A study of various biochemical parameters in patients with scrub typhus

Ashwini Manish Jantikar.1,*
1 Dept. of Biochemistry, American International Institute of Medical Sciences, Udaipur, Rajasthan, India

ABSTRACT

Background: scrub typhus is an infectious disease spreads by a bacterium called Orientiatsutsugamushi. The WHO has dubbed scrub typhus one of the world’s most under-diagnosed/underreported diseases. It’s so also in Rajasthan (India). Since antigen detection tests have low sensitivity/specificity and require biopsy specimens, in the clinical setting, serological assays are the mainstay of diagnosis.

Objectives: To study various biochemical parameters in Scrub typhus positive patients and to compare of the same with the control group (negative suspects) to elicit some specific diagnostic parameters through data mining.

Materials and Methods: This retrospective observational cohort study included Fever of unknown origin (FUO) cases classified into patients with (group 1; n = 35) or without (group 2; n= 65) scrub typhus. Various biochemical parameters were correlated and compared through data mining and analyzed on Apache Hive. We used Welch test to estimate normality of data, Fligner Killeen test to estimate the homogeneity of variance and Mann Whitney U test for comparing the two data sets.

Results: Both the groups were comparable with respect to age and sex. Platelet count, SGOT, SGPT and random blood sugar were significantly different in both the groups. (p<0.05) Other laboratory parameters viz. hemoglobin, WBC counts, RBC counts, bilirubin, urea and creatinine were not different significantly among group 1 and group 2. (p>0.05)

Conclusion: Thrombocytopenia and deranged SGOT-SGPT were found to be prime clinico-pathological indicator of the scrub typhus.

1. Introduction:

Scrub typhus (also known as Bush typhus) is an infectious disease spreads by a bacterium called Orientiatsutsugamushi (Japanese name of the species tsutsugamushi comprises tsutsuga i.e. illness and mushi i.e. insect) belonging to order Rickettsiales, family Rickettsiaceae (phylum proteobacteria, class alpha proteobacteria). It is a natural obligate intracelluar gram negative bacterial parasite in mites of Trombiculidae family.¹ This Rickettsial infection is a zoonotic acute febrile illness spread by the chigger mite (Leptotrombidiumdeliense).²

Orientiatsutsugamushi is the most commonly reported rickettsial infection in Indian subcontinent. The mortality of untreated epidemic can range from 30% to 60%.³,⁴ In untreated cases, fatality can rise to 70% as reported in a article.⁵ The disease is currently estimated to impact 1 billion populations globally with 1 million casualties.⁶

Scrub typhus is grossly under diagnosed in India because of their nonspecific presentations, less awareness in patients as well as in clinicians (esp. that serologic testing is not useful acutely) and lack of facilities for diagnosis and treatment in periphery (like PCR based diagnosis).⁷ The World Health Organization has dubbed scrub typhus one of the world’s most under-diagnosed/underreported diseases that often requires hospitalization.⁸ Currently, there is a widespread re-emergence of scrub typhus in India,
Micronesia, and the Maldives. Reports are there which say that this expansion of geographies in India is more likely to be due to increased awareness and heightened level of suspicion rather than change in agent, host or environment.9

Moreover, O. tsutsugamushi serotype distribution varies from region to region and strain types are identified by sequencing the 56 kDa gene. In India, based on a 56 kDa analysis, strains similar to Kato and Karp are common. Eschar, as widely accepted cornerstone of clinical diagnosis, is not universal too, Scrub typhus without the eschar is a febrile illness without any evidence of localization, and is hence termed “acute undifferentiated fever”.10 This illness is thus clinically indistinguishable from malaria, dengue fever, other rickettsioses, leptospirosis, and enteric fever, which are common causes of acute undifferentiated fever.10 Timely diagnosis of the disease can aid the clinician in instituting proper treatment in patients suspected of suffering from scrub typhus and may help to reduce the morbidity and mortality of this disease.11

Due to lack of awareness, a low index of suspicion among clinicians, paucity of confirmatory diagnostic facilities and clinical symptoms mimicking other more prevalent diseases such as dengue, malaria and leptospirosis, scrub typhus is under-diagnosed in India, especially in Rajasthan.12 Since antigen detection tests have low sensitivity/specificity and require biopsy specimens, in the clinical setting, serological assays are the mainstay of diagnosis.13 The cheapest and most easily available serological test is the Weil-Felix (WF) test which has a high specificity but a low sensitivity.14 Thus serologic testing alone is insufficient because of substantial background seropositivity in regions where it is endemic and because of the delay during a primary infection before antibodies are produced.15

This study aims to explore some correlation of ‘laboratory parameters (through data mining)’ to the finally diagnosed cases - for increasing the diagnostic efficacy. Instead of an already done prospective observational study,12 this retrospective study could be much more time/resource efficient. This retrospective study was based on a secondary data through which we have tried to elicit a pattern – but no pre-hypothesized pattern was rigged to reach the required p-value. Here lies the difference between our statistically valid data mining and malicious data dredging.

Contextually, the unbiased methodology of data mining (also called information harvesting or knowledge discovery) is different from data dredging (or data fishing) which joggles with data and hunt for a pattern till a correlation frame emerges (and thus the latter is derogatorily called p-hacking, data snooping or even data butchering).16 To the contrary, here the dark data (unused data that is otherwise discarded) is revisited (for cluster analysis, pattern hunt and anomaly location) to see if anything unexpected is left overlooked.17,18 In 1995, the First International Conference on Data Mining and Knowledge Discovery (KDD-95) was started in Montreal under AAAI sponsorship.16

Starting from Bayes’ theorem (1700s) and regression analysis (1800s); data mining has evolved through neural networks, cluster analysis, genetic algorithms (1950s), decision trees and decision rules (1960s), and support vector machines (1990s). Currently data mining involves the 6 steps of anomaly detection, association rule learning (dependency modeling), clustering, classification, regression and summarization.19 So, the present was planned to study various biochemical parameters in Scrub typhus positive patients and to compare of the same with the control group (negative suspects) to elicit some specific diagnostic parameters through data mining.

2. Materials and Methods

This observational study was conducted after approval from institutional ethics committee. In this retrospective observational cohort study (typhus versus non-typhus cohorts), already existing laboratory data of all scrub typhus suspect cases from out patients and in-patients departments of tertiary care teaching hospital of Udaipur, Rajasthan between July to December 2019 was collected. Fever of unknown origin (FUO) is said when the body temperature increases to 38.3°C (101°F) or more several times a day lasting longer than 3 weeks or failure to reach a diagnosis despite 1 week of inpatient evaluation.20 Patients admitted with a provisional diagnosis of this FUO were selected for this study, and based on case records, bifurcated finally into scrub typhus cohort (group 1, cases) versus non-typhus cohorts (group 2, controls).

Various biochemical tests were correlated and compared through data mining and analyzed on Apache Hive (technique used by Facebook and Netflix). From the hospital database, we retrieved relevant data series and their correlations through The Apache HiveTM data warehouse software which facilitates reading, writing, and managing large datasets residing in distributed storage and queried using SQL syntax.21 There is not a single "Hive format" in which data must be stored. Hive comes with built in connectors for comma and tab-separated values (CSV/TSV) text files, Apache ParquetTM, Apache ORC™, and other formats. We could extend Hive with connectors for other formats.21

Usually a schema is applied to a table in traditional databases. In such traditional databases, the table typically enforces the schema when the data is loaded into the table. This enables the database to make sure that the data entered follows the representation of the table as specified by the table definition. In comparison, our software Apache Hive does not verify the data against the table schema on write. Instead, it subsequently does run time checks when the data is read. This model is called ‘schema on read’ against other
Laboratory parameters viz. hemoglobin level, WBC count, RBC count, platelet count, total bilirubin, SGOT (=AST) and SGPT (=ALT), serum urea, serum creatinine and random blood sugar were noted down from existing laboratory data of both cases and control group. We used Welch test to estimate normality of data, Fligner Killeen test to estimate the homogeneity of variance and Mann Whitney U test for comparing the two data sets. P value < 0.05 was considered significant.

3. Results

Both the groups were comparable with respect to age and sex. Platelet count, SGOT, SGPT and random blood sugar were significantly different in both the groups. (p<0.05) Other laboratory parameters viz. hemoglobin, WBC counts, RBC counts, bilirubin, urea and creatinine were not different significantly among group 1 and group 2. (p>0.05)

Platelet count was found significantly diminished in case group as compared to control. (Figure 1 and Table 1) Deranged liver profile (SGOT, SGPT) was also significantly high in case group as compare to controls. (Figures 2 and 3 and Table 1) Serum creatinine and blood urea rise was found more in case group but that was non-significant. Significantly high random blood sugar was found in case group. (Figure 4 and Table 1) Seeing wide standard deviation, median (which is the basis of comparative box whisker plot) was preferred over mean based comparator.

4. Discussion

Although Scrub typhus is a neglected disease in India, but in recent years there are reports from Maharashtra, Tamil Nadu, Karnataka, Kerala, Jammu and Kashmir, Uttarakhand, Himachal Pradesh, Rajasthan, Assam and West Bengal indicating the resurgence. Thrombocytopenia is the single most common hematological abnormality; severe thrombocytopenia with platelet count less than 50x10^3 /μL is seen in up to half of patients. Lowering of platelet count (thrombocytopenia)> 60% is also on record. In our study it was 65%
(12/35) in case group against 49% (32/65) in the control group. Here exclusion of differential diagnosis of dengue is warranted.

It was also reported that scrub typhus is the most common cause of febrile jaundice in a tertiary care hospital. Hepatomegaly is reported in up to 70% cases while SGOT/SGPT is increased in 49% cases. Another study reported rise in SGOT/SGPT in 68.75% cases while 30% incidence of hepatospleno-megaly. Yet another study reported around 90% cases of scrub typhus having higher SGOT/SGPT as well as bilirubin. In our study SGOT/SGPT and bilirubin rise were parameters of the significant differences between case and control.

Like our significantly different serum creatinine in patient group, serum creatinine is reported raised in >40% cases. But in contrast, leucocytosis (which is highly non-specific and might be similarly increased in the suspects due to other unhygienic pathologies) was not significantly different in our study. Significantly higher incidence of high random blood sugar in case group was unique in this study which was epidemiologically less tested or established. Similar rise in random blood sugar was seen in some other researches (114 mg/dL and 165 mg/dL respectively) but no causation or correlation was reported.

Other parameters were not different significantly. For example gender propensity was favoring males or females in different studies while in our study, the difference was statistically insignificant. In contrast to this prospective study, as our study was retrospective, instead of incidence confirmation, we focused on comparator validation.

Acute undifferentiated febrile illness of three days or more with or without organ involvement during typical tropical rainy season should be suspected as a case of scrub typhus. Presence of eschar is pathognomonic and is a useful diagnostic clue. However, at acute care settings, several other tropical infections such as dengue, malaria, typhoid, leptospirosis and severe bacterial sepsis may present with overlapping clinical features and may be confused with scrub typhus. The challenge lies in distinguishing them at the time of presentation.

Both dengue and scrub typhus present with thrombocytopenia, signs of capillary leak and circulatory abnormalities, subtle laboratory features like presence of hemoconcentration or leucopenia may help in discriminating dengue to a certain extent. Dengue and malaria can be diagnosed at admission using point of care rapid diagnostic tests. However, if a definitive diagnosis is not possible at the outset, it is recommended to treat children empirically for scrub typhus till serological confirmation is available.

### 5. Limitations

There is a limitation of this data mining method too. Conventional tests of statistical significance are based on the probability that a particular result would arise if chance alone were at work, and necessarily accept some risk of mistaken conclusions of a certain type (mistaken rejections of the null hypothesis). This level of risk is called the significance. Though we have taken precaution to avoid over-fitting, a conventional null hypothesis method is necessary to be replicated on a newer larger data set before these results are clinically generalized.

### 6. Conclusion

In our study, low platelet count (thrombocytopenia) and deranged liver profile (increased SGOT-SGPT) were found to be prime clinico-pathological indicator of the scrub typhus. Random blood sugar was significantly different in scrub typhus patients; relevance of which could be as precipitating or prognostic factor (as diabetic profile has not been correlated directly to scrub typhus anyhow).

### Table 1: Comparison of demographic and biochemical parameters in both the groups

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Group 1 (Case)(N= 35) mean± SD</th>
<th>Group 2 (Control)(N= 65) mean± SD</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (Years)</td>
<td>38.11 ± 15.66</td>
<td>39 ± 7.79</td>
<td>0.741</td>
</tr>
<tr>
<td>Male/Female ratio</td>
<td>13/22</td>
<td>26/39</td>
<td>0.949</td>
</tr>
<tr>
<td>Hb (mg/ dL)</td>
<td>10.28 ± 2.68</td>
<td>11.4 ± 2.09</td>
<td>0.339</td>
</tr>
<tr>
<td>WBC (x10³/ dL)</td>
<td>7.11 ± 2.73</td>
<td>7.06 ± 3.92</td>
<td>0.397</td>
</tr>
<tr>
<td>RBC(x 10⁹/μL)</td>
<td>3.9 ± 0.992</td>
<td>4.2±0.64</td>
<td>0.128</td>
</tr>
<tr>
<td>Platelet (x10³/ dL)</td>
<td>133.26 ± 132.27</td>
<td>168.38 ± 114.45</td>
<td>0.023*</td>
</tr>
<tr>
<td>Bilirubin(mg/dL)</td>
<td>1.21 ± 1.44</td>
<td>1.098 ± 1.32</td>
<td>0.051</td>
</tr>
<tr>
<td>SGOT(units/ L)</td>
<td>139.4 ± 132.26</td>
<td>103.88 ± 140.73</td>
<td>0.039*</td>
</tr>
<tr>
<td>SGPT(units/ L)</td>
<td>105.89 ± 153.3</td>
<td>82.14 ± 108.03</td>
<td>0.032*</td>
</tr>
<tr>
<td>Urea (mg/dL)</td>
<td>49.743 ± 41.82</td>
<td>39.88 ± 38.14</td>
<td>0.054</td>
</tr>
<tr>
<td>Creatinine(mg/dL)</td>
<td>1.121 ± 0.564</td>
<td>1.054 ± 0.4255</td>
<td>0.053</td>
</tr>
<tr>
<td>RBS(mg/dL)</td>
<td>122.49 ± 58.31</td>
<td>103.5 ± 21.053</td>
<td>0.048*</td>
</tr>
</tbody>
</table>

*significant p value
7. Source of Funding
No financial support was received for the work within this manuscript.

8. Conflict of Interest
The authors declare they have no conflict of interest.

References


Author biography
Ashwini Manish Jantikar, Associate Professor