemorrhage is also thought to be caused by the direct effects of viral replication in the liver and kidneys (Barnett, 2007).

Following the finding that RB could be calculated and was present in over 70% of cases, I constructed a linear model that asked whether the proportion of days a patient was suggested to experience RB (a better measure of magnitude, not just the presence), patient age, and/or ethnicity, was associated with case outcome (Table 6).

Table 6: The output of an AICc used to compare the suitability of several possible models, in which it was asked if outcome was most likely to be associated with the level of RB, or patient age, and/or ethnicity. Maximal model: **Outcome ~ age + ethnicity + proportion of days with RB** ("proportion brady")**.** A delta of less than 2 indicates a competitive model, compared with 0, which is the most probable model of those used.

(Intercept)	age	ethnicity	proportion brady	family	df	logLik	AICc	delta	weight
0.9	NA	+	-0.49	gaussian(identity)	4	-23.56	55.46	0	0.66
0.8	0	+	-0.5	gaussian(identity)	5	-23.15	56.84	1.38	0.33
1.03	NA	NA	-0.54	gaussian(identity)	3	-30	66.2	10.74	0
1.01	0	NA	-0.55	gaussian(identity)	4	-29.98	68.31	12.85	0
0.72	NA	+	NA	gaussian(identity)	3	-35.86	77.93	22.47	0
0.68	0	+	NA	gaussian(identity)	4	-35.81	79.96	24.5	0
0.86	NA	NA	NA	gaussian(identity)	2	-43.85	91.8	36.34	0
0.92	0	NA	NA	gaussian(identity)	3	-43.73	93.67	38.21	0

The model most likely to explain variation in case outcome is one including patient ethnicity alongside the proportion of days a patient likely experienced RB (Table 6, top row, delta = 0). This suggests that the extent of RB in a case, and the ethnicity of a patient, may influence the chances of recovery.

This result is interesting as it poses several considerations. Old age is a factor commonly associated with cardiac health and abnormality, as well as with low immune function (Dai *et al.*, 2012; Montecino-Rodriguez, Berent-Maoz and Dorshkind, 2013), but also could be positively correlated with exposure to YF, which may be protective (Monath and Barrett, 2003). Age in this analysis was not part of the most competitive model, which might indicate a lack of any relationship, but may also be related to these antagonistic factors.

Records of patient ethnicity were not entirely reliable (see previous), but broadly were good indicators of the nationality of individuals, and of the level of exposure to YF. YF exposure and the resulting protective immunity may therefore introduce variation in the form of more favourable outcomes for native patients (Monath and Barrett, 2003). However, where native patients are exploited by white colonisers, or face difficulty acquiring nutritious foods or clean conditions, "native" ethnicity may no longer have a strong protective effect. Further, patients in some sub-sections (both native and "immigrant") were living/working/slaves on ships, which may influence diet and feeding rhythms associated with renal failure and disease severity, as well as general health.

As disease severity and outcome are associated with liver failure and haemorrhage (suggestive of disrupted rhythms in the liver and kidneys), and outcome is also associated with RB and ethnicity, it is possible that future work may find a relationship between RB and the disruption of normal, circadian liver functions. It could be considered that disassociation of temperature and pulse fluctuations (RB) is linked to liver rhythm disruptions, that all three become uncoupled relative to each other, or that there are separate processes acting in all three "situations" that produce apparent relationships.

Summary

Using a bespoke pipeline, the texts obtained from the WTL revealed almost 1350 days' worth of data on patient vitals, and over 50 high-resolution metadata variables. The hand-drawn plots were transformed into time-series data but could not be used to generate reliable period estimates for circadian analysis. However, evaluation of text contents showed that other features could be investigated: The outcome of a case may be related to the level of disassociation between daily changes in a patient's body temperature and their heart rate. This effect may be dependent on a patient's ethnicity. Also associated with outcome are symptoms related to the disruption of rhythms in the liver.

Future work could consider the following:

- The data protocol (evaluation, extraction, manipulation) developed here can be easily applied to other historical resources.
- Identifying resources with regular, and narrower, sampling intervals should be a priority.
- These resources should be considered as a whole, and metadata should be curated with an understanding of the author's biases and the regional contexts.
- An area of study may be to examine relationships between central rhythms and peripheral rhythms. What happens when central rhythms become uncoupled? Are peripheral rhythms disrupted by this? Does "uncoupling" occur between two, or three (or more) rhythms? Can the "type" and "magnitude" of disassociation be characterised quantitatively? If so, what relationships can be found?
- A core consideration in all data mining of historical texts should be their origins: many texts are the products of white Western colonialism and violence. Names should be anonymised, social scientists may need to be consulted, and discussion should not neglect to acknowledge these elements of our archived resources.

References

Abdallah El Bouzedi, M. A. D. (2015) 'Viral Haemorrhagic Fever in North Africa; an Evolving Emergency', *Journal of Clinical & Experimental Pathology*, 05(02). doi: 10.4172/2161-0681.1000215.

Adam, D. (2019) 'Emerging science of chronotherapy offers big opportunities to optimize drug delivery', *Proceedings of the National Academy of Sciences of the United States of America*, 116(44), pp. 21957–51959. doi: 10.1073/pnas.1916118116.

Ahmad, S. *et al.* (2009) 'Clinical review: A review and analysis of heart rate variability and the diagnosis and prognosis of infection', *Critical Care*. BioMed Central, pp. 1–7. doi: 10.1186/cc8132.

Barnett, E. D. (2007) 'Yellow fever: Epidemiology and prevention', *Clinical Infectious Diseases*, 44(6), pp. 850–856. doi: 10.1086/511869.

Borrmann, H., McKeating, J. A. and Zhuang, X. (2021) 'The Circadian Clock and Viral Infections', *Journal of Biological Rhythms*. J Biol Rhythms, pp. 9–22. doi: 10.1177/0748730420967768.

Buhr, E. D. and Takahashi, J. S. (2013) 'Molecular components of the mammalian circadian clock', *Handbook of Experimental Pharmacology*, 217(217), pp. 3–27. doi: 10.1007/978-3-642-25950-0_1.

Dai, D. F. *et al.* (2012) 'Cardiac aging: From molecular mechanisms to significance in human health and disease', *Antioxidants and Redox Signaling*. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, pp. 1492–1536. doi: 10.1089/ars.2011.4179.

Elliot, S. L., Blanford, S. and Thomas, M. B. (2002) 'Host-pathogen interactions in a varying environment: Temperature, behavioural fever and fitness', *Proceedings of the Royal Society B: Biological Sciences*,

269(1500), pp. 1599–1607. doi: 10.1098/rspb.2002.2067.

Feldmann, H. and Geisbert, T. W. (2011) 'Ebola haemorrhagic fever', *The Lancet*. NIH Public Access, pp. 849–862. doi: 10.1016/S0140-6736(10)60667-8.

Gardner, C. L. and Ryman, K. D. (2010) 'Yellow fever: A reemerging threat', *Clinics in Laboratory Medicine*. NIH Public Access, pp. 237–260. doi: 10.1016/j.cll.2010.01.001.

Geissmann, Q. *et al.* (2019) 'Rethomics: An R framework to analyse high-throughput behavioural data', *PLoS ONE*, 14(1), p. e0209331. doi: 10.1371/journal.pone.0209331.

Güereña-Burgueño, F. (2002) 'The centennial of the Yellow Fever Commission and the use of informed consent in medical research', *Salud Publica de Mexico*. Salud Publica Mex, pp. 140–144. doi: 10.1590/S0036-36342002000200009.

Kelly, R. M. *et al.* (2018) 'Clocks in the clinic: Circadian rhythms in health and disease', *Postgraduate Medical Journal*, 94(1117), pp. 653–658. doi: 10.1136/postgradmedj-2018-135719.

Lambin, E. F. *et al.* (2010) 'Pathogenic landscapes: Interactions between land, people, disease vectors, and their animal hosts', *International Journal of Health Geographics*, 9, p. 54. doi: 10.1186/1476-072X-9-54.

Levi, F. and Schibler, U. (2007) 'Circadian rhythms: Mechanisms and therapeutic implications', *Annual Review of Pharmacology and Toxicology*. Annual Reviews Inc., pp. 593–628. doi: 10.1146/annurev.pharmtox.47.120505.105208.

Mackowiak, P. A. (1998) 'Concepts of fever', *Archives of Internal Medicine*. American Medical Association, pp. 1870–1881. doi: 10.1001/archinte.158.17.1870.

Manella, G. *et al.* (2021) 'The liver-clock coordinates rhythmicity of peripheral tissues in response to feeding', *Nature Metabolism*, 3(6), pp. 829–842. doi: 10.1038/s42255-021-00395-7.

McCarthy, M. (2001) 'A century of the US Army yellow fever research', *Lancet*. Elsevier, p. 1772. doi: 10.1016/S0140-6736(00)04943-6.

Mohawk, J. A., Green, C. B. and Takahashi, J. S. (2012) 'Central and peripheral circadian clocks in mammals', *Annual Review of Neuroscience*, pp. 445–462. doi: 10.1146/annurev-neuro-060909-153128.

Monath, T. P. (2008) 'Treatment of yellow fever', *Antiviral Research*. Elsevier, pp. 116–124. doi: 10.1016/j.antiviral.2007.10.009.

Monath, T. P. and Barrett, A. D. T. (2003) 'Pathogenesis and Pathophysiology of Yellow Fever', *Advances in Virus Research*, 60, pp. 343–395. doi: 10.1016/S0065-3527(03)60009-6.

Montecino-Rodriguez, E., Berent-Maoz, B. and Dorshkind, K. (2013) 'Causes, consequences, and reversal of immune system aging', *Journal of Clinical Investigation*. American Society for Clinical Investigation, pp. 958–965. doi: 10.1172/JCI64096.

Nijhawan, R. I. and Alexis, A. F. (2011) 'Practical approaches to medical and cosmetic dermatology in skin of color patients', *Expert Review of Dermatology*, pp. 175–187. doi: 10.1586/edm.10.75.

Ogoina, D. (2011) 'Fever, fever patterns and diseases called "fever" - A review', *Journal of Infection and Public Health*. J Infect Public Health, pp. 108–124. doi: 10.1016/j.jiph.2011.05.002.

Plaza, J. J. G. *et al.* (2016) 'Fever as an important resource for infectious diseases research', *Intractable and Rare Diseases Research*. International Advancement Center for Medicine & Health Research Co., pp. 97–102. doi: 10.5582/irdr.2016.01009.

Quaresma, J. A. S. *et al.* (2006) 'Revisiting the liver in human yellow fever: Virus-induced apoptosis in hepatocytes associated with TGF-β, TNF-α and NK cells activity', *Virology*, 345(1), pp. 22–30. doi: 10.1016/j.virol.2005.09.058.

Quaresma, J. A. S., Duarte, M. I. S. and Vasconcelos, P. F. C. (2006) 'Midzonal lesions in yellow fever: A specific pattern of liver injury caused by direct virus action and in situ inflammatory response', *Medical Hypotheses*, 67(3), pp. 618–621. doi: 10.1016/j.mehy.2006.01.060.

Refinetti, R. and Menaker, M. (1992) 'The circadian rhythm of body temperature', *Physiology and Behavior*. Elsevier, pp. 613–637. doi: 10.1016/0031-9384(92)90188-8.

Sajadi, M. M. *et al.* (2012) 'Akhawaynī and the first fever curve', *Clinical Infectious Diseases*, 55(7), pp. 976–980. doi: 10.1093/cid/cis596.

Schibler, U. *et al.* (2016) 'Clock-Talk: Interactions between central and peripheral circadian oscillators in mammals', in *Cold Spring Harbor Symposia on Quantitative Biology*, pp. 223–232. doi: 10.1101/sqb.2015.80.027490.

Sengupta, S. *et al.* (2021) 'Clocks, Viruses, and Immunity: Lessons for the COVID-19 Pandemic', *Journal of Biological Rhythms*. SAGE Publications Inc., pp. 23–34. doi: 10.1177/0748730420987669.

Song, A. T. W. and Carneiro D'Albuquerque, L. A. (2019) 'Acute Liver Failure Secondary to

Yellow Fever: A Challenging Scenario', *Clinical Liver Disease*. John Wiley & Sons, Ltd, pp. 58–61. doi: 10.1002/cld.784.

Staples, J. E. *et al.* (2020) 'Review of data and knowledge gaps regarding yellow fever vaccine-induced immunity and duration of protection', *npj Vaccines*, 5(1), pp. 1–7. doi: 10.1038/s41541-020-0205-6.

VanderPlas, J. T. (2018) 'Understanding the Lomb–Scargle Periodogram', *The Astrophysical Journal Supplement Series*, 236(1), p. 16. doi: 10.3847/1538-4365/aab766.

Westwood, M. L. *et al.* (2019) 'The evolutionary ecology of circadian rhythms in infection', *Nature Ecology and Evolution*. Nat Ecol Evol, pp. 552–560. doi: 10.1038/s41559-019-0831-4.

Wiwanitkit, S. and Wiwanitkit, V. (2013) 'Excessive menstruation bleeding as a presentation of dengue hemorrhagic fever', *Archives of Gynecology and Obstetrics*. Springer, p. 1271. doi: 10.1007/s00404-012-2657-z.

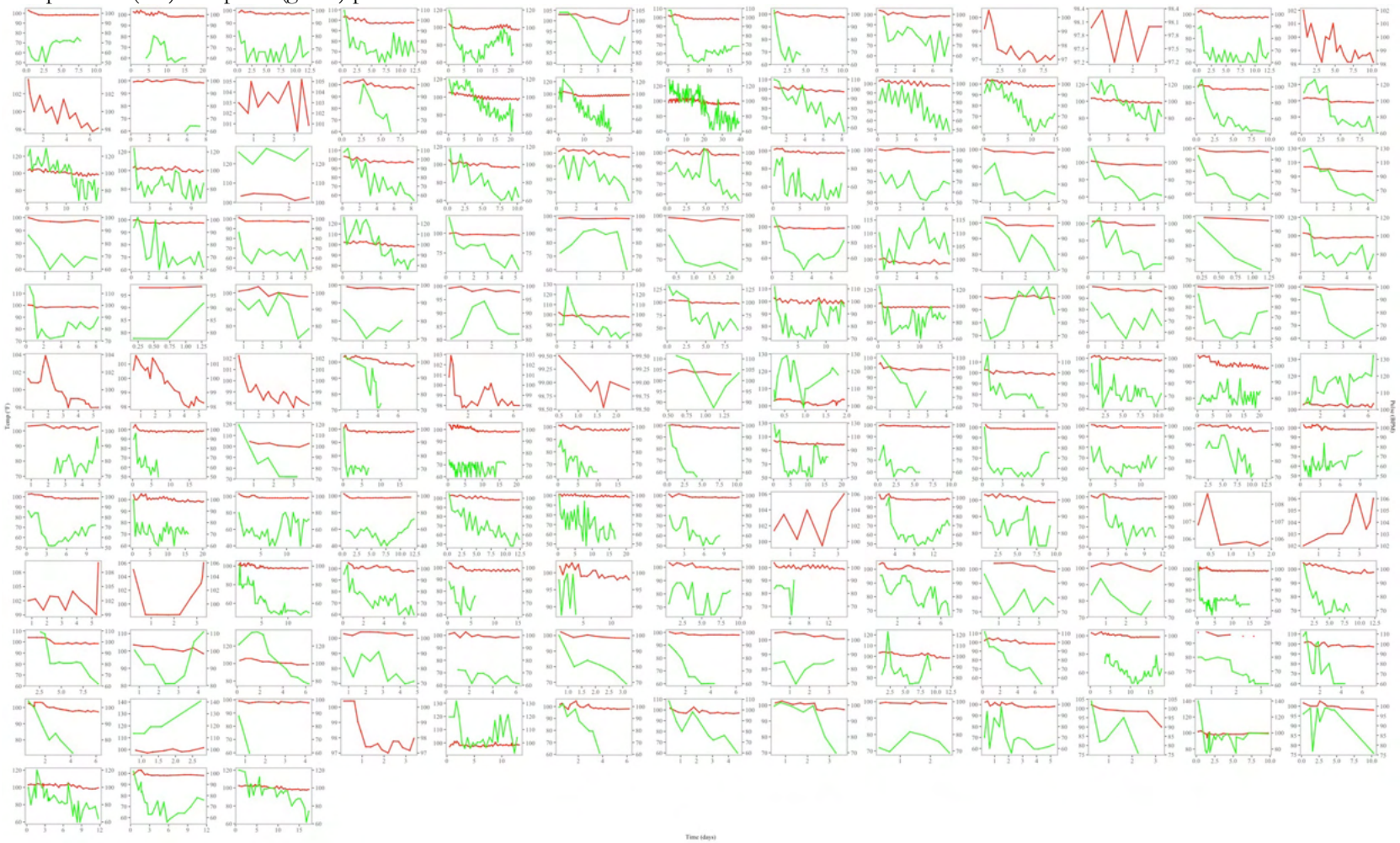
Yamamoto, Y. *et al.* (2020) 'Neisseria macacae bacteremia: Report of two cases with a literature review', *BMC Infectious Diseases*, 20(1), pp. 1–4. doi: 10.1186/s12879-020-05346-3.

Yang, S. L. *et al.* (2014) 'Hepatitis B virus X protein disrupts the balance of the expression of circadian rhythm genes in hepatocellular carcinoma', *Oncology Letters*, 8(6), pp. 2715–2720. doi: 10.3892/ol.2014.2570.

Ye, F. *et al.* (2018) 'Proposed mechanisms of relative bradycardia', *Medical Hypotheses*, 119, pp. 63–67. doi: 10.1016/j.mehy.2018.07.014.

Appendix

Temperature (red) and pulse (green) plots extracted from the three volumes:



DAVIS EXPEDITION FUND

Some of the variables in the "condensed" metadata gathered from the texts, with descriptions of type and values.

Variable	Values	Criteria	Notes
Ethnicity	native; immigrant	Native: the patient is described as West African or is listed as being black. Immigrant: the patient is described as European or not West African.	Native patients are considered to have a level of immunity to yellow fever.
Diagnosis	yellow fever; malaria; typhoid; other.	Three most common diagnoses named. If no diagnosis is described, NA is recorded. If a rare or ambiguous diagnosis is described, "other" is recorded to preserve sample size.	
Urine Albumen	Y/N	Y: Author reports albumen or "cloud" in the urine at any point, or "albuminuria".	Specific to YF
Urine Acid	Y/N	Y: the author has described a positive acid reaction of the urine.	Suggestive of YF
Tongue	Y/N	Y: Author has described 'fur' on the tongue, regardless of colour.	
Headache	Y/N	Y: Author reports patient experiences headache at any time.	
Quinine	Y/N	Y: Patient is given or has taken quinine at any time for the illness. This does not include mention of non-specific or historical quinine prophylaxis, which was commonplace.	
Other Drugs	Y/N	Y: Patient is given or has taken any drugs other than quinine. Not including "tonics", alcohols or "rubs".	
Neurological	Y/N	Y: Author describes coma, convulsion, delirium, loss of consciousness or fits. I excluded any mention of 'hysteria' or 'anxiety'.	Suggestive of severe YF
Coinfection	Y/N	Y: Any coinfection is described and/or confirmed, pre- or post-mortem.	
Lungs	Y/N	Y: Author or autopsy reports damage of the lungs.	
Liver	Y/N	Y: Author describes, either during treatment or confirms via autopsy, damage to or pain in the liver.	
Epigastralgia	Y/N	Y: Author reports specifically that the patient has epigastric pain.	Suggestive of severe YF
Dark Excrement	Y/N	Y: Vomit or stool is described as 'black', 'bloody', having dark 'debris', 'chocolate' or 'coffee grounds'.	Specific to severe YF
Bile in Excrement	Y/N	Y: Author explicitly describes stool, urine or vomit as containing bile or melaena, being 'bilious' 'green' or 'yellow'	Suggestive of severe YF
Gums Bleeding	Y/N	Y: Specifically stated by the author.	Suggestive of severe YF
Epistaxis	Y/N	Y: Specifically stated by the author.	Suggestive of severe YF
Other Haemorrhage	Y/N	Y: Author reports one or more of the following: 'epistaxis', 'nosebleed', 'gums bleeding'. In females, abnormal or unexpected menstruation.	Specific to severe YF
Leukocyte Count	Y/N	Y: Author has provided at least one leucocyte count from a blood smear.	Recorded for future usage
Jaundice	Y/N	Y: Author or autopsy reports jaundice or 'yellow' skin or eyes. Y is also recorded when an autopsy shows jaundice of the internal organs, even if not externally visible.	Specific to YF
Malaria Parasites	Y/N	Y: Author reports that a blood smear has been positive for malaria parasites. If no blood smear is taken, N is recorded as default.	
Loin Pain	Y/N	Y: Author reports that patient experiences pain in the loins.	Suggestive of YF
Rash	Y/N	Y: One or more from 'rash', 'petechiae', 'lividity', either during infection or post-mortem.	Suggestive of YF

DAVIS EXPEDITION FUND

"Condensed" metadata – the 30 variables remaining after those with insufficient or inconsistent coverage were removed – 1 of 2. ID is representative of: The first numeral of each ID indicates which of the three volumes the chart is contained by, and further numerals indicate sections and subsections.

id	page	has_temp	has_pulse	diagnosis	ethnicity	age	sex	locality	outcome	albumin_urine	acid_urine	tongue_furred	headache	quinine	other_drugs	neurological	coinfection	chest	liver	epigastralgia	dark_excrement	bile_excrement	gums_bleeding	leucocyte_count	nosebleed	jaundice	malaria_found	loin_pain	rash	bleeding
1_1_2_1	34	y	y	yellow	native	29	m	lagos	recovery	y	y	n	y	n	n	n	y	n	n	n	n	y	n	n	n	y	y	n	n	n
1_1_3_2	38	y	y	yellow	immigrant	26	m	abeokuta	recovery	y	n	y	y	y	y	y	n	n	n	y	y	y	y	n	n	y	n	n	y	y
1_1_4_1_3	46	y	y	yellow	immigrant	30	m	warri	death	y	n	n	y	n	y	n	n	n	y	n	y	n	n	n	n	y	n	n	n	n
1_1_4_1_4	47	y	y	yellow	immigrant	32	m	warri	recovery	y	n	y	y	y	y	n	n	n	n	n	y	n	n	y	y	y	n	n	n	y
1_1_4_1_5	54	y	y	malaria	native	25	m	warri	recovery	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_1_6	55	y	y	malaria	native	40	m	warri	recovery	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_1_7	61	y	n	malaria	immigrant	24	m	warri	recovery	n	n	n	n	y	y	n	n	n	n	n	n	n	n	n	n	n	y	n	n	n
1_1_4_1_8	62	y	n	NA	immigrant	NA	m	warri	recovery	n	n	n	n	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_1_9	64	y	y	malaria	immigrant	22	m	warri	recovery	y	n	n	n	y	n	n	y	n	n	n	n	n	n	n	n	n	y	n	n	n
1_1_4_1_10	65	y	y	malaria	immigrant	22	m	warri	recovery	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_1_11	66	y	y	malaria	immigrant	23	m	warri	recovery	y	n	n	n	y	n	n	y	n	n	n	n	n	n	n	n	n	y	n	n	n
1_1_4_1_12	67	y	y	malaria	immigrant	17	m	warri	recovery	y	n	n	n	y	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_13	83	y	n	malaria	immigrant	NA	m	forcados	recovery	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	y	n	n
1_1_4_2_14	83	y	n	NA	immigrant	NA	m	forcados	recovery	n	n	y	n	y	n	y	n	n	n	n	n	y	n	n	n	n	n	y	y	n
1_1_4_2_15	84	y	y	NA	immigrant	NA	m	forcados	recovery	n	n	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	y	y	n
1_1_4_2_16	85	y	n	malaria	immigrant	NA	m	forcados	death	n	n	n	y	n	y	y	n	n	n	n	n	y	n	n	n	n	n	n	n	n
1_1_4_2_17	90	y	y	NA	native	NA	m	buutu	recovery	y	n	y	n	n	n	y	n	n	y	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_18	91	y	y	NA	native	30	m	buutu	recovery	y	n	n	n	n	n	y	n	n	y	n	n	y	n	n	n	n	n	n	n	n
1_1_4_2_19	93	y	y	NA	native	26	m	buutu	recovery	y	n	y	n	n	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_20	94	y	y	NA	native	NA	m	buutu	death	y	n	n	n	n	n	n	y	y	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_21	95	y	y	NA	native	NA	m	buutu	recovery	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_22	96	y	y	NA	native	29	m	buutu	recovery	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_23	97	y	y	NA	native	26	m	buutu	recovery	y	n	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_24	98	y	y	NA	native	18	m	buutu	recovery	y	n	y	n	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_25	99	y	y	NA	native	NA	m	buutu	recovery	y	n	n	n	n	n	n	y	y	n	n	n	n	n	n	n	n	y	n	n	n
1_1_4_2_26	100	y	y	NA	native	NA	m	buutu	recovery	y	n	n	n	n	n	n	y	y	y	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_27	101	y	y	NA	native	NA	m	buutu	recovery	y	n	n	n	n	n	n	y	y	n	n	n	n	n	n	n	n	n	y	n	n
1_1_4_2_28	103	y	y	NA	native	29	m	buutu	recovery	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_29	103	y	y	NA	native	16	m	buutu	death	y	n	n	n	n	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_30	104	y	y	NA	native	NA	m	buutu	recovery	y	n	y	n	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_31	105	y	y	NA	native	NA	m	buutu	recovery	y	n	y	n	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_32	106	y	y	NA	native	26	m	buutu	recovery	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_33	107	y	y	NA	native	29	m	buutu	recovery	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_34	108	y	y	NA	native	NA	m	forcados	recovery	y	n	n	n	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_35	109	y	y	NA	native	27	m	forcados	recovery	y	n	n	y	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_36	110	y	y	NA	native	24	m	forcados	recovery	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_37	111	y	y	NA	native	12	m	forcados	recovery	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_38	112	y	y	NA	native	17	m	forcados	recovery	y	n	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_39	113	y	y	NA	native	13	m	forcados	recovery	y	n	y	y	n	n	n	n	n	y	y	n	n	n	n	n	n	n	n	n	n
1_1_4_2_40	114	y	y	NA	native	22	m	forcados	recovery	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_41	115	y	y	NA	native	16	m	forcados	recovery	y	n	n	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_42	116	y	y	NA	native	26	m	forcados	recovery	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_43	117	y	y	NA	native	24	m	forcados	recovery	y	n	n	y	n	n	n	n	y	n	y	n	n	n	n	n	n	n	n	n	n
1_1_4_2_44	118	y	y	NA	native	22	m	forcados	recovery	y	n	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_45	119	y	y	NA	native	23	m	forcados	recovery	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_46	120	y	y	NA	native	24	m	forcados	recovery	y	n	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_47	121	y	y	NA	native	NA	m	forcados	recovery	y	n	y	y	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_48	122	y	y	NA	native	24	f	forcados	recovery	y	n	y	y	n	n	n	y	n	y	n	n	n	n	n	n	n	n	n	y	y
1_1_4_2_49	123	y	y	NA	native	NA	m	forcados	recovery	n	n	y	y	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_50	124	y	y	NA	native	NA	m	forcados	recovery	y	n	y	y	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_51	125	y	y	NA	native	28	m	forcados	recovery	y	n	y	n	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_52	126	y	y	NA	native	26	m	forcados	recovery	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_53	127	y	y	NA	native	21	m	forcados	recovery	y	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_54	128	y	y	NA	native	24	m	forcados	recovery	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_55	129	y	y	NA	native	28	m	forcados	recovery	y	n	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_56	130	y	y	NA	native	28	m	forcados	recovery	y	n	n	n	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_57	131	y	y	NA	native	34	m	forcados	recovery	y	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_58	132	y	y	NA	native	16	m	forcados	recovery	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_59	133	y	y	NA	native	29	m	forcados	recovery	y	n	y	y	n	n	n	n	y	y	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_60	134	y	y	other	native	38	m	forcados	recovery	y	n	n	n	n	n	n	y	y	n	n	n	n	n	n	n	n	y	n	n	n
1_1_4_2_61	135	y	y	NA	native	24	m	forcados	recovery	y	n	y	y	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_62	136	y	y	other	native	25	m	forcados	recovery	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_63	137	y	y	NA	native	20	m	forcados	recovery	y	n	n	y	n	n	n	y	n	y	n	n	y	n	n	n	n	y	n	n	n
1_1_4_2_64	138	y	y	NA	native	23	m	forcados	recovery	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_2_65	139	y	y	NA	native	23	m	forcados	recovery	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_3_66	158	y	n	malaria	immigrant	28	m	lagos	recovery	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_3_67	159	y	n	malaria	immigrant	40	m	lagos	recovery	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_3_68	160	y	n	malaria	immigrant	34	m	lagos	recovery	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_3_69	162	y	y	malaria	immigrant	NA	m	lagos	recovery	y	n	n	y	y	n	n	n	n	n	n	n	n	n	n	n	n	y	n	n	n
1_1_4_3_70	163	y	n	malaria	immigrant	33	m	lagos	recovery	y	n	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_3_71	164	y	n	malaria	immigrant	41	m	lagos	recovery	y	n	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_3_72	165	y	y	malaria	immigrant	28	m	lagos	death	y	y	y	y	y	y	y	n	n	y	n	y	n	n	n	n	n	y	n	n	n

DAVIS EXPEDITION FUND

"condensed" metadata continued – 2 of 2.

1_1_4_3_73	181	y	y	malaria	immigrant	NA	m	lagos	death	y	n	n	n	y	y	y	n	y	y	n	y	y	n	n	n	n	n	n	n	n
1_1_4_3_74	187	y	y	NA	immigrant	NA	m	warri	recovery	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_3_75	188	y	y	NA	immigrant	NA	m	warri	recovery	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_3_76	189	y	y	NA	immigrant	NA	m	warri	recovery	n	n	n	n	y	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_3_77	191	y	y	typhoid	immigrant	28	m	sapele	NA	y	n	y	y	y	y	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n
1_1_4_3_78	193	y	y	typhoid	immigrant	40	m	sapele	death	y	n	y	y	y	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n
1_2_1	218	y	y	yellow	immigrant	28	m	lagos	death	y	y	y	y	n	n	y	n	y	y	y	y	n	y	n	y	n	y	y	n	n
1_2_2	223	y	y	yellow	native	25	m	lagos	recovery	y	y	y	y	y	y	n	n	n	y	y	n	y	n	y	n	y	n	y	n	n
1_2_3	225	y	y	yellow	native	29	m	lagos	recovery	y	y	y	y	n	n	n	n	n	n	y	n	y	n	y	n	y	n	y	n	n
1_2_4	226	y	y	yellow	native	22	m	lagos	recovery	y	y	y	y	n	n	n	n	n	n	y	y	y	n	y	n	y	n	y	n	n
1_2_5	228	y	y	yellow	native	26	m	lagos	recovery	n	y	y	y	n	n	n	n	n	n	y	n	y	n	y	n	y	n	y	n	n
1_2_6	230	y	y	malaria_yell	immigrant	21	m	lagos	recovery	y	y	y	y	n	n	n	y	n	n	y	y	n	n	y	n	y	y	y	n	n
1_2_7	232	y	y	yellow	immigrant	34	m	lagos	death	y	y	y	y	n	n	y	n	y	y	y	n	y	n	y	n	y	n	y	y	n
1_2_8	235	y	y	yellow	immigrant	40	m	lagos	death	y	y	n	y	n	y	y	n	y	n	y	y	n	n	y	y	y	n	y	y	y
1_2_9	238	y	y	yellow	immigrant	24	f	lagos	recovery	y	y	y	y	n	y	n	y	n	n	y	y	y	n	y	n	y	n	y	n	n
1_2_10	241	y	y	yellow	native	29	m	lagos	recovery	y	y	y	y	n	y	y	y	n	n	y	?	y	n	y	n	y	y	y	n	n
1_2_11	242	y	y	yellow	immigrant	35	m	lagos	death	y	y	y	y	n	n	y	y	y	n	y	y	y	n	y	n	y		n	n	n
1_2_12	245	y	y	yellow	native	26	m	lagos	recovery	y	y	y	y	n	n	n	n	y	n	y	n	y	n	y	n	y	n	y	n	n
1_2_13	247	y	y	yellow	native	36	m	lagos	recovery	y	y	y	y	n	n	n	y	n	n	y	n	y	n	y	n	y	y	y	n	n
1_2_14	250	y	y	yellow	immigrant	30	m	lagos	recovery	y	y	y	y	n	n	n	n	n	n	y	n	n	n	y	n	y	n	y	n	n
1_2_15	252	y	y	yellow	native	23	m	lagos	recovery	y	y	y	y	n	n	n	y	n	n	n	y	n	y	n	y	n	y	y	n	n
1_2_16	254	y	y	yellow	native	26	m	lagos	recovery	y	y	y	y	y	n	n	y	n	n	y	y	y	n	y	n	y	y	y	n	n
1_2_17	255	y	y	yellow	native	39	m	lagos	recovery	y	y	y	y	y	n	y	n	n	n	n	n	y	n	y	n	y	y	y	n	n
1_2_18	257	y	y	yellow	native	21	m	lagos	recovery	y	y	y	y	n	n	n	n	n	n	y	n	y	n	y	n	y	n	y	n	n
1_2_19	259	y	y	yellow	native	20	m	lagos	recovery	y	y	y	y	n	n	n	n	n	n	n	n	y	n	y	n	y	n	y	n	n
1_2_20	260	y	y	yellow	native	29	m	lagos	recovery	y	y	n	y	n	n	n	y	n	n	y	n	y	n	y	n	y	y	y	n	n
1_2_21	263	y	y	yellow	native	27	m	lagos	recovery	y	y	y	y	n	n	n	n	n	n	y	y	y	n	y	n	y	n	y	n	n
1_2_22	265	y	y	yellow	native	22	m	lagos	recovery	y	y	y	y	y	n	n	y	n	n	y	n	y	n	y	n	y	y	y	n	n
1_2_23	267	y	y	yellow	native	27	m	lagos	recovery	y	y	n	y	n	n	n	y	n	n	n	n	y	n	y	n	y	y	y	n	n
1_2_24	269	y	y	yellow	native	21	m	lagos	recovery	y	y	y	y	n	n	n	n	n	n	y	n	n	n	y	n	y	n	y	n	n
1_2_25	270	y	y	yellow	native	36	m	lagos	recovery	y	y	y	y	n	n	n	y	n	n	n	n	y	n	y	n	y	y	y	n	n
1_2_26	272	y	y	yellow	native	29	m	lagos	recovery	y	y	y	y	n	n	n	y	n	n	y	n	n	n	y	n	y	y	y	n	n
1_2_27	273	y	y	yellow	immigrant	21	m	lagos	recovery	y	y	y	y	n	n	n	n	n	n	n	n	y	n	y	n	y	n	y	n	n
1_2_28	275	y	n	yellow	immigrant	21	m	lagos	death	y	y	n	y	n	n	n	n	y	y	n	n	n	n	n	n	n	n	n	n	n
1_2_29	277	y	y	yellow	immigrant	20	m	lagos	recovery	y	y	y	y	n	n	n	n	n	n	n	n	y	n	y	n	y	n	y	n	n
1_2_30	279	y	y	yellow	immigrant	37	m	lagos	recovery	y	y	y	y	n	n	n	y	n	n	y	n	y	n	y	n	y	y	y	n	n
1_2_31	281	y	n	yellow	immigrant	27	m	lagos	death	y	y	n	y	n	n	y	y	n	n	n	n	y	y	n	y	n	y	y	n	y
1_2_32	285	y	n	yellow	immigrant	NA	m	lagos	death	y	n	n	n	n	n	y	n	n	n	n	y	n	n	n	n	n	n	n	y	n
1_2_33	286	y	n	NA	immigrant	32	m	lagos	death	n	n	n	n	y	n	y	n	y	y	n	n	n	n	n	n	y	n	n	y	n
1_2_34	288	y	y	malaria	immigrant	NA	m	lagos	death	n	n	n	n	y	y	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n
1_2_35	290	y	y	yellow	immigrant	21	m	lagos	recovery	y	n	y	n	n	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n
1_2_36	309	y	y	yellow	immigrant	30	m	lagos	recovery	y	n	y	y	n	n	n	n	n	n	y	y	n	n	n	n	n	n	n	n	n
1_2_37	313	y	y	yellow	immigrant	30	m	lagos	recovery	y	y	y	y	y	n	y	n	n	n	n	n	y	n	y	n	y	y	n	n	n
1_2_38	315	y	y	yellow	immigrant	27	m	lagos	recovery	y	y	y	y	y	n	n	n	n	n	y	n	n	n	y	n	y	n	y	n	n
1_3_1	328	y	y	yellow	immigrant	19	m	guayaquil	recovery	y	n	y	y	n	n	n	n	n	n	y	y	y	y	y	y	n	n	n	n	y
1_3_2	331	y	y	yellow	immigrant	25	f	guayaquil	recovery	y	n	y	y	n	n	n	n	n	n	y	y	y	n	y	y	y	n	n	y	y
1_3_3	334	y	y	yellow	immigrant	20	m	guayaquil	death	y	n	y	n	n	n	n	n	n	n	y	y	y	n	y	y	y	n	n	n	y
1_3_4	336	y	y	yellow	immigrant	19	m	guayaquil	recovery	y	n	y	n	n	n	n	n	n	n	y	y	y	y	y	y	y	n	n	n	y
1_3_5	337	y	y	yellow	immigrant	18	f	guayaquil	recovery	y	n	y	n	n	n	n	n	n	n	y	n	n	n	y	n	y	n	n	n	n
1_3_6	340	y	y	yellow	immigrant	33	m	guayaquil	recovery	y	n	y	n	n	n	y	n	n	n	y	y	y	n	y	n	y	n	y	n	n
1_3_7	342	y	y	yellow	immigrant	26	m	guayaquil	death	y	n	y	y	n	y	n	n	n	n	y	y	y	n	y	n	y	n	y	n	n
1_3_8	344	y	y	yellow	immigrant	28	f	guayaquil	recovery	y	n	y	y	n	n	n	n	n	n	y	n	y	n	y	n	y	n	y	n	y
1_3_9	347	y	y	yellow	immigrant	20	m	guayaquil	recovery	y	n	y	y	n	n	n	n	n	n	y	y	y	n	y	n	y	n	n	n	n
1_3_10	349	y	y	yellow	immigrant	20	m	guayaquil	recovery	y	n	y	n	n	n	n	y	n	n	y	y	y	n	y	y	y	n	n	y	y
1_3_11	352	y	y	yellow	immigrant	18	m	guayaquil	death	y	n	y	y	n	y	y	n	n	n	y	y	y	n	y	n	y	n	n	y	n
2_10_1	576	y	y	yellow	immigrant	32	m	freetown	recovery	y	n	y	y	y	y	y	n	n	n	y	y	y	n	n	y	y	n	n	n	y
2_11_p589	589	y	y	yellow	immigrant	28	f	qaittah	death	y	y	y	y	n	y	y	n	y	n	y	y	y	n	y	n	y	n	y	y	y
2_13_1	663	y	y	malaria	immigrant	37	m	sekondi	recovery	y	n	y	y	y	y	n	y	n	y	y	n	y	n	y	n	y	y	n	n	n
2_13_2	665	y	n	malaria	immigrant	34	m	sekondi	recovery	y	n	y	y	y	n	n	y	n	n	y	n	y	n	y	y	y	y	n	n	y
2_13_3	666	y	y	malaria	immigrant	29	m	sekondi	recovery	y	n	y	y	y	n	n	y	n	n	y	n	y	n	y	n	y	y	n	n	n
2_13_4	668	y	y	malaria	immigrant	26	m	sekondi	death	y	n	y	y	y	y	y	y	n	n	y	n	y	n	y	n	y	y	n	n	n
2_13_5	670	y	y	malaria	immigrant	31	m	sekondi	recovery	y	n	y	y	y	y	n	y	n	n	n	y	y	n	y	n	y	y	n	n	n
2_13_6	672	y	y	malaria	immigrant	47	m	sekondi	recovery	y	n	y	y	y	n	n	y	y	n	y	y	y	n	y	n	y	y	n	n	n
2_13_7	674	y	y	malaria	immigrant	37	m	sekondi	recovery	y	n	y	y	y	y	n	y	n	n	y	n	y	n	y	y	y	y	n	n	y
2_13_8	675	y	y	malaria	native	30	m	sekondi	recovery	y	n	y	y	n	n	n	y	n	n	y	n	y	n	y	n	y	y	n	n	n
2_13_9	677	y	y	NA	immigrant	25	m	sekondi	recovery	n	n	y	y	n	y	n	n	y	n	y	n	y	y	y	n	n	n	n	n	y
2_13_10	680	y	y	NA	immigrant	39	m	sekondi	recovery	y	n	y	y	y	y	n	n	y	n	n	n	y	n	y	n	y	n	n	n	n
2_13_11	719	y	y	yellow	native	19	m	sekondi	death	y	y	y	y	n	n	n	y	y	y	y	n	y	y	y	n	y	n	n	n	y
3_7_1	279	y	y	other	native	25	m	qaittah	NA	y	n	n	n	n	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n
3_7_2	282	y	y	yellow	native	26	m	qaittah	recovery	y	y	y	y	n	y	y	y	n	y	y	n	n	n	n	y	n	y	y	n	n
3_7_3	284	y	y	other	native	26	m	qaittah	recovery	y	y	y	y	n	y	n	y	n	y	y	y	n	y	n	n	n	n	n	n	n
3_7_4	286	y	y	typhoid	native	45	m	qaittah	recovery	y	y	n	y	n	n	y	y	n	n	n	n	n	n	n	n	n	n	n	y	n
3_7_5	289	y	y	yellow	native	26	m	qaittah	recovery	y	n	y	y	y	n	n	y	n	n	y	n	n	n	n	n	n	y	n	y	n
3_7_6	291	y	y	typhoid	native	17	f	qaittah	recovery	y	n	n	y	n	y	n	y	n	n	n	n	n	n	n	n	n	n	y	n	n

DAVIS EXPEDITION FUND

Graphs of "binned" amplitudes – the changes in temperatures and pulses of each case on a 24-hour basis across the case duration.