



universität
wien

DISSERTATION

Titel der Dissertation

**Evaluation of Rapid Diagnostic Tests for Typhoid Fever and Cholera in Sub
Saharan Africa**

Verfasser

Mag. rer. nat. Benedikt Nikolaus Ley

angestrebter akademischer Grad

Doctor of Philosophy (PhD)

Wien, 2012

Studienkennzahl lt. Studienblatt: A 094 437

Dissertationsgebiet lt.
Studienblatt: Biologie

Betreuer: ao. Univ.-Prof. Dr. Harald Wilfing

To Kamala, David, and Sophie

Acknowledgement

This work would not have been possible without the ongoing support of my supervisor Harald Wilfing.

Lorenz v. Seidlein and Jacqueline Deen have extensively contributed to all phases of the research described below and have provided their expertise and thoughts throughout the entire period of my PhD.

Franz Jirsa has contributed to this work through his critical review of this thesis.

My sister Johanna Trankovits has, with great efficiency and reliability, supported this project in all practical aspects related to this thesis.

My beautiful wife, Kamala Thriemer, has been on the ground with me throughout the last years, has been the best research partner possible and the most critical reviewer.

Thanks a lot to all of you!

Table of Contents

Introduction:	6
Global Situation	6
Favoring circumstances and possible interventions	7
Funding Dilemma	8
Laboratory Confirmation	8
Aim of this Thesis	9
Background.....	10
Rapid Diagnostic Tests.....	10
Performance Indicators.....	13
Gold Standard and control groups.....	14
Results:	15
Evaluation of the Widal tube agglutination test	15
Evaluation of the TUBEX test	32
Evaluation of the Crystal VC.....	46
Discussion:.....	64
Conclusion	68
References:.....	69
Summary	74
Zusammenfassung.....	77
Curriculum Vitae.....	81
Appendix.....	84
All three publications in online format:	84

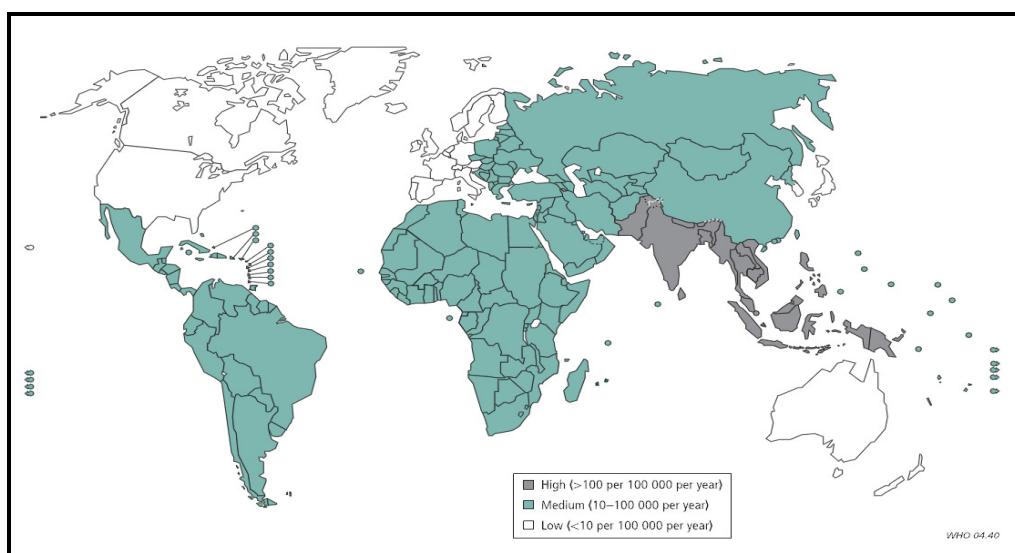
Introduction:

Global Situation

Despite substantial progress in hygiene and sanitation and a significant reduction of typhoid fever and cholera in most developed countries, both diseases remain of immanent importance and a permanent threat in many developing countries.

Crump et al (2004) reviewed 22 studies on burden of disease of typhoid fever published in peer reviewed journals from around the globe and conclude that more than 21.000.000 episodes of typhoid fever occurred in the year 2000 (see fig. 1), with more than 215.000 deaths. The same study highlights the differences in burden of disease between developing and developed countries, categorizing Europe, North America, Australia, and New Zealand as areas with <10 episodes of typhoid fever per 100.000 population, whereas south- central and south – east Asia where considered as high endemic (>100 / 100.000) and the rest of the world as medium incidence (10-100 / 100.000) (Crump et al, 2004). More recent publications suggest that these numbers may be an underestimate for Africa. Thriemer et al. (2012) report incidence rates as high as 110/ 100.000 population from Zanzibar, Tanzania; Breiman et al (2012) observed crude incidence rates of 247 / 100.000 person years of observation in an urban slum in Kenya (Breiman et al, 2012).

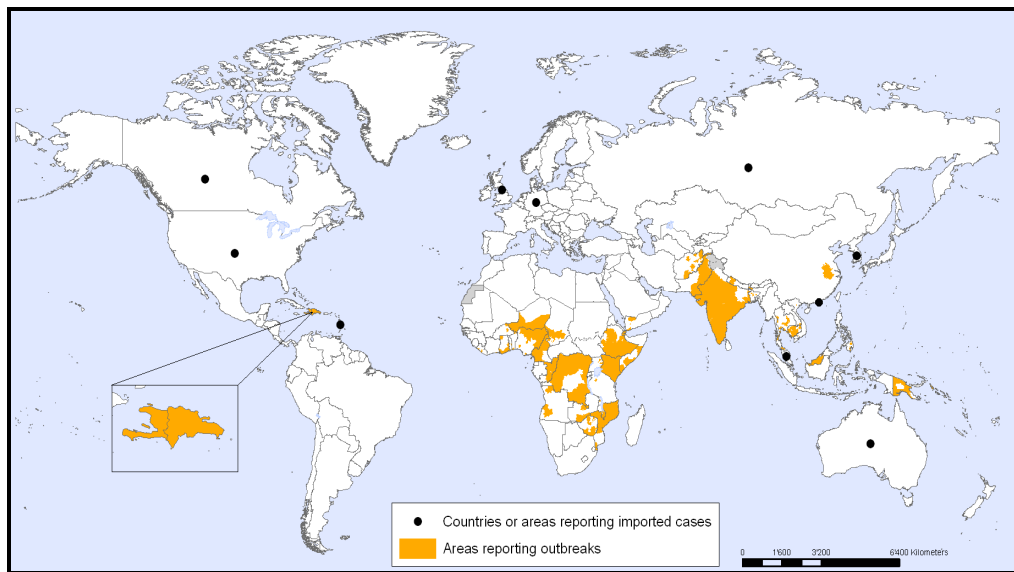
Fig. 1: Geographical distribution of typhoid fever, 2000



From: Crump, J. A., S.P. Luby, E. D. Mintz. (2004): The global burden of typhoid fever. *Bulletin of the World Health Organization* 82 (5):346—353.

Similar to typhoid fever, cholera is endemic in developing countries only (see fig. 2). While neither Europe, the Northern Americas, Australia, or New Zealand report non-imported, local cholera cases, the main burden of disease lies within Africa, the Indian subcontinent and Haiti. David Sack (2012) estimated 2.55 mio. cases globally / year with a case fatality rate of >8.2%. He concludes that case fatality rates are almost threefold higher in Africa (11.9%), compared to Asia (4.1%).

Fig.2: Areas reporting Cholera outbreaks 2010-2011 to the WHO



From: World Health Organization (2012): *Cholera, areas reporting outbreaks, 2010-2011* [online]:

http://gamapserver.who.int/mapLibrary/Files/Maps/Global_Cholera_ITHRisk_20120118.png (last accessed on 19.03.2012)

Favoring circumstances and possible interventions

Typhoid fever and cholera are transmitted via fecal-oral infection (WHO, 2009a; WHO 2010) and are poverty associated diseases (Miller et al, 2000; Zuckerman et al, 2007). Informal settlements, refugee camps, and disaster zones are prone to outbreaks of either disease, due to the lack of access to safe and clean water and a functioning sewage system. The cholera outbreaks in 2011 / 2012 in Haiti have dramatically shown how the destruction of local infrastructure may result in an outbreak situation; studies conducted in the years 2007 – 2009 in Kibera, Kenya (Breiman et al, 2012), one of Africa's biggest slums, demonstrate the permanent danger of a typhoid fever infection for people with limited access to drinking and no sewage water system.

While improving local infrastructure is a sustainable long term solution, this is often not possible due to budget limitations in developing countries or the character of a sudden disaster. Since both diseases,

typhoid fever and cholera, are preventable through vaccination (Lesser and Miller in Kasper et al, 2005a; Waldor and Keusch in Kasper et al, 2005b) reactive vaccination campaigns may be considered effective in case of an outbreak as has been shown through modeling for cholera (Reyburn et al, 2011). Since both diseases are curable (Ghenghesh et al, 2009; WHO, 2009a) by standard antibiotics (typhoid fever) and rehydration (cholera) if detected soon after onset of symptoms (Lesser and Miller in Kasper et al, 2005a; Waldor and Keusch in Kasper et al, 2005b; WHO, 2009b; WHO 2010; Baker et al, 2010), an early and precise diagnosis of disease is essential. Precise and accurate diagnosis will support planning of preventive and reactive measures by providing reliable data on disease burden and trends.

Funding Dilemma

Many areas suffering from high incidence rates of typhoid fever and cholera are facing the dilemma that both diseases have a negative impact on the affected populations' macro-economic situation, resulting in insufficient funding for diagnostic facilities and staff. Insufficient funding for diagnosis and treatment in return results in high disease incidence rates, prolonged episodes of disease and high case fatality rates (CFR) due to treatment errors. These factors result in a negative impact on the macro-economy of the affected society (Aguayo-Rico et al, 2005), resulting in low GDP and in turn poor funding for health systems.

One possible solution to this diagnostic dilemma is treatment and reporting based on clinical picture only. However as the clinical picture especially for typhoid fever is subject to a broad variety of symptoms (Lesser and Miller in Kasper et al, 2005a; Thriemer et al, in preparation) empirical treatment based exclusively on the clinical picture is often insufficient, may lead to treatment errors and thereby increases cost of illness.

Since cholera as any episode of acute watery diarrhea is mainly treated symptomatically through fluid replacement (Waldor and Keusch in Kasper et al, 2005b), an accurate and immediate diagnosis is not essential. However, since the route of infection for both diseases is fecal – oral, appropriate and early diagnosis is likely to reduce or inhibit spreading of disease and will support early identification of outbreaks and on-setting epidemics. The impact on public health is therefore significant.

Laboratory Confirmation

In both cases confirmation of disease requires a well equipped laboratory, knowledge on laboratory procedures and sufficient funding for equipment and consumables. While the gold standard for diagnosis of typhoid fever is automated blood culture and agglutination testing, confirmation of *Vibrio cholerae* is

performed by stool culture and agglutination testing (Cheesborough, 2006a; Cheesborough 2006b; Perilla et al, 2003).

Automated blood culture requires a blood culture machine, an incubator as well as further standard laboratory equipment such as a fridge / freezer for storage of consumables, a safety cabinet for sample handling and access to distilled water for media preparation. The process of pathogen isolation and identification requires a number of selective and differential media as well as antisera and may take up to seven days. Excluding running costs such as electricity, labor costs, and purchase of machinery, we found that processing a single sample for blood culture cost approximately 20-30 US Dollars (USD) in our study setting.

Stool culture for *Vibrio cholerae* requires less machinery, apart from an incubator and a fridge for storage as well as access to distilled water and can be completed within 24 – 48 hours. If laboratory procedures are reduced to a minimum, including only selective media, enrichment media and gelatine agar for later agglutination testing as well as antisera, processing one stool sample cost approximately 4 USD in the below described study.

Aim of this Thesis

Considering that many of the above addressed issues can't be provided in a sustainable manner in a number of developing country settings this thesis addresses the problem of accurate diagnosis by evaluating three different rapid diagnostic tests (RDTs) under developing country conditions. The outcome of this thesis may provide a basis for decision makers to include / exclude the assessed RDTs for the future in the diagnostic process.

Background

Rapid Diagnostic Tests

Any rapid diagnostic test should fulfill the following requirements if introduced in public health systems:

- High reliability
- High accuracy
- Convenient and simple operational characteristics
- Cost effectiveness
- Good availability

Widal Test

A number of rapid diagnostic tests exist and are available for typhoid fever (Olopoenia & King, 2000). The Widal test, possibly the first rapid diagnostic test developed was first introduced in 1896 by F. Widal (Olopoenia & King, 2000). The test is based on a macroscopically visible agglutination reaction between anti - *Salmonella* Typhi specific antibodies of patient's serum and somatic *S. Typhi* O and flagellar H antigens in a solution (Olopoenia & King, 2000).

Agglutination may be performed on a slide, whereby patient serum is separately mixed with O or H antigen and observed for macroscopically visible agglutination. Of better diagnostic value is the tube agglutination approach, whereby patient serum is diluted in doubling steps, mixed with the respective antigen, incubated in a water bath and then observed for agglutination (Olopoenia & King, 2000). Albeit this approach is more labor intensive specificity is better than with the slide agglutination test (Olopoenia & King, 2000). One of the major obstacles with the Widal test is its poor performance due to cross reactivity and the requirement to establish local cut off titers, which will have to be re-evaluated in regular time intervals and can't be transferred to other settings. Despite its shortcomings the test is widely in use in Sub-Saharan Africa (Olopoenia & King, 2000; Mweu & English, 2008) due to its very competitive price of less than 50 US cents / test.

Tubex Test

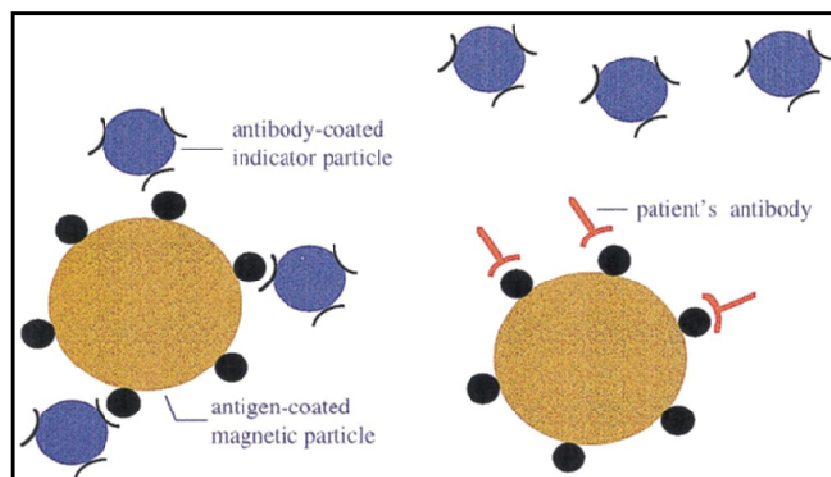
A number of other test kits have been developed over the last twenty years that do not require a local cut off titer, though at a higher price, one of them being the TUBEX (IDL, Sweden). The mode of action for all of these tests may be considered as a variation of the Widal test with techniques adapted from an

Enzyme Linked Immuno Sorbent Assay (ELISA) (Kawano et al, 2007). The test is based on an inhibition reaction between patient antibodies (pAB) and monoclonal antibodies (mAB):

Magnetic beads coated with antigen (Ag) are added to patient serum in a reaction well. If the patient serum contains pAB these will bind to the magnetic beads and will inhibit binding of blue colored beads coated with mAB. If placed on a magnet, the magnetic beads will sink to the bottom of the reaction well, while the blue beads will remain in solution, coloring the solution blue. If no pAB are present, mAB beads will bind to Ag beads, both will sink to the wells bottom and the solution will decolorize.

The blue intensity of the final solution can be compared to a provided color scale, in order to grade the obtained result on a scale from 1-10, whereby any result >3 should be considered positive according to manufacturer.

Figure 3: Mechanism of the Tubex Test.



Lim et al, One-Step 2-Minute Test To Detect Typhoid-Specific Antibodies Based on Particle Separation in Tubes, JOURNAL OF CLINICAL MICROBIOLOGY, Vol. 36 (8) Aug. 1998, p. 2271–2278

The test has three major advantages over the Widal test:

1. The test measures only immunoglobulin M (IgM) and hence only recent infections (Tam and Lim, 2003)
2. The used Ag (O9) is highly specific, cross reactions are less than for the Widal (Lim et al, 1998)
3. The test has a universally applicable cut – off titre.

However the test is significantly more expensive than the Widal test at approximately 2.15 USD / test and does not work on hemolyzed serum samples with the standard procedures described above.

Crystal VC

In 2003 the Institute Pasteur developed a dipstick rapid diagnostic test for cholera. Through licensure agreement the test was made available to the Indian company Span Diagnostics and is currently produced and distributed through said company under the name Crystal VC.

The test uses immunochromatographic methods, and detects Ag of *V. cholerae* O1 and O139 (Kalluri et al, 2006). In detail the test strip is inserted into (liquid) stool. A pad on the test strip contains mobile mAB conjugated to gold nanoparticles. Through capillary forces the sample is dragged across the mobile mAB, which bind to any Ag present in the stool. When passing over a test line with stationary mAB this second set of mAB binds to a different epitope of the same Ag, immobilizing the Ag-mAB complex. The conjugated gold nanoparticles appear as a positive test line. Further upstream is a third set of mAB, which bind to the mobile mAB, whereby a test line appears, and confirms correct test procedures. This design (with slight variations) is used for a number of rapid diagnostic tests including pregnancy (urine) and malaria (blood) tests and is considered to be easy in use and handling.

After development by the Institute Pasteur a prototype of the current test was evaluated in 2003 by Nato et al. in Madagascar, successive evaluation studies of the prototype followed (Bhuiyan et al, 2003, Wang et al, 2006, Kalluri et al, 2006). First evaluation studies of the commercial version were performed in 2009 in Guinea Bissau (Harris et al) and 2010 India (Mukherjee et al), however with varying test strips and procedures and small sample sizes of <220 samples. While sensitivity remained >90% in all studies, a decrease in specificities had been observed after Span Diagnostics took up production and it had been suggested that test performance may be influenced by reader qualification and experience (Kalluri et al, 2006).

Performance Indicators

The performance of a diagnostic test can be assessed using a set of standardized indicators (Bruce et al, 2008):

- Sensitivity describes the percentage of correctly identified true positive (TP) samples from all true positives as well as false negative (FN) samples. $(TP/TP+FN)$
- Specificity describes the number of correctly identified true negative samples (TN) among all samples that are truly negative. $(TN/TN+FP)$
- The positive predictive value (PPV) describes the likelihood that an individual with a positive test result has the disease $(TP/TP+FP)$
- The negative predictive value (NPV) indicates the likelihood in percent, that the test correctly excludes a respective diagnosis $(TN/TN+FN)$

Sensitivity and specificity are *retrospective* (Altman and Bland, 1994a) indicators and hence of importance for surveillance while not being applicable for routine diagnosis. Sensitivity and specificity describe the likelihood that a disease positive individual will achieve a positive result (sensitivity) or that a disease negative individual will be tested negative (specificity). Both indicators are independent of disease prevalence within the study population and hence provide the basis for test comparison in between different settings (Altman and Bland, 1994a, Loong, 2003).

PPV and NPV on the other hand are often considered the most important indicators for the practical application of a test. These indicators describe the likelihood that if a test result is obtained this result describes the true presence or absence of disease within the test subject. However both indicators are dependent on prevalence of disease (Altman and Bland, 1994b, Loong, 2003). While PPV will improve with disease prevalence, since the number of FP become a smaller fraction in the presence of high numbers of TP, whereas the NPV will decrease since the fraction of TN among all negative test results is less, whereas this is not the case for FN results. In the presence of low prevalence on the other hand, PPV decreases, while NPV improves.

Both indicators may therefore also be calculated as a function of sensitivity, specificity, and prevalence (Altman and Bland, 1994b):

- $PPV = (sensitivity * prevalence) / (sensitivity * prevalence + (1 - specificity) * (1 - prevalence))$
- $NPV = (specificity * (1 - prevalence)) / ((1 - sensitivity) * prevalence + specificity * (1 - prevalence))$

Gold Standard and control groups

In order to evaluate a rapid diagnostic test a method needs to be defined that confirms or contradicts the rapid test result and is considered as 100% correct. This assumption in itself introduces a bias, the sensitivity of the gold standard will later need to be taken into consideration for analysis. A gold standard hence needs to be a widely accepted method, which is considered superior to other methods and follows a standardized protocol.

For the purpose of the below described studies, automated blood culture was considered as the gold standard for typhoid fever. Using blood culture as a gold standard poses the problem of defining adequate control groups against which the performance indicators can be measured. Blood culture will in the case of typhoid fever identify three categories of samples: *Salmonella* Typhi positives samples, samples with another human pathogenic bacterial blood stream infection and samples with no positive test result. Since blood culture has a sensitivity of 40-80% (Bhutta, 2006; Farooqui et al; 1991; Gilman et al; 1975; Wain et al, 2001; Parry et al, 2002; Wilke et al, 2002), the last category is likely to consist of TN as well as FN results which will affect performance calculations. To address this problem different control groups have been defined for below projects, namely all blood culture negative results as well as all cases where a human pathogenic bacterial infection had been isolated. Since samples were collected in the course of a fever study at local hospitals we assumed that it is unlikely that if a likely bacterial cause for the febrile episode could be identified, the participant in addition suffered from an unrecognized *S. Typhi* infection.

Since stool culture was chosen as the gold standard for the diagnosis of cholera, it was not possible to define control groups in a similar manner since all samples are streaked out on highly selective media (TCBS) initially (Murray et al, 2007). It was hence not possible to define comparable categories, since TCBS inhibits most other growth apart from *Vibrio spp.* (Murray et al, 2007). In contrast to collected blood samples, stool samples for this study were selected in the course of a cholera surveillance project, inclusion criteria were defined to exclude as many other possible diarrheal pathogens. Categories applied here therefore were growth or no growth. To reduce the number of FN enrichment media was used prior to plating.

Results:

Results of the performance assessment for Widal, Tubex, and Crystal VC are presented in the below manuscripts. All manuscripts have been published in international peer reviewed journals.

BMC Infectious Diseases 2010, **10**:180 doi:10.1186/1471-2334-10-180 (Highly accessed)

Evaluation of the Widal tube agglutination test for the diagnosis of typhoid fever among children admitted to a rural hospital in Tanzania and a comparison with previous studies

Benedikt Ley^{1*}, George Mtove^{2,3}, Kamala Thriemer¹, Ben Amos^{3,4}, Lorenz von Seidlein^{1,3}, Ilse Hendriksen^{3,5}, Abraham Mwambuli³, Aikande Shoo^{3,4}, Rajabu Malahiyo⁴, Shaali M. Ame⁶, Deok R. Kim¹, Leon R. Ochiai¹, John D. Clemens¹, Hugh Reyburn^{3,7}, Harald Wilfing⁸, Stephen Magesa^{2,3}, Jacqueline L. Deen^{1,3}

1 International Vaccine Institute, Seoul, Korea

2 National Institute for Medical Research - Amani Centre, Tanga, Tanzania

3 Joint Malaria Program, Tanga, Tanzania

4 Teule Hospital, Muheza, Tanga, Tanzania

5 Mahidol Oxford Research Unit, Bangkok, Thailand

6 Public Health Laboratory (Pemba) – Ivo de Carneri, Chake Chake, Tanzania

7 London School of Hygiene and Tropical Medicine, Keppel St, London, UK

8 University of Vienna, Biocenter, Vienna, Austria

***Corresponding author:**

Benedikt Ley

International Vaccine Institute

San 4-8 Bongcheon-7-dong, Kwanak-gu

Seoul, Korea 151-818

Email: bnley@ivi.int

Abstract

Background: The diagnosis of typhoid fever is confirmed by culture of *Salmonella enterica* serotype Typhi (*S. typhi*). However, a more rapid, simpler, and cheaper diagnostic method would be very useful especially in developing countries. The Widal test is widely used in Africa but little information exists about its reliability.

Methods: We assessed the performance of the Widal tube agglutination test among febrile hospitalized Tanzanian children. We calculated the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of various anti-TH and –TO titers using culture-confirmed typhoid fever cases as the “true positives” and all other febrile children with blood culture negative for *S. typhi* as the “true negatives.”

Results: We found that 16 (1%) of 1,680 children had culture-proven typhoid fever. A single anti-TH titer of 1:80 and higher was the optimal indicator of typhoid fever. This had a sensitivity of 75%, specificity of 98%, NPV of 100%, but PPV was only 26%. We compared our main findings with those from previous studies.

Conclusion: Among febrile hospitalized Tanzanian children with a low prevalence of typhoid fever, a Widal titer of $\geq 1:80$ performed well in terms of sensitivity, specificity, and NPV. However a test with improved PPV that is similarly easy to apply and cost-efficient is desirable.

Background

Salmonella enterica serotype Typhi (*S. typhi*), the causative agent of typhoid fever, was calculated to have caused approximately 200,000 deaths globally in 2000 [1]. The clinical picture of typhoid fever is nonspecific; confirmed diagnosis through blood or bone-marrow culture requires expensive and labor-intensive isolation and identification of the organism, which may take up to seven days. A cheap and rapid alternative laboratory test is desirable, especially for developing country settings where typhoid fever is a major public health burden.

Various agglutination tests have been developed [2] of which the Widal method is the oldest and remains the most widely used. The test was first introduced by F. Widal in 1896 [2] and is based on a macroscopically visible serum – mediated agglutination reaction between *S. typhi* somatic lipopolysaccharide O antigens (TO) and flagellar H antigens (TH). Laboratories in industrialized countries have stopped performing the assay. In Africa the Widal test is still widely used [3] because typhoid fever

is perceived to be endemic in the area [3] and the Widal test is the only rapid diagnostic assay that is available and affordable. The Widal test is commonly performed when children and adults present with fever to treatment centers, as few centers have the capacity to perform micro-bacterial culture [4]. Despite this widespread use, little has been published on its performance in Africa.

We assessed the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the Widal tube agglutination test among Tanzanian children hospitalized with febrile illness and compared our results with those from previous studies.

Methods

Study site and population. The study was conducted as part of a childhood fever surveillance study at Teule Hospital in Muheza district of Northeastern Tanzania from 2008 to 2009. Muheza district is located between the foothills of Kilimanjaro and the coastal town of Tanga. The area is highly endemic for *Plasmodium falciparum* malaria with perennial transmission and two seasonal peaks [5]. HIV sero-prevalence among antenatal clinic attendees was about 7% in 2007 [6]. Teule Hospital is a busy 330-bed district-level general hospital, serving a surrounding population of 277,000. It has two 35-bed in-patient pediatric wards receiving approximately 5,000 child admissions per year (2008).

Inclusion Criteria: Children aged 2 months to 14 years were screened for eligibility during study hours from 7am to 7pm, Monday to Sunday. Children with fever of 3 or more days prior to admission, or fever of less than 3 days but with at least one severity criteria (respiratory distress, deep breathing, respiratory distress in combination with severe pallor, prostration, capillary refill ≥ 3 seconds, temperature gradient, systolic blood pressure < 70 mm Hg, coma defined by Glasgow Coma Scale ≤ 10 or Blantyre Coma Scale ≤ 2 , severe jaundice, history of 2 or more convulsions in the last 24 hours, blood glucose < 3 mmol associated with clinical signs, neck stiffness, bulging fontanel, or oxygen saturation $< 90\%$) were recruited into the study. All clinical information was recorded on a standard case record form. Treatment was provided according to national guidelines. On admission we collected 3 to 5 milliliters (ml) of blood (depending on body weight) from each eligible child for the Widal test and a single blood culture. All clinical procedures were performed by trained study clinical officers and nurses under the supervision of study physicians.

Laboratory. Blood for culture was inoculated into BacT/ALERT™ Pediatric-fan bottles (bioMérieux, Marcy l'Etoile, France). Inoculated blood culture bottles were transported immediately to the hospital laboratory and incubated in the BacT/ALERT 3D automated microbial detection system. Blood cultures were processed according to standard methods. Colonies with biochemical reactions on API20E suggestive of *Salmonellae* were confirmed serologically by slide and tube agglutination testing using specific O and H antisera (Becton Dickinson, NJ, USA).

A minimum of 0.5 ml of blood was separated to obtain serum samples. All serum samples were frozen at -70°C until Widal testing was done in three batches. Widal testing was performed using standardized TO (IgM and IgG) and TH (IgG) antigens (Stained *Salmonella* Antigens kit, Span Diagnostics, India) according to standard methods as described on the package insert. In brief, each sample was diluted to a concentration of 1:40 with 0.9% NaCl in two separate plastic tubes. A single drop of antigen was added to the respective tube. Incubation times for both O and H agglutinations were 16 to 20 hours at 37°C in a water bath. Evaluation of test results was performed by at least two lab technicians on an independent basis under standardized light conditions. If agglutination was detected in a sample, testing was done on that sample diluted serially from 1:80 to 1:1280 for both O and H antigens. All laboratory procedures were performed by trained laboratory technicians under the supervision of microbiologists. Technicians performing the Widal tests were blinded to the participants' clinical picture and blood culture results.

Data management. Case report forms were double-entered into custom-made data entry programs using MS-Access (Microsoft Corp.). Data management programs included error, range, and consistency check programs. Analyses were performed using EpiInfo v 3.4.3 (Centers for Disease Control and Prevention, USA) and Stata TM v 10.0 (Stata Corp., USA).

Definitions and analysis. Fever was defined as stated history or presence of fever of $\geq 37.5^{\circ}\text{C}$. Bacteremia was defined as fever with isolation of pathogenic bacteria from blood culture. Children with a febrile illness were classified as follows: those with *S. typhi* subsequently isolated from blood culture (group 1), those with non-Typhi serotypes of *S. enterica* (NTS) subsequently isolated from blood culture (group 2), those with pathogenic bacteria other than *Salmonellae* subsequently isolated from blood culture (group 3), and those whose blood culture yielded no bacterial pathogen (group 4). Malaria status was not considered in the classification. In areas of high transmission of *P. falciparum* where individuals develop immunity from previous episodes of malaria starting at a young age, asymptomatic parasitemia is common and may be detected during febrile episodes caused by another infection [7,8].

For the primary analysis, sensitivity (true-positive rate) was defined as the probability that the Widal test result would be positive when blood culture confirmed that typhoid fever was present (group 1) and specificity (true-negative rate) was the probability that the Widal test result would be negative when *S. typhi* was not isolated from blood culture (groups 2, 3, and 4). The positive predictive value was the probability that culture-confirmed typhoid was present when the test was positive, and the negative predictive value was the probability that culture-confirmed typhoid was not present when the test was negative. Since serological tests detect antibody response and perform better after a period of time from the onset of the illness, sensitivity, specificity, PPV, and NPV were also calculated separately for cases presenting with fever for 5 days or less and for more than 5 days. Because controversy exists about what is the most appropriate control group to use [9,10,11], we conducted a secondary analysis using two alternative “true-negative” control groups as follows: those with NTS and other pathogenic bacteria isolated from blood culture (groups 2 and 3), and those with pathogenic bacteria other than *Salmonellae* isolated from blood culture (group 3).

Comparisons were made using the Chi square or Fishers’ Exact test, as appropriate. Sensitivity, specificity, PPV, and NPV were calculated according to standard methods. The 95% confidence interval for sensitivity and specificity was calculated using the Wilson’s Score method [12]. Analyses were performed using EpiInfo v 3.4.3 (Centers for Disease Control and Prevention, USA) and Stata TM v 10.0 (Stata Corp., USA).

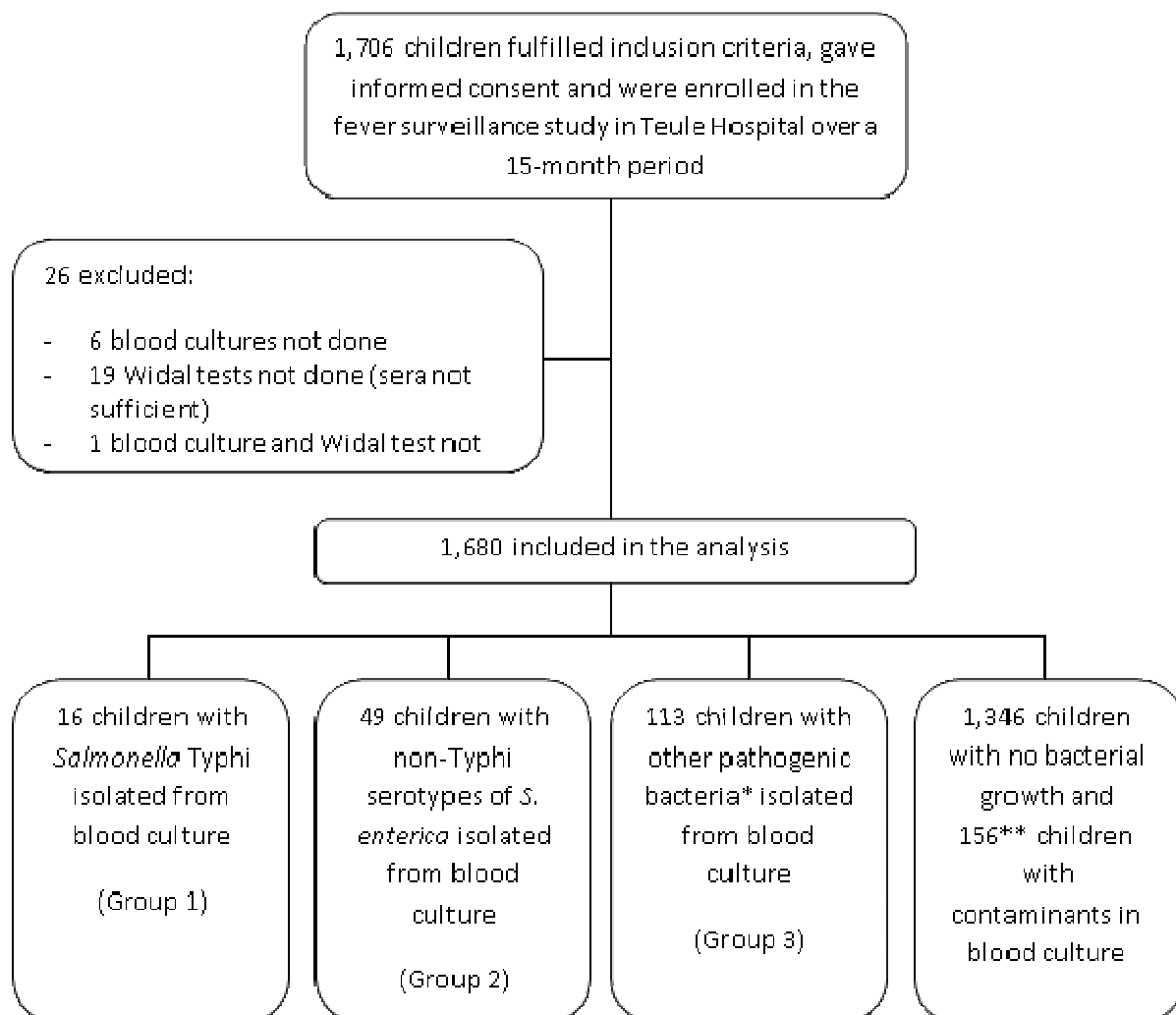
Literature review. We conducted a literature review to compare our main findings with those from previous studies of similar character. We included studies of the Widal test which were identified by direct searches of the MEDLINE database through PubMed. The searches were restricted to publications from 1993 to date. We also conducted supplementary searches of the references in retrieved articles. Abstracts were reviewed and if relevant, the article was included.

Ethics. The fever surveillance was conducted following the principles governing biomedical research involving human subjects. Prior written informed consent was obtained from the parent or guardian of all study participants. The study protocol was approved by the National Institute of Medical Research, Tanzania, and the International Vaccine Institute Institutional Review Board.

Results

The flow of patients is shown in Figure 1. A total of 1,706 febrile children were enrolled out of which 26 (1.5%) were excluded for the following reasons: 19 or 1.1% had no blood culture done, 6 or 0.4% had no Widal testing done due to insufficient quantities of sera, and one case or 0.1% had no blood culture nor Widal testing done. A total of 1,680 (98.5%) samples were included in the analysis. There were 16 or 1.0% culture-confirmed typhoid fever cases (group 1), 49 or 2.9% with NTS infection (group 2), and 113 or 6.7% with non-*Salmonella* bacteremia (group 3). From 1,502 (89.4%) children, no pathogenic bacteria were isolated from blood culture (group 4).

FIGURE 1. Flow of patients



* Species included: *Streptococcus pneumoniae* (n=11), beta hemolytic *Streptococci* (n=10), *Staphylococcus aureus* (n=5), *Haemophilus influenzae* type b (n=26), *Escherichia coli* (n=31), *Acinetobacter species* (n=6), Non-fermenters (n=12), Others: (n=5), *Haemophilus parainfluenzae* (n=2), and Gram negative rods not identified (n=5).

** Species included: *Bacillus* (n=22), Diphtheroids (n=7), *Micrococcus* (n=9), alpha-hemolytic *Streptococcus viridans* (n=4), coagulase negative *Staphylococcus* (n=104), yeast (n=5), mixed bacterial species (n=4), Gram positive rods not identified (n=1).

We assessed the age distribution and highest anti-TH and -TO titer by blood culture-confirmed diagnosis (Table 1). Children with typhoid fever were significantly older compared to the other groups. Anti-TH agglutination titers of 1:80 and higher were detected among 12/16 (75.0%) culture-confirmed typhoid fever cases compared to 7/49 (14.3%) of those with NTS infection, and 1/113 (0.9%) with other bacteremia (p values = 0.001 and <0.001, respectively). Similarly, anti-TO agglutination titers of 1:80 and higher were detected among 11/16 (68.8%) cases of typhoid fever compared to 7/49 (14.3%) of those with NTS infection and 0/113 (0%) with other bacteremia (p values <0.001 and <0.001, respectively).

TABLE 1. Number and cumulative frequencies of anti-TH and anti-TO levels overall and by blood culture isolates

Highest titer reached; Number (%)	All (n=1,680)	Children with culture-confirmed typhoid fever (n=16)	Children with non-Typhi serotypes of <i>S. enterica</i> (n=49)	Children with other pathogenic bacteria (n=113)	Children with no pathogenic bacteria isolated (n=1,502)
Median age (Range)*	1.83 (14.81)	7.21 (11.96)	1.58 (6.73)	1.43 (11.83)	1.84 (14.81)
Anti-TH					
≥1:640	15 (0.9)	3 (18.8)	5 (10.2)	0 (0)	7 (0.5)
1:320	24 (1.4)	6 (37.5)	6 (12.2)	0 (0)	12 (0.8)
1:160	36 (2.1)	11 (68.8)	7 (14.3)	0 (0)	18 (1.2)
1:80	46 (2.7)	12 (75.0)	7 (14.3)	1 (0.9)	26 (1.7)
1:40	85 (5.1)	12 (75.0)	9 (18.4)	3 (2.7)	61 (4.1)
No agglutination	1,595 (94.9)	4 (25.0)	40 (81.6)	110 (97.3)	1,441 (95.9)
Anti-TO					
≥1:640	6 (0.4)	3 (18.8)	2 (4.1)	0 (0)	1 (0.1)
1:320	18 (1.1)	6 (37.5)	3 (6.1)	0 (0)	9 (0.6)
1:160	34 (2.0)	10 (62.5)	6 (12.2)	0 (0)	18 (1.2)
1:80	44 (2.6)	11 (68.8)	7 (14.3)	0 (0)	26 (1.7)
1:40	95 (5.7)	12 (75.0)	10 (20.4)	3 (2.7)	70 (4.7)
No agglutination	1,585 (94.3)	4 (25.0)	39 (79.6)	110 (97.3)	1,432 (95.3)

* P values: Group 1 vs 4 <0.001; Group 2 vs 4 = 0.209; Group 3 vs 4 = 0.013

Primary analysis. We calculated the sensitivity, specificity, PPV, and NPV of various Widal test cut-offs for the diagnosis of typhoid fever (Table 2). The sensitivity, specificity, PPV, and NPV of an anti-TH titer of 1:80 were 75, 98, 26, and 100%, respectively. The sensitivity, specificity, PPV, and NPV of an anti-TO titer of 1:80 were 69, 98, 25, and 100%, respectively.

TABLE 2. Primary analysis for the performance* of the Widal test for typhoid fever diagnosis (n=1680)

Widal titer	Sensitivity (95%CI)	Specificity (95%CI)	Positive Predictive Value	Negative Predictive Value
TH \geq 1:80	12/16, 0.75 (0.51-0.90)	1630/1664, 0.98 (0.97-0.99)	12/46, 0.26	1630/1634, 1.00
TH \geq 1:160	11/16, 0.69 (0.44-0.86)	1639/1664, 0.98 (0.98-0.99)	11/36, 0.31	1639/1644, 1.00
TH \geq 1:320	6/16, 0.38 (0.18-0.61)	1646/1664, 0.99 (0.98-0.99)	6/24, 0.25	1646/1656, 0.99
TO \geq 1:80	11/16, 0.69 (0.44-0.86)	1631/1664, 0.98 (0.97-0.99)	11/44, 0.25	1631/1636, 1.00
TO \geq 1:160	10/16, 0.63 (0.39-0.82)	1640/1664, 0.99 (0.98-0.99)	10/34, 0.29	1640/1646, 1.00
TO \geq 1:320	6/16, 0.38 (0.18-0.61)	1652/1664, 0.99 (0.99-1.00)	6/18, 0.33	1652/1662, 0.99

*The values were calculated using culture-confirmed typhoid fever cases (group 1; n=16) as the true positives and those cases from which *S. typhi* were not isolated from blood culture (groups 2, 3, and 4; n=1664) as the true negatives.

We compared the performance of the Widal test between patients who presented with fever of 5 days or less and those who presented with more than 5 days of fever (Table 3). Of the 16 typhoid fever cases, 6 (37.5%) presented with fever of 5 days or less, and 10 (62.5%) with more than 5 days of fever. Of the 1,664 children in the control group, 1,117 (67%) presented with fever of 5 days or less and 544 (33%) with more than 5 days of fever. Three control cases, whose fever duration was unknown, were excluded from the analysis. The sensitivity of an anti-TH and -TO titer of 1:80 increased, however not significantly, from 67% to 80% and 67% to 70%, respectively, with the longer duration of fever prior to admission (both $p>0.05$). The PPV of an anti-TH and -TO titer of 1:80 increased from 21% to 30% and 19% to 30% (both $p>0.05$), respectively, with the longer duration of fever prior to admission. But the change was also not statistically significant.

TABLE 3. Comparison of the performance* of the Widal test for typhoid fever diagnosis by number of days of fever** prior to admission

Widal titer	Sensitivity (95%CI)	Specificity (95%CI)	Positive Predictive Value	Negative Predictive Value
≤ 5 days fever (n=1123)				
TH ≥1:80	4/6, 0.67 (0.30-0.90)	1102/1117, 0.99 (0.98-0.99)	4/19, 0.21	1102/1104, 1.00
TH ≥1:160	3/6, 0.50 (0.19-0.81)	1106/1117, 0.99 (0.98-0.99)	3/14, 0.21	1106/1109, 1.00
TH ≥1:320	2/6, 0.33 (0.10-0.70)	1110/1117, 0.99 (0.99-1.00)	2/9, 0.22	1110/1114, 1.00
TO ≥1:80	4/6, 0.67 (0.30-0.90)	1100/1117, 0.98 (0.98-0.99)	4/21, 0.19	1100/1102, 1.00
TO ≥1:160	3/6, 0.50 (0.19-0.81)	1107/1117, 0.99 (0.98-1.00)	3/13, 0.23	1107/1110, 1.00
TO ≥1:320	1/6, 0.17 (0.03-0.56)	1112/1117, 1.00 (0.99-1.00)	1/6, 0.17	1112/1117, 1.00
> 5 days fever (n=554)				
TH ≥1:80	8/10, 0.80 (0.49-0.94)	525/544, 0.97 (0.95-0.98)	8/27, 0.30	525/527, 1.00
TH ≥1:160	8/10, 0.80 (0.49-0.94)	530/544, 0.97 (0.96-0.98)	8/22, 0.36	530/532, 1.00
TH ≥1:320	4/10, 0.40 (0.17-0.69)	533/544, 0.98 (0.96-0.99)	4/15, 0.27	533/539, 0.99
TO ≥1:80	7/10, 0.70 (0.40-0.89)	528/544, 0.97 (0.95-0.98)	7/23, 0.30	528/531, 0.99
TO ≥1:160	7/10, 0.70 (0.40-0.89)	530/544, 0.97 (0.96-0.98)	7/21, 0.33	530/533, 0.99
TO ≥1:320	5/10, 0.50 (0.24-0.76)	537/544, 0.99 (0.97-0.99)	5/12, 0.42	537/542, 0.99

*The values were calculated using culture-confirmed typhoid fever cases (group 1; n=16) as the true positives and those cases from which *S. typhi* were not isolated from blood culture (groups 2, 3, and 4; n=1664) as the true negatives.

** Three patients whose fever duration was unknown were excluded from the analysis.

Secondary analysis. Using different control groups, we compared the resulting sensitivity, specificity, PPV, and NPV of a Widal test cut-off of an anti-TH and -TO titer of ≥1:80 for the diagnosis of typhoid fever (Table 4). Changing the control group had no significant effect on the sensitivity, specificity, and NPV but markedly increased the PPV of an anti-TH titer of ≥1:80 from 26% to 92% and the PPV of an anti-TO titer of ≥1:80 from 25% to 100%.

TABLE 4. Secondary analysis of the performance of a Widal anti-TH and -TO titer of $\geq 1:80$ for typhoid fever diagnosis using group 1 as true positives and three different control groups as true negatives*

Control group	Sensitivity (95%CI)	Specificity (95%CI)	Positive Predictive Value	Negative Predictive Value
TH titer $\geq 1:80$:				
Groups 2, 3, and 4 (n=1664)	12/16, 0.75 (0.51-0.90)	1630/1664, 0.98 (0.97-0.99)	12/46, 0.26	1630/1634, 1.00
Groups 2 and 3 (n=162)	12/16, 0.75 (0.51-0.90)	154/162, 0.95 (0.91-0.97)	12/20, 0.60	154/158, 0.97
Group 3 (n=113)	12/16, 0.75 (0.51-0.90)	112/113, 0.99 (0.95-1.00)	12/13, 0.92	112/116, 0.97
TO titer $\geq 1:80$:				
Groups 2, 3, and 4 (n=1664)	11/16, 0.69 (0.44-0.86)	1631/1664, 0.98 (0.97-0.99)	11/44, 0.25	1631/1636, 1.00
Groups 2 and 3 (n=162)	11/16, 0.69 (0.44-0.86)	155/162, 0.96 (0.91-0.98)	11/18, 0.61	155/160, 0.97
Group 3 (n=113)	11/16, 0.69 (0.44-0.86)	113/113, 1.00 (0.97-1.00)	11/11, 1.00	113/118, 0.96

* Group 1 were those with *S. typhi* isolated from blood culture (n=16), group 2 were those with non-typhi serotypes of *S. enterica* (NTS) isolated from blood culture (n=49), group 3 were those with pathogenic bacteria other than *Salmonellae* isolated from blood culture (n=113), and group 4 were those whose blood culture yielded no bacterial pathogen (n=1502).

Comparison with earlier studies. We found 4 articles from 3 countries. In this series, the age group included and prevalence of blood-culture confirmed typhoid fever varied considerably. The cut-off titer used ranged from $\geq 1:20$ to $\geq 1:200$ and the resulting sensitivity, specificity, PPV and NPV varied considerably (Table 5).

Table 5 Summary of Widal performances in earlier studies

Authors	Date	Study Country	Sample Size	Age classes included	Prevalence of <i>S. Typhi</i> in participants	Sensitivity	Specificity	PPV	NPV	Cut Off Titer	Control Group(s)	Gold Standard
Choo et al.	1993	Malaysia	2382	Children	6.1%	89%	89%	<50%	99.2%	O or H ≥1:40	Non-typhoid febrile children admitted to hospital	Blood Culture
Parry et al.	1999	Vietnam	2000	Children & Adults	30.8%	O: 49% H: 67%; O or H ≥1:100:88%	O: 97% H: 96%; O or H ≥1:100:87%	O: 88% H: 88%; O or H ≥1:100:74%	O: 82% H: 87%; O or H ≥1:100:94%	O: ≥1:200 H: ≥1:100; O or H ≥1:100	Lab confirmed malaria, dengue or bacteremia	Blood Culture
Wilke et al.	2002	Turkey	410	≥18 y	13.2%	52% Post 7-10 d: 90%	88% Post 7-10 d: 90%	76% Post 7-10 d: 88%	71% Post 7-10 d: 93%	O: ≥1:200 H: ≥1:200	Healthy controls, nontyphoidal febrile patients, blood culture negative febrile cases	Blood Culture, Stool Culture
Olsen et al.	2004	Vietnam	80	≥3y	73.8%	64% (field) 61% (lab)	76% (field) 100% (lab)	88% (field) 100% (lab)	43% (field) 48% (lab)	O or H ≥1:100	Lab confirmed bacteremia, AFB, dengue, malaria, pos. stool culture, pos. urine culture	Blood Culture
Ley et al.	This study	Tanzania	1680	2m. – 14y	1%	75%	98%	26%	100%	H: ≥1:80	Non-typhoid febrile children admitted to hospital	Blood Culture

Discussion

We found that a Widal titer of $\geq 1:80$ was the optimal indicator of typhoid fever in our study population. The PPV, NPV and specificity in the primary analysis was more-or-less unchanged from the cut-off titers of $\geq 1:80$ to $\geq 1:320$, whereas the sensitivity was highest at a cut-off titer of $\geq 1:80$. Although the Widal test at this cut-off titer performed relatively well in terms of sensitivity, specificity and NPV, its PPV was low. It has been argued that PPV is the most important measure of a clinical diagnostic method since it represents the proportion of patients with positive test results that are correctly diagnosed [13]. The PPV is not intrinsic to the test; it is affected by prevalence of the disease. In our setting, where 16 (1%) out of 1,680 febrile patients admitted to the pediatric ward had culture-proven typhoid fever, a negative Widal test result would have a good predictive value for the absence of disease but a positive result would have a low predictive value for typhoid fever, making the use of the Widal test in our setting questionable.

In a previous paper describing the clinical aspects of the children included in this study [14] older age and long duration of fever were predictive of typhoid fever in this group.

There are several difficulties associated with evaluation of the Widal test. Firstly, levels of agglutinins detectable in the non-infected populations of different areas vary considerably by time and place depending on the endemicity of the disease, which affects test performance. For example, the sensitivity and specificity of a Widal test anti-TO titer of 1:80 in Kolkata, India was 58% and 85% [10] compared to our findings of 69% and 98%. Secondly, test performance is also affected by cross-reacting infections. In our study, none of the 113 children with non-*Salmonella* bacteremia exhibited titers above 1:80 for both O and H, although cross-reactions with *Klebsiella spp.* and *Staphylococcus aureus* [15] have been reported. In contrast, 7 (14.3%) of the 49 children with NTS had titers above 1:80 for both O and H. There is also the possibility of cross-reactivity with non-bacterial infections such as malaria, dengue, hepatitis A, and infectious mononucleosis [2, 9, 16]. The third limitation is the choice of a satisfactory gold standard for diagnosis. We used blood culture-positive patients as our true positives. Although bone marrow culture would be the ideal gold standard, this test is difficult to perform in small rural hospitals in Africa. We found that 26 (1.7%) of 1,502 children from whom pathogenic bacteria were not isolated showed agglutination at 1:80 or higher, both for O and H antigens. These may be Widal false positive results due to cross-reaction. Alternatively, since the reported sensitivity of a single blood culture is only 40% to 60% [16, 17, 18, 19], some of these are likely to be false negative blood culture results. The final, and what we found to be the most contentious issue, is the selection of the most appropriate control group. It is difficult to choose

patients with febrile illness who are blood culture-negative and who definitely do not have typhoid fever. Furthermore, there were relatively few hospitalized children with no bacteremia in the same age range as those with typhoid fever. Thus, the control children were significantly younger than the cases. For our primary analysis, we used groups 2, 3 and 4 (i.e., all children admitted for a febrile illness who were subsequently culture-negative for *S. typhi*). These would be the most conservative controls for specificity since blood culture picks up only a fraction of typhoid cases, resulting in a control group that is likely contaminated with culture-negative typhoid cases. Despite this, the specificity of the Widal test was high. Using the more exclusive control groups as others had done previously [9, 10, 11] did not appreciably alter the sensitivity, specificity, and NPV but they increased the PPV.

The previous studies included in our review (Table 5) had not been performed in Africa hence different cut – off titers were applied, and the resulting sensitivity, specificity, PPV and NPV varied considerably. PPV as well as NPV are dependent on the prevalence of disease within the group of participants; the selection process of study participants has therefore direct influence on the results. The difficulty of choosing the correct control group has been noted earlier [9]. While the gold standard, blood culture, is applied in most studies, the true negatives may be defined as febrile patients with a non-typhi laboratory-confirmed diagnosis as done by Parry et al. and Olsen et al. [9, 20]. Alternatively, some studies use healthy controls. Choo et al [21]. considered all febrile cases with an *S.typhi* negative blood culture as the control group which is problematic as a number of blood culture-negative results are likely to be false-negative due to the poor sensitivity of the blood culture [17, 18, 19, 22]. Furthermore, it is difficult to compare the different test kits, as varying antigens perform differently [23].

Conclusion

In summary, a Widal titer of $\geq 1:80$ performed relatively well in terms of sensitivity and specificity. However, the low prevalence of typhoid fever of approximately 1% amongst children at Teule Hospital meant that the Widal test was only useful for excluding the disease.

Considering the low cost of Widal testing and the absence of comparably cheap tests, Widal testing is likely to remain the test of choice in many developing country settings. But the need for rapid and cheap diagnostic tools with superior performance remains high.

Competing interests

We declare that we have no conflict of interest.

Authors contribution

BL performed the Widal tests, analyzed and compared results and wrote the manuscript; GM was in charge of the implementation and management of the study; KT performed Widal tests, analyzed results and contributed to the manuscript; BA supervised the laboratory where blood cultures were performed and contributed to the manuscript; LvS provided scientific support to study staff and manuscript and was involved in clinical care of participants; IH was involved in clinical care of participants; AM was in charge of data management; AS performed blood culture procedures, RM facilitated activities to make data collection possible, SA provided laboratory support; DRK performed the statistical analysis; RLO provided scientific support to the manuscript; JDC provided scientific support to the manuscript; HR provided major contributions to the manuscript; HW provided scientific support to the manuscript; SM facilitated activities to make data collection possible; JLD provided major scientific support to the manuscript and was involved in clinical care of participants.

All authors have read and approved the final manuscript

Acknowledgement

This work was supported by a grant from the Korean International Cooperation Agency through the International Vaccine Institute.

The study is published with the permission of the Director General of the Tanzanian National Institute for Medical Research, Dar-Es-Salaam. We are grateful to the patients and their parents who made this work possible. We thank all technical staff and research assistants who were involved in the study.

References

1. Crump JA, Luby SP, Mintz ED: **The global burden of typhoid fever.** *Bulletin of the World Health Organization* 2004, **82** [5]:346—353.
2. Olopoenia LA, and King AI: **Widal agglutination test – 100 years later: still plagued by controversy.** *Postgrad Med J.* 2000, 76:80 – 84.
3. Mweu E. and English M: **Typhoid Fever in Children in Africa.** *Trop Med Int Health* 2008 **13**[4]:532—540.
4. Chart H, Cheesbrough J and Waghorn D: **The serodiagnosis of infection with Salmonella typhi.** *Journal of Clinical Pathology* 2000, **53**: 851–853
5. Maxwell CA, Chambo W, Mwaimu M, Magogo F, Carneiro IA, Curtis CF: **Variation of malaria transmission and morbidity with altitude in Tanzania and with introduction of alphacypermethrin treated nets.** *Malar J.* 2003, 10:28.
6. Edmonds SM, Meadway J: **Getting Pregnant Women onto HAART (Highly active antiretroviral therapy); developing a strategy for advanced prevention of mother-to-child transmission of HIV (PMTCT+) in rural Tanzania.** *HIV Medicine* 2009, 10(Suppl.1):26.
7. Smith T, Schellenberg JA, Hayes R: **Attributable fraction estimates and case definitions for malaria in endemic areas.** *Stat Med.* 1994. 13(22):2345–2358.
8. Mwangi TW, Ross A, Snow RW, Marsh K: **Case definitions of clinical malaria under different transmission conditions in Kilifi District, Kenya.** *J Infect Dis.* 2005. 191:1932–1939.
9. Parry CM, Hoa, Diep TS, Wain J, Chinh NT, Vinh H, Hien TT, White NJ, Farrar JJ: **Value of a Single-tube Widal Test in Diagnosis of Typhoid Fever in Vietnam.** *J Clin Microbiol.* 1999: **37**[9]:2882-2886.
10. Dutta S, Sur D, Manna B, Sen B, Deb AK, Deen JL, Wain J, von Seidlein L, Ochiai RL, Clemens JD, Bhattacharya JK: **Evaluation of a new-generation serologic test for the diagnosis of**

typhoid fever: data from a community – based surveillance in Calcutta, India. *Diagnostic Microbiology and Infectious Disease* 2006, 56:359 – 365.

11. Dong B, Gallindo CM, Shin E, Acosta CJ, Page AL, Wang M, Kim D, Ochiai RL, Park J, Ali M, von Seidlein L, Xu Z, Clemens, JD: **Optimizing typhoid fever case definitions by combining serological tests in a large population study in Hechi City, China.** *Epidemiol Infect* 2007, **135** [6]:1014 – 1020.
12. Newcombe RG: **Two-Sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods.** *Stat Med.* 1998. 17[8]:857 – 872.
13. Altman DG, and Bland JM: **Diagnostic tests 2: Predictive values.** *BMJ* 1994, 309:102.
14. Mtove G, Amos B, von Seidlein L, Hendriksen I, Mwambuli A, Kimera J, Mallahiyo R, Kim DR, Ochiai RL, Clemens JD, Reyburn H, Magesa S, Deen JL: **Invasive salmonellosis among children admitted to a rural Tanzanian hospital and a comparison with previous studies.** *PLoS One.* 2010 Feb 16;5(2):e9244
15. Pokhrel BM, Karmacharya R, Mishra SK, Koirala J: **Distribution of antibody titer against *Salmonella enteric* among healthy individuals in Nepal.** *Ann Clin Microbiol Antimicrob.* 2009, 7;8:1
16. Bhutta ZA: **Current concepts in the diagnosis and treatment of typhoid fever.** *BMJ* 2006, 333:78 – 82.
17. Farooqui BJ, Khurshid M, Ashfaq MK, Khan MA: **Comparative yield of *Salmonella typhi* from blood and bone marrow cultures in patients with fever of unknown origin.** *J Clin Pathol.* 1991, 44:258-259.
18. Gilman RH, Terminel M, Levine, Hernandez-Mendoza P, Hornick RB: **Relative efficacy of blood, urine, rectal swab, bone marrow, and rose spot cultures for recovery of *Salmonella typhi* in typhoid fever.** *Lancet* 1975, 31;1(7918):1211-1213.

19. Wain J, Pham VB, Ha V, Nguyen NM, To SD, Walsh AL, Parry CM, Hasserjian RP, HoHo VA, Tran TH, Farrar J, White NJ, Day NP: **Quantitation of bacteria in bone marrow from patients with typhoid fever: relationship between counts and clinical features.** *J Clin Microbiol.* 2001, **39**:1571-1576.
20. Olsen SJ, Pruckler J, Bibbb W, Thanh NTM, Trinh TM, Minh NT, Sivapalsingam S, Gupta A, Phuong PT, Chinh NT, Chau NV, Cam PD, Mintz ED: **Evaluation of Rapid Diagnostic Tests for Typhoid Fever.** *Journal of Clinical Microbiology.* 2004, DOI: 10.1128/JCM.42.5.1885–1889.2004
21. Hoffman SL, Edman DC, Punjabi NH, Lesmana M, Cholid A, Sundah S, Harahap J: **Bone marrow aspirate culture superior to streptokinase clot culture and 8 ml 1:10 blood to broth ratio blood culture for diagnosis of typhoid fever.** *Am J Trop Med Hyg.* 1986, 35:836-839.
22. Choo KE, Razif AR, Oppenheimer SJ, Ariffin WA, Lau J, Abraham T: **Usefulness of the Widal test in diagnosing childhood typhoid fever in endemic areas.** *J Paediatr Child Health.* 1993, **29**(1):36-9.
23. Handojo I, Edijanto SP, Proboboesodo MY, Mahartini NN: **Comparisson of the Diagnostic Value of Local Widal slide Test with Imported Widal Slide Test.** *Southeast Asian J Trop Med Public Health.* 2004, **35** (2): 366-70

Assessment and comparative analysis of a rapid diagnostic test (Tubex®) for the diagnosis of typhoid fever among hospitalized children in rural Tanzania

Benedikt Ley^{1,8}, Kamala Thriemer¹, Shaali M Ame², George M Mtove^{3,4}, Lorenz von Seidlein^{1,4,5}, Ben Amos^{4,6}, Ilse CE Hendriksen^{4,7}, Abraham Mwambuli⁴, Aikande Shoo^{4,6}, Deok R Kim¹, Leon R Ochiai¹, Michael Favorov¹, John D Clemens¹, Harald Wilfing⁸, Jacqueline L Deen^{1,4}, Said M Ali²

1. Translational Research Division, International Vaccine Institute, Seoul, Korea
2. Laboratory Division, Public Health Laboratory (Pemba) – Ivo de Carneri, Chake Chake, Tanzania
3. Amani Centre, National Institute for Medical Research, Tanga, Tanzania
4. Joint Malaria Program, Tanga, Tanzania
5. Asia Pacific Malaria Elimination Network (APMEN), Menzies School of Health Research, Casuarina, Australia
6. Teule Hospital, Muheza, Tanga, Tanzania
7. Oxford Research Unit, Mahidol University, Bangkok, Thailand
8. Biocenter, University of Vienna, Vienna, Austria

Keywords: Salmonella, Tubex®, Widal, Africa, Rapid Diagnostic Test

***Corresponding author:**

Benedikt Ley
International Vaccine Institute
San 4-8 Nakseongdae-dong, Kwanak-gu
Seoul, Korea 151-818
Email: bnley@ivi.int

Abstract

Background: Typhoid fever remains a significant health problem in many developing countries. A rapid test with a performance comparable to that of blood culture would be highly useful. A rapid diagnostic test for typhoid fever, Tubex®, is commercially available that uses particle separation to detect immunoglobulin M directed towards *Salmonella* Typhi O9 lipopolysaccharide in sera.

Methods: We assessed the sensitivity and specificity of the Tubex test among Tanzanian children hospitalized with febrile illness using blood culture as gold standard. Evaluation was done considering blood culture confirmed *S. Typhi* with non-typhi salmonella (NTS) and non-salmonella isolates as controls as well as with non-salmonella isolates only.

Results: Of 139 samples tested with Tubex, 33 were positive for *S. Typhi* in blood culture, 49 were culture-confirmed NTS infections, and 57 were other non-salmonella infections. Thirteen hemolyzed samples were excluded. Using all non-*S. Typhi* isolates as controls, we showed a sensitivity of 79% and a specificity of 89%. When the analysis was repeated excluding NTS from the pool of controls we showed a sensitivity of 79% and a specificity of 97%. There was no significant difference in the test performance using the two different control groups ($p>0.05$).

Conclusion: This first evaluation of the Tubex test in an African setting showed a similar performance to those seen in some Asian settings. Comparison with the earlier results of a Widal test using the same samples showed no significant difference ($p>0.05$) for any of the performance indicators, irrespective of the applied control group.

Background

Typhoid fever remains a significant health problem in many developing countries. Estimates suggest an incidence rate of more than 21.5 million cases globally in the year 2000 [1]. Recent data from Tanzania mainland have found a strong variation of prevalence rates among blood culture positive isolates collected in local hospitals, ranging from 9% [2] to 21.4% [3] for *Salmonella enterica* serovar Typhi (*S. Typhi*), no data from Zanzibar are available to date. As the clinical picture of typhoid fever is often unspecific, misdiagnosis and insufficient or inadequate treatment are potential risks associated with the disease. In the absence of difficult-to-obtain bone marrow specimens, microbiologic culture of a blood sample is considered to be the current state-of-the art test for the diagnosis of typhoid fever even though its sensitivity may be as low as 40% [4; 5]. Culture may take up to seven days and requires a well-run and equipped laboratory, which is often not available in settings with endemic typhoid fever. The widely in use Widal test provides a cost efficient alternative [6] for serological diagnosis, however its performance remains unsatisfying with sensitivity reported from Tanzania of 75% using blood culture as the gold standard and applying a cut off titer of 1:80 [7]. The test further requires the establishment of a local cut off titer prior to use which is complicated. Therefore, a rapid test with a performance comparable to that of blood culture would be desirable.

A rapid diagnostic test for typhoid fever, Tubex® is commercially available that uses particle separation to detect immunoglobulin M (IgM) directed towards *Salmonella enterica* serovar Typhi (*S. Typhi*) O9 lipopolysaccharide in patient sera. Performance of the test has previously been evaluated in a number of studies in Asia but none in Africa. Using blood culture results for comparison, we assessed the sensitivity and specificity of the Tubex test among Tanzanian children hospitalized with febrile illness and compared our results with those from previous studies.

Methods

For evaluation of the Tubex test, we used a selected subset of serum samples that was obtained for a fever surveillance study [2] from Teule Hospital in Muheza District, Tanzania. In order to accommodate the required sample size for the test validation, we included randomly selected and age-matched *Salmonella enterica* serotype Typhi (*S. Typhi*) positive serum samples from a second fever surveillance study conducted at Chake Chake Hospital in Pemba, Zanzibar. All samples were collected from children between the ages of 2 months to 14 years from 2008 to 2009.

At Teule Hospital in Muheza, sera and blood was collected for culture from children with a history of three days of fever, or a history of less than three days of fever but with at least one of the following

severity criteria: respiratory distress; deep breathing; respiratory distress in combination with severe pallor; prostration; capillary refill ≥ 3 seconds; temperature gradient; systolic blood pressure < 70 mm Hg; coma defined by Glasgow Coma Scale (GCS) ≤ 10 or Blantyre Coma Scale (BCS) ≤ 2 ; severe jaundice; history of two or more convulsions in the last 24 hours; blood glucose < 3 mmol associated with clinical signs; neck stiffness; bulging fontanel; or oxygen saturation $< 90\%$ [2].

At Chake Chake Hospital in Pemba, sera and blood was collected for culture from children with a recorded body temperature of $> 37.5^{\circ}\text{C}$ for outpatients and any history of fever for inpatients. Duration of fever was not considered for study recruitment.

About 3 to 5 milliliters (ml) of blood (depending on body weight) was collected and inoculated in a BactALERT™ Pediatric-fan bottle (Teule Hospital) or a BacTec Peds PLUS™/F bottle (Chake Chake Hospital) and incubated in the respective machine (BacT/ALERT 3D or BacTec 9050). Bacterial growth was evaluated following standard procedures.

The Tubex® test (IDL – Sweden) was conducted according to the manufacturer's instructions, which are as follows. Forty-five microliters (μl) of antigen covered particles were added to the Tubex Reaction Well Strip and $45\mu\text{l}$ of non-hemolyzed serum was added. After two minutes of incubation time, $90\mu\text{l}$ of magnetic antibody coated solution was added, and the strip was sealed and shaken for two minutes. The strip was then placed on a magnetic tray for five minutes, separating the particles if a positive sample had been added. The resulting color change of the solution was read and categorized on a scale from 0 to 10. The results were interpreted as positive for scores of 4 or greater and as negative for scores of 2 or below as per the manufacturer's instructions. Samples with a color corresponding to the value of 3 were interpreted as indeterminate. All blood culture isolates from individuals that matched the inclusion criteria and that were not considered a contaminant were included in the analysis.

We performed the Tubex test on non-hemolyzed serum samples from the patients of the two surveillance studies who had blood culture-confirmed *S. Typhi* (defined as group 1), randomly selected cases of non-Typhi serotypes of *S. enterica* (NTS) (defined as group 2), and randomly selected cases with other (non-*Salmonellae*) pathogenic bacteria (defined as group 3). Staff members performing the Tubex test were blinded to the blood culture results.

For the analysis, sensitivity (true-positive rate) was defined as the probability that the Tubex test result will be positive when there is blood culture-confirmed typhoid fever (group 1) and specificity (true-negative rate) was defined as the probability that the Tubex® test result will be negative when *S. Typhi* is not isolated from blood culture (groups 2 and 3). We conducted a secondary analysis using only group 3 as the control group. Comparison of test performance using different control groups was done using the Yates Chi-Square Test corrected for continuity.

We conducted a literature review in order to compare our findings with those from previous studies. We included studies of the Tubex test, which were identified by directly searching the MEDLINE database through PubMed. All articles since the first publication of the test [8] were included. We also conducted a supplementary search of references in retrieved articles. Abstracts were reviewed, and if relevant, the article was included.

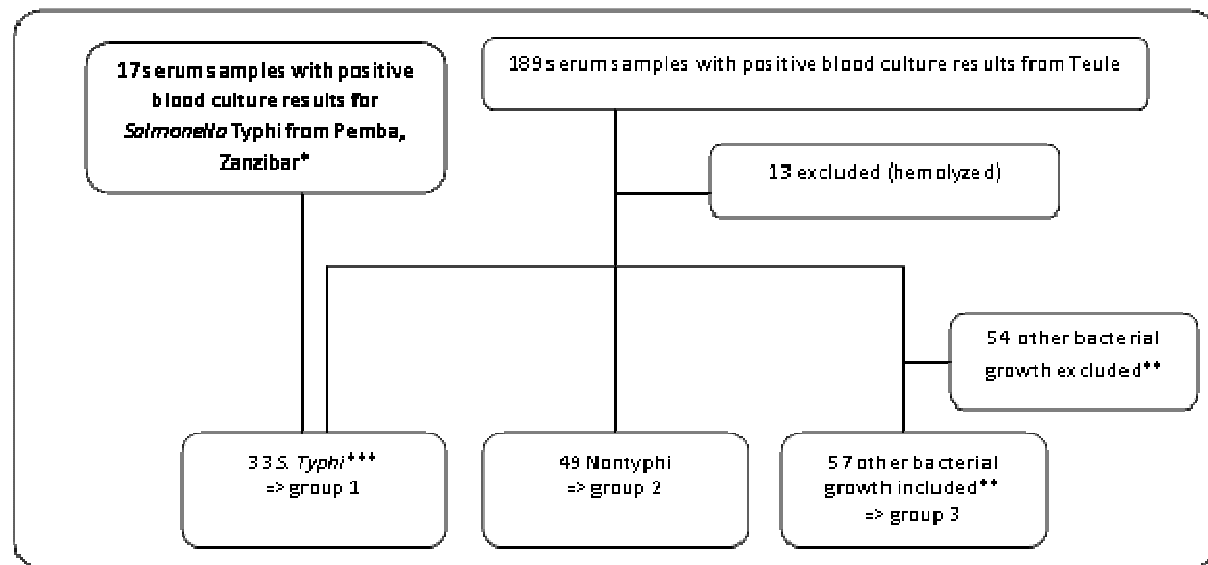
A comparison of performance of the Tubex® test with earlier published Widal test results obtained from the same samples was done using McNemar's Test for Correlated Proportions (<http://faculty.vassar.edu>).

The fever surveillance studies at Chake-Chake and Teule Hospitals were approved by their respective local ethical review boards (Tanzania and Zanzibar), as well as by the International Vaccine Institute's Institutional Review Board. Written informed consent was obtained from legal guardians of all participants prior to any sample or data collection.

Results

A total of 139 samples were tested with Tubex. Thirty -three were found positive for *S. Typhi* in blood culture (group 1), 49 were culture-confirmed non-*S. Typhi* (NTS) infection (group 2), and 57 were other non-*Salmonella* infections that were not contaminants (group 3). Thirteen hemolyzed samples were excluded (Figure 1).

Figure 1: Specimen assembly



*Age-matched, randomly selected

**Randomly selected

*** 17 from Zanzibar and 16 from Teule

+ Other bacterial growth included: 11x *Escherichia coli*, 10x *Haemophilus influenzae* type B, 8x *Streptococcus pneumoniae*, 6x *Pseudomonas* spp., 5x *Acinetobacter baumannii*, 3x *Streptococcus* beta – hemolytic group A, 3x *Haemophilus influenzae*, 2x *Staphylococcus aureus*, 1x *Burkholderia cepacia*, 1x *Streptococcus* beta-hemolytic group C, 1x *Campylobacter* spp., 1x *Pasteurella multocida*, 1x gram negative rods (not identified further), 1x *Citrobacter braakii*, 1x *Stenotrophomonas*, 1x *Haemophilus parainfluenzae*, 1x *Streptococcus* beta - hemolytic

Of the 33 blood culture-positive *S. Typhi* cases, 26 had a positive Tubex result and were considered as true positives. Of the 106 blood culture confirmed NTS and non-salmonella cases (groups 2 and 3), 94 yielded a negative Tubex result and were considered as true negatives. Considering only the 57 non-*Salmonella* cases (group 3) as controls, resulted in 54 true negative cases.

Using groups 2 and 3 as controls showed a sensitivity of 79% and a specificity of 89% (Table 1). The same analysis was repeated excluding NTS from the pool of controls and showed 79% and 97% for sensitivity and specificity, respectively. There was no significant difference in the test performance using the two different control groups (all were $p > 0.05$ using the Chi square test).

Table 1: Performance of Tubex® using group 1 as true positives and two different control groups as true negatives

	Control Group	
	Group 2 + 3*	Group 3*
Sensitivity (95% CI)	0.79 (0.52-0.81)	0.79 (0.62-0.90)
(absolute numbers)	(26/33)	(26/33)
Specificity (95% CI)	0.89 (0.81-0.94)	0.97 (0.85-0.99)
(absolute numbers)	(94/106)	(94/97)

*Group 1 = *S. Typhi* (n=33), Group 2 = all non-typhi *Salmonella* (n=49), Group 3 = all blood culture-positive non-*Salmonella* cases (n=57)

A total of 14 articles were retrieved and evaluated for inclusion into the review. All of the reported studies were performed in Asia; none in Africa. A total of six articles were excluded: two evaluated the test for non-typhoidal *Salmonella* [9] or *S. Paratyphi* [10], three did not evaluate the sensitivity and specificity of the test [11; 12; 13], and one was a letter to the editor [14]. Thus, eight publications were included in the review (Table 2). Five of the included articles reported findings of test performance that were similar to our results [15; 16; 17; 18; 19]. Two publications showed considerably lower sensitivity and specificity [20; 6], and one reported higher values [8] (Table 2).

Table 2: Comparison of the performance of the Tubex® test from published reports

Author	Year	Journal	Sam ple size	Location	Tubex ® cut off	Sens	Spec	True neg. definition	Reader	Gold standard	Study population
Ley, B. et al	This pape r	This Journal	139	Tanzania	>4	79%	89%	All non- typhi bacteriamia	Investigator	Blood Culture (BACTEC)	>2months + >37.5° (inpatients) & history of fever (outpatients)
			88			79%	97%	All non- salmonella bacteriamia			
Naheed, A. et al	2008	Diagn Microbiol Infect Dis.	867	Bangladesh	≥4	60%	58%	Other confirmed bacteremia	ICDDRb lab	Manual Blood Culture	Active surveillance Temp ≥38°C
						60%	64%	Blood culture neg & other bacteremia			
Rahman, M. et al	2007	Diagn Microbiol Infect Dis.	243	Bangladesh	> 4	91.2%	82.3%	Other febrile patients	ICDDRb lab, min. 2 independe nt lab techs	Manual Blood Culture	Outpatients, all ages with history of fever
						No pos.	89.5%	Healthy subjects			Healthy subjects
Dong, B. et al	2007	Epidemio l. Infect.	1732	China	≥2	100%	43%	Paratyphoid cases	-	Blood culture (BACTEC)	Age 5-60 with reported history of fever for 3 days
					≥4	69%	95%				
					≥6	62%	95%				
					≥8	23%	100%				
					≥10	15%	100%				
Kawano, R. L. et al	2006	JCM.	177	Philippines	≥2	94.7%	80.4%	Blood culture neg.	n/A	Manual Blood Culture &BACTEC	Clinically suspected typhoid cases
Dutta, S. et al	2006	Diagn Microbiol Infect Dis.	495	India	≥4	56%	88%	Paratyphoid and malaria cases	n/A	Blood Culture BACTEC	Outpatients, all ages, Pat with history of fever for 3 days
Ohlsen, S. J. et al	2004	JCM	79	Vietnam	Accor ding to proto col	78%	94%	Other lab-confirmed febrile illnesses	n/A	Manual Blood Culture/ BACTEC	Pat ≥3year and history of ≥ 4 day fever
House, D. et al	2001	JCM.	127	Vietnam	>2	87%	76%	Febrile hospitalized patients	labtech	Culture	Children and adults
Lim et al	1998	JCM	105	Hong Kong & Malaysia	>2	100%	100%	Healthy individuals and pat with other bacterial diseases and autoimmune disease	labtechs	Culture confirmed (56% of pos.), clinical picture, various other tests	Clinical picture, culture confirmed,

Discussion

We found Tubex has a sensitivity of 79% using either control group (95%CI: 52-81% for groups 2 and 3, and 62-90% for group 3 only) and a specificity of 89-97% (95%CI: 81-94% for groups 2 and 3 and 85-99% for group 3 only) irrespective of control group. To our knowledge, this is the first evaluation of the test in an African population. Our results were similar to those observed in previous studies (five out of eight studies) in Asia assessing the performance of the test [15; 16; 17; 18; 19], though Kawano *et al.* [17] and House *et al.* [19] used a lower cut-off titer than is recommended in the manual. In contrast, two studies [20; 6] found the performance of Tubex to be poorer than our findings, despite using a similar cut-off value, gold standard, and inclusion criteria. The extremely good performance of Tubex observed by Lim *et al.* [8] has not been reproduced since.

An important limitation of this study is that the sera are combined from two different patient populations and the purposeful selection of samples included in the three groups. During the preparation of the study, we calculated the sample sizes of true positive sera and true negative sera that are required for validation of the Tubex test and for comparison with the Widal test performance. The number of true positive sera from either hospital alone was insufficient for the validation. Thus, we included *S. Typhi* blood culture-confirmed sera from Pemba. Analysis of the results by hospital was not possible because of insufficient sample size.

In a sub-analysis in assessing cross reactivity with NTS, blood culture-confirmed NTS cases were considered as true positives, and all other positive isolates, excluding *S. Typhi*, were considered as true negatives. In this sub-analysis Tubex had a sensitivity of 18% and a specificity of 95% (analysis not shown).

Comparison with a Widal test that was earlier conducted using the same samples [7] revealed no significant difference ($p>0.05$) for any of the performance indicators, irrespective of the applied control group. But compared to the Widal test, Tubex is easier and quicker to perform. The Widal test requires 16 – 20 hours until the results are obtained while the complete procedure for the Tubex test is approximately 20 minutes. Tubex is more expensive at approximately 2.15 USD per test compared to <0.80 USD per test for the Widal tube agglutination test [6].

Interpreting the Tubex test results was found to be difficult and the results were prone to inter – reader variation. Assessing the color change according to the provided color scale requires experience and standardized good lighting conditions. The Tubex test can only be applied to non-hemolyzed and non-icteric serum samples, thus limiting its general application. However Tam *et al.*

[12] have described a method that includes a washing step and thereby addresses the problem of turbid serum. This method requires double the amount of antigen-coated particles as well as glycine buffered saline (GBS), thereby increasing the price per sample to approximately 4.50 USD and reducing its feasibility as an easy-to-perform test. While neither of the tests can be performed by untrained staff, interpretation of results is considered easier for the Widal test compared to Tubex.

Conclusion

The advantages of Tubex over the Widal test and the gold standard of blood culture is the short time it requires to obtain a result, and it does not require establishing a local cut-off value as with the Widal. In settings that can afford the relatively high cost of Tubex and that require instant individual diagnoses to support the clinical diagnosis of typhoid fever, Tubex is superior to the Widal tube agglutination test. For screening and surveillance purposes, as well as in settings with limited financial and technical resources, the Widal tube agglutination test is a possible alternative that can provide a similar performance as Tubex at a lower cost though it requires more time. Our evaluation of Tubex showed that any result must be handled with precaution. Results should be considered as indicative, not confirmatory. The test may be used to exclude disease though. In conclusion, the need for a reliable, fast, cheap, and easy-to-apply rapid diagnostic test for typhoid fever remains in high demand.

Competing interest

We declare that we have no conflict of interest.

Authors contributions

BL performed the TUBEX test, analyzed results and wrote the manuscript, KT performed TUBEX tests, literature search and contributed to the manuscript, SMA supervised the laboratory work in Pemba, GM was in charge of the implementation and study management in Teule, LvS provided scientific support to study staff and contributed to the manuscript, BA supervised the laboratory work in Teule, ICEH was involved in the clinical care of patients, AM was in charge of data management, AS performed blood culture procedures, DRK performed statistical analyses, RLO provided scientific support to the manuscript, MF provided scientific support to the manuscript, JDC provided scientific support to the manuscript, HW provided scientific support to the manuscript, JLD provided major scientific support to the manuscript and was involved in the clinical care of patients, SMA provided scientific support to the manuscript and the study in Pemba. All authors have read and approved the final manuscript.

Acknowledgements

We thank Hugh Reyburn of the London School of Hygiene and Tropical Medicine who supported this project and provided his scientific expertise. We also thank Rajabu Malahiyo and Steven Magesa for their support.

This work was supported by a grant from the Korea International Cooperation Agency (KOICA) and the Swedish International Development Cooperation Agency (SIDA) to the International Vaccine Institute.

This study is published with permission from the Director General of the Tanzanian National Institute for Medical Research, Dar-Es-Salaam. We are grateful to the patients and their parents who made this work possible. We thank all of the technical staff and research assistants who were involved in the study.

References

1. Crump JA, Luby SP, Mintz ED. **The global burden of typhoid fever.** *Bulletin of the World Health Organization.* 2004, **82**(5):346—353.
2. Mtove G, Amos B, von Seidlein L, Hendriksen I, Mwambuli A, Kimera J, Mallahiyo R, Kim DR, Ochiai RL, Clemens JD, Reyburn H, Magesa S, Deen JL. **Invasive salmonellosis among children admitted to a rural Tanzanian hospital and a comparison with previous studies.** *PLoS One.* 2010, 16; **5**(2):e9244.
3. Crump JC, Ramadhani HO, Morrissey AB, Saganda W, Mwako MS, Yang LY, Chow SC, Morpeth SC, Reyburn H, Njau BN, Shaw AV, Diefenthal HC, Shao JF, Bartett JA, Maro VP. **Invasive bacterial and fungal infections among hospitalized HIV-infected and HIV-uninfected adults and adolescents in northern Tanzania.** *CID.* 2011, **52** (3): 341-348.
4. Parry CM, Hien TT, Dougan G, White N, Farrar JJ. **Typhoid Fever.** *The New England Journal of Medicine.* 2002, **347**(22): 1770 – 1782.
5. Wilke A, Ergonul O, Bayar B. **Widal Test in Diagnosis of Typhoid Fever in Turkey.** *Clin Diagn Lab Immunol.* 2002, Jul;**9**(4):938-41.
6. Dutta, S, Sur D, Manna B, Sen B, Deb AK, Deen JL, Wain J, von Seidlein L, Ochiai RL, Clemens JD, Bhattacharya SK. **Evaluation of a new-generation serologic test for the diagnosis of typhoid fever: data from a community – based surveillance in Calcutta, India.** *Diagnostic Microbiology and Infectious Disease.* 2006, **56**:359 – 365.
7. Ley B, Mtove G, Thriemer K, Amos B, von Seidlein L, Hendriksen I, Mwambuli A, Shoo A, Malahiyo R, Ame S, Kim DR, Ochiai RL, Clemens JD, Reyburn H, Wilfing H, Magesa S, Deen JL. **Evaluation of the Widal tube agglutination test for the diagnosis of typhoid fever among children admitted to a rural hospital in Tanzania and a comparison with previous studies.** *BMC.* 2010, **22**; 10:180.
8. Lim PL, Tam FC, Cheong YM, Jegathesan M. **One-step 2-minute test to detect typhoid-specific antibodies based on particle separation in tubes.** *J Clin Microbiol.* 1998, **36**(8):2271-8.

9. Oracz G, Feleszko W, Golicka D, Maksymiuk J, Klonowska A, Szajewska H. **Rapid diagnosis of acute *Salmonella* gastrointestinal infection.** *Clin Infect Dis.* 2003, 1;**36**(1):112-5.
10. Tam FC, Wang M, Dong B, Leung DT, Ma CH, Lim PL. **New rapid test for paratyphoid a fever: usefulness, cross-detection, and solution.** *Diagn Microbiol Infect Dis.* 2008 Oct;**62**(2):142-50.
11. Tam FC, Lim PL. **The TUBEX typhoid test based on particle-inhibition immunoassay detects IgM but not IgG anti-O9 antibodies.** *J Immunol Methods.* 2003, **282**(1-2):83-91.
12. Tam FC, Leung DT, Ma CH, Lim PL. **Modification of the TUBEX typhoid test to detect antibodies directly from haemolytic and whole blood,** *Journal of Medical Microbiology.* 2008, **57**: 1349 – 1353, DOI 10.1099/jmm.0.2008/002683-0.
13. Tam FC, Ling TK, Wong KT, Leung DT, Chan RC, Lim PL. **The TUBEX test detects not only typhoid-specific antibodies but also soluble antigens and whole bacteria.** *J Med Microbiol.* 2008, **57**(Pt 3):316-23.
14. Feleszko W, Maksymiuk J, Oracz G, Golicka D, Szajewska H. **The TUBEX typhoid test detects current *Salmonella* infections.** *J Immunol Methods.* 2004, **1**;**285**(1):137-8.
15. Rahman M, Siddique AK, Tam FC, Sharmin S, Rashid H, Iqbal A, Ahmed S, Nair GB, Chaignat CL, Lim PL. **Rapid detection of early typhoid fever in endemic community children by the TUBEX O9-antibody test.** *Diagn Microbiol Infect Dis.* 2007, **58**(3):275-81.
16. Dong B, Galindo CM, Shin E, Acosta CJ, Page AL, Wang M, Kim D, Ochiai RL, Park J, Ali M, Seidlein LV, Xu Z, Yang J, Clemens JD. **Optimizing typhoid fever case definitions by combining serological tests in a large population study in Hechi City, China.** *Epidemiol Infect;* 2007, **135**(6):1014.
17. Kawano RL, Leano SA, Agdamag DMA. **Comparison of serological Test Kits for Diagnosis of Typhoid Fever in the Philippines.** *Journal of Clinical Microbiology.* 2007, **45** (1): 246 – 247, doi:10.1128/JCM.01403-06.

18. Olsen SJ, Pruckler J, Bibb W, Thanh NTM, Trinh TM, Minh NT, Sivapalsingam S, Gupta A, Phuong PT, Chinh NT, Chau NV, Cam PD, Mintz ED. **Evaluation of Rapid Diagnostic Test for Typhoid Fever.** *Journal of Clinical Microbiology*. 2004, **42 (5)**: 1885 – 1889, doi: 10.1128/JCM.42.5.1885–1889.2004.
19. House D, Wain J, Ho VA, Diep TS, Chinh NT, Bay PV, Vinh H, Duc M, Parry CM, Dougan G, White NJ, Hien TT, Farrar JJ. **Serology Of Typhoid Fever in an Area of Endemicity and Its Relevance to Diagnosis.** *Journal of Clinical Microbiology*. 2001, **39 (2)**: 1002 – 1007, doi: 10.1128/JCM.39.3.1002–1007.2001.
20. Naheed A, Ram PK, Brooks WA, Mintz ED, Hossain MA, Parsons MM, Luby SP, Breiman RF. **Clinical value of Tubex® and Typhidot rapid diagnostic tests for typhoid fever in an urban community clinic in Bangladesh.** *Diagn Microbiol Infect Dis*. 2008, **61(4)**:381-6.

Evaluation of a rapid dipstick (Crystal VC) for the diagnosis of cholera in Zanzibar and a comparison with previous studies

Benedikt Ley^{1,2}, Ahmed M. Khatib³, Kamala Thriemer¹, Lorenz von Seidlein⁴, Jacqueline Deen⁴, Asish Mukhopadhyay⁵, Na-Yoon Chang¹, Ramadhan Hashim¹, Wolfgang Schmied¹, Clara J-L. Busch⁶, Rita Reyburn¹, Thomas Wierzba¹, John D. Clemens¹, Harald Wilfing², Godwin Enwere⁷, Theresa Aguado⁷, Mohammad S. Jiddawi³, David Sack⁸, Said M. Ali^{3,9}

1. Translational Research Division, International Vaccine Institute, Seoul, Korea
2. Biocenter, University of Vienna, Vienna, Austria
3. Ministry of Health and Social Welfare, Zanzibar, Tanzania
4. Global Health Division, Menzies School of Health Research, Casuarina, NT, Australia
5. National Institute of Cholera and Enteric Diseases, Kolkata, India
6. Institute of Molecular Biotechnology, Vienna, Austria
7. World Health Organization, Geneva, Switzerland
8. John Hopkins Bloomberg School of Public Health, Baltimore, USA
9. Public Health Laboratory (Pemba) – Ivo de Carneri, Chake Chake, Tanzania

Keywords: Cholera, Rapid Diagnostic Test, Zanzibar, Crystal VC

***Corresponding author:**

Benedikt Ley

Email: ley.benedikt@gmail.com

Abstract

Background. The gold standard for the diagnosis of cholera is stool culture, but this requires laboratory facilities and takes at least 24 hours. A rapid diagnostic test (RDT) that can be used by minimally trained staff at treatment centers could potentially improve the reporting and management of cholera outbreaks.

Methods. We evaluated the Crystal VC™ RDT under field conditions in Zanzibar in 2009. Patients presenting to treatment centers with watery diarrhea provided a stool sample for rapid diagnostic testing. Results were compared to stool culture performed in a reference laboratory. We assessed the overall performance of the RDT and evaluated whether previous intake of antibiotics, intravenous fluids, location of testing, and skill level of the technician affected the RDT results.

Results. We included stool samples from 624 patients. Compared to culture, the overall sensitivity of the RDT was 93.1% (95%CI: 88.7 to 96.2%), specificity was 49.2% (95%CI: 44.3 to 54.1%), the positive predictive value was 47.0% (95%CI: 42.1 to 52.0%) and the negative predictive value was 93.6% (95%CI: 89.6 to 96.5%). The overall false positivity rate was 50.8% (213/419); fieldworkers frequently misread very faint test lines as positive.

Conclusion. The observed sensitivity of the Crystal VC RDT evaluated was similar compared to earlier versions, while specificity was poorer. The current version of the RDT could potentially be used as a screening tool in the field. Because of the high proportion of false positive results when field workers test stool specimens, positive results will need to be confirmed with stool culture.

Introduction

Cholera remains a very common and potentially lethal disease in Asia and Africa. Globally, more than 220,000 cases were reported to the World Health Organization (WHO) in 2009 [1], however the true number of cases, including unreported cases, is likely to be much higher – perhaps 3 -5 million cases / year [2]. Cholera occurs mainly in areas with poor infrastructure and limited access to clean water. The etiologic organisms, *Vibrio cholerae* O1 and O139, are highly transmissible and can cause explosive outbreaks. While many of those affected experience only mild symptoms, some suffer from severe disease characterized by profuse diarrhea, electrolyte imbalance, coma and death if prompt rehydration is not provided [3, 4]. Cholera cases have been reported from Zanzibar since 1978 with regular outbreaks documented since then [5, 6].

The gold standard for laboratory confirmation of cholera is stool culture [7]. This is a routine procedure but requires laboratory infrastructure including trained staff. A single stool culture costs approximately 4 USD / case [8] and requires about 24 to 72 hours and transport to the closest sufficiently equipped laboratory, which may create additional costs. Furthermore, microbiologic facilities are often not available in locations where cholera occurs. A rapid diagnostic test (RDT) that is simple, easy to use and interpret, can be stored without refrigeration and is reasonably priced so that it can be deployed widely would be useful for the early confirmation of cholera outbreaks. Ideally, the RDT should be highly sensitive so as not to miss the diagnosis of cases and be sufficiently specific when used under actual field conditions [9]. Cholera confirmation would enable immediate implementation of control measures such as reactive vaccination [6], as well as more accurate reporting of the burden of the disease.

A cholera RDT based on the detection of lipopolysaccharide (LPS) using gold particles was developed by the Institute Pasteur (IP). The RDT is a lateral flow immunochromatographic test for the qualitative determination of lipopolysaccharide antigen of both *Vibrio cholerae* O1 and O139 serogroups from stool specimens using monoclonal antibodies specific to *V. cholerae* O1 and O139 LPS. Through a licensure agreement, the RDT is now being produced by Span Diagnostics (Surat, India) under the trade name Crystal VC™ at a price of 19.00 USD / test kit (10 test strips). The test kit is stable at temperatures between 4°C to 30°C, and test strips are packed in waterproof pouches, allowing storage under high humidity conditions. Previous evaluations have been performed on the prototype and commercial versions of the RDT [10-15]. The primary objective of this study is to validate the current version of the Crystal VC™ RDT when performed by health workers in first-level treatment centers in Zanzibar. We also sought to assess if the RDT results were affected by the skill level of the reader and previous intake of antibiotics or intravenous fluids.

Methods

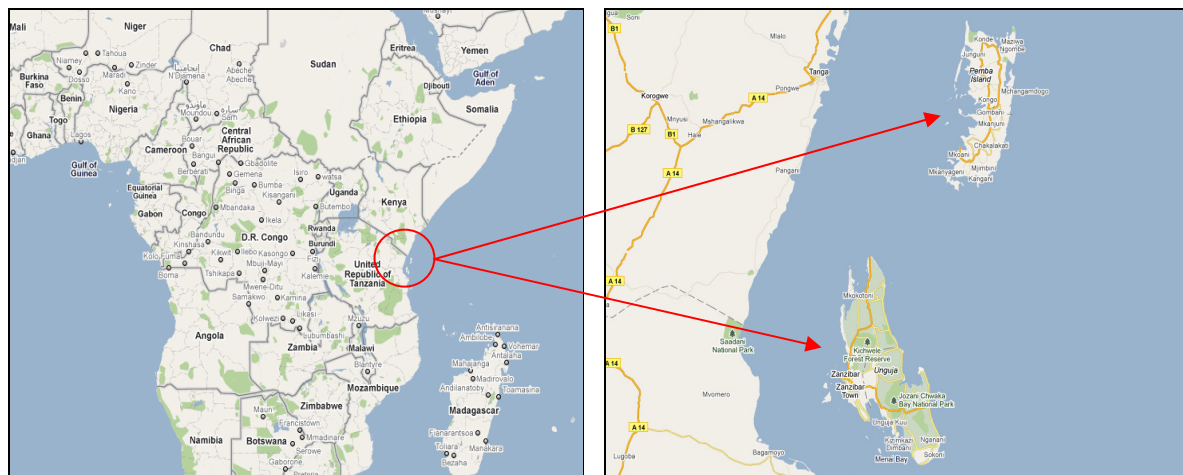
Ethics

The study was conducted according to the principles expressed in the Declaration of Helsinki. Written informed consent was obtained from all participants. The Zanzibar Research Council Ethics Committee, the Institutional Review Board of the International Vaccine Institute, Seoul, Korea, and the Research Ethics Review Committee of the World Health Organization, Geneva, Switzerland approved this project.

Study site

The archipelago of Zanzibar lies about 50 kilometers east of mainland Tanzania and consists of two main islands, Unguja and Pemba, as well as smaller islets (Figure 1). In 2009, Zanzibar had a population of about 1.22 million [16]. Stool samples were collected at cholera treatment camps that were set up during outbreaks on the two main islands, Unguja and Pemba, in 2009. Treatment of patients was provided according to national guidelines.

Figure 1: Study site



Acute watery diarrhea was defined as a minimum of three liquid, non-bloody, stools within 24 hours. Prior to presentation, no further inclusion / exclusion criteria were applied. Patients presenting with acute watery diarrhea were requested to provide a stool sample in a disposable plastic container. A swab was inserted into the stool sample and used to inoculate a tube of pre-packaged Cary Blair medium (EIKEN, Japan) for transport to the laboratory. About 200µl of stool from each sample was used for dipstick testing on site. A case report form (CRF) was completed to record frequency of

bowel movement over the previous 24 hours, antibiotics received, and fluid management (intravenous (IV) or oral rehydration solution (ORS)) provided at the health center. Bulk stool from a subset of patients attending a camp close to one of the participating laboratories was transported to the lab for additional testing on the same day as described below.

Dipstick test

The RDT was stored at room temperature and performed according to the package insert. Liquid stool was collected in a disposable plastic container. Approximately 200µl (4 drops) of stool were transferred with a disposable pipette to a disposable test tube provided with the kit. One drop of dilution buffer was added. The dipstick was inserted into the diluted stool and results were read within 15 – 20 minutes. The appearance of two bands on the dipstick, one control and one test, indicated that the stool sample was positive for *V. cholerae*. The appearance of only the control band indicated a negative sample. The non-appearance of the control band indicated a procedural error. Stool samples were tested under field conditions in the cholera treatment centers and in the laboratory as described below. Dipstick results were recorded on the CRF, whereas laboratory results were recorded in separate laboratory forms.

Performed under field conditions

A local health worker in each cholera treatment camp performed the RDT after training and a copy of the English test kit manual containing illustrations on test procedure and interpretation had been provided. Training consisted of a theoretical session using a Power Point Presentation containing information on test procedures and schematic pictures of positive and negative test results based on the package insert. This was followed by a practical session during which the test was performed a number of times. All field workers were visited frequently in the field to ensure correct handling of the test. All local health workers had completed at least primary education and delivered basic medical services to attending diarrheal patients. Fieldworkers performed the test outdoors in daylight.

Performed under laboratory conditions

In order to assess the potential influence of environmental and light conditions, laboratory technicians were asked to repeat the test on bulk stool collected at the camps and to read the result independently. Stool culture results were not yet available at the time of the performance of the test and laboratory technicians were blinded as to the results of the RDT performed in the field, as well as the clinical picture of the patient. Two laboratory technicians performed the RDT after receiving training similar to the field workers and had received a copy of the test kit manual. The laboratory

technicians performed the test indoors using electric light sources. All participating technicians had a diploma in laboratory sciences, which requires a minimum of one year of education, and had a minimum of three years working experience.

Stool culture

Upon arrival in the laboratory, samples from the Cary Blair media were streaked out on Thiosulphate Citrate Bile Sucrose Agar (TCBS; EIKEN, Japan), inoculated in alkaline peptone water (APW) and incubated at 37°C for 12 – 24 hours. If samples arrived as bulk stool, the samples were diluted in APW. An aliquot was streaked out on TCBS and the samples in APW and on TCBS were incubated for 12 – 24 hours at 37°C. If no growth on TCBS was detected after incubation, an aliquot of the sample in APW was streaked out on TCBS and incubated again. If yellow colonies indicative of *V. cholerae* were detected on TCBS, motility indole ornithine agar (MIO) and triple sugar iron agar (TSI) were inoculated with colonies from TCBS and incubated for 18 hours at 37°C. In addition, a colony from TCBS was sub-cultivated on gelatin agar for later serological confirmation and incubated at 37°C overnight. If colonies indicative of *V. cholerae* were observed on TSI and MIO after incubation, colonies from gelatin agar were tested for agglutination reactions with O1 polyvalent, O1 Inaba, O1 Ogawa and O139 antiserum (Beckton Dickinson, USA) as described elsewhere [17]. *V. cholerae* strains were transported to the National Institute of Cholera and Enteric Diseases in Kolkata, India where identification of the isolates was confirmed.

Definitions, data management and analysis

The CRF and laboratory results of each patient were computerized and linked using unique study identification numbers. The primary endpoint was the assessment of the performance of the RDT (done in the field) using microbiological stool culture result as the gold standard for comparison. Sensitivity (true-positive or TP rate) was defined as the probability that patients with laboratory-confirmed cholera had a positive RDT. Specificity (true-negative or TN rate) was the probability that patients with no laboratory-confirmed cholera had a negative RDT. The positive predictive value (PPV) was the probability that patients with a positive RDT had *V.cholerae* isolated from stool culture. The negative predictive value (NPV) was the probability that patients with a negative RDT had no *V.cholerae* isolated from a stool culture. The false positivity or FP rate ($= \text{FP} / [\text{FP} + \text{TN}]$ or $1 - \text{specificity}$) was the proportion of stool samples with no *V. cholerae* isolated on culture but showed a positive RDT result. The false negativity or FN rate ($= \text{FN} / [\text{TP} + \text{FN}]$ or $1 - \text{sensitivity}$) was the proportion of stool samples with *V. cholerae* isolated on culture but showed a negative RDT result.

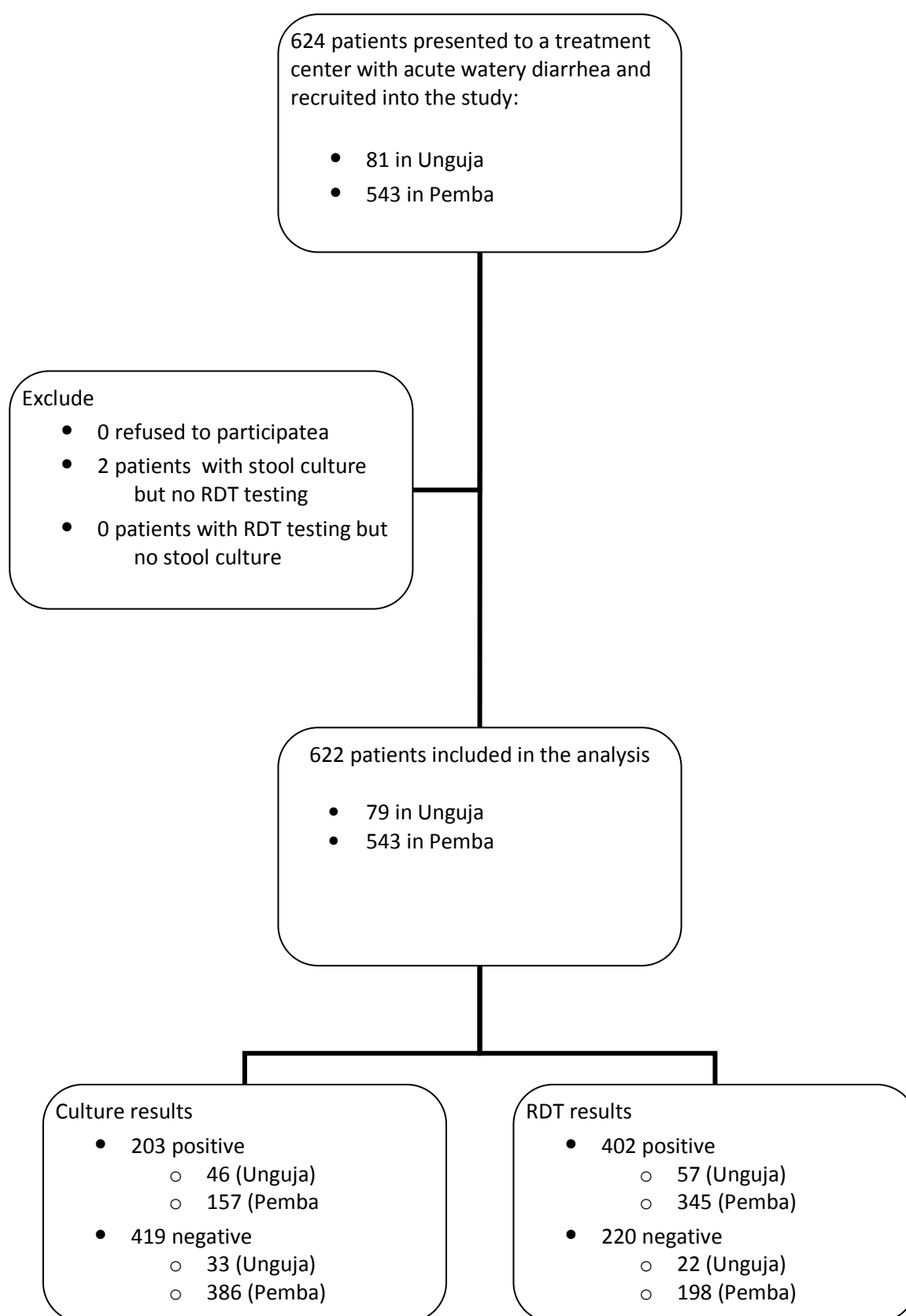
We performed sub-group analyses by island (Pemba or Unguja), by previous recent intake versus non-intake of antibiotics and by receipt of intravenous fluids following previously published studies showing differences in RDT performance [13]. We defined recent intake of antibiotics as receipt of any oral or parenteral antimicrobial for the current illness prior to the collection of a stool sample. We classified fluid management at the treatment center as oral rehydration solution (ORS) or intravenous fluids (IVF) with or without ORS. To assess whether test validity was related to skill level of the reader and location of the testing, we compared the performance of the RDT when done in the field versus in the laboratory on a subset of samples.

Comparison of unpaired samples was done using chi-square test; comparison of paired samples was done using McNemars test. Confidence intervals were calculated using exact method. Level of agreement was calculated using Cohen's Kappa test for unweighted proportions. Calculations were done using Stata, version 10 (StataCorp, College Station, TX, USA). Performance indicators were calculated using Excel 2010 (Microsoft, WA, USA).

Results:

There were 624 patients who presented to a cholera treatment camp with acute watery diarrhea and were recruited into the study: 81 in Unguja and 543 in Pemba. We excluded 2 samples sent for culture but on which no RDT was done. A total of 622 stool samples were included in the analysis, 79 (13%) from Unguja and 543 (87%) from Pemba residents (Figure 2).

Figure 2: Flow of study participants



Performance of the RDT in the field

Of the 622 stool samples, 203 (32.6%) yielded *V. cholerae* O1. No *V. cholerae* O139 was isolated. Using culture results as the gold standard, we calculated the sensitivity, specificity, PPV, and NPV of the RDT performed in the field for the diagnosis of cholera (Table 1). Overall sensitivity was 93.1% (95%CI: 88.7 to 96.2%), specificity was 49.2% (95%CI: 44.3 to 54.1%), the positive predictive value (PPV) was 47.0% (95%CI: 42.1 to 52.0%) and the negative predictive value (NPV) was 93.6% (95%CI: 89.6 to 96.5%). The overall false positivity rate was 50.8% (213/419).

Sub-group analyses of performance of the RDT in the field

We evaluated the RDT performance by island (Table 1). *V. cholerae* was isolated from 46/79 or 58.2% of stool samples from Unguja compared to a significantly lower proportion of 157/543 or 28.9%, from Pemba ($p<0.01$). No significant differences in sensitivity, specificity and NPV of the RDT were observed between Unguja and Pemba, as well as between each island with the overall results (all $p>0.05$). However, we found a significant difference in PPV between Unguja and Pemba (71.9, 95%CI: 58.5-83.0 versus 42.9, 95%CI: 37.6-48.3; $p=0.02$).

Table 1: Performance of the cholera rapid diagnostic test, Pemba and Unguja, Zanzibar

	Sensitivity% (95%CI*) (TP/(TP+FN))	Specificity% (95%CI*) (TN/(TN+FP))	PPV% (95%CI*) (TP/(TP+FP))	NPV% (95%CI*) (TN/(TN+FN))	No. pos. / total (%) with <i>V. cholera</i> isolated on stool culture
Total (n=622)	93.1 (88.7-96.2) (189/203)	49.2 (44.3-54.1) (206/419)	47.0 (42.1-52.0) (189/402)	93.6 (89.6-96.5) (206/220)	203 / 622 (32.6)
Unguja (n=79)	89.1 (76.4-96.4) (41/46)	51.5 (33.5-69.2) (17/33)	71.9 (58.5-83.0) (41/57)	77.3 (54.6-92.2) (17/22)	46 / 79 (58.2)
Pemba (n=543)	94.3 (89.4-97.4) (148/157)	49.0 (43.9-54.1) (189/386)	42.9 (37.6-48.3) (148/345)	95.5 (91.5-97.9) (189/198)	157 / 543 (28.9)
p	0.82	0.87	0.02	0.53	<0.01

*using exact method

We compared the RDT performance by fluid management of patients: Rehydration treatment at the cholera camp was recorded for 592 / 622 (95.2%) participants (Table 2). Only 15.2% (32/210) of participants who received oral rehydration had a positive stool culture, compared with 43.2% (165/382) of those who received IV fluids ($p<0.01$). There were no statistically significant differences in sensitivity, specificity and NPV of RDT performance by fluids received (all $p>0.05$). However we again found a significant difference in PPV among those who received oral compared to IV

rehydration (26.8%, 95%CI: 18.9 to 36.0% versus 55.2%, 95%CI: 49.2 to 61.2%; $p<0.01$). To evaluate whether the provided fluid treatment biased the field workers' interpretation of the RDT results, we compared the false positivity rate by fluid management. The false positivity rate was 46.1% among those who were orally rehydrated and 57.1% among those who were intravenously rehydrated ($p=0.22$).

Information on prior antibiotic treatment was recorded for 576 / 622 (92.6%) participants. The percentage with a positive stool culture was 27.6% among those who had received antibiotics and 32.2% among those who had not ($p=0.60$). We assessed whether previous antibiotic treatment affected the RDT performance (Table 2). Sensitivity, specificity, PPV and NPV did not vary significantly among recipients and non-recipients of antibiotics. The false positivity rate was 40.5% among those who had received antibiotics and 50.7% among those who had not ($p=0.45$).

Table 2: Stratified analysis of the performance of cholera dipstick test according to fluid management (oral rehydration or intravenous fluids) in 592 patients and recent antibiotic intake (yes or no) in 576 patients

	Sensitivity (95%CI)* (TP/(TP+FN))	Specificity (95%CI)* (TN/(TN+FP))	PPV (95%CI)* (TP/(TP+FP))	NPV (95%CI)* (TN/(TN+FN))	No. pos. / total (%) with <i>V. cholerae</i> isolated on stool culture
Fluid management (n=592)					
Oral rehydration n=210	93.8 (79.2-99.2) (30/32)	53.9 (46.3-61.4) (96/178)	26.8 (18.9-36.0) (30/112)	98.0 (92.8-99.8) (96/98)	32** / 210(15.2)
Intravenous rehydration n=382	92.7 (87.6-96.1) (153/165)	42.9 (36.2-49.7) (93/217)	55.2 (49.2-61.2) (153/277)	88.6 (80.9-94.0) (93/105)	165** / 382 (43.2)
p	0.97	0.19	<0.01	0.62	<0.01
Recent antibiotic (AB) intake (n=576)					
AB received n=58	93.8 (69.7-99.8) (15/16)	59.5 (43.3-74.4) (25/42)	46.9 (29.1-65.3) (15/32)	96.2 (80.4-99.9) (25/26)	16 / 58 (27.6)
No AB received n=519	92.8 (87.8-96.2) (155/167)	49.3 (43.9-54.7) (173/351)	46.5 (41.1-52.1) (155/333)	93.5 (86.1-94.7) (173/190)	167 / 518 (32.2)
p	0.98	0.48	0.98	0.86	0.60

*using exact method

** in 6/203 culture positive cases no rehydration treatment was provided

Comparison of the performance of the RDT in the field versus in the laboratory

We compared the performance of RDT on a subset of 67/79 (84.8%) stool samples from Unguja tested both in the field and in the laboratory (Table 3). In this subset, 40/67 (59.7%) samples yielded *V. cholerae* on culture. There was no statistically significant difference in the sensitivity, specificity PPV and NPV of the RDT's performance (all $p > 0.05$). The false positivity rate of the RDT was 45.4% in the field and 25.9% in the laboratory (Cohen's kappa 0.8).

Table 3: Comparison of the field and laboratory performance of the cholera dipstick test

(n= 67)	Sensitivity% (95%CI*) (TP/TP+FN)	Specificity% (95%CI*) (TN/TN+FP)	PPV% (95%CI*) (TP/TP+FP)	NPV% (95%CI*) (TN/TN+FN)	No. (%) with <i>V. cholera</i> isolated on stool culture**
Field	90.0 (76.3-97.2) (36/40)	55.6 (35.3-74.5) (15/27)	75.0 (60.4-86.4) (36/48)	78.9 (54.4-93.9) (15/19)	40 (59.7)
Laboratory	87.5 (73.2-95.8) (35/40)	74.1 (53.7-88.9) (20/27)	83.3 (68.6-93.0) (35/42)	80.0 (59.3-93.2) (20/25)	40 (59.7)
p***	0.931	0.510	0.740	0.977	

*using exact method

**rapid diagnostic testing was done both in the field and in the laboratory on specimens from the same stool samples

***using McNemars test

Operational characteristics

The test procedure, excluding sample collection, requires 20-25 minutes. The test kit manual provides clear instructions, and handling of the test was considered simple by field workers. Field workers found it easy to distinguish between valid and non-valid test results, based on the appearance of a control line. However, very faint positive test lines were interpreted as a positive result, but could not be confirmed by culture.

Discussion

We found an overall high sensitivity (93.1%) of the current version of the cholera RDT consistent with previous reports but a much poorer specificity (49.2%). Earlier studies were performed using a prototype dipstick – developed by the Institute Pasteur – and earlier versions of the commercial kit. Studies using the prototype versions of the RDT reported sensitivities in the range of 93% to 99% and specificities of 67%-97% [12-15]], whereas more recent reports on earlier versions of the commercial kit showed sensitivities of 92-97% and specificities of 71-76% [10, 11]. These studies not only varied by the RDT version used but also by the methodology and qualifications of the study personnel

performing the RDT (Table 4). There were also variations in the test procedure such as addition or non-addition of a buffer solution to the sample. In some studies, a 4-hour incubation step in alkaline peptone water was added. Overall, the prototype and precursor commercial kits performed better than the current version tested.

The poor specificity of the current version of the commercial kit was associated with an overall false positive rate of 50.8%. The RDT's false positivity rate when the RDT was read in the field was 44% versus 26% when done by laboratory technicians; possibly faint test lines on the dipsticks were over-read by field workers as positive. We hypothesized whether patient characteristics (fluid management or receipt of antibiotics) biased the field workers' interpretation of the RDT results. However, we found no significant differences in false positivity in these sub-group analyses. More likely the fieldworkers over-interpreted faint test lines which could be recognized in daylight but not in the indoor laboratory setting.

Previously, Kalluri et al. assessed the impact of the reader's qualification on the performance of the prototype test [12]. Laboratory technicians with several years of working experience as well as field workers with at least a college degree but no laboratory experience were asked to perform the test on 304 stool samples. The reported RDT sensitivities of 94% and 93% when done by laboratory technicians and field workers, respectively, were similar, but RDT specificity was higher when performed by the technicians (76% versus 67%) [12]. Harris et al. report a sensitivity of 97% and a specificity of 71-76%, when staff with graduate-level laboratory training performed the test in Guinea Bissau [10]. Mukherjee et al. reported a similar sensitivity and specificity (92% and 73%, respectively) when the test was done by graduate-level staff during a surveillance study at a hospital in Kolkata [11].

In contrast to Wang et al., we did not find a higher sensitivity of the RDT when testing stool samples from patients receiving IVF compared to samples from patients who did not receive IVF [13]. However, we noted that the PPV in this study varied according to the proportion of culture-positive specimens. It has been argued that PPV is the most important measure of a clinical diagnostic method since it represents the proportion of patients with positive test results that are correctly diagnosed [18]. The PPV is not intrinsic to the test; it is affected by prevalence of the disease. For example, the PPV was 55% for samples from patients given IVF (43% of whom had *V. cholerae* isolated) while it was 27% for samples from patients managed with ORS (15% of whom had *V. cholerae* isolated). The PPV was 71.9% for the Unguja sub-sample with 58% cholera confirmation while for the Pemba sub-sample it was 43% with 29% cholera confirmation. In outbreak settings,

when a large proportion of patients presenting with acute watery diarrhea have cholera, a positive RDT result would have a good predictive value. In other situations (e.g. areas with seasonal cholera but also high rates of diarrheal diseases from other pathogens), the RDT may be less useful.

Our study has several limitations. Firstly, a large sample from 622 study participants was available for the overall evaluation, but only 67 stool samples were used for sub-analyses. Secondly, while confirmation of *V. cholerae* isolates was performed at a reference laboratory, culture-negative stool samples were not validated further. In particular, we did not perform PCR testing on our RDT-positive, culture-negative samples. Bhuiyan, et al. [14] analyzed five stool samples collected in Bangladesh by multiplex PCR that were O1 dipstick positive but culture-negative and found that all five were negative by PCR, indicating that the five dipstick-positive results were false positives. This is reassuring but does not entirely exclude the possibility of false negativity by stool culture. Thirdly, Alam et al. pointed out that the dipstick may detect non-culturable forms of *V. cholerae* that have transformed into a coccoid form due to unfavorable intra-host conditions, such as antibiotic treatment prior to testing [7]. We tried to assess the influence antibiotic treatment prior to sample collection may have had on our results but found no significant difference ($p>0.05$) in the false positive rates among participants who had taken antibiotics prior to sample collection and those who had not. However, further research is needed to rule out the possibility that the RDT may detect *V. cholerae* antigen in some specimens which are culture negative.

Table 4: Validation studies of the Institut Pasteur prototype and Crystal VC™ test for diagnosis of *Vibrio cholerae* (O1 samples only considered)

RDT	Institut Pasteur prototype				Crystal VC™ (Span Diagnostics, India)		
Authors	F Nato, et.al. [14]	NA Bhuiyan, et.al. [13]	X-Y Wang, et.al. [12]	Kalluri, et.al. [11]	J R Harris, et.al. [9]	P. Mukherjee et al. [10]	This Paper
Journal, Year	Clinical and Diagnostic Laboratory Immunology, 2003	Journal of Clinical Microbiology, 2003	BMC Infectious Diseases, 2006	Tropical Medicine and International Health, 2006	Tropical Medicine and International Health, 2009	Japanese Journal of Infectious Diseases, 2010	This Journal
Samples from	Madagascar	Bangladesh	Mozambique	Bangladesh	Guinea-Bissau	India	Zanzibar / Tanzania
Gold standard; type of stool samples used	Culture of frozen stool or rapid cultures of stool collected on filter paper	Culture of rectal swab in C-B media (and multiplex PCR when RDT+/culture -)	Culture of rectal swabs in C-B media and bulk stool	Culture of bulk stool	PCR of swab in C-B media	Culture of bulk stool	Culture of rectal swab in C-B media
Type of stool samples for RDT testing	frozen stool/rapid cultures of stool collected on filter paper	Rectal swab incubated in APW at 37°C for 4 h	Rectal swab incubated in APW at 37°C / 4 h and fresh, bulk stool	Fresh, bulk stool	Fresh, bulk stool	Fresh, bulk stool	Fresh, bulk stool
RDT done by	?	?	Trained technician at field laboratory	Trained technician at ICDDR,B and trained field 'paramedics'	Individuals with graduate-level laboratory technical training	Post graduate investigator	<ul style="list-style-type: none"> – Field worker at cholera treatment center (total) – Field worker at cholera treatment center (subset) – Laboratory technician (subset)
No of samples	140	134	<ul style="list-style-type: none"> – 219 rectal swabs in APW – 172 bulk stool 	304	101	212	<ul style="list-style-type: none"> – 622 (total) – 67(subset)
No (%) positive for VC O1 by gold standard	65 (46%)	68 (51%)	138 (35%)	116 (38%)	65 (64%)	72 (34%)	<ul style="list-style-type: none"> – 203 (33%) (total) – 40 (60%) (subset)
Sensitivity	98.5%	96%	<ul style="list-style-type: none"> – 97% rectal swabs in APW – 93% bulk stool 	<ul style="list-style-type: none"> – 94% lab technicians – 93% field 'paramedics' 	97%	92%	<ul style="list-style-type: none"> – 93% (total) – 88% lab technician (subset) – 90% field worker (subset)
Specificity	96%	92%	<ul style="list-style-type: none"> – 97% rectal swabs in APW – 77% bulk stool 	<ul style="list-style-type: none"> – 76% lab technicians – 67% by field 'paramedics' 	71–76%;	73%	<ul style="list-style-type: none"> – 49% (total) – 74% lab technician (subset) – 56% field worker (subset)
PPV	95.6%	93%	<ul style="list-style-type: none"> – 94% rectal swabs in APW – 74% bulk stool 	<ul style="list-style-type: none"> – 70% lab technicians – 63% field 'paramedics' 	87–89%	64%	<ul style="list-style-type: none"> – 47% (total) – 83% lab technician (subset) – 75% field worker (subset)
NPV	99%	95%	?	<ul style="list-style-type: none"> – 95% lab technicians – 94% field 'paramedics' 	92–93%	95%	<ul style="list-style-type: none"> – 94% (total) – 80% lab technician (subset) – 79% field worker (subset)

Conclusion

We found that field workers in this study who had basic general education but were not familiar with laboratory work experienced difficulties in interpreting the RDT performed in the cholera camps. If the RDT is to be deployed more widely, more extensive and repeated training may be required to improve the current RDT's specificity. The test cannot replace stool culture and due to the high number of false positive results observed is not suitable to trigger an outbreak response in a resource poor setting. However the test may be potentially used as a screening tool. During cholera outbreaks, especially when several samples test positive, the test has an enhanced predictive value. Further research is needed to evaluate the accuracy of the RDT with specimens which have been incubated in APW for 4 to 6 hours prior to testing in the RDT since this procedure should dilute out the materials in stool samples which are causing the false positive results while amplifying the antigen signal from the *V. cholerae*.

Acknowledgement

We would like to thank Marie Kieney and Raymond Hutubessy from the World Health Organization who have supported this project throughout the entire study period.

We are also grateful to all patients who made this work possible. We thank all of the technical staff and research assistants who were involved in the study. This project was coordinated by the WHO Initiative for Vaccine Research, Geneva, Switzerland and the International Vaccine Institute, Seoul, Korea and received financial support from the Bill & Melinda Gates Foundation. Additional funding was provided by the Swedish International Development Cooperation Agency and the Republic of Korea.

Financial Disclosure

This study has been funded by the Bill and Melinda Gates Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. World Health Organization (2011) Number of Cholera Cases. Available from Global Health Observatory (GHO):
http://www.who.int/gho/epidemic_diseases/cholera/situation_trends_cases/en/index.html. Accessed 15 August 2011.
2. World Health Organization (2011) Cholera: Fact sheet No. 107. Available from WHO Media Centre: <http://www.who.int/mediacentre/factsheets/fs107/en/>. Accessed 09. November 2011
3. Center for Disease Control (2010). Cholera. Available from CDC Homepage:
<http://www.cdc.gov/cholera/disease.html>. Accessed 15 August 2011.
4. Global Task Force on Cholera Control (2004) Case management: treatment. In: Cholera Outbreak - Assessing The Outbreak Response And Improving Preparedness. Geneva: World Health Organization. pp. 26-30
5. Chaignat C-L (2006). Cholera outbreak response in Pemba, 2006 And Pertinence of using Oral Cholera Vaccines for Control. Available from WHO Homepage:
<http://www.who.int/cholera/technical/CholeraZanzibarReviewJuly2006.pdf>. Accessed 15 August 2011
6. Reyburn R, Deen JL, Grais RF, Bhattacharya SK, Sur D, et al. (2011) The case for Reactive Mass Oral Cholera Vaccinations. PLoS Negl. Trop. Dis. 5(1): e952.
doi:10.1371/journal.pntd.0000952.
7. Alam M, Hasan NA, Sultana M, Nair GB, Sadique A, et al. (2010) Diagnostic Limitations to Accurate Diagnosis of Cholera. Journal of Clinical Microbiology, 48(11): 3918-3922.
doi:10.1128/JCM.00616-10
8. Riewpaiboon A, Intrapakan K, Phoungkatesunthorn S. (2008) Predicting Treatment Cost for Bacterial Diarrhoea at a Regional Hospital in Thailand. J. Health Popul. Nutr., 26(4): 442-450.

9. The TDR Diagnostics Evaluation Expert Panel (2010) Evaluation of diagnostic tests for infectious diseases: general principles. *Nature Reviews – Microbiology*, pp17-29 doi: 10.1038/nrmicro1523

10. Harris JR, Cavallaro EC, de Nobrega AA, dos S. Barrado JC, Bopp C, et al. (2009) Field Evaluation of Crystal VC Rapid Dipstick test for cholera during a cholera outbreak in Guinea - Bissau. *Tropical Medicine and International Health*, 14(9): 1117-1121. doi:10.1111/j.1365-3156.2009.02335.x

11. Mukherjee P, Ghosh S, Ramamurthy T, Bhattacharya MK, Nandy RK, et al. (2010) Evaluation of a Rapid Immunochromatographic Dipstick Kit for Diagnosis of Cholera Emphasizes Its Outbreak Utility. *Jpn. J. Infect. Dis.*, 63(4): 234-238.

12. Kalluri P, Naheed A, Rahman S, Ansaruzzaman M, Faruque AS, Bird, et al. (2006) Evaluation of three rapid diagnostic tests for cholera: does the skill level of the technician matter? *Tropical Medicine and International Health*, 11(1): 49-55. doi:10.1111/j.1365-3156.2005.01539.x

13. Wang X-Y, Ansaruzzaman M, Vaz R, Mondlane C, Lucas ME, et al. (2006) Field evaluation of a rapid immunochromatographic dipstick test for the diagnosis of cholera in a high-risk population. *BMC Infectious Diseases*, 6(17). doi:10.1186/1471-2334-6-17

14. Bhuiyan NA, Qadri F, Faruque AS, Malek MA, Salam MA, et al. (2003) Use of Dipsticks for Rapid Diagnosis of Cholera Caused by *Vibrio cholerae* O1 and O139 from Rectal Swabs. *Journal of Clinical Microbiology*, 41(8): 3939-3941. doi: 10.1128/JCM.41.8.3939-3941.2003

15. Nato F, Boutonnier A, Rajerison M, Grosjean P, Darteville S, et al. (2003) One-Step Immunochromatographic Dipstick Tests for Rapid Detection of *Vibrio cholerae* O1 and O139 in Stool Samples. *Clinical And Diagnostic Laboratory Immunology*, 10(3): 476-478. doi: 10.1128/CDLI.10.3.476-478.2003

16. National Bureau of Statistics, Ministry of Planning, Economy and Empowerment (2006) Analytical Report, Volume X, p2. Available from National Bureau of Standards (NBS) homepage: <http://www.nbs.go.tz/pdf/2002popcensus.pdf> Accessed on 29 August 2011

17. Bopp C, Ries AA, Wells JG, Koplan JP Hughes, JM et al. (1999) Laboratory Methods for the Diagnosis of Epidemic Dysentery and Cholera. Geneva. Available from CDC homepage: <http://www.cdc.gov/ncidod/dbmd/diseaseinfo/cholera/complete.pdf> Accessed on 29 August 2011
18. Altman DG, and Bland JM (1994) Diagnostic tests 2: Predictive values. *BMJ*, 309 (6947):102.

Discussion:

Neither of the evaluated tests performed at sensitivity and specificity $\geq 95\%$, the recommended level of performance for Malaria RDTs as suggested by the WHO (WHO, 2006).

Table 1: Overview of test performance

Test	Sensitivity (%; 95% CI)	Specificity (%; 95% CI)	PPV	NPV
Widal	75 (51-90)	98 (97-99)	26*	100*
Tubex	79 (52-81)	89 (81-94)	7**	100**
Crystal VC	93 (89-96)	49 (44-54)	47***	94***

*observed prevalence of 1%

**calculated at 1% prevalence

***observed prevalence of 33%

Considering the main operational areas for RDTs, three fields for application can be identified: prospective diagnosis for treatment, retrospective diagnosis for surveillance purposes, and as a replacement for complicated and expensive laboratory procedures.

In view of these areas, our findings do not support use of the Widal tube agglutination test for direct diagnostic purposes. The test performed best at a cut off titer of 1:80 for TH with an observed PPV of 26% and an NPV of 100%. While a prevalence of 1% must be considered high (Crump et al, 2004) even at a higher prevalence of 5%, the PPV remains at 66%. The NPV is high at 100% and 99% respectively.

Sensitivity and specificity of the Widal test were 75% and 98% respectively, figures which may be considered when creating correcting factors for surveillance data of samples from the region of Teule, Tanzania. However samples from Zanzibar, Tanzania, tested a year later with a different batch of the same test at the same laboratory and using a comparable gold standard showed sensitivity and specificity of 48% and 99% respectively at a cut off titre of 1:80 (TO and TH) (Thriemer et al, in preparation), which could suggest inter-batch variations.

The TUBEX test showed a PPV of 7% at a prevalence of 1%, increasing to 27% at 5% prevalence, with an NPV of 100% and 81% respectively. The test is suitable to exclude typhoid fever, however chances of a correct positive result are low. Predictive values of the Tubex test were inferior to the Widal test at a higher price (2.15 USD vs. 0.50 USD). However the test provides a shorter time to result and does not require a setting and time specific validation study.

We found sensitivity and specificity of the Tubex test to be comparable ($p>0.05$) to the Widal test when tested on the same samples. Despite the superior operational characteristics of the Tubex, the Widal is preferable for RDT based surveillance in developing countries.

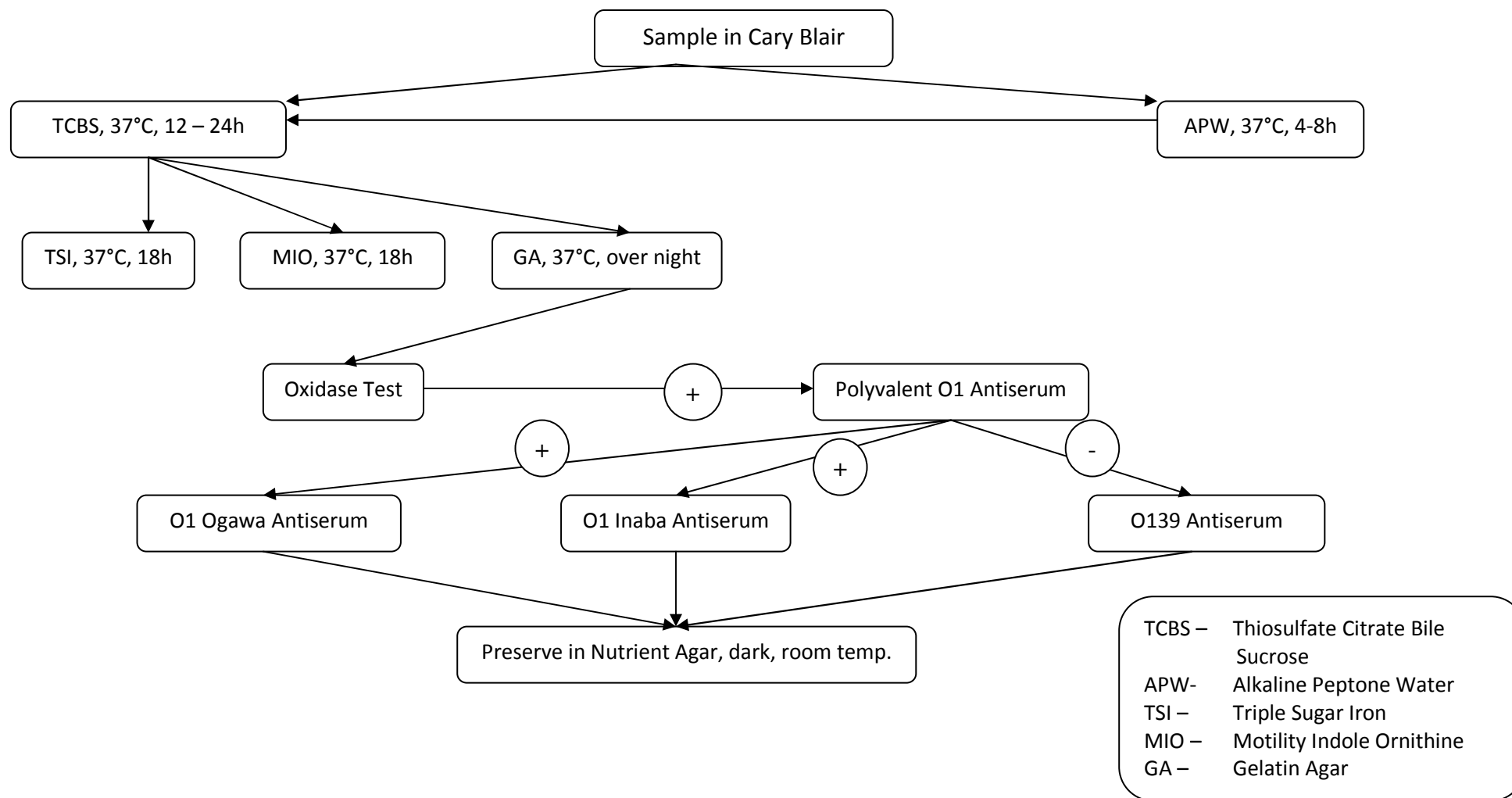
Introducing either the Widal tube agglutination test or the Tubex test into routine microbiologic testing does not seem reasonable. Neither test performed at a level that would allow confident diagnosis. Both tests detect antibodies, serum samples tested can't be analyzed for additional bacterial pathogens to the knowledge of the author. Since typhoid fever does not have a specific clinical picture (Thriemer et al, in preparation) which would allow to narrow down the number of samples assessed and thereby increase prevalence within the set of samples tested, neither test can be used as a final confirmation of disease due to the poor PPV observed .

Since treatment of cholera and other watery diarrheal diseases is mostly through fluid replacement (Kasper et al, 2005b), the use of a cholera RDT in an outbreak setting does not provide a significant benefit to patient or caretaker. However a good diagnostic tool to detect the index case at the beginning of an outbreak would be of great public health benefit. Applying observed sensitivity and specificity at a low prevalence of 1% we calculated a PPV and NPV of 2% and 100% respectively; detecting the index case at the beginning of an outbreak with the Crystal VC is unlikely.

Sensitivity and specificity of the Crystal VC were 93% and 49% respectively. In agreement with earlier studies (Kalluri et al, 2006) we observed a decrease in false positive test results when the test was performed by well trained staff. This favors using the test for surveillance purposes in a laboratory rather than as a diagnostic tool.

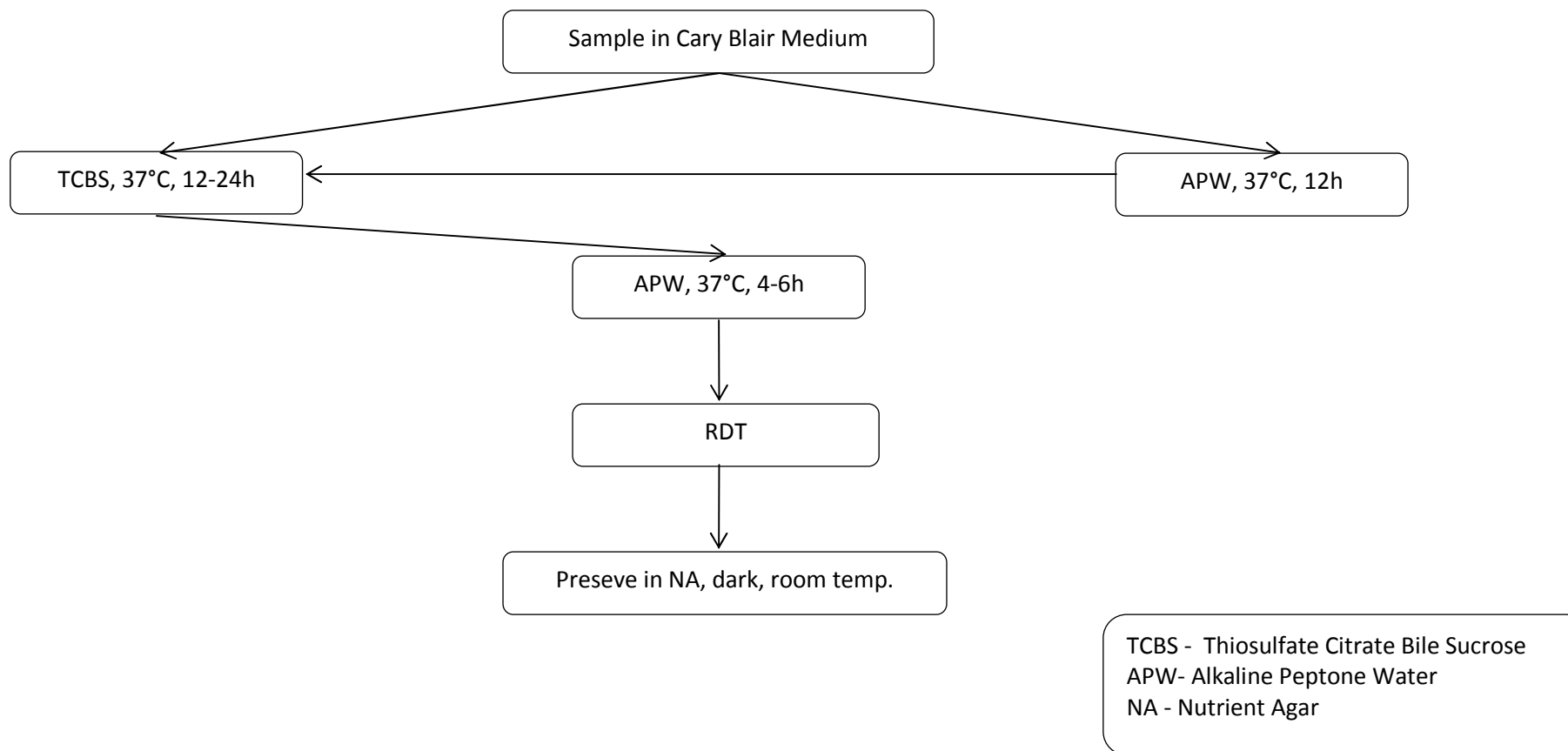
Despite the low performance indicators observed for the Crystal VC, the test can be included into standard lab procedures using TCBS (Fig.1, 2).

Figure 1: Isolation of *V. cholerae* using serology for confirmation



Adapted and modified from: Perilla, MJ. Ajello, G. Bopp, C. Elliott, J. Facklam, R. Knapp, JS. Popovic, T. Wells, J. Dowell, SF. (2003): *Vibrio cholerae*, *Manual for the Laboratory Identification and Antimicrobial Susceptibility Testing of Bacterial Pathogens of Public Health Importance in the Developing World*, WHO & CDC, p.143

Figure 2: Isolation of *V. cholerae* using RDT for confirmation



Since TCBS is a selective and differential medium that mainly supports growth of *V. cholerae* (Murray et al, 2007), any growth detected can be confirmed or excluded with the test. Prevalence within the sample set is likely to be high, resulting in a good PPV. Introducing the test into routine laboratory work in developing countries may reduce costs per sample by at least 2 USD. The test is easier and faster to perform than agglutination testing and will require very little quality control since the test is stable at room temperatures and to our knowledge no strong inter-batch variations have been reported to date.

Conclusion

None of the tests assessed performed well enough to be considered for diagnosis on its own. Widal tube agglutination test, Tubex and the Crystal VC may however be considered for surveillance purposes if validation studies, including establishing a cut off titer for the Widal tube agglutination test, are conducted. A possible application for the Crystal VC is the confirmation of suspected *V. cholerae* colonies from TCBS.

References:

- Altman, DG. Bland, MJ. (1994a): Diagnostic tests 1: sensitivity and specificity, *BMJ*, 308(11), pp.:1552-1553
- Altman, DG. Bland, MJ. (1994b): Diagnostic tests 2: predictive values, *BMJ*, 308(11), pp.:102-103
- Andrés Aguayo-Rico, Iris A. Guerra-Turrubiates, Ricardo Montes de Oca-Hernández (2005): Empirical Evidence of the Impact of Health on Economic Growth, *Issues in Political Economy*, Vol. 14
- Baker, S. Favorov, M. Dougan, G. (2010): Searching for the elusive typhoid diagnostic, *BMC Infectious Disease*, 10:45
- Bhuiyan NA, Qadri F, Faruque AS, Malek MA, Salam MA, et al. (2003) Use of Dipticks for Rapid Diagnosis of Cholera Caused by *Vibrio cholerae* O1 and O139 from Rectal Swabs. *Journal of Clinical Microbiology*, 41(8): 3939-3941. doi: 10.1128/JCM.41.8.3939–3941.2003
- Bhutta ZA: Current concepts in the diagnosis and treatment of typhoid fever. *BMJ* 2006, 333:78 – 82.
- Breiman, RF. Cosmas, L. Njuguna, H. Audi, A. Olack, B. Ochieng, JB. Wamola, N. Bigogo, GM. Awiti, G. Tabu, CW. Burke, H. Williamson, J. Oundo, JO. Mintz, ED. Feikin, DR. (2012): Population-based incidence of typhoid fever in an urban informal settlement and a rural area in Kenya: implications for typhoid vaccine use in Africa, *PLoS One*, 7 (1):e29119
- Bruce, N. Pope, D. Stanistreet, D. (2008): *Quantitative methods for Health Research*, John Wiley & Sons, Ltd, UK, pp.448-452
- Cheesborough, M. (2006a): *Salmonellae, Laboratory Practice in Tropical Countries*, 2nd Edition, Vol. II, pp. 189-194, Cambridge University Press, UK
- Cheesborough, M. (2006b): *Vibrio cholerae, Laboratory Practice in Tropical Countries*, 2nd Edition, Vol. II, pp. 182-186, Cambridge University Press, UK

Crump, J. A., S.P. Luby, E. D. Mintz. (2004): The global burden of typhoid fever, *Bulletin of the World Health Organization*, 82 (5):346—353.

Farooqui BJ, Khurshid M, Ashfaq MK, Khan MA: Comparative yield of *Salmonella typhi* from blood and bone marrow cultures in patients with fever of unknown origin. *J Clin Pathol*. 1991, 44:258-259.

Ghenghesh, KS. Franka, E. Tawil, K. Wasfy, MO. Ahmed, SF. Rubino, S. Klena, JD. (2009): Enteric Fever in Mediterranean North Africa, *Journal of Infections in Developing Countries*, 3(10):753-761

Gilman RH, Terminel M, Levine, Hernandez-Mendoza P, Hornick RB: Relative efficacy of blood, urine, rectal swab, bone marrow, and rose spot cultures for recovery of *Salmonella typhi* in typhoid fever. *Lancet* 1975, 31;1(7918):1211-1213.

Harris JR, Cavallaro EC, de Nobrega AA, dos S. Barrado JC, Bopp C, et al. (2009) Field Evaluation of Crystal VC Rapid Dipstick test for cholera during a cholera outbreak in Guinea - Bissau. *Tropical Medicine and International Health*, 14(9): 1117-1121. doi:10.1111/j.1365-3156.2009.02335.x

Kalluri, P. Naheed, A. Rahman, S. Ansaruzzaman, M. Faruque, ASG. Bird, M. Khatun, F. Bhuyina, NA. Nato, F. Fournier, J-M. Bopp, C. Breiman, RF. Nair, GB. Mintz, ED. (2006): Evaluation of three rapid diagnostic test for cholera: does the skill level of the technician matter?, *Tropical Medicine and International Health*, doi: 10.1111/j1365-3156.2005.01539.x

Kasper, DL. Braunwald, E. Fauci, A. Hauser, SL. Longo, DL. Jameson JL. (2005a): Salmonellosis, *Harrison's Principles of Internal Medicine, 16th Ed*, Vol. I, pp897-902, McGrawHill, Medical Publishing Division, USA

Kasper, DL. Braunwald, E. Fauci, A. Hauser, SL. Longo, DL. Jameson JL. (2005b): Cholera and other Vibrioses, *Harrison's Principles of Internal Medicine, 16th Ed*, Vol. I, pp909-914, McGrawHill, Medical Publishing Division, USA

Kawano, RI. Leano, SA. Agdamag, DMA. (2007): Comparisson of serological Test Kits for Diagnosis of Typhoid fever in the Philippines, *Journal of Clinical Microbiology*, Vol. 45, No.1 pp246 – 7

Lim, P-L, Tam, FCH. Cheong, Y-M, Jegathesan, M. (1998): One-Step 2-Minute Test To Detect Typhoid – Specific Antibodies Based on Particle Separation in Tubes, *Journal of Clinical Microbiology*, 36(8), pp2271-2278

Loong, Tze-Wey (2003): Clinical review – Understanding sensitivity and specificity with the right side of the brain, *BMJ*, doi:10.1136/bmj.327.7417.716

Miller SI, Pegues DA. (2000): Salmonella Species, Including Salmonella typhi In: Madell GL., Bennett JE, Dolin R. *Principles and Practice of Infectious Disease*. 5th edition. New York: Churchill Livingstone.

Mukherjee P, Ghosh S, Ramamurthy T, Bhattacharya MK, Nandy RK, et al. (2010) Evaluation of a Rapid Immunochromatographic Dipstick Kit for Diagnosis of Cholera Emphasizes Its Outbreak Utility. *Jpn. J. Infect. Dis.*, 63(4): 234-238.

Murray, PR. (Editor) (2007): Reagents, Stains and Media: Bacteriology, *Manual of CLINICAL MICROBIOLOGY*, Vol.I, p360, ASM Press, USA

Mweu, E. and M. English. (2008): Typhoid Fever in Children in Africa. *Trop Med Int Health* 13(4):532—540.

Nato F, Boutonnier A, Rajerison M, Grosjean P, Darteville S, et al. (2003) One-Step Immunochromatographic Dipstick Tests for Rapid Detection of *Vibrio cholerae* O1 and O139 in Stool Samples. *Clinical And Diagnostic Laboratory Immunology*, 10(3): 476-478. doi: 10.1128/CDLI.10.3.476–478.2003

Olopoenia, L. A., and A. I. King. (2000): Widal agglutination test – 100 years later: still plagued by controversy. *Postgrad Med J*. 76:80 – 84.

Parry, C. M. Hien, T. T. Dougan, G. White, N. Farrar, J. J. (2002): Typhoid Fever, *The New England Journal of Medicine* 347 (22): 1770 – 1782

Perilla, MJ. Ajello, G. Bopp, C. Elliott, J. Facklam, R. Knapp, JS. Popovic, T. Wells, J. Dowell, SF. (2003): *Manual for the Laboratory Identification and Antimicrobial Susceptibility Testing of Bacterial Pathogens of Public Health Importance in the Developing World*, WHO & CDC

Reyburn, R. Deen, JL. Grais, RF. Bhattacharya, SK. Sur, D. Lopez, AL. Jiddawi, MS. Clemens, JD. von Seidlein, L. (2011): The case for reactive mass oral cholera vaccinations, *PLoS Negl. Trop. Dis.*, 5(1):e952

Sack, D. (2012): *Global Cholera Estimates* [online]:
<http://www.jhsph.edu/dept/ih/globalhealthresearch/cholera/global.html> (last accessed on 25.01.2011)

Tam, FC. Lim, PL. (2003): The TUBEX typhoid test based on particle-inhibition immunoassays detects IgM but not IgG anti-O9 antibodies, *J. Immunol. Methods*, 282(1-2):83-91

Thriemer K, Ley B, Ame S, Deen J, , Pak GD, Chang NY, Hashim R Schmied WH, Busch CJ, Nixon S, Morrissey A, Puri MK, Ali M, Ochiai RL, Wierzba T, Jiddawi MS, Clemens JD, Ali SM, von Seidlein L. Clinical and Epidemiological Features of Typhoid Fever in Pemba, Zanzibar. *In preparation*

Thriemer, K. Ley, B. Ame, S. von Seidlein, L. Pak, GD. Chang, NY. Hashim, R. Schmied, WH. Busch, CJ. Nixon, S. Morrissey, A. Puri, MK. Ali, M. Ochiai, RL. Wierzba, T. Jiddawi, MS. Clemens, JD. Ali, SM. Deen, JL. (2012): The burden of invasive bacterial infections in Pemba, Zanzibar. *PLoS One*, 7(2):e30350

Wain J, Pham VB, Ha V, Nguyen NM, To SD, Walsh AL, Parry CM, Hasserjian RP, HoHo VA, Tran TH, Farrar J, White NJ, Day NP: Quantitation of bacteria in bone marrow from patients with typhoid fever: relationship between counts and clinical features. *J Clin Microbiol.* 2001, 39:1571-1576.

Wang X-Y, Ansaruzzaman M, Vaz R, Mondlane C, Lucas ME, et al. (2006) Field evaluation of a rapid immunochromatographic dipstick test for the diagnosis of cholera in a high-risk population. *BMC Infectious Diseases*, 6(17). doi:10.1186/1471-2334-6-17

WHO (2006): *TOWARDS QUALITY TESTING OF MALARIA RAPID DIAGNOSTIC TESTS:EVIDENCE AND METHODS*, p4

Wilke, A. Ergonul, O. Bayar, B. (2002): Widal Test in Diagnosis of Typhoid Fever in Turkey, *Clinical and Diagnostic Laboratory Immunology*, doi: 10.1128/CDLI.9.4.938–941.2002

World Health Organization (2009a): *Initiative for Vaccine Research – Typhoid Fever*, [online]: http://www.who.int/vaccine_research/diseases/diarrhoeal/en/index7.html (last accessed on 27.07.2010)

World Health Organization (2009b): An old enemy returns, *Bull World Health Organ.*, 87(2): 85 – 86

World Health Organization (2010): Cholera, *Fact sheet No. 107*, [online]: <http://www.who.int/mediacentre/factsheets/fs107/en/index.html> (last accessed on 27.07.2010)

Zuckerman, JN, Rombo, L. Fisch, A. (2007): The true burden and risk of cholera: implications for prevention and control, *Lancet*, 7(8):521-30

Summary

Introduction

Typhoid fever and Cholera are poverty associated diseases and of major public health concern in developing countries. For typhoid fever more than 21.000.000 cases at 1% fatality rate have been calculated for the year 2000, while cholera has an estimated incidence of 2.550.000 cases per year at a fatality rate of >8.2%. Lack of functioning infrastructure for fresh and waste water as arise in disaster zones and areas of high poverty are favoring conditions for both diseases.

Effective vaccinations for both diseases exist, treatment through antibiotics (Typhoid fever) and fluid replacement (Cholera) can significantly reduce fatality rates. However neither is permanently available in many developing countries due to lack of funding as a result of reduced tax income due to disease. Lack of funding also results in lack of surveillance data for decision makers and lack of reliable diagnostic tools for caretakers. Time consuming and expensive laboratory based diagnosis is the gold standard for either disease, however these require well trained laboratory technicians. A possible solution are rapid diagnostic tests, if they are reliable, have convenient operational characteristics and are cheap.

The aim of this cumulative thesis is to evaluate rapid diagnostic tests under developing country conditions and in outbreak settings. For this purpose we have assessed the widely in Africa in use Widal tube agglutination test (Span Diagnostics, India) and the more recently developed Tubex test (IDL, Sweden) for typhoid fever as well as a rapid diagnostic test for Cholera (Span Diagnostics, India).

Background

Possibly the oldest rapid diagnostic test is the Widal test, based on a macroscopically visible agglutination reaction between patient antibodies and added antigen. The more recently developed Tubex test is based on an antibody inhibition reaction, resulting in a color change of a test-solution when serum is added and placed on a magnetic tray. The Crystal VC is a lateral flow immuno-chromatographic test, detecting *Vibrio cholera* antigen.

All three tests were evaluated by calculating the retrospective, prevalence independent, performance indicators sensitivity and specificity as well as the prospective, prevalence dependent, performance indicators positive and negative predictive value (PPV and NPV). In order to address poor sensitivity of the chosen gold standard, blood culture, different control groups were evaluated for the Widal and Tubex test. The Crystal VC was tested during cholera outbreaks in cholera camps. We also assessed the influence of degree of education of the reader for the Crystal VC.

Results

Evaluation of the Widal tube agglutination test for the diagnosis of typhoid fever among children admitted to a rural hospital in Tanzania and a comparison with previous studies

The Widal test is widely used in Africa but little information exists about its reliability. We therefore assessed the performance of the Widal tube agglutination test among febrile hospitalized Tanzanian children. We calculated sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of various anti-TH and –TO titers using culture-confirmed typhoid fever cases as the “true positives” and all other febrile children with blood culture negative for *S. Typhi* as the “true negatives”. We found a 1% prevalence of culture proven typhoid fever within all participants (n=1,680). A single anti-TH titer of 1:80 and higher was the optimal indicator of typhoid fever at a sensitivity, specificity, NPV and PPV of 75%, 98%, 100%, and 26% respectively.

We conclude that a Widal titer of $\geq 1:80$ performed best in terms of sensitivity, specificity, and NPV. However a test with improved PPV that is similarly easy to apply and cost-efficient is desirable

Assessment and comparative analysis of a rapid diagnostic test (Tubex®) for the diagnosis of typhoid fever among hospitalized children in rural Tanzania

A rapid diagnostic test for typhoid fever, Tubex®, is commercially available that uses particle separation to detect immunoglobulin M directed towards *Salmonella Typhi* O9 lipopolysaccharide in sera.

We assessed the sensitivity and specificity of the Tubex test among Tanzanian children hospitalized (n=139) with febrile illness using blood culture as gold standard. Evaluation was done considering blood culture confirmed *S. Typhi* with non-typhi salmonella (NTS) and non-salmonella isolates as controls as well as with non-salmonella isolates only.

Considering all non-*S. Typhi* isolates controls, we observed sensitivity and specificity of 79% and 89% respectively; considering only non-salmonella isolates as controls we did not observe a significant difference in performance ($p>0.05$).

We found that this first evaluation of the Tubex test in an African setting provided similar performance indicators as have been reported earlier from Asia. Comparison with the earlier results of a Widal test using the same samples showed no significant difference ($p>0.05$) for any of the performance indicators, irrespective of the applied control group.

Evaluation of a rapid dipstick (Crystal VC) for the diagnosis of cholera in Zanzibar and a comparison with previous studies

We evaluated a rapid diagnostic test for cholera (Crystal VC™) under field conditions in Zanzibar on patients presenting to treatment centers with watery diarrhea. The test was conducted on site, results were compared to stool culture results from the same samples performed in a nearby laboratory. We included 624 samples and found sensitivity of the RDT to be 93.1% (95%CI: 88.7 to 96.2%), specificity was 49.2% (95%CI: 44.3 to 54.1%), the positive predictive value was 47.0% (95%CI: 42.1 to 52.0%) and the negative predictive value was 93.6% (95%CI: 89.6 to 96.5%). The overall false positivity rate was 50.8% (213/419).

While the observed sensitivity was similar compared to earlier versions of the test, specificity was poorer. The current version of the RDT could potentially be used as a screening tool in the field. Because of the high proportion of false positive results when field workers test stool specimens, positive results will need to be confirmed with stool culture.

Discussion

Neither of the evaluated tests performed at a level sufficient to base treatment and therapy exclusively on the test result. While surveillance data can be based on test results for the Tubex test, a different batch of the same Widal test performed after the herein presented study showed poorer results (Thriemer et al, in preparation), which may be due to inter – batch variations. Neither test is suitable for introduction into routine laboratory work due to their underlying mechanism.

PPV of the Crystal VC was found to be very poor, the test is not suitable for direct diagnostic purposes. The test can be integrated into surveillance systems if the very poor specificity is considered. A possible application of the test is in laboratory work, where the test can be used to confirm / exclude growth observed on the differential and selective medium TCBS.

Conclusion

None of the tests can be used for direct diagnosis. Widal, Tubex and the Crystal VC can be applied for surveillance purposes if validated on site. The Crystal VC may be a useful addition to routine laboratory work.

Zusammenfassung

Einführung

Sowohl Typhus als auch Cholera sind Armutsassoziierte Erkrankungen und stellen für die öffentliche Gesundheit vieler Entwicklungsländer eine ernste Bedrohung dar. Für das Jahr 2000 wurden mehr als 21.000.000 Typhusfälle bei einer Sterblichkeitsrate von 1% berechnet; für Cholera gehen Schätzungen von einer jährlichen globalen Inzidenz von 2.550.000 Fällen bei einer Sterblichkeitsrate von >8.2% aus. Mangel an Trinkwasser und eine funktionierende Abwasserentsorgung, Bedingungen wie sie in Katastrophengebieten und vielen Entwicklungsländern herrschen, gelten als Hauptursache für die Verbreitung beider Krankheiten..

Wirksame Impfungen gegen beide Krankheiten existieren, effektive Therapien mit Antibiotika (Typhus) und Flüssigkeitssubstitution (Cholera) können die Sterblichkeitsraten in beiden Fällen signifikant senken. Allerdings sind weder Impfungen noch die benötigten Medikamente wegen chronischer Unterfinanzierung der Gesundheitssysteme in vielen Entwicklungsländern dauerhaft erhältlich. Eine mangelhafte Finanzierung ergibt sich teilweise aus geringem Steueraufkommen auf Grund verringerter krankheitsbedingter Arbeitskraft der Bevölkerung. Der Geldmangel hat außerdem eine Wissenslücke der Seuchenschutzbehörden über die eigentliche Höhe der Krankheitsfälle zur Folge und führt zu unzureichender Finanzierung diagnostischer Einrichtungen.

Eine laborbasierte Diagnose für Typhus und Cholera ist zur Zeit der Gold Standard. Die dafür nötigen Verfahren sind jedoch teuer und können nur von gut ausgebildeten Labortechnikern durchgeführt werden. Diagnostische Schnelltests stellen eine Option dar, die diagnostischen Möglichkeiten in Entwicklungsländern zu verbessern. Dazu müssen diese Tests verlässliche Ergebnisse liefern, einfach in der Handhabung und kostengünstig sein.

Das Ziel dieser kumulativen Dissertation war es daher, diagnostische Schnelltests in Entwicklungsländern und während mehrere Choleraausbrüche vor Ort auszuwerten. Dazu wurde der in Afrika weit verbreitete Widaltest (Span Diagnostics, Indien), der neuere Tubextest (IDL, Schweden) für die Diagnose von Typhus (*S. Typhi*) sowie ein Schnelltest für Cholera ausgewertet (Crystal VC, Span Diagnostics, Indien).

Hintergrund

Möglicherweise der älteste noch verwendete diagnostische Schnelltest ist der Widal Test. Dieser Test basiert auf einer makroskopisch sichtbaren Agglutinationsreaktion zwischen Antikörpern des Patienten in Serum und zugesetztem Antigen. Der wesentlich neuere Tubex Test basiert auf einer

Antikörper Inhibitionsreaktion, die zur Entfärbung einer Serum-Testlösung mittels Magnetismus führt. Der Crystal VC ist ein lateral flow immunochromatographischer Test zum Nachweis von *Vibrio cholerae* Antigen.

Alle Tests wurden mit den retrospektiven, Prävalenz unabhängigen Indikatoren Sensitivität und Spezifität, sowie den prospektiven, Prävalenz abhängigen positiven und negativen prognostischen Werten (PPV und NPV) bewertet. Da für den Widal- und Tubextest Blutkultur als Gold Standard gewählt wurde, wurden beide Tests mehrfach mit unterschiedlichen Kontrollgruppen ausgewertet um der geringen Sensitivität dieses Verfahrens Rechnung zu tragen.

Der Crystal VC wurde während mehrerer Cholera Ausbrüche in staatlichen Krankenhäusern getestet. Es wurde außerdem der Einfluss der Ausbildung des Testers auf das Testergebnis analysiert.

Ergebnisse

Beurteilung des Widal Tube Agglutination Tests zur Diagnose von Typhus (S. enterica Typhi) getestet an Proben von Kindern die in einem ländlichen Krankenhaus in Tansania aufgenommen wurden, sowie ein Vergleich mit früheren Studien

Der Widal Test ist in Afrika weit verbreitet, obwohl wenig über seine Genauigkeit bekannt ist. Wir haben daher die Verlässlichkeit des Widal Tube Agglutination Tests an Serumproben von tansanischen Kindern getestet, die stationär aufgenommen worden waren. Wir haben Sensitivität, Spezifität, positiven prognostischen (PPV) und negativen prognostischen Wert (NPV) von verschiedenen TH und TO Titern berechnet. Als Berechnungsgrundlage haben wir durch Blutkultur bestätigte Typhus Fälle als „echte Positive Fälle“ und alle anderen fieberhaften Kinder die nicht an Typhus litten als „echte negative Fälle“ definiert.

1% von 1680 Kindern hatte Typhus, bestätigt durch Blutkultur. Ein TH Titer von $\geq 1:80$ war der beste Indikator für Typhus und zeigte eine Sensitivität von 75%, eine Spezifität von 98%, einen NPV von 100% und einen PPV von 26%. Ein Test mit einem verbesserten PPV der ähnlich einfach in Handhabung und im Preis ist wäre jedoch wünschenswert.

Beurteilung und vergleichende Analyse eines diagnostischen Schnelltests zur Diagnose von Typhus (S. Typhi) bei stationär aufgenommenen Kindern im ländlichen Tansania

Ein diagnostischer Schnelltest für Typhus, Tubex®, der eine Partikel Separierung zum Nachweis von Immunoglobulin M gegen *Salmonella* Typhi O9 Lipopolysaccharid nutzt ist kommerziell erhältlich.

Wir haben die Sensitivität und Spezifität des Tubex Test an Serumproben von stationär aufgenommenen Kindern (n=139) getestet die an einem fiebrigen Infekt litten. Gold Standard war Blutkultur. Durch Blutkultur bestätigte *S. Typhi* Fälle und nicht – Typhi Salmonellosen (NTS), sowie nicht – *Salmonella* Isolate wurden zur Auswertung herangezogen.

Wir haben eine Sensitivität und Spezifität von 79% und 89% festgestellt. Bei einer zweiten Analyse, in der alle NTS Fälle aus der Kontrollgruppe ausgeschlossen wurden, ergab sich kein signifikanter Unterschied ($p>0,05$). Die erste Auswertung des Tests in Afrika ergab eine ähnliche Verlässlichkeit wie bereits früher beschrieben worden ist. Ein Vergleich mit den Ergebnissen eines Widal Tests der an den gleichen Proben früher durchgeführt wurde ergab keinen signifikanten Unterschied unabhängig von der Definition der Kontrollgruppen.

Beurteilung eines Schnelltests (Crystal VC) zur Diagnose von Cholera und ein Vergleich mit früheren Studien

Wir haben einen diagnostischen Schnelltest für Cholera (Crystal VC™) in Sansibar 2009 unter realen Bedingungen an Patienten mit wässrigem Durchfall getestet, die eine der örtlichen Gesundheitsstationen aufgesucht hatten. Der Schnelltest wurde vor Ort durchgeführt und mit den Ergebnissen einer Stuhlkultur verglichen die an den gleichen Proben durchgeführt wurde. Es wurden 624 Stuhlproben in die Analyse mit einbezogen. Im Vergleich zur Stuhlkultur betrug die Sensitivität des RDT 93.1% (95%CI: 88.7 bis 96.2%), Spezifität war 49.2% (95%CI: 44.3 bis 54.1%), der Positive Prognostische Wert war 47.0% (95%CI: 42.1 bis 52.0%) und der Negative Prognostische Wert betrug 93.6% (95%CI: 89.6 bis 96.5%). Insgesamt waren 50.8% (213/419) Testergebnisse falsch positiv.

Die ermittelte Sensitivität entspricht den Werten von Vorgängerversionen, die Spezifität war jedoch geringer. Eine mögliche Anwendung der aktuellen Version besteht in der Seuchenüberwachung vor Ort. Auf Grund der hohen Anzahl an falsch positiven Ergebnissen müssen positive Ergebnisse durch Stuhlkultur bestätigt werden.

Diskussion:

Keiner der bewerteten Tests war verlässlich genug um eine entsprechende Behandlung oder Therapie auf Grundlage des Testergebnisses zu beginnen. Während eine Seuchenüberwachung basierend auf dem Tubex Test möglich ist, gilt dies nicht für den Widal Test. Nach Studienende wurde eine andere Charge des gleichen Widal Tests für eine ähnliche Studie eingesetzt und ergab wesentlich schlechtere Ergebnisse (Thriemer et al, in Vorbereitung). Ein möglicher Grund könnten Qualitätsschwankungen bei der Produktion sein. Keiner der beiden Tests eignet sich zur Integration in die Labordiagnostik.

Der für den Crystal VC ermittelte PPV war zu niedrig um für die prognostische Diagnostik von Nutzen zu sein. Als Bestandteil einer Seuchenüberwachung und Erfassung von Cholera Fällen zu statistischen Zwecken stellt der Test eine mögliche Alternative dar, wenn die schlechte Spezifität berücksichtigt wird. Ein Anwendungsgebiet für den Test besteht in der Labordiagnostik. Bakterienkolonien die auf TCBS wachsen können mit dem Test als *V. cholerae* bestätigt werden.

Conclusio

Keiner der Tests stellt eine sinnvolle Ergänzung für die prognostische Diagnostik dar. Tubex und Crystal VC können in der Seuchenüberwachung eingesetzt werden. Der Crystal VC kann sinnvoll in der Labordiagnostik eingesetzt werden.

Curriculum Vitae

Name	Benedikt Nikolaus Ley
Address	Maria Lei 72 2018 Antwerp Belgium
Contact:	ley.benedikt@gmail.com Tel.: +32-494 429 435
Degree	Master of Science (MSc) Master of Public Health (MPH) PhD candidate (2012)
Summary	<p>Epidemiologist and Biologist with a focus on public health. Experienced in design, implementation, monitoring and reporting of health surveillance and intervention projects in developing countries. Competence in grant application writing, protocol design, preparation and implementation of standard operating procedures and collaboration with other non – governmental as well as governmental organizations at national and international level.</p> <p>Work experience in South East Asian and East African public health projects.</p> <p>Skills in working with an international team, problem solving in low income settings and statistical evaluation and analysis of data.</p>
References	<p>Dr. Lorenz von Seidlein (Former Supervisor, IVI): lseidlein@gmail.com</p> <p>Prof. H. Wilfing (PhD Supervisor, Univ. Vienna): harald.wilfing@univie.ac.at</p> <p>Dr. Robert Breiman (Collaborator, CDC): rbreiman@ke.cdc.gov</p>

Work Experience

March 2012 – Present, ZanHealth, Consultant

- Co-Author on Protocol and Budget
- Design of Study Specific Procedures
- Design of Laboratory Standard Operating Procedures and development of relevant paperwork
- Training of Staff

April 2010 – Feb. 2012, International Vaccine Institute, Associate Research Scientist

- Laboratory Field Coordinator for typhoid surveillance in several African countries
- Grant proposal writing (shortlisted with DFID / MRC)
- Creation of white papers
- Evaluation of possible collaborating sites (10 countries)
- Funded by Bill and Melinda Gates Foundation (7 mio. USD)

May 2008 – March 2010, International Vaccine Institute, Coordinator, Pemba – Tanzania

- Implementation of a cholera intervention program
- Census & re-census of study population
- Implementation of lab diagnostics and lab management
- Project coordinator for evaluation of cholera rapid diagnostic test
- Implementation of typhoid surveillance
- Implementation of lab diagnostics and lab management
- Implementation of a socio economic and behavioral typhoid study
- Funded by the Swedish International Development Agency
- Successful grant application to Hanako foundation and preparation of protocol for a stool culture based surveillance
- Monitoring of all participating sites

June 2007 – April 2008, Medical University of Vienna, Research Fellow, Bandarban – Bangladesh

- Community based active surveillance on malaria and other febrile diseases
- Clinical malaria trial
- Establishing and managing a microbiological laboratory

April – May 2007, Medical University of Vienna, Research Fellow, Vienna – Austria

- Research Management

May – July 2006, University of Vienna, MSc Student, Mae Sot – Thailand

- *In vitro* drug resistance of *P. falciparum*

University	<p>February 2010 – 2012 (?), PhD program with the University of Vienna, Austria</p> <p>June 2011 – Master of Public Health (distance learning), University of Liverpool, UK</p> <p>March 2007 – Master of Science, Human Biology, University of Vienna, Austria</p>
Reviews	<ul style="list-style-type: none"> • WHO Bulletin • PLoS One • Journal of Infectious Disease in Tropical Countries • Transactions of the Royal Society of Tropical Medicine and Hygiene
Publications	<p>Khatib AM, Ali M, von Seidlein L, Kim DR, Hashim R, Reyburn R, Ley BN, Thriemer K, Enwere G, Hutubessy R, Aguado MT, Kieny MP, Wierzba TF, Ali SM, Saleh AA, Clemens JD, Jiddawi MS, Deen JL. Direct and indirect effectiveness of the oral cholera vaccine in Zanzibar, East Africa: findings from a large mass vaccination campaign. <i>In preparation</i></p> <p>Kaljee LM, Pach A, Thriemer K, Ley B, Ali SM, Jiddawi M, Puri MK, von Seidlein L, Deen JL, Ochiai RL, Wierzba T, Clemens J. Healthcare utilization and accessibility on Pemba Island, Tanzania: implications for health outcomes for typhoid fever and disease burden surveillance. <i>In preparation</i></p> <p>Kaljee LM, Pach A, Thriemer K, Ley B, Jiddawi M, Puri MK, von Seidlein L, Deen JL, Ochiai RL, Wierzba T, Clemens J, Ali SM. Desirability for a typhoid fever vaccine among rural residents, Pemba Island Tanzania. <i>In preparation</i></p> <p>Riewpaiboon A, Piatti M, Ley B, Deen JL, Thriemer K, von Seidlein L, Jiddawi MS, Busch JL, Schmied WH, Ali SM and Typhoid Economic Study Group. Cost of Illness due to typhoid fever in Pemba, Zanzibar, East Africa. <i>In preparation</i></p> <p>Thriemer K, Ley B, Ame S, Deen J, , Pak GD, Chang NY, Hashim R, Schmied WH, Busch CJ, Nixon S, Morrissey A, Puri MK, Ali M, Ochiai RL, Wierzba T, Jiddawi MS, Clemens JD, Ali SM, von Seidlein L. Clinical and Epidemiological Features of Typhoid Fever in Pemba, Zanzibar. <i>In preparation</i></p> <p>Naha A, Chowdhury G, Ghosh-Banerjee J, Senoh M, Takahashi T, Ley BN, Thriemer K, Deen J, Seidlein L, Ali SM, Khatib A, Ramamurthy T, Nandy RK, Nair GB, Takeda Y, Mukhopadhyay AK. Molecular Characterization of High Cholera Toxin-Producing El Tor Variant <i>Vibrio cholerae</i> in Zanzibar, East Africa. <i>Submitted to Journal of Clin Microbiol</i></p> <p>B Ley, K Thriemer, L von Seidlein, J Deen, A Mukhopadhyay, NY Chang, R Hashim, W Schmidt, J Busch, R Reyburn, GD Pak, T Wierzba, JD Clemens, H Wilfing, G Enwere, R Hutubessy, T Aguado, MP Kieny, AM Khatib, AA Saleh, MS Jiddawi, D Sack, SM Ali: Evaluation of a rapid dipstick (Crystal VC) for the diagnosis of cholera in Zanzibar and a comparison with previous studies; <i>PLoS ONE</i> PLoS ONE (2012), 7(5): e36930. doi:10.1371/journal.pone.0036930</p> <p>Thriemer K, Ley B, Ame S, Puri MK, Hashim R, Chang NY, Salim LA, Ochiai RL, Wierzba T, Clemens JD, von Seidlein L, Deen JL, Ali SM, Ali M. Replacing paper data collection forms with electronic data entry in the field: findings from a study of community-acquired bloodstream infections in Pemba, Zanzibar. <i>BMC Res Notes</i>. 2012 Feb 21;5(1):113.</p> <p>Thriemer K, Ley B, Ame S, von Seidlein L, Pak GD, Chang NY, Hashim R, Schmied WH, Busch CJ, Nixon S, Morrissey A, Puri MK, Ali M, Ochiai RL, Wierzba T, Jiddawi MS, Clemens JD, Ali SM, Deen J. The burden of invasive bacterial infections in Pemba, Zanzibar, <i>PLoS ONE</i> 2012; 7(2): e30350.</p>

doi:10.1371/journal.pone.0030350

Ley B, Thriemer K, Ame S, Mtove G, Amos B, von Seidlein L, Hendriksen I, Mwambuli A, Shoo A, Malahiyo R, Kim DR, Ochiai RL, Favorov M, Clemens JD, Reyburn H, Wilfing H, Magesa S, Deen JL, Ali S: Evaluation of a Rapid Diagnostic Test (Tubex®) for the diagnosis of typhoid fever among children admitted to two Tanzanian rural hospitals, *BMC – Infect. Dis.* 2011 11(1):147

Ley B, Mtove G, Thriemer K, Amos B, von Seidlein L, Hendriksen I, Mwambuli A, Shoo A, Malahiyo R, Ame S, Kim DR, Ochiai RL, Clemens JD, Reyburn H, Magesa S, Deen JL: Evaluation of the Widal tube agglutination test for the diagnosis of typhoid fever among children admitted to a rural hospital in Tanzania, *BMC – Infect. Dis.* 2010; 10(1):180

Ley B, Wernsdorfer G, Frank C, Sirischaisinthop J, Congpuong K, Wernsdorfer WH: Pharmacodynamic interaction between 4-aminoquinolines and retinol in *Plasmodium falciparum* in vitro. *Wien Klin Wochenschr* 120 (Suppl 4) 2008: 74-79.

Starzengruber P, Thriemer K, Haque R, Khan WA, Fuehrer HP, Siedl A, Hofecker V, Ley B, Wernsdorfer WH, Noedl H.: Antimalarial activity of tigecycline, a novel glycylcycline antibiotic, *Antimicrob Agents Chemother.* 2009; 53(9):4040-2.

Ali M, Deen JL, Khatib A, Enwere G, von Seidlein L, Reyburn R, Ali SM, Chang NY, Perroud V, Marodon F, Saleh AA, Hashim R, Lopez AL, Beard J, Ley BN, Thriemer K, Puri MK, Sah B, Jiddawi MS, Clemens JD. Paperless registration during survey enumerations and large oral cholera mass vaccination in Zanzibar, the United Republic of Tanzania. *Bull World Health Organ* 2010;88:556–559
doi:10.2471/BLT.09.070334

Thriemer K, Starzengruber P, Khan WA, Haque R, Marma ASP, Ley B, Vossen M, Swoboda P, Akter A, Noedl H. (2010): Azithromycin combination therapy for the treatment of uncomplicated falciparum malaria in Bangladesh. An open label randomized, controlled clinical trial. *J Infect Dis.* 2010 Jun 17.

Fuehrer HP, Starzengruber P, Swoboda P, Khan WA, Matt J, Ley B, Thriemer K, Haque R, Yunus EB, Hossain SM, Walochnik J, Noedl H. (2010): Indigenous *Plasmodium ovale* malaria in Bangladesh. *Am J Trop Med Hyg.*; 83(1):75-78

Antwerp, 04.10.2012

Benedikt Ley

Appendix

All three publications in online format:

RESEARCH ARTICLE

Open Access

Evaluation of the Widal tube agglutination test for the diagnosis of typhoid fever among children admitted to a rural hospital in Tanzania and a comparison with previous studies

Benedikt Ley^{*1}, George Mtove^{2,3}, Kamala Thriemer¹, Ben Amos^{3,4}, Lorenz von Seidlein^{1,3}, Ilse Hendriksen^{3,5}, Abraham Mwambuli³, Aikande Shoo^{3,4}, Rajabu Malahiyo⁴, Shaali M Ame⁶, Deok R Kim¹, Leon R Ochiai¹, John D Clemens¹, Hugh Reyburn^{3,7}, Harald Wilfing⁸, Stephen Magesa^{2,3} and Jacqueline L Deen^{1,3}

Abstract

Background: The diagnosis of typhoid fever is confirmed by culture of *Salmonella enterica* serotype Typhi (*S. typhi*). However, a more rapid, simpler, and cheaper diagnostic method would be very useful especially in developing countries. The Widal test is widely used in Africa but little information exists about its reliability.

Methods: We assessed the performance of the Widal tube agglutination test among febrile hospitalized Tanzanian children. We calculated the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of various anti-TH and -TO titers using culture-confirmed typhoid fever cases as the "true positives" and all other febrile children with blood culture negative for *S. typhi* as the "true negatives."

Results: We found that 16 (1%) of 1,680 children had culture-proven typhoid fever. A single anti-TH titer of 1:80 and higher was the optimal indicator of typhoid fever. This had a sensitivity of 75%, specificity of 98%, NPV of 100%, but PPV was only 26%. We compared our main findings with those from previous studies.

Conclusion: Among febrile hospitalized Tanzanian children with a low prevalence of typhoid fever, a Widal titer of $\geq 1:80$ performed well in terms of sensitivity, specificity, and NPV. However a test with improved PPV that is similarly easy to apply and cost-efficient is desirable.

Background

Salmonella enterica serotype Typhi (*S. typhi*), the causative agent of typhoid fever, was calculated to have caused approximately 200,000 deaths globally in 2000 [1]. The clinical picture of typhoid fever is nonspecific; confirmed diagnosis through blood or bone-marrow culture requires expensive and labor-intensive isolation and identification of the organism, which may take up to seven days. A cheap and rapid alternative laboratory test is desirable, especially for developing country settings where typhoid fever is a major public health burden.

Various agglutination tests have been developed [2] of which the Widal method is the oldest and remains the

most widely used. The test was first introduced by F. Widal in 1896 [2] and is based on a macroscopically visible serum - mediated agglutination reaction between *S. typhi* somatic lipopolysaccharide O antigens (TO) and flagellar H antigens (TH). Laboratories in industrialized countries have stopped performing the assay. In Africa the Widal test is still widely used [3] because typhoid fever is perceived to be endemic in the area [3] and the Widal test is the only rapid diagnostic assay that is available and affordable. The Widal test is commonly performed when children and adults present with fever to treatment centers, as few centers have the capacity to perform micro-bacterial culture [4]. Despite this widespread use, little has been published on its performance in Africa.

* Correspondence: bnley@ivi.int

¹ International Vaccine Institute, Seoul, Korea

Full list of author information is available at the end of the article

We assessed the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the Widal tube agglutination test among Tanzanian children hospitalized with febrile illness and compared our results with those from previous studies.

Methods

Study site and population

The study was conducted as part of a childhood fever surveillance study at Teule Hospital in Muheza district of Northeastern Tanzania from 2008 to 2009. Muheza district is located between the foothills of Kilimanjaro and the coastal town of Tanga. The area is highly endemic for *Plasmodium falciparum* malaria with perennial transmission and two seasonal peaks [5]. HIV sero-prevalence among antenatal clinic attendees was about 7% in 2007 [6]. Teule Hospital is a busy 330-bed district-level general hospital, serving a surrounding population of 277,000. It has two 35-bed in-patient pediatric wards receiving approximately 5,000 child admissions per year (2008).

Inclusion Criteria

Children aged 2 months to 14 years were screened for eligibility during study hours from 7am to 7pm, Monday to Sunday. Children with fever of 3 or more days prior to admission, or fever of less than 3 days but with at least one severity criteria (respiratory distress, deep breathing, respiratory distress in combination with severe pallor, prostration, capillary refill ≥ 3 seconds, temperature gradient, systolic blood pressure < 70 mm Hg, coma defined by Glasgow Coma Scale ≤ 10 or Blantyre Coma Scale ≤ 2 , severe jaundice, history of 2 or more convulsions in the last 24 hours, blood glucose < 3 mmol associated with clinical signs, neck stiffness, bulging fontanel, or oxygen saturation $< 90\%$) were recruited into the study. All clinical information was recorded on a standard case record form. Treatment was provided according to national guidelines. On admission we collected 3 to 5 milliliters (ml) of blood (depending on body weight) from each eligible child for the Widal test and a single blood culture. All clinical procedures were performed by trained study clinical officers and nurses under the supervision of study physicians.

Laboratory

Blood for culture was inoculated into BacT/ALERT™ Pediatric-fan bottles (bioMérieux, Marcy l'Etoile, France). Inoculated blood culture bottles were transported immediately to the hospital laboratory and incubated in the BacT/ALERT 3 D automated microbial detection system. Blood cultures were processed according to standard methods. Colonies with biochemical reactions on API20E suggestive of *Salmonellae* were confirmed serologically

by slide and tube agglutination testing using specific O and H antisera (Becton Dickinson, NJ, USA).

A minimum of 0.5 ml of blood was separated to obtain serum samples. All serum samples were frozen at -70°C until Widal testing was done in three batches. Widal testing was performed using standardized TO (IgM and IgG) and TH (IgG) antigens (Stained *Salmonella* Antigens kit, Span Diagnostics, India) according to standard methods as described on the package insert. In brief, each sample was diluted to a concentration of 1:40 with 0.9% NaCl in two separate plastic tubes. A single drop of antigen was added to the respective tube. Incubation times for both O and H agglutinations were 16 to 20 hours at 37°C in a water bath. Evaluation of test results was performed by at least two lab technicians on an independent basis under standardized light conditions. If agglutination was detected in a sample, testing was done on that sample diluted serially from 1:80 to 1:1280 for both O and H antigens. All laboratory procedures were performed by trained laboratory technicians under the supervision of microbiologists. Technicians performing the Widal tests were blinded to the participants' clinical picture and blood culture results.

Data management

Case report forms were double-entered into custom-made data entry programs using MS-Access (Microsoft Corp.). Data management programs included error, range, and consistency check programs. Analyses were performed using EpiInfo v 3.4.3 (Centers for Disease Control and Prevention, USA) and Stata TM v 10.0 (Stata Corp., USA).

Definitions and analysis

Fever was defined as stated history or presence of fever of $\geq 37.5^{\circ}\text{C}$. Bacteremia was defined as fever with isolation of pathogenic bacteria from blood culture. Children with a febrile illness were classified as follows: those with *S. typhi* subsequently isolated from blood culture (group 1), those with non-Typhi serotypes of *S. enterica* (NTS) subsequently isolated from blood culture (group 2), those with pathogenic bacteria other than *Salmonellae* subsequently isolated from blood culture (group 3), and those whose blood culture yielded no bacterial pathogen (group 4). Malaria status was not considered in the classification. In areas of high transmission of *P. falciparum* where individuals develop immunity from previous episodes of malaria starting at a young age, asymptomatic parasitemia is common and may be detected during febrile episodes caused by another infection [7,8].

For the primary analysis, sensitivity (true-positive rate) was defined as the probability that the Widal test result would be positive when blood culture confirmed that typhoid fever was present (group 1) and specificity (true-

negative rate) was the probability that the Widal test result would be negative when *S. typhi* was not isolated from blood culture (groups 2, 3, and 4). The positive predictive value was the probability that culture-confirmed typhoid was present when the test was positive, and the negative predictive value was the probability that culture-confirmed typhoid was not present when the test was negative. Since serological tests detect antibody response and perform better after a period of time from the onset of the illness, sensitivity, specificity, PPV, and NPV were also calculated separately for cases presenting with fever for 5 days or less and for more than 5 days. Because controversy exists about what is the most appropriate control group to use [9-11], we conducted a secondary analysis using two alternative "true-negative" control groups as follows: those with NTS and other pathogenic bacteria isolated from blood culture (groups 2 and 3), and those with pathogenic bacteria other than *Salmonellae* isolated from blood culture (group 3).

Comparisons were made using the Chi square or Fishers' Exact test, as appropriate. Sensitivity, specificity, PPV, and NPV were calculated according to standard methods. The 95% confidence interval for sensitivity and specificity was calculated using the Wilson's Score method [12]. Analyses were performed using EpiInfo v 3.4.3 (Centers for Disease Control and Prevention, USA) and Stata TM v 10.0 (Stata Corp., USA).

Literature review

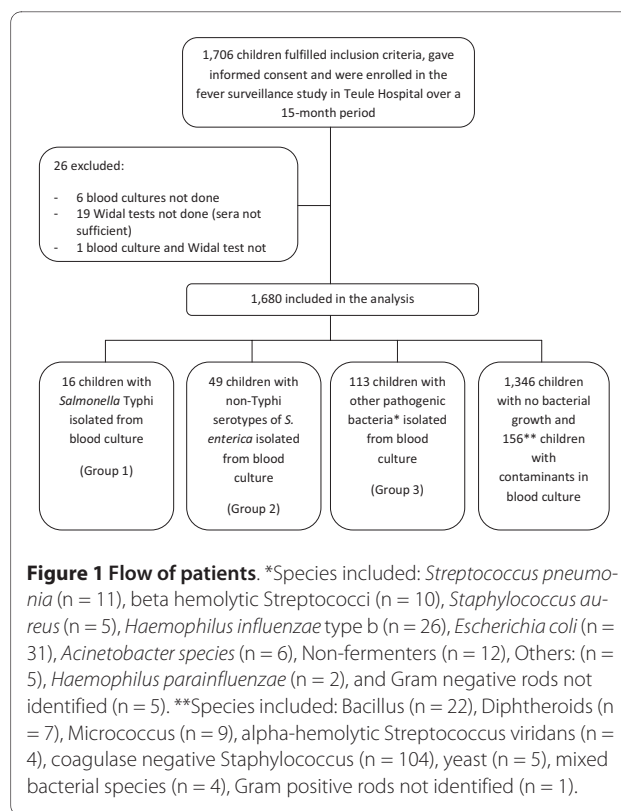
We conducted a literature review to compare our main findings with those from previous studies of similar character. We included studies of the Widal test which were identified by direct searches of the MEDLINE database through PubMed. The searches were restricted to publications from 1993 to date. We also conducted supplementary searches of the references in retrieved articles. Abstracts were reviewed and if relevant, the article was included.

Ethics

The fever surveillance was conducted following the principles governing biomedical research involving human subjects. Prior written informed consent was obtained from the parent or guardian of all study participants. The study protocol was approved by the National Institute of Medical Research, Tanzania, and the International Vaccine Institute Institutional Review Board.

Results

The flow of patients is shown in Figure 1. A total of 1,706 febrile children were enrolled out of which 26 (1.5%) were excluded for the following reasons: 19 or 1.1% had no blood culture done, 6 or 0.4% had no Widal testing done due to insufficient quantities of sera, and one case or 0.1%



had no blood culture nor Widal testing done. A total of 1,680 (98.5%) samples were included in the analysis. There were 16 or 1.0% culture-confirmed typhoid fever cases (group 1), 49 or 2.9% with NTS infection (group 2), and 113 or 6.7% with non-*Salmonella* bacteremia (group 3). From 1,502 (89.4%) children, no pathogenic bacteria were isolated from blood culture (group 4).

We assessed the age distribution and highest anti-TH and -TO titer by blood culture-confirmed diagnosis (Table 1). Children with typhoid fever were significantly older compared to the other groups. Anti-TH agglutination titers of 1:80 and higher were detected among 12/16 (75.0%) culture-confirmed typhoid fever cases compared to 7/49 (14.3%) of those with NTS infection, and 1/113 (0.9%) with other bacteremia (p values = 0.001 and < 0.001, respectively). Similarly, anti-TO agglutination titers of 1:80 and higher were detected among 11/16 (68.8%) cases of typhoid fever compared to 7/49 (14.3%) of those with NTS infection and 0/113 (0%) with other bacteremia (p values < 0.001 and < 0.001, respectively).

Primary analysis

We calculated the sensitivity, specificity, PPV, and NPV of various Widal test cut-offs for the diagnosis of typhoid fever (Table 2). The sensitivity, specificity, PPV, and NPV of an anti-TH titer of 1:80 were 75, 98, 26, and 100%, respectively. The sensitivity, specificity, PPV, and NPV of

Table 1: Number and cumulative frequencies of anti-TH and anti-TO levels overall and by blood culture isolates

Highest titer reached; Number (%)	All (n = 1,680)	Children with culture-confirmed typhoid fever (n = 16)	Children with non-Typhi serotypes of <i>S. enterica</i> (n = 49)	Children with other pathogenic bacteria (n = 113)	Children with no pathogenic bacteria isolated (n = 1,502)
Median age (Range)*	1.83 (14.81)	7.21 (11.96)	1.58 (6.73)	1.43 (11.83)	1.84 (14.81)
Anti-TH					
≥ 1:640	15 (0.9)	3 (18.8)	5 (10.2)	0 (0)	7 (0.5)
1:320	24 (1.4)	6 (37.5)	6 (12.2)	0 (0)	12 (0.8)
1:160	36 (2.1)	11 (68.8)	7 (14.3)	0 (0)	18 (1.2)
1:80	46 (2.7)	12 (75.0)	7 (14.3)	1 (0.9)	26 (1.7)
1:40	85 (5.1)	12 (75.0)	9 (18.4)	3 (2.7)	61 (4.1)
No agglutination	1,595 (94.9)	4 (25.0)	40 (81.6)	110 (97.3)	1,441 (95.9)
Anti-TO					
≥ 1:640	6 (0.4)	3 (18.8)	2 (4.1)	0 (0)	1 (0.1)
1:320	18 (1.1)	6 (37.5)	3 (6.1)	0 (0)	9 (0.6)
1:160	34 (2.0)	10 (62.5)	6 (12.2)	0 (0)	18 (1.2)
1:80	44 (2.6)	11 (68.8)	7 (14.3)	0 (0)	26 (1.7)
1:40	95 (5.7)	12 (75.0)	10 (20.4)	3 (2.7)	70 (4.7)
No agglutination	1,585 (94.3)	4 (25.0)	39 (79.6)	110 (97.3)	1,432 (95.3)

*P values: Group 1 vs 4 < 0.001; Group 2 vs 4 = 0.209; Group 3 vs 4 = 0.013

Table 2: Primary analysis for the performance* of the Widal test for typhoid fever diagnosis (n = 1680)

Widal titer	Sensitivity (95%CI)	Specificity (95%CI)	Positive Predictive Value	Negative Predictive Value
TH \geq 1:80	12/16, 0.75 (0.51-0.90)	1630/1664, 0.98 (0.97-0.99)	12/46, 0.26	1630/1634, 1.00
TH \geq 1:160	11/16, 0.69 (0.44-0.86)	1639/1664, 0.98 (0.98-0.99)	11/36, 0.31	1639/1644, 1.00
TH \geq 1:320	6/16, 0.38 (0.18-0.61)	1646/1664, 0.99 (0.98-0.99)	6/24, 0.25	1646/1656, 0.99
TO \geq 1:80	11/16, 0.69 (0.44-0.86)	1631/1664, 0.98 (0.97-0.99)	11/44, 0.25	1631/1636, 1.00
TO \geq 1:160	10/16, 0.63 (0.39-0.82)	1640/1664, 0.99 (0.98-0.99)	10/34, 0.29	1640/1646, 1.00
TO \geq 1:320	6/16, 0.38 (0.18-0.61)	1652/1664, 0.99 (0.99-1.00)	6/18, 0.33	1652/1662, 0.99

*The values were calculated using culture-confirmed typhoid fever cases (group 1; n = 16) as the true positives and those cases from which *S. typhi* were not isolated from blood culture (groups 2, 3, and 4; n = 1664) as the true negatives.

an anti-TO titer of 1:80 were 69, 98, 25, and 100%, respectively.

We compared the performance of the Widal test between patients who presented with fever of 5 days or less and those who presented with more than 5 days of fever (Table 3). Of the 16 typhoid fever cases, 6 (37.5%) presented with fever of 5 days or less, and 10 (62.5%) with more than 5 days of fever. Of the 1,664 children in the control group, 1,117 (67%) presented with fever of 5 days or less and 544 (33%) with more than 5 days of fever. Three control cases, whose fever duration was unknown, were excluded from the analysis. The sensitivity of an anti-TH and -TO titer of 1:80 increased, however not significantly, from 67% to 80% and 67% to 70%, respectively, with the longer duration of fever prior to admission (both $p > 0.05$). The PPV of an anti-TH and -TO titer of 1:80 increased from 21% to 30% and 19% to 30% (both $p > 0.05$), respectively, with the longer duration of fever prior to admission. But the change was also not statistically significant.

Secondary analysis

Using different control groups, we compared the resulting sensitivity, specificity, PPV, and NPV of a Widal test cut-off of an anti-TH and -TO titer of \geq 1:80 for the diagnosis of typhoid fever (Table 4). Changing the control group had no significant effect on the sensitivity, specificity, and NPV but markedly increased the PPV of an anti-TH titer of \geq 1:80 from 26% to 92% and the PPV of an anti-TO titer of \geq 1:80 from 25% to 100%.

Comparison with earlier studies

We found 4 articles from 3 countries. In this series, the age group included and prevalence of blood-culture confirmed typhoid fever varied considerably. The cut-off titer

used ranged from \geq 1:20 to \geq 1:200 and the resulting sensitivity, specificity, PPV and NPV varied considerably (Table 5).

Discussion

We found that a Widal titer of \geq 1:80 was the optimal indicator of typhoid fever in our study population. The PPV, NPV and specificity in the primary analysis was more-or-less unchanged from the cut-off titers of \geq 1:80 to \geq 1:320, whereas the sensitivity was highest at a cut-off titer of \geq 1:80. Although the Widal test at this cut-off titer performed relatively well in terms of sensitivity, specificity and NPV, its PPV was low. It has been argued that PPV is the most important measure of a clinical diagnostic method since it represents the proportion of patients with positive test results that are correctly diagnosed [13]. The PPV is not intrinsic to the test; it is affected by prevalence of the disease. In our setting, where 16 (1%) out of 1,680 febrile patients admitted to the pediatric ward had culture-proven typhoid fever, a negative Widal test result would have a good predictive value for the absence of disease but a positive result would have a low predictive value for typhoid fever, making the use of the Widal test in our setting questionable.

In a previous paper describing the clinical aspects of the children included in this study [14] older age and long duration of fever were predictive of typhoid fever in this group.

There are several difficulties associated with evaluation of the Widal test. Firstly, levels of agglutinins detectable in the non-infected populations of different areas vary considerably by time and place depending on the endemicity of the disease, which affects test performance. For example, the sensitivity and specificity of a Widal test

Table 3: Comparison of the performance* of the Widal test for typhoid fever diagnosis by number of days of fever prior to admission**

Widal titer	Sensitivity (95%CI)	Specificity (95%CI)	Positive Predictive Value	Negative Predictive Value
≤ 5 days fever (n = 1123)				
TH ≥1:80	4/6, 0.67 (0.30-0.90)	1102/1117, 0.99 (0.98-0.99)	4/19, 0.21	1102/1104, 1.00
TH ≥1:160	3/6, 0.50 (0.19-0.81)	1106/1117, 0.99 (0.98-0.99)	3/14, 0.21	1106/1109, 1.00
TH ≥1:320	2/6, 0.33 (0.10-0.70)	1110/1117, 0.99 (0.99-1.00)	2/9, 0.22	1110/1114, 1.00
TO ≥1:80	4/6, 0.67 (0.30-0.90)	1100/1117, 0.98 (0.98-0.99)	4/21, 0.19	1100/1102, 1.00
TO ≥1:160	3/6, 0.50 (0.19-0.81)	1107/1117, 0.99 (0.98-1.00)	3/13, 0.23	1107/1110, 1.00
TO ≥1:320	1/6, 0.17 (0.03-0.56)	1112/1117, 1.00 (0.99-1.00)	1/6, 0.17	1112/1117, 1.00
> 5 days fever (n = 554)				
TH ≥1:80	8/10, 0.80 (0.49-0.94)	525/544, 0.97 (0.95-0.98)	8/27, 0.30	525/527, 1.00
TH ≥1:160	8/10, 0.80 (0.49-0.94)	530/544, 0.97 (0.96-0.98)	8/22, 0.36	530/532, 1.00
TH ≥1:320	4/10, 0.40 (0.17-0.69)	533/544, 0.98 (0.96-0.99)	4/15, 0.27	533/539, 0.99
TO ≥1:80	7/10, 0.70 (0.40-0.89)	528/544, 0.97 (0.95-0.98)	7/23, 0.30	528/531, 0.99
TO ≥1:160	7/10, 0.70 (0.40-0.89)	530/544, 0.97 (0.96-0.98)	7/21, 0.33	530/533, 0.99
TO ≥1:320	5/10, 0.50 (0.24-0.76)	537/544, 0.99 (0.97-0.99)	5/12, 0.42	537/542, 0.99

*The values were calculated using culture-confirmed typhoid fever cases (group 1; n = 16) as the true positives and those cases from which *S. typhi* were not isolated from blood culture (groups 2, 3, and 4; n = 1664) as the true negatives. **Three patients whose fever duration was unknown were excluded from the analysis.

anti-TO titer of 1:80 in Kolkata, India was 58% and 85% [10] compared to our findings of 69% and 98%. Secondly, test performance is also affected by cross-reacting infections. In our study, none of the 113 children with non-*Salmonella* bacteremia exhibited titers above 1:80 for both O and H, although cross-reactions with *Klebsiella spp.* and *Staphylococcus aureus* [15] have been reported. In contrast, 7 (14.3%) of the 49 children with NTS had titers above 1:80 for both O and H. There is also the possibility of cross-reactivity with non-bacterial infections such as malaria, dengue, hepatitis A, and infectious mononucleosis [2,9,16]. The third limitation is the choice of a satisfactory gold standard for diagnosis. We used blood culture-positive patients as our true positives. Although bone marrow culture would be the ideal gold standard, this test is difficult to perform in small rural

hospitals in Africa. We found that 26 (1.7%) of 1,502 children from whom pathogenic bacteria were not isolated showed agglutination at 1:80 or higher, both for O and H antigens. These may be Widal false positive results due to cross-reaction. Alternatively, since the reported sensitivity of a single blood culture is only 40% to 60% [16-19], some of these are likely to be false negative blood culture results. The final, and what we found to be the most contentious issue, is the selection of the most appropriate control group. It is difficult to choose patients with febrile illness who are blood culture-negative and who definitely do not have typhoid fever. Furthermore, there were relatively few hospitalized children with no bacteremia in the same age range as those with typhoid fever. Thus, the control children were significantly younger than the cases. For our primary analysis, we used groups 2, 3 and 4

Table 4: Secondary analysis of the performance of a Widal anti-TH and -TO titer of $\geq 1:80$ for typhoid fever diagnosis using group 1 as true positives and three different control groups as true negatives*

Control group	Sensitivity (95%CI)	Specificity (95%CI)	Positive Predictive Value	Negative Predictive Value
TH titer $\geq 1:80$:				
Groups 2, 3, and 4 (n = 1664)	12/16, 0.75 (0.51-0.90)	1630/1664, 0.98 (0.97-0.99)	12/46, 0.26	1630/1634, 1.00
Groups 2 and 3 (n = 162)	12/16, 0.75 (0.51-0.90)	154/162, 0.95 (0.91-0.97)	12/20, 0.60	154/158, 0.97
Group 3 (n = 113)	12/16, 0.75 (0.51-0.90)	112/113, 0.99 (0.95-1.00)	12/13, 0.92	112/116, 0.97
TO titer $\geq 1:80$:				
Groups 2, 3, and 4 (n = 1664)	11/16, 0.69 (0.44-0.86)	1631/1664, 0.98 (0.97-0.99)	11/44, 0.25	1631/1636, 1.00
Groups 2 and 3 (n = 162)	11/16, 0.69 (0.44-0.86)	155/162, 0.96 (0.91-0.98)	11/18, 0.61	155/160, 0.97
Group 3 (n = 113)	11/16, 0.69 (0.44-0.86)	113/113, 1.00 (0.97-1.00)	11/11, 1.00	113/118, 0.96

*Group 1 were those with *S. typhi* isolated from blood culture (n = 16), group 2 were those with non-typhi serotypes of *S. enterica* (NTS) isolated from blood culture (n = 49), group 3 were those with pathogenic bacteria other than *Salmonellae* isolated from blood culture (n = 113), and group 4 were those whose blood culture yielded no bacterial pathogen (n = 1502).

(i.e., all children admitted for a febrile illness who were subsequently culture-negative for *S. typhi*). These would be the most conservative controls for specificity since blood culture picks up only a fraction of typhoid cases, resulting in a control group that is likely contaminated with culture-negative typhoid cases. Despite this, the specificity of the Widal test was high. Using the more exclusive control groups as others had done previously [9-11] did not appreciably alter the sensitivity, specificity, and NPV but they increased the PPV.

The previous studies included in our review (Table 5) had not been performed in Africa hence different cut - off titers were applied, and the resulting sensitivity, specificity, PPV and NPV varied considerably. PPV as well as NPV are dependent on the prevalence of disease within the group of participants; the selection process of study participants has therefore direct influence on the results. The difficulty of choosing the correct control group has been noted earlier [9]. While the gold standard, blood culture, is applied in most studies, the true negatives may be defined as febrile patients with a non-typhi laboratory-confirmed diagnosis as done by Parry et al. and Olsen et al. [9,20]. Alternatively, some studies use healthy controls. Choo et al [21]. considered all febrile cases with an *S. typhi* negative blood culture as the control group which is problematic as a number of blood culture-negative results are likely to be false-negative due to the poor sensitivity of the blood culture [17-19,22]. Furthermore, it is

difficult to compare the different test kits, as varying antigens perform differently [23].

Conclusion

In summary, a Widal titer of $\geq 1:80$ performed relatively well in terms of sensitivity and specificity. However, the low prevalence of typhoid fever of approximately 1% amongst children at Teule Hospital meant that the Widal test was only useful for excluding the disease.

Considering the low cost of Widal testing and the absence of comparably cheap tests, Widal testing is likely to remain the test of choice in many developing country settings. But the need for rapid and cheap diagnostic tools with superior performance remains high.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

BL performed the Widal tests, analyzed and compared results and wrote the manuscript; GM was in charge of the implementation and management of the study; KT performed Widal tests, analyzed results and contributed to the manuscript; BA supervised the laboratory where blood cultures were performed and contributed to the manuscript; LvS provided scientific support to study staff and manuscript and was involved in clinical care of participants; IH was involved in clinical care of participants; AM was in charge of data management; AS performed blood culture procedures, RM facilitated activities to make data collection possible, SA provided laboratory support; DRK performed the statistical analysis; RLO provided scientific support to the manuscript; JDC provided scientific support to the manuscript; HR provided major contributions to the manuscript; HW provided scientific support to the manuscript; SM facilitated activities to make data collection possible; JLD provided major scientific support to the manuscript and was involved in clinical care of participants.

Table 5: Summary of Widal performances in earlier studies

Authors	Date	Study Country	Sample Size	Ageclasses included	Prevalence of <i>S. typhi</i> in participants	Sensitivity	Specificity	PPV	NPV	Cut Off Titer	Control Group(s)	Gold Standard
Choo et al.	1993	Malaysia	2382	Children	6.1%	89%	89%	< 50%	99.2%	O or H $\geq 1:40$	Non-typhoidfebrile children admitted to hospital	Blood Culture
Parry et al.	1999	Vietnam	2000	Children & Adults	30.8%	O: 49% H: 67% O or H $\geq 1:100:88\%$	O: 97% H: 96% O or H $\geq 1:100:87\%$	O: 88% H: 88% O or H $\geq 1:100:74\%$	O: 82% H: 87% O or H $\geq 1:100:94\%$	O $\geq 1:200$ H: $\geq 1:100$ O or H $\geq 1:100$	Lab confirmed malaria, dengue or bacteremia	Blood Culture
Wilke et al.	2002	Turkey	410	≥ 18 y	13.2%	52% Post 7-10 d: 90%	88% Post 7-10 d: 90%	76% Post 7-10 d: 88%	71% Post 7-10 d: 93%	O: $\geq 1:200$ H: $\geq 1:200$	Healthy controls, nontyphoidal febrile patients, blood culture negative febrile cases	Blood Culture, Stool Culture
Olsen et al.	2004	Vietnam	80	≥ 3 y	73.8%	64% (field) 61% (lab)	76% (field) 100% (lab)	88% (field) 100% (lab)	43% (field) 48% (lab)	O or H $\geq 1:100$	Lab confirmed bacteremia, AFB, dengue, malaria, pos. stool culture, pos. urine culture	Blood Culture
Ley et al.	This study	Tanzania	1680	2 m. - 14y	1%	75%	98%	26%	100%	H: $\geq 1:80$	Non-typhoidfebrile children admitted to hospital	Blood Culture

All authors have read and approved the final manuscript

Acknowledgements

This work was supported by a grant from the Korean International Cooperation Agency through the International Vaccine Institute.

The study is published with the permission of the Director General of the Tanzanian National Institute for Medical Research, Dar-Es-Salaam. We are grateful to the patients and their parents who made this work possible. We thank all technical staff and research assistants who were involved in the study.

Author Details

¹International Vaccine Institute, Seoul, Korea, ²National Institute for Medical Research - Amani Centre, Tanga, Tanzania, ³Joint Malaria Program, Tanga, Tanzania, ⁴Teule Hospital, Muheza, Tanga, Tanzania, ⁵Mahidol Oxford Research Unit, Bangkok, Thailand, ⁶Public Health Laboratory (Pemba) - Ivo de Carneri, Chake Chake, Tanzania, ⁷London School of Hygiene and Tropical Medicine, Keppel St, London, UK and ⁸University of Vienna, Biocenter, Vienna, Austria

Received: 23 March 2010 Accepted: 22 June 2010

Published: 22 June 2010

References

- Crump JA, Luby SP, Mintz ED: **The global burden of typhoid fever.** *Bulletin of the World Health Organization* 2004, **82**(5):346-353.
- Olopoenia LA, King AL: **Widal agglutination test - 100 years later: still plagued by controversy.** *Postgrad Med J* 2000, **76**:80-84.
- Mweu E, English M: **Typhoid Fever in Children in Africa.** *Trop Med Int Health* 2008, **13**(4):532-540.
- Chart H, Cheesbrough J, Waghorn D: **The serodiagnosis of infection with Salmonella typhi.** *Journal of Clinical Pathology* 2000, **53**:851-853.
- Maxwell CA, Chambo W, Mwaimu M, Magogo F, Carneiro IA, Curtis CF: **Variation of malaria transmission and morbidity with altitude in Tanzania and with introduction of alphacypermethrin treated nets.** *Malar J* 2003, **10**:28.
- Edmonds SM, Meadway J: **Getting Pregnant Women onto HAART (Highly active antiretroviral therapy): developing a strategy for advanced prevention of mother-to-child transmission of HIV (PMTCT+) in rural Tanzania.** *HIV Medicine* 2009, **10**(Suppl 1):26.
- Smith T, Schellenberg JA, Hayes R: **Attributable fraction estimates and case definitions for malaria in endemic areas.** *Stat Med* 1994, **13**(22):2345-2358.
- Mwangi TW, Ross A, Snow RW, Marsh K: **Case definitions of clinical malaria under different transmission conditions in Kilifi District, Kenya.** *J Infect Dis* 2005, **191**:1932-1939.
- Parry CM, Diep TS, Wain J, Chinh NT, Vinh H, Hien TT, White NJ, Farrar JJ: **Value of a Single-tube Widal Test in Diagnosis of Typhoid Fever in Vietnam.** *J Clin Microbiol* 1999, **37**(9):2882-2886.
- Dutta S, Sur D, Manna B, Sen B, Deb AK, Deen JL, Wain J, von Seidlein L, Ochial RL, Clemens JD, Bhattacharya JK: **Evaluation of a new-generation serologic test for the diagnosis of typhoid fever: data from a community - based surveillance in Calcutta, India.** *Diagnostic Microbiology and Infectious Disease* 2006, **56**:359-365.
- Dong B, Gallindo CM, Shin E, Acosta CJ, Page AL, Wang M, Kim D, Ochial RL, Park J, Ali M, von Seidlein L, Xu Z: **Optimizing typhoid fever case definitions by combining serological tests in a large population study in Hechi City, China.** *Epidemiol Infect* 2007, **135**(6):1014-1020.
- Newcombe RG: **Two-Sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods.** *Stat Med* 1998, **17**(8):857-872.
- Altman DG, Bland JM: **Diagnostic tests 2: Predictive values.** *BMJ* 1994, **309**:102.
- Mtove G, Amos B, von Seidlein L, Hendriksen I, Mwambuli A, Kimera J, Mallahiyo R, Kim DR, Ochial RL, Clemens JD, Reyburn H, Magesa S, Deen JL: **Invasive salmonellosis among children admitted to a rural Tanzanian hospital and a comparison with previous studies.** *PLoS One* 2010, **5**(2):e9244.
- Pokhrel BM, Karmacharya R, Mishra SK, Koirala J: **Distribution of antibody titer against Salmonella enteric among healthy individuals in Nepal.** *Ann Clin Microbiol Antimicrob* 2009, **7**:8-1.
- Bhutta ZA: **Current concepts in the diagnosis and treatment of typhoid fever.** *BMJ* 2006, **333**:78-82.

- Farooqui BJ, Khurshid M, Ashfaq MK, Khan MA: **Comparative yield of Salmonella typhi from blood and bone marrow cultures in patients with fever of unknown origin.** *J Clin Pathol* 1991, **44**:258-259.
- Gilman RH, Termini M, Levine , Hernandez-Mendoza P, Hornick RB: **Relative efficacy of blood, urine, rectal swab, bone marrow, and rose spot cultures for recovery of Salmonella typhi in typhoid fever.** *Lancet* 1975, **31**(17918):1211-1213.
- Wain J, Pham VB, Ha V, Nguyen NM, To SD, Walsh AL, Parry CM, Hasserjian RP, HoHo VA, Tran TH, Farrar J, White NJ, Day NP: **Quantitation of bacteria in bone marrow from patients with typhoid fever: relationship between counts and clinical features.** *J Clin Microbiol* 2001, **39**:1571-1576.
- Olsen SJ, Pruckler J, Bibbb W, Thanh NT, Trinh TM, Minh NT, Sivapalsingam S, Gupta A, Phuong PT, Chinh NT, Chau NV, Cam PD, Mintz ED: **Evaluation of Rapid Diagnostic Tests for Typhoid Fever.** *Journal of Clinical Microbiology* 2004. DOI: 10.1128/JCM.42.5.1885-1889.2004
- Hoffman SL, Edman DC, Punjabi NH, Lesmana M, Cholid A, Sundah S, Harahap J: **Bone marrow aspirate culture superior to streptokinase clot culture and 8 ml 1:10 blood to broth ratio blood culture for diagnosis of typhoid fever.** *Am J Trop Med Hyg* 1986, **35**:836-839.
- Choo KE, Razif AR, Oppenheimer SJ, Ariffin WA, Lau J, Abraham T: **Usefulness of the Widal test in diagnosing childhood typhoid fever in endemic areas.** *J Paediatr Child Health* 1993, **29**(1):36-9.
- Handojo I, Edijanto SP, Probahoesodo MY, Mahartini NN: **Comparisson of the Diagnostic Value of Local Widal slide Test with Imported Widal Slide Test.** *Southeast Asian J Trop Med Public Health* 2004, **35**(2):366-70.

Pre-publication history

The pre-publication history for this paper can be accessed here:
<http://www.biomedcentral.com/1471-2334/10/180/prepub>

doi: 10.1186/1471-2334-10-180

Cite this article as: Ley et al., Evaluation of the Widal tube agglutination test for the diagnosis of typhoid fever among children admitted to a rural hospital in Tanzania and a comparison with previous studies *BMC Infectious Diseases* 2010, **10**:180

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



RESEARCH ARTICLE

Open Access

Assessment and comparative analysis of a rapid diagnostic test (Tubex[®]) for the diagnosis of typhoid fever among hospitalized children in rural Tanzania

Benedikt Ley^{1,8*}, Kamala Thriemer¹, Shaali M Ame², George M Mtove^{3,4}, Lorenz von Seidlein^{1,4,5}, Ben Amos^{4,6}, Ilse CE Hendriksen^{4,7}, Abraham Mwambuli⁴, Aikande Shoo^{4,6}, Deok R Kim¹, Leon R Ochiai¹, Michael Favorov¹, John D Clemens¹, Harald Wilfing⁸, Jacqueline L Deen^{1,4} and Said M Ali²

Abstract

Background: Typhoid fever remains a significant health problem in many developing countries. A rapid test with a performance comparable to that of blood culture would be highly useful. A rapid diagnostic test for typhoid fever, Tubex[®], is commercially available that uses particle separation to detect immunoglobulin M directed towards *Salmonella* Typhi O9 lipopolysaccharide in sera.

Methods: We assessed the sensitivity and specificity of the Tubex test among Tanzanian children hospitalized with febrile illness using blood culture as gold standard. Evaluation was done considering blood culture confirmed *S. Typhi* with non-typhi salmonella (NTS) and non - salmonella isolates as controls as well as with non-salmonella isolates only.

Results: Of 139 samples tested with Tubex, 33 were positive for *S. Typhi* in blood culture, 49 were culture-confirmed NTS infections, and 57 were other non-salmonella infections. Thirteen hemolyzed samples were excluded. Using all non - *S. Typhi* isolates as controls, we showed a sensitivity of 79% and a specificity of 89%. When the analysis was repeated excluding NTS from the pool of controls we showed a sensitivity of 79% and a specificity of 97%. There was no significant difference in the test performance using the two different control groups ($p > 0.05$).

Conclusion: This first evaluation of the Tubex test in an African setting showed a similar performance to those seen in some Asian settings. Comparison with the earlier results of a Widal test using the same samples showed no significant difference ($p > 0.05$) for any of the performance indicators, irrespective of the applied control group.

Keywords: *Salmonella*, Tubex[®], Widal, Africa, Rapid Diagnostic Test

Background

Typhoid fever remains a significant health problem in many developing countries. Estimates suggest an incidence rate of more than 21.5 million cases globally in the year 2000 [1]. Recent data from Tanzania mainland have found a strong variation of prevalence rates among blood culture positive isolates collected in local hospitals, ranging from

9% [2] to 21.4% [3] for *Salmonella* enterica serovar Typhi (*S. Typhi*), no data from Zanzibar are available to date. As the clinical picture of typhoid fever is often unspecific, misdiagnosis and insufficient or inadequate treatment are potential risks associated with the disease. In the absence of difficult-to-obtain bone marrow specimens, microbiologic culture of a blood sample is considered to be the current state-of-the art test for the diagnosis of typhoid fever even though its sensitivity may be as low as 40% [4,5]. Culture may take up to seven days and requires a well-run and equipped laboratory, which is often not available in

* Correspondence: bnley@ivi.int

¹Translational Research Division, International Vaccine Institute, Seoul, Korea
Full list of author information is available at the end of the article

settings with endemic typhoid fever. The widely in use Widal test provides a cost efficient alternative [6] for serological diagnosis, however its performance remains unsatisfying with sensitivity reported from Tanzania of 75% using blood culture as the gold standard and applying a cut off titer of 1:80 [7]. The test further requires the establishment of a local cut off titer prior to use which is complicated. Therefore, a rapid test with a performance comparable to that of blood culture would be desirable.

A rapid diagnostic test for typhoid fever, Tubex[®] is commercially available that uses particle separation to detect immunoglobulin M (IgM) directed towards *Salmonella enterica* serovar Typhi (*S. Typhi*) O9 lipopolysaccharide in patient sera. Performance of the test has previously been evaluated in a number of studies in Asia but none in Africa. Using blood culture results for comparison, we assessed the sensitivity and specificity of the Tubex test among Tanzanian children hospitalized with febrile illness and compared our results with those from previous studies.

Methods

For evaluation of the Tubex test, we used a selected subset of serum samples that was obtained for a fever surveillance study [2] from Teule Hospital in Muheza District, Tanzania. In order to accommodate the required sample size for the test validation, we included randomly selected and age-matched *Salmonella enterica* serotype Typhi (*S. Typhi*) positive serum samples from a second fever surveillance study conducted at Chake Chake Hospital in Pemba, Zanzibar. All samples were collected from children between the ages of 2 months to 14 years from 2008 to 2009.

At Teule Hospital in Muheza, sera and blood was collected for culture from children with a history of three days of fever, or a history of less than three days of fever but with at least one of the following severity criteria: respiratory distress; deep breathing; respiratory distress in combination with severe pallor; prostration; capillary refill ≥ 3 seconds; temperature gradient; systolic blood pressure < 70 mm Hg; coma defined by Glasgow Coma Scale (GCS) ≤ 10 or Blantyre Coma Scale (BCS) ≤ 2 ; severe jaundice; history of two or more convulsions in the last 24 hours; blood glucose < 3 mmol associated with clinical signs; neck stiffness; bulging fontanel; or oxygen saturation $< 90\%$ [2].

At Chake Chake Hospital in Pemba, sera and blood was collected for culture from children with a recorded body temperature of $> 37.5^{\circ}\text{C}$ for outpatients and any history of fever for inpatients. Duration of fever was not considered for study recruitment.

About 3 to 5 milliliters (ml) of blood (depending on body weight) was collected and inoculated in a BactALERT[™] Pediatric-fan bottle (Teule Hospital) or a

BacTec Peds PLUS [™]/F bottle (Chake Chake Hospital) and incubated in the respective machine (BacT/ALERT 3D or BacTec 9050). Bacterial growth was evaluated following standard procedures.

The Tubex[®] test (IDL - Sweden) was conducted according to the manufacturer's instructions, which are as follows. Forty-five microliters (μl) of antigen covered particles were added to the Tubex Reaction Well Strip and 45 μl of non-hemolyzed serum was added. After two minutes of incubation time, 90 μl of magnetic antibody coated solution was added, and the strip was sealed and shaken for two minutes. The strip was then placed on a magnetic tray for five minutes, separating the particles if a positive sample had been added. The resulting color change of the solution was read and categorized on a scale from 0 to 10. The results were interpreted as positive for scores of 4 or greater and as negative for scores of 2 or below as per the manufacturer's instructions. Samples with a color corresponding to the value of 3 were interpreted as indeterminate. All blood culture isolates from individuals that matched the inclusion criteria and that were not considered a contaminant were included in the analysis.

We performed the Tubex test on non-hemolyzed serum samples from the patients of the two surveillance studies who had blood culture-confirmed *S. Typhi* (defined as group 1), randomly selected cases of non-Typhi serotypes of *S. enterica* (NTS) (defined as group 2), and randomly selected cases with other (non-*Salmonellae*) pathogenic bacteria (defined as group 3). Staff members performing the Tubex test were blinded to the blood culture results.

For the analysis, sensitivity (true-positive rate) was defined as the probability that the Tubex test result will be positive when there is blood culture-confirmed typhoid fever (group 1) and specificity (true-negative rate) was defined as the probability that the Tubex[®] test result will be negative when *S. Typhi* is not isolated from blood culture (groups 2 and 3). We conducted a secondary analysis using only group 3 as the control group. Comparison of test performance using different control groups was done using the Yates Chi-Square Test corrected for continuity.

We conducted a literature review in order to compare our findings with those from previous studies. We included studies of the Tubex test, which were identified by directly searching the MEDLINE database through PubMed. All articles since the first publication of the test [8] were included. We also conducted a supplementary search of references in retrieved articles. Abstracts were reviewed, and if relevant, the article was included.

A comparison of performance of the Tubex[®] test with earlier published Widal test results obtained from the same samples was done using McNemar's Test for

Correlated Proportions <http://faculty.vassar.edu/lowry/VassarStats.html>.

The fever surveillance studies at Chake-Chake and Teule Hospitals were approved by their respective local ethical review boards (Tanzania and Zanzibar), as well as by the International Vaccine Institute's Institutional Review Board. Written informed consent was obtained from legal guardians of all participants prior to any sample or data collection.

Results

A total of 139 samples were tested with Tubex. Thirty-three were found positive for *S. Typhi* in blood culture (group 1), 49 were culture-confirmed non-*S. Typhi* (NTS) infection (group 2), and 57 were other non-*Salmonella* infections that were not contaminants (group 3). Thirteen hemolyzed samples were excluded (Figure 1).

Of the 33 blood culture-positive *S. Typhi* cases, 26 had a positive Tubex result and were considered as true positives. Of the 106 blood culture confirmed NTS and non-salmonella cases (groups 2 and 3), 94 yielded a negative Tubex result and were considered as true negatives. Considering only the 57 non-*Salmonella* cases (group 3) as controls, resulted in 54 true negative cases.

Using groups 2 and 3 as controls showed a sensitivity of 79% and a specificity of 89% (Table 1). The same

Table 1 Performance of Tubex® using group 1 as true positives and two different control groups as true negatives

	Control Group	
	Group 2 + 3*	Group 3*
Sensitivity (95% CI)	0.79 (0.52-0.81)	0.79 (0.62-0.90)
(absolute numbers)	(26/33)	(26/33)
Specificity (95% CI)	0.89 (0.81-0.94)	0.97 (0.85-0.99)
(absolute numbers)	(94/106)	(94/97)

*Group 1 = *S. Typhi* (n = 33), Group 2 = all non-*yphi* *Salmonella* (n = 49), Group 3 = all blood culture-positive non-*Salmonella* cases (n = 57).

analysis was repeated excluding NTS from the pool of controls and showed 79% and 97% for sensitivity and specificity, respectively. There was no significant difference in the test performance using the two different control groups (all were $p > 0.05$ using the Chi square test).

A total of 14 articles were retrieved and evaluated for inclusion into the review. All of the reported studies were performed in Asia; none in Africa. A total of six articles were excluded: two evaluated the test for non-typhoidal *Salmonella* [9] or *S. Paratyphi* [10], three did not evaluate the sensitivity and specificity of the test [11-13], and one was a letter to the editor [14]. Thus, eight publications were included in the review (Table 2).

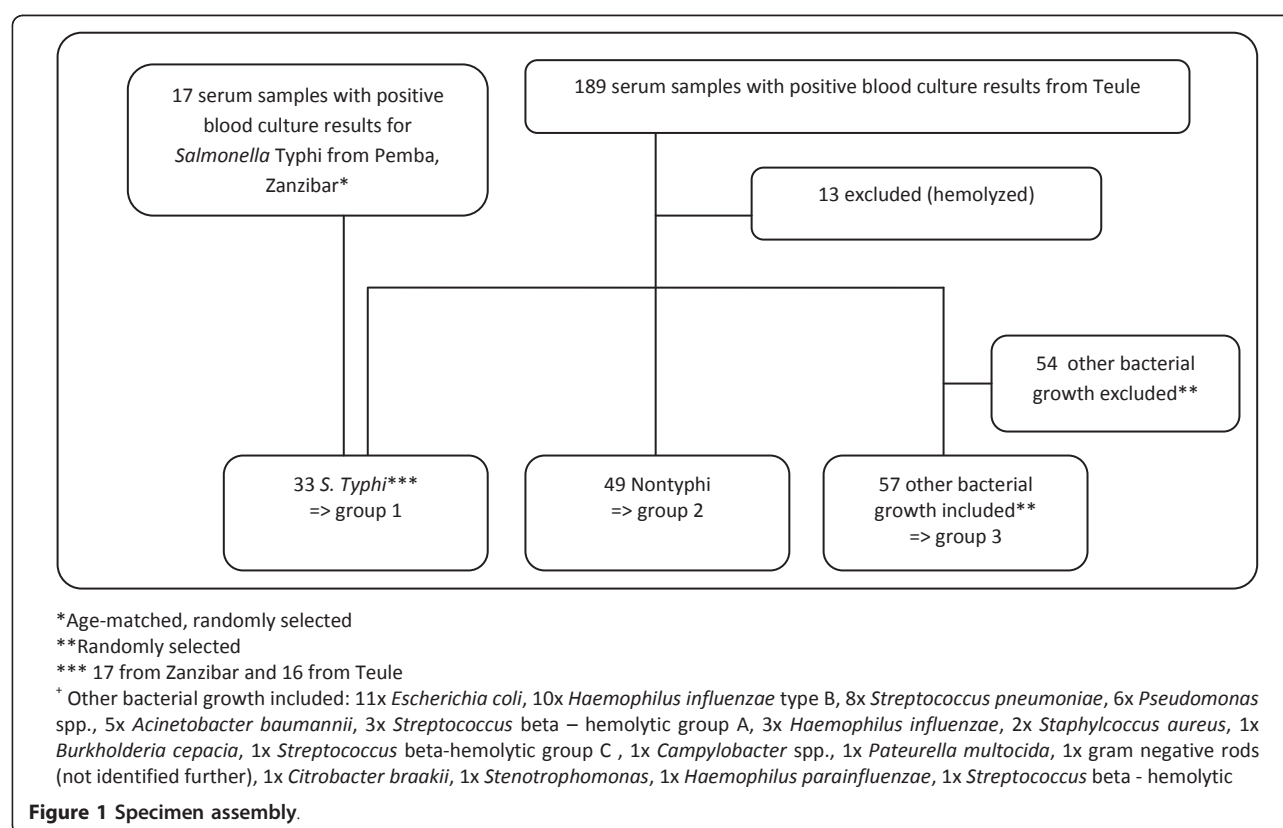


Table 2 Comparison of the performance of the Tubex® test from published reports

Author	Year	Journal	Sample Size	Location	Tubex® cut off	Sens	Spec	True neg. definition	Reader	Gold standard	Study population
Ley, B. et al	This paper	This Journal	139	Tanzania	>4	79%	89%	All non-typhi bacteremia	Investigator	Blood Culture (BACTEC)	>2 months + >37.5° (inpatients) & history of fever (outpatients)
			88			79%	97%	All non-salmonella bacteremia			
Naheed, A. et al	2008	Diagn Microbiol Infect Dis.	867	Bangladesh	≥4	60%	58%	Other confirmed bacteremia	ICDDR lab	Manual Blood Culture	Active surveillance Temp ≥38°C
						60%	64%	Blood culture neg & other bacteremia			
Rahman, M. et al	2007	Diagn Microbiol Infect Dis.	243	Bangladesh	>4	91.2%	82.3%	Other febrile patients	ICDDR lab, min. 2 independent lab techs	Manual Blood Culture	Outpatients, all ages with history of fever
						No. Pos	89.5%	Healthy subjects			
Dong, B. et al	2007	Epidemiol. Infect.	1732	China	≥2	100%	43%	Paratyphoid cases	-	Blood culture (BACTEC)	Age 5-60 with reported history of fever for 3 days
					≥4	69%	95%				
					≥6	62%	95%				
					≥8	23%	100%				
					≥10	15%	100%				
Kawano, R. L. et al	2006	JCM.	177	Philippines	≥2	94.7%	80.4%	Blood culture neg.	n/A	Manual Blood Culture & BACTEC	Clinically suspected typhoid cases
Dutta, S. et al	2006	Diagn Microbiol Infect Dis.	495	India	≥4	56%	88%	Paratyphoid and malaria cases	n/A	Blood Culture BACTEC	Outpatients, all ages, Pat with history of fever for 3 days
Ohlsen, S. J. et al	2004	JCM.	79	Vietnam	According to protocol	78%	94%	Other lab-confirmed febrile illnesses	n/A	Manual Blood Culture/ BACTEC	Pat ≥3 year and history of ≥4 day fever
House, D. et al	2001	JCM.	127	Vietnam	>2	87%	76%	Febrile hospitalized patients	labtech	Culture	Children and adults
Lim et al	1998	JCM.	105	Hong Kong & Malaysia	>2	100%	100%	Healthy individuals and pat with other bacterial diseases and autoimmune disease	labtechs	Culture confirmed (56% of pos.), clinical picture, various other tests	Clinical picture, culture confirmed,

Five of the included articles reported findings of test performance that were similar to our results [15-19]. Two publications showed considerably lower sensitivity and specificity [20,6], and one reported higher values [8] (Table 2).

Discussion

We found Tubex has a sensitivity of 79% using either control group (95%CI: 52-81% for groups 2 and 3, and 62-90% for group 3 only) and a specificity of 89-97% (95%CI: 81-94% for groups 2 and 3 and 85-99% for

group 3 only) irrespective of control group. To our knowledge, this is the first evaluation of the test in an African population. Our results were similar to those observed in previous studies (five out of eight studies) in Asia assessing the performance of the test [15-19], though Kawano *et al.* [17] and House *et al.* [19] used a lower cut-off titer than is recommended in the manual. In contrast, two studies [20,6] found the performance of Tubex to be poorer than our findings, despite using a similar cut-off value, gold standard, and inclusion criteria. The extremely good performance of Tubex observed by Lim *et al.* [8] has not been reproduced since.

An important limitation of this study is that the sera are combined from two different patient populations and the purposeful selection of samples included in the three groups. During the preparation of the study, we calculated the sample sizes of true positive sera and true negative sera that are required for validation of the Tubex test and for comparison with the Widal test performance. The number of true positive sera from either hospital alone was insufficient for the validation. Thus, we included *S. Typhi* blood culture-confirmed sera from Pemba. Analysis of the results by hospital was not possible because of insufficient sample size.

In a sub-analysis in assessing cross reactivity with NTS, blood culture-confirmed NTS cases were considered as true positives, and all other positive isolates, excluding *S. Typhi*, were considered as true negatives. In this sub-analysis Tubex had a sensitivity of 18% and a specificity of 95% (analysis not shown).

Comparison with a Widal test that was earlier conducted using the same samples [7] revealed no significant difference ($p > 0.05$) for any of the performance indicators, irrespective of the applied control group. But compared to the Widal test, Tubex is easier and quicker to perform. The Widal test requires 16 - 20 hours until the results are obtained while the complete procedure for the Tubex test is approximately 20 minutes. Tubex is more expensive at approximately 2.15 USD per test compared to <0.80 USD per test for the Widal tube agglutination test [6].

Interpreting the Tubex test results was found to be difficult and the results were prone to inter - reader variation. Assessing the color change according to the provided color scale requires experience and standardized good lighting conditions. The Tubex test can only be applied to non-hemolyzed and non-icteric serum samples, thus limiting its general application. However Tam *et al.* [12] have described a method that includes a washing step and thereby addresses the problem of turbid serum. This method requires double the amount of antigen-coated particles as well as glycine buffered saline (GBS), thereby increasing the price per sample to

approximately 4.50 USD and reducing its feasibility as an easy-to-perform test. While neither of the tests can be performed by untrained staff, interpretation of results is considered easier for the Widal test compared to Tubex.

Conclusion

The advantages of Tubex over the Widal test and the gold standard of blood culture is the short time it requires to obtain a result, and it does not require establishing a local cut-off value as with the Widal. In settings that can afford the relatively high cost of Tubex and that require instant individual diagnoses to support the clinical diagnosis of typhoid fever, Tubex is superior to the Widal tube agglutination test. For screening and surveillance purposes, as well as in settings with limited financial and technical resources, the Widal tube agglutination test is a possible alternative that can provide a similar performance as Tubex at a lower cost though it requires more time. Our evaluation of Tubex showed that any result must be handled with precaution. Results should be considered as indicative, not confirmatory. The test may be used to exclude disease though. In conclusion, the need for a reliable, fast, cheap, and easy-to-apply rapid diagnostic test for typhoid fever remains in high demand.

Acknowledgements

We thank Hugh Reyburn of the London School of Hygiene and Tropical Medicine who supported this project and provided his scientific expertise. We also thank Rajabu Malahiyo and Steven Magesa for their support. This work was supported by a grant from the Korea International Cooperation Agency (KOICA) and the Swedish International Development Cooperation Agency (SIDA) to the International Vaccine Institute. This study is published with permission from the Director General of the Tanzanian National Institute for Medical Research, Dar-Es-Salaam. We are grateful to the patients and their parents who made this work possible. We thank all of the technical staff and research assistants who were involved in the study.

Author details

¹Translational Research Division, International Vaccine Institute, Seoul, Korea. ²Laboratory Division, Public Health Laboratory (Pemba) - Ivo de Carneri, Chake Chake, Tanzania. ³Amani Centre, National Institute for Medical Research, Tanga, Tanzania. ⁴Joint Malaria Program, Tanga, Tanzania. ⁵Asia Pacific Malaria Elimination Network (APMEN), Menzies School of Health Research, Casuarina, Australia. ⁶Teule Hospital, Muheza, Tanga, Tanzania. ⁷Oxford Research Unit, Mahidol University, Bangkok, Thailand. ⁸Biocenter, University of Vienna, Vienna, Austria.

Authors' contributions

BL performed the TUBEX test, analyzed results and wrote the manuscript, KT performed TUBEX tests, literature search and contributed to the manuscript, SMA supervised the laboratory work in Pemba, GM was in charge of the implementation and study management in Teule, LvS provided scientific support to study staff and contributed to the manuscript, BA supervised the laboratory work in Teule, ICEH was involved in the clinical care of patients, AM was in charge of data management, AS performed blood culture procedures, DRK performed statistical analyses, RLO provided scientific support to the manuscript, MF provided scientific support to the manuscript, JDC provided scientific support to the manuscript, HW provided scientific support to the manuscript, JLD provided major scientific support to the

manuscript and was involved in the clinical care of patients, SMA provided scientific support to the manuscript and the study in Pemba. All authors have read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 25 January 2011 Accepted: 24 May 2011

Published: 24 May 2011

References

- Crump JA, Luby SP, Mintz ED: **The global burden of typhoid fever.** *Bulletin of the World Health Organization* 2004, **82**(5):346-353.
- Mtove G, Amos B, von Seidlein L, Hendriksen I, Mwambuli A, Kimera J, Mallahiyo R, Kim DR, Ochiai RL, Clemens JD, Reyburn H, Magesa S, Deen JL: **Invasive salmonellosis among children admitted to a rural Tanzanian hospital and a comparison with previous studies.** *PLoS One* 2010, **5**(2):e9244.
- Crump JC, Ramadhani HO, Morrissey AB, Saganda W, Mwako MS, Yang LY, Chow SC, Morpeth SC, Reyburn H, Njau BN, Shaw AV, Diefenthal HC, Shao JF, Bartlett JA, Maro VP: **Invasive bacterial and fungal infections among hospitalized HIV-infected and HIV-uninfected adults and adolescents in northern Tanzania.** *CID* 2011, **52**(3):341-348.
- Parry CM, Hien TT, Dougan G, White N, Farrar JJ: **Typhoid Fever.** *The New England Journal of Medicine* 2002, **347**(22):1770-1782.
- Wilke A, Ergonul O, Bayar B: **Widal Test in Diagnosis of Typhoid Fever in Turkey.** *Clin Diagn Lab Immunol* 2002, **9**(4):938-41.
- Sur D, Manna B, Sen B, Deb AK, Deen JL, Wain J, von Seidlein L, Ochiai RL, Clemens JD, Bhattacharya SK: **Evaluation of a new-generation serologic test for the diagnosis of typhoid fever: data from a community - based surveillance in Calcutta, India.** *Diagnostic Microbiology and Infectious Disease* 2006, **56**:359-365.
- Ley B, Mtove G, Thriemer K, Amos B, von Seidlein L, Hendriksen I, Mwambuli A, Shoo A, Malahiyo R, Ame S, Kim DR, Ochiai RL, Clemens JD, Reyburn H, Wilfing H, Magesa S, Deen JL: **Evaluation of the Widal tube agglutination test for the diagnosis of typhoid fever among children admitted to a rural hospital in Tanzania and a comparison with previous studies.** *BMC* 2010, **10**:180.
- Lim PL, Tam FC, Cheong YM, Jegathesan M: **One-step 2-minute test to detect typhoid-specific antibodies based on particle separation in tubes.** *J Clin Microbiol* 1998, **36**(8):2271-8.
- Oracz G, Feleszko W, Golicka D, Maksymiuk J, Klonowska A, Szajewska H: **Rapid diagnosis of acute Salmonella gastrointestinal infection.** *Clin Infect Dis* 2003, **36**(1):112-5.
- Tam FC, Wang M, Dong B, Leung DT, Ma CH, Lim PLL: **New rapid test for paratyphoid a fever: usefulness, cross-detection, and solution.** *Diagn Microbiol Infect Dis* 2008, **62**(2):142-50.
- Tam FC, Lim PL: **The TUBEX typhoid test based on particle-inhibition immunoassay detects IgM but not IgG anti-O9 antibodies.** *J Immunol Methods* 2003, **282**(1-2):83-91.
- Tam FC, Leung DT, Ma CH, Lim PL: **Modification of the TUBEX typhoid test to detect antibodies directly from haemolytic and whole blood.** *Journal of Medical Microbiology* 2008, **57**:1349-1353.
- Tam FC, Ling TK, Wong KT, Leung DT, Chan RC, Lim PL: **The TUBEX test detects not only typhoid-specific antibodies but also soluble antigens and whole bacteria.** *J Med Microbiol* 2008, **57**(Pt 3):316-23.
- Feleszko W, Maksymiuk J, Oracz G, Golicka D, Szajewska H: **The TUBEX typhoid test detects current Salmonella infections.** *J Immunol Methods* 2004, **285**(1):137-8.
- Rahman M, Siddique AK, Tam FC, Sharmin S, Rashid H, Iqbal A, Ahmed S, Nair GB, Chaignat CL, Lim PL: **Rapid detection of early typhoid fever in endemic community children by the TUBEX O9-antibody test.** *Diagn Microbiol Infect Dis* 2007, **58**(3):275-81.
- Dong B, Galindo CM, Shin E, Acosta CJ, Page AL, Wang M, Kim D, Ochiai RL, Park J, Ali M, Seidlein LV, Xu Z, Yang J, Clemens JD: **Optimizing typhoid fever case definitions by combining serological tests in a large population study in Hechi City, China.** *Epidemiol Infect* 2007, **135**(6):1014.
- Kawano RL, Leano SA, Agdamag DMA: **Comparison of serological Test Kits for Diagnosis of Typhoid Fever in the Philippines.** *Journal of Clinical Microbiology* 2007, **45**(1):246-247.
- Olsen SJ, Pruckler J, Bibb W, Thanh NTM, Trinh TM, Minh NT, Sivapalingam S, Gupta A, Phuong PT, Chinh NT, Chau NV, Cam PD, Mintz ED: **Evaluation of Rapid Diagnostic Test for Typhoid Fever.** *Journal of Clinical Microbiology* 2004, **42**(5):1885-1889.
- House D, Wain J, Ho VA, Diep TS, Chinh NT, Bay PV, Vinh H, Duc M, Parry CM, Dougan G, White NJ, Hien TT, Farrar JJ: **Serology Of Typhoid Fever in an Area of Endemicity and Its Relevance to Diagnosis.** *Journal of Clinical Microbiology* 2001, **39**(2):1002-1007.
- Naheed A, Ram PK, Brooks WA, Mintz ED, Hossain MA, Parsons MM, Luby SP, Breiman RF: **Clinical value of Tubex® and Typhidot rapid diagnostic tests for typhoid fever in an urban community clinic in Bangladesh.** *Diagn Microbiol Infect Dis* 2008, **61**(4):381-6.

Pre-publication history

The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2334/11/147/prepub

doi:10.1186/1471-2334-11-147

Cite this article as: Ley et al.: Assessment and comparative analysis of a rapid diagnostic test (Tubex®) for the diagnosis of typhoid fever among hospitalized children in rural Tanzania. *BMC Infectious Diseases* 2011 **11**:147.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



Evaluation of a Rapid Dipstick (Crystal VC) for the Diagnosis of Cholera in Zanzibar and a Comparison with Previous Studies

Benedikt Ley^{1,2*}, Ahmed M. Khatib³, Kamala Thriemer¹, Lorenz von Seidlein⁴, Jacqueline Deen⁴, Asish Mukhopadhyay⁵, Na-Yoon Chang¹, Ramadhan Hashim¹, Wolfgang Schmied¹, Clara J-L. Busch⁶, Rita Reyburn¹, Thomas Wierzb¹, John D. Clemens¹, Harald Wilfing², Godwin Enwere⁷, Theresa Aguado⁷, Mohammad S. Jiddawi³, David Sack⁸, Said M. Ali^{3,9}

1 Translational Research Division, International Vaccine Institute, Seoul, Korea, **2** Biocenter, University of Vienna, Vienna, Austria, **3** Ministry of Health and Social Welfare, Zanzibar, Tanzania, **4** Global Health Division, Menzies School of Health Research, Casuarina, Australia, **5** National Institute of Cholera and Enteric Diseases, Kolkata, India, **6** Institute of Molecular Biotechnology, Vienna, Austria, **7** World Health Organization, Geneva, Switzerland, **8** John Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States of America, **9** Public Health Laboratory (Pemba) – Ivo de Carneri, Chake Chake, Tanzania

Abstract

Background: The gold standard for the diagnosis of cholera is stool culture, but this requires laboratory facilities and takes at least 24 hours. A rapid diagnostic test (RDT) that can be used by minimally trained staff at treatment centers could potentially improve the reporting and management of cholera outbreaks.

Methods: We evaluated the Crystal VCTM RDT under field conditions in Zanzibar in 2009. Patients presenting to treatment centers with watery diarrhea provided a stool sample for rapid diagnostic testing. Results were compared to stool culture performed in a reference laboratory. We assessed the overall performance of the RDT and evaluated whether previous intake of antibiotics, intravenous fluids, location of testing, and skill level of the technician affected the RDT results.

Results: We included stool samples from 624 patients. Compared to culture, the overall sensitivity of the RDT was 93.1% (95%CI: 88.7 to 96.2%), specificity was 49.2% (95%CI: 44.3 to 54.1%), the positive predictive value was 47.0% (95%CI: 42.1 to 52.0%) and the negative predictive value was 93.6% (95%CI: 89.6 to 96.5%). The overall false positivity rate was 50.8% (213/419); fieldworkers frequently misread very faint test lines as positive.

Conclusion: The observed sensitivity of the Crystal VC RDT evaluated was similar compared to earlier versions, while specificity was poorer. The current version of the RDT could potentially be used as a screening tool in the field. Because of the high proportion of false positive results when field workers test stool specimens, positive results will need to be confirmed with stool culture.

Citation: Ley B, Khatib AM, Thriemer K, von Seidlein L, Deen J, et al. (2012) Evaluation of a Rapid Dipstick (Crystal VC) for the Diagnosis of Cholera in Zanzibar and a Comparison with Previous Studies. PLoS ONE 7(5): e36930. doi:10.1371/journal.pone.0036930

Editor: Philip Campbell Hill, University of Otago, New Zealand

Received: February 15, 2012; **Accepted:** April 17, 2012; **Published:** May 25, 2012

Copyright: © 2012 Ley et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study has been funded by the Bill and Melinda Gates Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ley.benedikt@gmail.com

Introduction

Cholera remains a very common and potentially lethal disease in Asia and Africa. Globally, more than 220,000 cases were reported to the World Health Organization (WHO) in 2009 [1], however the true number of cases, including unreported cases, is likely to be much higher – perhaps 3–5 million cases/year [2]. Cholera occurs mainly in areas with poor infrastructure and limited access to clean water. The etiologic organisms, *Vibrio cholerae* O1 and O139, are highly transmissible and can cause explosive outbreaks. While many of those affected experience only mild symptoms, some suffer from severe disease characterized by profuse diarrhea, electrolyte imbalance, coma and death if prompt rehydration is not provided [3,4]. Cholera cases have been

reported from Zanzibar since 1978 with regular outbreaks documented since then [5,6].

The gold standard for laboratory confirmation of cholera is stool culture [7]. This is a routine procedure but requires laboratory infrastructure including trained staff. A single stool culture costs approximately 4 USD/case [8] and requires about 24 to 72 hours and transport to the closest sufficiently equipped laboratory, which may create additional costs. Furthermore, microbiologic facilities are often not available in locations where cholera occurs. A rapid diagnostic test (RDT) that is simple, easy to use and interpret, can be stored without refrigeration and is reasonably priced so that it can be deployed widely would be useful for the early confirmation of cholera outbreaks. Ideally, the RDT should be highly sensitive so as not to miss the diagnosis of

cases and be sufficiently specific when used under actual field conditions [9]. Cholera confirmation would enable immediate implementation of control measures such as reactive vaccination [6], as well as more accurate reporting of the burden of the disease.

A cholera RDT based on the detection of lipopolysaccharide (LPS) using gold particles was developed by the Institute Pasteur (IP). The RDT is a lateral flow immunochromatographic test for the qualitative determination of lipopolysaccharide antigen of both *Vibrio cholerae* O1 and O139 serogroups from stool specimens using monoclonal antibodies specific to *V. cholerae* O1 and O139 LPS. Through a licensure agreement, the RDT is now being produced by Span Diagnostics (Surat, India) under the trade name Crystal VC™ at a price of 19.00 USD/test kit (10 test strips). The test kit is stable at temperatures between 4°C to 30°C, and test strips are packed in waterproof pouches, allowing storage under high humidity conditions. Previous evaluations have been performed on the prototype and commercial versions of the RDT [10–15]. The primary objective of this study is to validate the current version of the Crystal VC™ RDT when performed by health workers in first-level treatment centers in Zanzibar. We also sought to assess if the RDT results were affected by the skill level of the reader and previous intake of antibiotics or intravenous fluids.

Methods

Ethics

The study was conducted according to the principles expressed in the Declaration of Helsinki. Written informed consent was obtained from all participants. The Zanzibar Research Council Ethics Committee, the Institutional Review Board of the International Vaccine Institute, Seoul, Korea, and the Research Ethics Review Committee of the World Health Organization, Geneva, Switzerland approved this project.

Study site

The archipelago of Zanzibar lies about 50 kilometers east of mainland Tanzania and consists of two main islands, Unguja and Pemba, as well as smaller islets (Figure 1). In 2009, Zanzibar had a population of about 1.22 million [16]. Stool samples were collected at cholera treatment camps that were set up during outbreaks on the two main islands, Unguja and Pemba, in 2009. Treatment of patients was provided according to national guidelines.

Study procedures

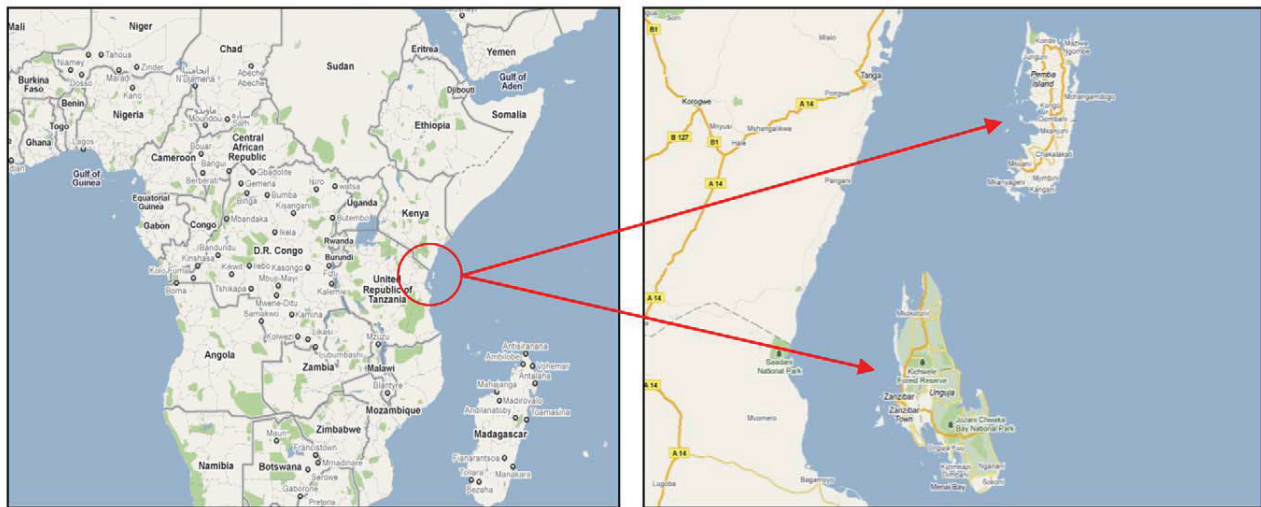
Acute watery diarrhea was defined as a minimum of three liquid, non-bloody, stools within 24 hours. Prior to presentation, no further inclusion/exclusion criteria were applied. Patients presenting with acute watery diarrhea were requested to provide a stool sample in a disposable plastic container. A swab was inserted into the stool sample and used to inoculate a tube of pre-packaged Cary Blair medium (EIKEN, Japan) for transport to the laboratory. About 200 µl of stool from each sample was used for dipstick testing on site. A case report form (CRF) was completed to record frequency of bowel movement over the previous 24 hours, antibiotics received, and fluid management (intravenous (IV) or oral rehydration solution (ORS)) provided at the health center. Bulk stool from a subset of patients attending a camp close to one of the participating laboratories was transported to the lab for additional testing on the same day as described below.

Dipstick test. The RDT was stored at room temperature and performed according to the package insert. Liquid stool was collected in a disposable plastic container. Approximately 200 µl (4 drops) of stool were transferred with a disposable pipette to a disposable test tube provided with the kit. One drop of dilution

buffer was added. The dipstick was inserted into the diluted stool and results were read within 15–20 minutes. The appearance of two bands on the dipstick, one control and one test, indicated that the stool sample was positive for *V. cholerae*. The appearance of only the control band indicated a negative sample. The non-appearance of the control band indicated a procedural error. Stool samples were tested under field conditions in the cholera treatment centers and in the laboratory as described below. Dipstick results were recorded on the CRF, whereas laboratory results were recorded in separate laboratory forms.

- Performed under field conditions. A local health worker in each cholera treatment camp performed the RDT after training and a copy of the English test kit manual containing illustrations on test procedure and interpretation had been provided. Training consisted of a theoretical session using a Power Point Presentation containing information on test procedures and schematic pictures of positive and negative test results based on the package insert. This was followed by a practical session during which the test was performed a number of times. All field workers were visited frequently in the field to ensure correct handling of the test. All local health workers had completed at least primary education and delivered basic medical services to attending diarrheal patients. Fieldworkers performed the test outdoors in daylight.
- Performed under laboratory conditions. In order to assess the potential influence of environmental and light conditions, laboratory technicians were asked to repeat the test on bulk stool collected at the camps and to read the result independently. Stool culture results were not yet available at the time of the performance of the test and laboratory technicians were blinded as to the results of the RDT performed in the field, as well as the clinical picture of the patient. Two laboratory technicians performed the RDT after receiving training similar to the field workers and had received a copy of the test kit manual. The laboratory technicians performed the test indoors using electric light sources. All participating technicians had a diploma in laboratory sciences, which requires a minimum of one year of education, and had a minimum of three years working experience.

Stool culture. Upon arrival in the laboratory, samples from the Cary Blair media were streaked out on Thiosulphate Citrate Bile Sucrose Agar (TCBS; EIKEN, Japan), inoculated in alkaline peptone water (APW) and incubated at 37°C for 12–24 hours. If samples arrived as bulk stool, the samples were diluted in APW. An aliquot was streaked out on TCBS and the samples in APW and on TCBS were incubated for 12–24 hours at 37°C. If no growth on TCBS was detected after incubation, an aliquot of the sample in APW was streaked out on TCBS and incubated again. If yellow colonies indicative of *V. cholerae* were detected on TCBS, motility indole ornithine agar (MIO) and triple sugar iron agar (TSI) were inoculated with colonies from TCBS and incubated for 18 hours at 37°C. In addition, a colony from TCBS was sub-cultivated on gelatin agar for later serological confirmation and incubated at 37°C overnight. If colonies indicative of *V. cholerae* were observed on TSI and MIO after incubation, colonies from gelatin agar were tested for agglutination reactions with O1 polyvalent, O1 Inaba, O1 Ogawa and O139 antiserum (Beckton Dickinson, USA) as described elsewhere [17]. *V. cholerae* strains were transported to the National Institute of Cholera and Enteric Diseases in Kolkata, India where identification of the isolates was confirmed.



Source: Google Maps

Figure 1. Study site.
doi:10.1371/journal.pone.0036930.g001

Definitions, data management and analysis

The CRF and laboratory results of each patient were computerized and linked using unique study identification numbers. The primary endpoint was the assessment of the performance of the RDT (done in the field) using microbiological stool culture result as the gold standard for comparison. Sensitivity (true-positive or TP rate) was defined as the probability that patients with laboratory-confirmed cholera had a positive RDT. Specificity (true-negative or TN rate) was the probability that patients with no laboratory-confirmed cholera had a negative RDT. The positive predictive value (PPV) was the probability that patients with a positive RDT had *V.cholerae* isolated from stool culture. The negative predictive value (NPV) was the probability that patients with a negative RDT had no *V.cholerae* isolated from a stool culture. The false positivity or FP rate ($= FP/[FP+TN]$ or $1 - \text{specificity}$) was the proportion of stool samples with no *V.cholerae* isolated on culture but showed a positive RDT result. The false negativity or FN rate ($= FN/[TP+FN]$ or $1 - \text{sensitivity}$) was the proportion of stool samples with *V.cholerae* isolated on culture but showed a negative RDT result.

We performed sub-group analyses by island (Pemba or Unguja), by previous recent intake versus non-intake of antibiotics and by receipt of intravenous fluids following previously published studies showing differences in RDT performance [13]. We defined recent intake of antibiotics as receipt of any oral or parenteral antimicrobial for the current illness prior to the collection of a stool sample. We classified fluid management at the treatment center as oral rehydration solution (ORS) or intravenous fluids (IVF) with or without ORS. To assess whether test validity was related to skill level of the reader and location of the testing, we compared the performance of the RDT when done in the field versus in the laboratory on a subset of samples.

Comparison of unpaired samples was done using chi-square test; comparison of paired samples was done using McNemars test. Confidence intervals were calculated using exact method. Level of agreement was calculated using Cohen's Kappa test for unweighted proportions. Calculations were done using Stata, version 10

(StataCorp, College Station, TX, USA). Performance indicators were calculated using Excel 2010 (Microsoft, WA, USA).

Results

There were 624 patients who presented to a cholera treatment camp with acute watery diarrhea and were recruited into the study: 81 in Unguja and 543 in Pemba. We excluded 2 samples sent for culture but on which no RDT was done. A total of 622 stool samples were included in the analysis, 79 (13%) from Unguja and 543 (87%) from Pemba residents (Figure 2).

Performance of the RDT in the field

Of the 622 stool samples, 203 (32.6%) yielded *V.cholerae* O1. No *V.cholerae* O139 was isolated. Using culture results as the gold standard, we calculated the sensitivity, specificity, PPV, and NPV of the RDT performed in the field for the diagnosis of cholera (Table 1). Overall sensitivity was 93.1% (95%CI: 88.7 to 96.2%), specificity was 49.2% (95%CI: 44.3 to 54.1%), the positive predictive value (PPV) was 47.0% (95%CI: 42.1 to 52.0%) and the negative predictive value (NPV) was 93.6% (95%CI: 89.6 to 96.5%). The overall false positivity rate was 50.8% (213/419).

Sub-group analyses of performance of the RDT in the field

We evaluated the RDT performance by island (Table 1). *V.cholerae* was isolated from 46/79 or 58.2% of stool samples from Unguja compared to a significantly lower proportion of 157/543 or 28.9%, from Pemba ($p < 0.01$). No significant differences in sensitivity, specificity and NPV of the RDT were observed between Unguja and Pemba, as well as between each island with the overall results (all $p > 0.05$). However, we found a significant difference in PPV between Unguja and Pemba (71.9, 95%CI: 58.5–83.0 versus 42.9, 95%CI: 37.6–48.3; $p = 0.02$).

We compared the RDT performance by fluid management of patients: Rehydration treatment at the cholera camp was recorded for 592/622 (95.2%) participants (Table 2). Only 15.2% (32/210) of participants who received oral rehydration had a positive stool

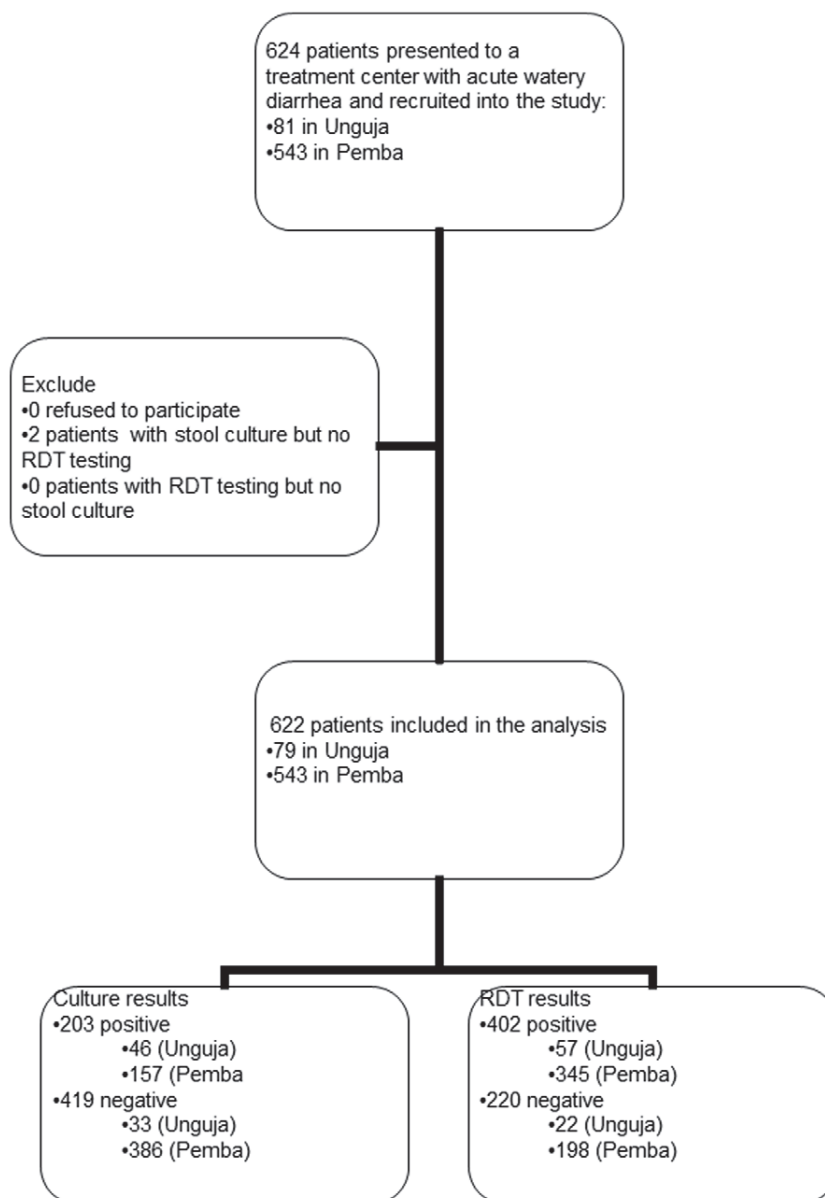


Figure 2. Flow of study participants.
doi:10.1371/journal.pone.0036930.g002

Table 1. Performance of the cholera rapid diagnostic test, Pemba and Unguja, Zanzibar.

	Sensitivity% (95%CI*) (TP/(TP+FN))	Specificity% (95%CI*) (TN/(TN+FP))	PPV% (95%CI*) (TP/ (TP+FP))	NPV% (95%CI*) (TN/ (TN+FN))	No. pos./total (%) with <i>V. cholera</i> isolated on stool culture
Total (n = 622)	93.1 (88.7–96.2) (189/203)	49.2 (44.3–54.1) (206/419)	47.0 (42.1–52.0) (189/402)	93.6 (89.6–96.5) (206/220)	203/622 (32.6)
Unguja (n = 79)	89.1 (76.4–96.4) (41/46)	51.5 (33.5–69.2) (17/33)	71.9 (58.5–83.0) (41/57)	77.3 (54.6–92.2) (17/22)	46/79 (58.2)
Pemba (n = 543)	94.3 (89.4–97.4) (148/157)	49.0 (43.9–54.1) (189/386)	42.9 (37.6–48.3) (148/345)	95.5 (91.5–97.9) (189/198)	157/543 (28.9)
p	0.82	0.87	0.02	0.53	<0.01

*using exact method.

doi:10.1371/journal.pone.0036930.t001

Table 2. Stratified analysis of the performance of cholera dipstick test according to fluid management (oral rehydration or intravenous fluids) in 592 patients and recent antibiotic intake (yes or no) in 576 patients.

	Sensitivity (95%CI)* (TP/(TP+FN))	Specificity (95%CI)* (TN/(TN+FP))	PPV (95%CI)* (TP/ (TP+FP))	NPV (95%CI)* (TN/ (TN+FN))	No. pos./total (%) with <i>V. cholerae</i> isolated on stool culture
Fluid management (n = 592)					
Oral rehydration n = 210	93.8 (79.2–99.2) (30/32)	53.9 (46.3–61.4) (96/178)	26.8 (18.9–36.0) (30/112)	98.0 (92.8–99.8) (96/98)	32**/210 (15.2)
Intravenous rehydration n = 382	92.7 (87.6–96.1) (153/165)	42.9 (36.2–49.7) (93/217)	55.2 (49.2–61.2) (153/277)	88.6 (80.9–94.0) (93/105)	165**/382 (43.2)
p	0.97	0.19	<0.01	0.62	<0.01
Recent antibiotic (AB) intake (n = 576)					
AB received n = 58	93.8 (69.7–99.8) (15/16)	59.5 (43.3–74.4) (25/42)	46.9 (29.1–65.3) (15/32)	96.2 (80.4–99.9) (25/26)	16/58 (27.6)
No AB received n = 519	92.8 (87.8–96.2) (155/167)	49.3 (43.9–54.7) (173/351)	46.5 (41.1–52.1) (155/333)	93.5 (86.1–94.7) (173/190)	167/518 (32.2)
p	0.98	0.48	0.98	0.86	0.60

*using exact method.

**in 6/203 culture positive cases no rehydration treatment was provided.

doi:10.1371/journal.pone.0036930.t002

culture, compared with 43.2% (165/382) of those who received IV fluids ($p < 0.01$). There were no statistically significant differences in sensitivity, specificity and NPV of RDT performance by fluids received (all $p > 0.05$). However we again found a significant difference in PPV among those who received oral compared to IV rehydration (26.8%, 95%CI: 18.9 to 36.0% versus 55.2%, 95%CI: 49.2 to 61.2%; $p < 0.01$). To evaluate whether the provided fluid treatment biased the field workers' interpretation of the RDT results, we compared the false positivity rate by fluid management. The false positivity rate was 46.1% among those who were orally rehydrated and 57.1% among those who were intravenously rehydrated ($p = 0.22$).

Information on prior antibiotic treatment was recorded for 576/622 (92.6%) participants. The percentage with a positive stool culture was 27.6% among those who had received antibiotics and 32.2% among those who had not ($p = 0.60$). We assessed whether previous antibiotic treatment affected the RDT performance (Table 2). Sensitivity, specificity, PPV and NPV did not vary significantly among recipients and non-recipients of antibiotics. The false positivity rate was 40.5% among those who had received antibiotics and 50.7% among those who had not ($p = 0.45$).

Comparison of the performance of the RDT in the field versus in the laboratory

We compared the performance of RDT on a subset of 67/79 (84.8%) stool samples from Unguja tested both in the field and in

the laboratory (Table 3). In this subset, 40/67 (59.7%) samples yielded *V. cholerae* on culture. There was no statistically significant difference in the sensitivity, specificity PPV and NPV of the RDT's performance (all $p > 0.05$). The false positivity rate of the RDT was 45.4% in the field and 25.9% in the laboratory (Cohen's kappa 0.8).

Operational characteristics

The test procedure, excluding sample collection, requires 20–25 minutes. The test kit manual provides clear instructions, and handling of the test was considered simple by field workers. Field workers found it easy to distinguish between valid and non-valid test results, based on the appearance of a control line. However, very faint positive test lines were interpreted as a positive result, but could not be confirmed by culture.

Discussion

We found an overall high sensitivity (93.1%) of the current version of the cholera RDT consistent with previous reports but a much poorer specificity (49.2%). Earlier studies were performed using a prototype dipstick – developed by the Institute Pasteur – and earlier versions of the commercial kit. Studies using the prototype versions of the RDT reported sensitivities in the range of 93% to 99% and specificities of 67%–97% [12–15], whereas more recent reports on earlier versions of the commercial kit

Table 3. Comparison of the field and laboratory performance of the cholera dipstick test.

(n = 67)	Sensitivity% (95%CI)* (TP/TP+FN)	Specificity% (95%CI)* (TN/TN+FP)	PPV% (95%CI)* (TP/ TP+FP)	NPV% (95%CI)* (TN/ TN+FN)	No. (%) with <i>V. cholera</i> isolated on stool culture**
Field	90.0 (76.3–97.2) (36/40)	55.6 (35.3–74.5) (15/27)	75.0 (60.4–86.4) (36/48)	78.9 (54.4–93.9) (15/19)	40 (59.7)
Laboratory	87.5 (73.2–95.8) (35/40)	74.1 (53.7–88.9) (20/27)	83.3 (68.6–93.0) (35/42)	80.0 (59.3–93.2) (20/25)	40 (59.7)
p***	0.931	0.510	0.740	0.977	

*using exact method.

**rapid diagnostic testing was done both in the field and in the laboratory on specimens from the same stool samples.

***using McNemars test.

doi:10.1371/journal.pone.0036930.t003

Table 4. Validation studies of the Institut Pasteur prototype and Crystal VC™ test for diagnosis of *Vibrio cholerae*(O1 samples only considered).

RDT	Institut Pasteur prototype				Crystal VC™ (Span Diagnostics, India)		
Authors	Nato, et.al. [14]	Bhuiyan, et.al. [13]	Wang, et.al. [12]	Kalluri, et.al. [11]	Harris, et.al. [9]	Mukherjee et al. [10]	This Paper
Journal, Year	CDLI, 2003	JCM, 2003	BMC Inf. Dis., 2006	TM&IH, 2006	TM&IH, 2009	JJD, 2010	This Journal
Samples from	Madagascar	Bangladesh	Mozambique	Bangladesh	Guinea-Bissau	India	Zanzibar/Tanzania
Gold standard; type of stool samples used	Culture of frozen stool or rapid cultures of stool collected on filter paper	Culture of rectal swab in C-B media (and multiplex PCR when RDT+/culture -)	Culture of rectal swabs in C-B media and bulk stool	Culture of bulk stool	PCR of swab in C-B media	Culture of bulk stool	Culture of rectal swab in C-B media
Type of stool samples for RDT testing	frozen stool/rapid cultures of stool collected on filter paper	Rectal swab incubated in APW at 37°C for 4 h	Rectal swab incubated in APW at 37°C/4 h and fresh, bulk stool	Fresh, bulk stool	Fresh, bulk stool	Fresh, bulk stool	Fresh, bulk stool
RDT done by	?	?	Trained technician at field laboratory	Trained technician at ICDDR&B and trained field 'paramedics'	Individuals with graduate-level laboratory technical training	Post graduate investigator	Field worker at cholera treatment center (total) Field worker at cholera treatment center (subset) Lab. tech. (subset)
No of samples	140	134	219 rectal swabs in APW 172 bulk stool	304	101	212	622 (total) 67(subset)
No (%) positive for VC O1 by gold standard	65 (46%)	68 (51%)	138 (35%)	116 (38%)	65 (64%)	72 (34%)	203 (33%) (total) 40 (60%) (subset)
Sensitivity	98.5%	96%	97% rectal swabs in APW 93% bulk stool	94% lab tech. 93% field 'paramedics'	97%	92%	93% (total) 88% lab tech (subset) 90% field worker (subset)
Specificity	96%	92%	97% rectal swabs in APW 77% bulk stool	76% lab tech. 67% by field 'paramedics'	71–76%;	73%	49% (total) 74% lab tech (subset) 56% field worker (subset)
PPV	95.6%	93%	94% rectal swabs in APW 74% bulk stool	70% lab tech. 63% field 'paramedics'	87–89%	64%	47% (total) 83% lab tech (subset) 75% field worker (subset)
NPV	99%	95%	?	95% lab tech. 94% field 'paramedics'	92–93%	95%	94% (total) 80% lab tech (subset) 79% field worker (subset)

doi:10.1371/journal.pone.0036930.t004

showed sensitivities of 92–97% and specificities of 71–76% [10,11]. These studies not only varied by the RDT version used but also by the methodology and qualifications of the study personnel performing the RDT (Table 4). There were also variations in the test procedure such as addition or non-addition of a buffer solution to the sample. In some studies, a 4-hour incubation step in alkaline peptone water was added. Overall, the prototype and precursor commercial kits performed better than the current version tested.

The poor specificity of the current version of the commercial kit was associated with an overall false positive rate of 50.8%. The RDT's false positivity rate when the RDT was read in the field was 44% versus 26% when done by laboratory technicians; possibly faint test lines on the dipsticks were over-read by field workers as positive. We hypothesized whether patient characteristics (fluid management or receipt of antibiotics) biased the field workers' interpretation of the RDT results. However, we found no significant differences in false positivity in these sub-group analyses. More likely the fieldworkers over-interpreted faint test lines which could be recognized in daylight but not in the indoor laboratory setting.

Previously, Kalluri et al. assessed the impact of the reader's qualification on the performance of the prototype test [12]. Laboratory technicians with several years of working experience as well as field workers with at least a college degree but no laboratory experience were asked to perform the test on 304 stool samples. The reported RDT sensitivities of 94% and 93% when done by laboratory technicians and field workers, respectively, were similar, but RDT specificity was higher when performed by the technicians (76% versus 67%) [12]. Harris et al. report a sensitivity of 97% and a specificity of 71–76%, when staff with graduate-level laboratory training performed the test in Guinea Bissau [10]. Mukherjee et al. reported a similar sensitivity and specificity (92% and 73%, respectively) when the test was done by graduate-level staff during a surveillance study at a hospital in Kolkata [11].

In contrast to Wang et al., we did not find a higher sensitivity of the RDT when testing stool samples from patients receiving IVF compared to samples from patients who did not receive IVF [13]. However, we noted that the PPV in this study varied according to the proportion of culture-positive specimens. It has been argued that PPV is the most important measure of a clinical diagnostic method since it represents the proportion of patients with positive test results that are correctly diagnosed [18]. The PPV is not intrinsic to the test; it is affected by prevalence of the disease. For example, the PPV was 55% for samples from patients given IVF (43% of whom had *V. cholerae* isolated) while it was 27% for samples from patients managed with ORS (15% of whom had *V. cholerae* isolated). The PPV was 71.9% for the Unguja sub-sample with 58% cholera confirmation while for the Pemba sub-sample it was 43% with 29% cholera confirmation. In outbreak settings, when a large proportion of patients presenting with acute watery diarrhea have cholera, a positive RDT result would have a good predictive value. In other situations (e.g. areas with seasonal cholera but also high rates of diarrheal diseases from other pathogens), the RDT may be less useful.

Our study has several limitations. Firstly, a large sample from 622 study participants was available for the overall evaluation, but

only 67 stool samples were used for sub-analyses. Secondly, while confirmation of *V. cholerae* isolates was performed at a reference laboratory, culture-negative stool samples were not validated further. In particular, we did not perform PCR testing on our RDT-positive, culture-negative samples. Bhuiyan, et al. [14] analyzed five stool samples collected in Bangladesh by multiplex PCR that were O1 dipstick positive but culture-negative and found that all five were negative by PCR, indicating that the five dipstick-positive results were false positives. This is reassuring but does not entirely exclude the possibility of false negativity by stool culture. Thirdly, Alam et al. pointed out that the dipstick may detect non-culturable forms of *V. cholerae* that have transformed into a coccoid form due to unfavorable intra-host conditions, such as antibiotic treatment prior to testing [7]. We tried to assess the influence antibiotic treatment prior to sample collection may have had on our results but found no significant difference ($p>0.05$) in the false positive rates among participants who had taken antibiotics prior to sample collection and those who had not. However, further research is needed to rule out the possibility that the RDT may detect *V. cholerae* antigen in some specimens which are culture negative.

Conclusion

We found that field workers in this study who had basic general education but were not familiar with laboratory work experienced difficulties in interpreting the RDT performed in the cholera camps. If the RDT is to be deployed more widely, more extensive and repeated training may be required to improve the current RDT's specificity. The test cannot replace stool culture and due to the high number of false positive results observed is not suitable to trigger an outbreak response in a resource poor setting. However the test may be potentially used as a screening tool. During cholera outbreaks, especially when several samples test positive, the test has an enhanced predictive value. Further research is needed to evaluate the accuracy of the RDT with specimens which have been incubated in APW for 4 to 6 hours prior to testing in the RDT since this procedure should dilute out the materials in stool samples which are causing the false positive results while amplifying the antigen signal from the *V. cholerae*.

Acknowledgment

We would like to thank Marie Kieny and Raymond Hutubessy from the World Health Organization who have supported this project throughout the entire study period.

We are also grateful to all patients who made this work possible. We thank all of the technical staff and research assistants who were involved in the study. This project was coordinated by the WHO Initiative for Vaccine Research, Geneva, Switzerland and the International Vaccine Institute, Seoul, Korea.

Author Contributions

Conceived and designed the experiments: LVS JD DS. Performed the experiments: BL KT WS CJB. Analyzed the data: BL KT LVS JD NYC RH DS. Contributed reagents/materials/analysis tools: AM SMA. Wrote the paper: BL KT. Provided significant scientific support: AMK. Provided scientific support: LVS JD AM TW JC HW GE TA MSJ DS SMA. Collected data: RR.

References

1. World Health Organization (2011) Number of Cholera Cases. Available: http://www.who.int/gho/epidemic_diseases/cholera/situation_trends_cases/en/index.html. Accessed 15 August 2011.
2. World Health Organization (2011) Cholera: Fact sheet No. 107. Available: <http://www.who.int/mediacentre/factsheets/fs107/en/>. Accessed 09. November 2011.

3. Center for Disease Control (2010) Cholera. Available: <http://www.cdc.gov/cholera/disease.html>. Accessed 15 August 2011.
4. Global Task Force on Cholera Control (2004) Case management: treatment. In: Cholera Outbreak - Assessing The Outbreak Response And Improving Preparedness. Geneva: World Health Organization. pp 26–30.
5. Chaignat C-L (2006) Cholera outbreak response in Pemba, 2006 And Pertinence of using Oral Cholera Vaccines for Control. Available from WHO Homepage: <http://www.who.int/cholera/technical/CholeraZanzibarReviewJuly2006.pdf>. Accessed 15 August 2011.
6. Reyburn R, Deen JL, Grais RF, Bhattacharya SK, Sur D, et al. (2011) The case for Reactive Mass Oral Cholera Vaccinations. *PLoS Negl Trop Dis* 5(1): e952. doi:10.1371/journal.pntd.0000952.
7. Alam M, Hasan NA, Sultana M, Nair GB, Sadique A, et al. (2010) Diagnostic Limitations to Accurate Diagnosis of Cholera. *Journal of Clinical Microbiology* 48(11): 3918–3922. doi:10.1128/JCM.00616-10.
8. Riewpaiboon A, Intrapakan K, Phoungkatesunthorn S (2008) Predicting Treatment Cost for Bacterial Diarrhoea at a Regional Hospital in Thailand. *J Health Popul Nutr* 26(4): 442–450.
9. The TDR Diagnostics Evaluation Expert Panel (2010) Evaluation of diagnostic tests for infectious diseases: general principles. *Nature Reviews – Microbiology*. pp 17–29. doi: 10.1038/nrmicro1523.
10. Harris JR, Cavallaro EC, de Nobrega AA, dos S Barrado JC, Bopp C, et al. (2009) Field Evaluation of Crystal VC Rapid Dipstick test for cholera during a cholera outbreak in Guinea - Bissau. *Tropical Medicine and International Health* 14(9): 1117–1121. doi:10.1111/j.1365-3156.2009.02335.x.
11. Mukherjee P, Ghosh S, Ramamurthy T, Bhattacharya MK, Nandy RK, et al. (2010) Evaluation of a Rapid Immunochromatographic Dipstick Kit for Diagnosis of Cholera Emphasizes Its Outbreak Utility. *Jpn J Infect Dis* 63(4): 234–238.
12. Kalluri P, Naheed A, Rahman S, Ansaruzzaman M, Faruque AS, et al. (2006) Evaluation of three rapid diagnostic tests for cholera: does the skill level of the technician matter? *Tropical Medicine and International Health* 11(1): 49–55. doi:10.1111/j.1365-3156.2005.01539.x.
13. Wang X-Y, Ansaruzzaman M, Vaz R, Mondlane C, Lucas ME, et al. (2006) Field evaluation of a rapid immunochromatographic dipstick test for the diagnosis of cholera in a high-risk population. *BMC Infectious Diseases* 6(17), doi:10.1186/1471-2334-6-17.
14. Bhuiyan NA, Qadri F, Faruque AS, Malek MA, Salam MA, et al. (2003) Use of Dipsticks for Rapid Diagnosis of Cholera Caused by *Vibrio cholerae* O1 and O139 from Rectal Swabs. *Journal of Clinical Microbiology* 41(8): 3939–3941. doi: 10.1128/JCM.41.8.3939–3941.2003.
15. Nato F, Boutonnier A, Rajerison M, Grosjean P, Darteville S, et al. (2003) One-Step Immunochromatographic Dipstick Tests for Rapid Detection of *Vibrio cholerae* O1 and O139 in Stool Samples. *Clinical And Diagnostic Laboratory Immunology* 10(3): 476–478. doi: 10.1128/CDLI.10.3.476–478.2003.
16. National Bureau of Statistics, Ministry of Planning, Economy and Empowerment (2006) Analytical Report, Volume X, p2. Available: <http://www.nbs.go.tz/pdf/2002popcensus.pdf> Accessed on 29 August 2011.
17. Bopp C, Ries AA, Wells JG, Koplan JP, Hughes JM, et al. (1999) Laboratory Methods for the Diagnosis of Epidemic Dysentery and Cholera. Geneva. Available: <http://www.cdc.gov/ncidod/dbmd/diseaseinfo/cholera/complete.pdf>. Accessed 29 August 2011.
18. Altman DG, Bland JM (1994) Diagnostic tests 2: Predictive values. *BMJ* 309(6947): 102.