Diagnostic Performance of a Rapid Syphilis Test Among Pregnant Women in Lima, Peru

by

Kristen M. Roehl

Duke Global Health Institute
Duke University

Date: _____

Approved:.

_____
Brian Pence, Supervisor

_____
David Boyd

_____
David Walmer

Thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science in the Department of Global Health
in the Graduate School of Duke University

2011

ABSTRACT

Diagnostic Performance of a Rapid Syphilis Test
Among Pregnant Women in Lima, Peru

by

Kristen M. Roehl

Duke Global Health Institute
Duke University

Date: _____

Approved:

_____
Brian Pence, Supervisor

_____
David Boyd

_____
David Walmer

An abstract of a thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science in the Department of Global Health
in the Graduate School of Duke University

2011

ABSTRACT

Background: Maternal and congenital syphilis are pressing concerns in Latin America, with consequences ranging from newborn mental retardation to perinatal death. Widespread, accurate screening and timely penicillin treatment can help. Simple, affordable, point-of-care rapid syphilis tests (RSTs) promise to improve screening coverage among pregnant women.

Methods: From September 2009 to November 2010, Project CISNE implemented the SD Bioline Syphilis 3.0 RST into two health networks, offering the test to pregnant women aged 16-55 who attended antenatal care, delivery/postpartum, and abortion services. The performance analysis compared Bioline RST results with reference standards TPPA and RPR+TPPA, adjusting estimates according to sampling realities.

Results: 17,147 rapid syphilis tests were performed in the field and 11,169 were screened in the central laboratory. Syphilis prevalence was 1.05% (0.73% adjusted) according to the gold standard vs. 0.90% according to the field RST. The Bioline RST displayed an unadjusted sensitivity of 91.0% (95% CI 86.4-95.0) and specificity of 99.1% (98.1-99.6) compared to TPPA, and an unadjusted sensitivity of 91.5% (84.8-95.8) and specificity of 99.6% (99.4-99.7) compared to RPR+TPPA. When adjusted, overall sensitivity and specificity compared to RPR+TPPA were 86.5% (78.8-92.0) and 99.7% (99.6-99.8), respectively. The Bioline RST yielded more false-positive than false-negative results due to the observed low prevalence.

Discussion: Despite limitations, this study displays the field RST to be reliable, reproducible, as valid as previous studies, and diagnostically apt for implementation in maternal care services in Peru.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION & BACKGROUND

In December 2010, the Pan American Health Organization (PAHO) awarded its sixth Latin American Competition in Health Journalism *(Concurso Latinoamericano de Periodismo en Salud*) to three illuminators of the region's continuing challenges to eliminating congenital syphilis [1]. Indeed, congenital syphilis is a pressing concern in Latin America's current health landscape. According to PAHO's Dr. Fernando Gonzales, 164,000 infants were born with congenital syphilis in 2007 alone, representing 25 times the number born with HIV in the same year [2]. Despite this disparity, syphilis has not triggered an equivalent degree of institutional response in Latin America as HIV [3]. Journalists and health officials from around Latin America have finally sounded the alarm.

Perhaps the gravest consequence of untreated syphilis infection during pregnancy, congenital syphilis ranges in incidence from 1.4 to 12 per 1000 live births in Latin America[1] [3], compared to 0.01 case per 1000 live births in the United States (2008) [4]. Among other symptoms, the condition causes mental retardation, deafness, blindness, growth retardation, and skeletal deformities in newborns [5]. Furthermore, untreated maternal syphilis limits the potential of pregnancies *in utero*, causing spontaneous abortion, stillbirth, low birth-weight, preterm labor, and perinatal death [6]. In general, an estimated one-third of untreated pregnancies end in fetal loss, another one-third in the birth of a child with congenital syphilis, and the remaining one-third in the birth of a healthy child [3].

Fortunately, these adverse events can be prevented with widespread, accurate screening and timely penicillin treatment. Such efforts are especially indicated in areas with high incidence of infection. On the other hand, the discontinuity between syphilis burden and diagnostic capacity (as well as public health priority) in many of

[1] Range is from 1.40 per 1000 live births in El Salvador to 12 per 1000 live births in Honduras, according to 2002 data.

these regions has often stalled efforts to meet syphilis control guidelines that call for more widespread screening and treatment [7]. For instance, an estimated 330,000 infected pregnant women in Latin America were not treated for syphilis during prenatal monitoring in 2004 alone [3, 8]. The number of undetected cases of maternal syphilis depends primarily on the lack of screening and secondarily on prevalence.

Since syphilis screening is one of the more highly cost-effective and immediately useful health interventions in resource-poor settings [9], universal screening for pregnant women during prenatal care visits is now standard policy in several Latin-American countries [10, 11]. However, factors such as difficult access to and poor organization of maternal care and laboratory screening services; lack of infrastructure, basic equipment, and trained personnel; low demand for the test among users; and the stigma associated with sexually transmitted infections (STIs) have resulted in suboptimal screening coverage for pregnant women [3].

According to data submitted to PAHO by nationwide programs against STIs, HIV, and AIDS (2002), the overall prevalence of syphilis among pregnant women in Latin America is approximately 3.1%, with country-specific screening coverage ranging from 30% to 80% [3]. PAHO's Regional Initiative for the Elimination of Mother-to-Child Transmission of HIV and Congenital Syphilis in Latin America and the Caribbean aims to increase this coverage to 95% or more throughout the region by the year 2015 [12]. In Peru, the Ministry of Health (MINSA) requires syphilis screening for all pregnant women attending antenatal care (upon first visit), delivery, or abortion services [13], yet actual coverage is currently unknown. Reports from the Maternal-Perinatal National Institute (1995), the General Directorate of Epidemiology (1996-2000), and the STI/AIDS control program PROCETSS (2000-2004) have claimed a syphilis prevalence of 1.6% among screened pregnant women and an average 4.9 cases of congenital syphilis per 1000 live births annually [13].

Fortunately, these rates are below the regional average, indicating a lesser extent of undetected disease due to deficiencies in screening in Peru.

Despite this silver lining, realizing the PAHO target of under 0.5 cases of congenital syphilis per 1000 live births [14] requires improvement in screening coverage among pregnant women. The simple, affordable, point-of-care rapid syphilis tests (RSTs) that can be performed by personnel with minimal training currently being incorporated into MINSA guidelines promise to work toward this goal. However, test availability and government recommendations mean little without concrete proof of diagnostic accuracy. In order to ensure the reliability of the RSTs poised for implementation throughout the country, therefore, it is essential to analyze their performance in comparison to current gold standards in the international community and in clinical practice in Peru.

1.1 Epidemiology of Syphilis

Syphilis is a systemic disease caused by the spirochete *Treponema pallidum*, with an average incubation period of 21 days (range: 10-90). The course of the disease is divided into four stages: primary, secondary, latent, and tertiary. Primary syphilis is characterized by the appearance of an initial ulcer (chancre) lasting two to six weeks. Early symptoms of infection are difficult to recognize because the chancre is generally painless and inconspicuous. Unfortunately, like HIV, the principal mechanisms of syphilis transmission are sexual and blood contact during this inconspicuous early stage, complicating efforts to reduce the spread of infection. Secondary syphilis manifestations include skin rash, condylomatalata, and mucocutaneous lesions. When secondary syphilis is untreated, it either enters latency or progresses into a more serious condition. Approximately one-third of people with secondary syphilis develop tertiary syphilis, causing complications of the heart, joints,

brain, and nervous system. Treatment depends upon the stage of disease, but generally involves repeated doses of penicillin [3, 15]. Fortunately, penicillin is a relatively inexpensive and available medication in Latin America. In Peru, the treatment regimen for maternal syphilis is three subcutaneous injections of penicillin, begun immediately upon detection of an active case and spaced evenly over 60 days [13].

1.2 Laboratory Detection of Syphilis

Like treatment, detection of syphilis varies depending upon stage of disease and screening method. In general, laboratory tests for syphilis fall into four categories: 1) direct microscopic examination, used when lesions are present, 2) nontreponemal tests, used for primary screening, 3) treponemal tests, used for confirmatory screening, and 4) direct antigen detection tests, often used as gold standards for research. Women who present without visible symptoms in Peru are screened via a combination of nontreponemal and treponemal methods.

Nontreponemal tests utilize VDRL-derived antigens to detect *T. pallidum* antibodies [16]. The most commonly used are the Venereal Disease Research Laboratory (VDRL) test and the Rapid Plasma Reagin (RPR) test, which exists in two versions: a teardrop card test using whole blood in the field and a circle card test using serum in the laboratory. Although the teardrop RPR is increasingly being replaced by more sensitive and user-friendly immunochromatographic strip (ICS) RSTs, the circle RPR[2] is still widely used for primary syphilis detection [17]. In real-world practice, positive identification of *T. pallidum* antibodies using human serum or plasma by a nontreponemal test is a presumptive diagnosis for syphilis [16].

[2] BioMérieux's RPR-nosticon II slide flocculation test uses a modified VDRL antigen that contains microparticulate carbon, which enhances the visual difference between a positive and a negative result for the qualitative and semi-quantitative determination of antibodies to VDRL antigen.

While qualitative nontreponemal tests use undiluted serum to measure the presence or absence of antibodies, their quantitative counterparts qualify the diagnosis using serial two-fold dilutions, reporting the highest reactive dilution. Nontreponemal tests like RPR are generally high in sensitivity and low in specificity for secondary and latent syphilis. However, they may lack sensitivity (in addition to specificity) for primary and tertiary syphilis, when antibodies are not as readily present [17]. RPR has other disadvantages, as well. Nonspecific prozone (false-negative) reactions can also occur, usually due to host tissue damage from infection, immunization, or autoimmune disease [17].

Because of the risk for false results, treponemal tests are often used for confirmation of nontreponemal test results. Treponemal tests vary from UV techniques (FTA-Abs) to hemagglutination techniques (TPPA/TPHA), but are distinct from nontreponemal tests in that they use suspensions of inactive treponemes as conjugates for *T. pallidum* antibodies. Although used primarily to confirm reactive RPR, they are also used to verify cases in which clinical evidence contradicts a nonreactive nontreponemal test, as might occur in tertiary syphilis. The principal disadvantage of treponemal tests, aside from their higher cost, is that previously treated cases remain reactive for years, if not a lifetime. In a different way than with nontreponemal tests, therefore, false-positive results also occur with treponemal tests. Indeed, if a treponemal test is used as a primary screening method, about 1% of the general population can be expected to display false-positive results [17]. The advantages and disadvantages of nontreponemal and treponemal tests are summarized in Table 1 [17, 18].

In order to more accurately screen out false-positive results, the gold standard for both clinical practice and research is generally a combination of primary nontreponemal and secondary treponemal tests, with a positive result defined as positive for *both* tests.

Table 1: Advantages and Disadvantages of Common Laboratory Tests

Test Type	Examples	Sensitivity (RPR)	Specificity (TPPA)	Advantages	Disadvantages
Nontre-ponemal	RPR, VDRL, ELISA	86-100%	85-100%	▪ Processed rapidly (<45 minutes) [12] ▪ Require minimal equipment	▪ Risk for false-positive results ▪ Results take too long in real-world practice ▪ Reagents require refrigeration
Trepo-nemal	FTA-ABs, TPPA, TPHA	93-98%	98-100%	▪ High specificity makes tests well-suited for secondary confirmation	▪ More difficult and costly to perform ▪ Cannot be used to monitor treatment ▪ Remain reactive when disease has been treated

1.3 Field Detection of Syphilis (Rapid Syphilis Tests)

In developing settings, the principal disadvantage of laboratory tests is precisely that they require a laboratory, limiting screening access in remote locations and, for those with screening access, inserting a gap between screening and treatment. Due to their different yet complementary strengths in diagnostic performance, the combination of RPR and a treponemal test – Treponema pallidum particle agglutination assay (TPPA) or fluorescent treponemal antibody absorption (FTA-ABs) – has become the gold standard for syphilis diagnosis [11]. However, although the RPR test is a procedure requiring minimal equipment and few steps, it usually takes anywhere from a few hours to a few days to obtain results in clinical practice. In response to this reality, scientists have formulated over 20 available Rapid Syphilis Tests (RST) [18], several of which have been reviewed by the UNICEF/UNDP/World Bank/WHO Sexually Transmitted Diseases Diagnostics Initiative (SDI) and tested in countries such as Bangladesh, Bolivia, Brazil, China, Gambia, Haiti, Italy, Mexico, Peru, Tanzania, South Africa, and Vietnam [7, 11].

Stored in normal conditions and used at the point of care, RSTs have the potential to increase the availability and efficiency of screening and more widely prevent the adverse effects of untreated syphilis during pregnancy. In contrast to the nontreponemal and nonspecific RPR test, rapid tests use recombinant treponemal antigens to specifically detect antibodies in whole blood, serum, or plasma samples. Most use ICSs coated with antigens of *T. pallidum* and pre-packaged in cassettes, as shown in Figure 1 below. Antigen-antibody reactions (positive results) appear as colored lines or spots on the membrane [18]. Advantages of rapid tests are their affordability (US$ 0.19-3.00 per test), simplicity (3-4 steps), quick results (20-minute turnaround), robustness (8-30°C), independence (no external equipment needed), and deliverability [19]. RSTs also eliminate the false-positive results and prozone effect of the RPR test. However, true to their treponemal nature, they cannot differentiate between active and previously treated infections, and therefore exhibit different false-positive results. Because of this, they are better suited to screening in low-risk populations such as pregnant women than to screening in the general population. On a practical level, RSTs allow for screening and treatment in the same appointment, eliminating the time gap between the two and – in a research context – avoiding losses to follow-up [11].

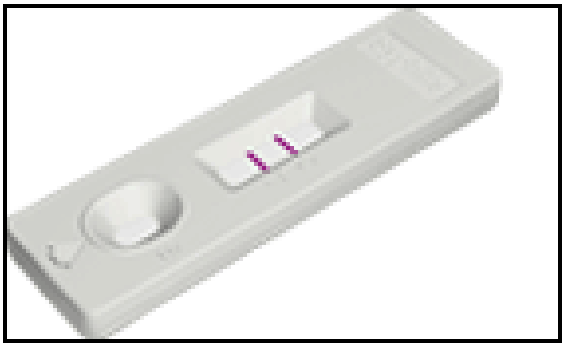

Figure 1: Positive Bioline Rapid Syphilis Test

Several RSTs can be used with whole blood (via fingerstick) and without external equipment at the point of care [18]. In 2004, SDI developed criteria to determine the suitability of marketed RSTs for increasing screening coverage in developing settings. Now known as the ASSURED criteria, SDI stated that tests should be:

- Affordable,
- Sensitive,
- Specific,
- User-friendly (simple to perform with minimal training),
- Rapid and robust (results available in less than 30 minutes),
- Equipment-free, and
- Deliverable to developing countries [18-20].

Among these criteria are the two principal indicators of diagnostic performance: sensitivity and specificity. According to real-world applicability and SDI guidance, therefore, research into rapid test implementation must include diagnostic performance as a central focus, ideally testing RSTs using whole blood in the field rather than serum in the laboratory. For research to align with efforts to increase screening and treatment among pregnant women, it should focus specifically on pregnant women in appropriate healthcare sites and use substantial enough sample sizes to yield performance results with satisfactory precision.

1.4 Diagnostic Performance of Rapid Syphilis Tests

In 2006, SDI evaluated nine RSTs against the ASSURED criteria in urban field sites in Russia, USA (Alabama), Haiti, China, Tanzania, Sri Lanka, South Africa, and Gambia. Overall, the tests displayed sensitivities of 85-99% and specificities of 93-100%, with Abbott Labs' Determine Syphilis TP, Fujirebio Diagnostics' Espiline TP, Biotech Inc's Syphilis On Site Rapid CTK, and Standard Diagnostics' SD Bioline 3.0 earning the highest summed (*z*) performance scores [19, 21]. The SDI evaluation provided an interesting comparison of the performance of each test under "optimum" conditions, but its use of reconstituted serum from archived samples rather than whole blood limited its clinical relevance.

Indeed, largely because RSTs are a relatively new set of marketed diagnostics, not much large-scale research has tested them in their intended field settings. A 2010 review of ICS rapid tests discovered that 8 studies of 823 citations had taken place within the context of antenatal care, and only 7 overall had used whole blood [7]. Only one study in Bolivia (2006) satisfied the major requirements for congenital syphilis prevention, using whole blood in the field, focusing on pregnant women, *and* using a substantial sample size [22].

Along with a recent study in the Brazilian Amazon and a relevant study in Peru, Table 2 shows the performance of the foremost RSTs during these whole-blood field studies. To date, the Bolivian study is the most robust, relevant research published in Latin America – indeed, in the developing world overall. A large-scale study in Peru is both timely and necessary.

Table 2: Performance of Whole-Blood RSTs in Developing Settings [7, 22-32]

Site (Year)	Study Population	Type of RST	Reference Standard	N	Syphilis Prevalence	% Sensitivity (95%CI)	% Specificity (95%CI)
Brazil (2010)	Remote (Amazon) antenatal clinics	Visitect	VDRL + FTA-Abs	712	2.2%	62.5 (39-82)	99.1 (98-100)
Bolivia (2009)	Urban antenatal clinic	Determine	RPR + FTA-Abs	489	4.5%	98.0	99.8
Brazil (2008)	Urban red-light district STI clinics	Visitect	VDRL + FTA-Abs	506	18.0%	57.0 (46-67)	99.0 (97-100)
Bangladesh (2008)	Urban female sex workers	Syphilis On Site Rapid	RPR + TPHA	684	20.8%	94.4	92.6
South Africa (2007)	Rural antenatal clinics	Determine	RPR + TPHA	341	6.5%	85.7 (57-98)	90.9 (87-94)
Mexico (2006)	Urban antenatal clinics	Determine	VDRL + FTA-Abs	1322	0.3%	100.0	100.0
Mozambique (2006)	Rural antenatal clinics	SD Bioline 3.0	RPR + TPHA	326	8.4%	86.0 (82-89)	96.8 (96-97)
Bolivia (2006)	Urban maternity hospitals	Determine	RPR + TPPA	8892	3.9%	91.8 (88-95)	98.5 (98-99)
Tanzania (2006)	Large government maternity hospital	Determine Visitect Qualpro SD Bioline 3.0	TPPA, TPHA, etc.	528 528 582 582	10.8% 10.8% 9.5% 11.3%	59.6 (47-72) 75.0 (64-86) 78.6 (68-89) 85.7 (77-94)	99.4 (99-100) 99.8 (99-100) 99.1 (98-100) 98.1 (97-99)

Table 2: Performance of Whole-Blood RSTs in Developing Settings (Continued)

Site (Year)	Study Population	Type of RST	Reference Standard	N	Syphilis Prevalence	% Sensitivity (95%CI)	% Specificity (95%CI)
Brazil (2006)	Urban STI clinics	Determine	TPPA, TPHA, etc.	247	21.1%	88.5 (80-97)	97.9 (96-100)
		Visitect		244	20.9%	96.1 (91-101)	98.5 (97-100)
		Qualpro		542	9.2%	84.3 (74-94)	99.6 (99-100)
		SD Bioline 3.0		542	9.2%	88.2 (79-97)	99.4 (99-100)
China (2006)	Urban STI clinics	Determine	TPPA, TPHA, etc.	445	18.7%	81.9 (74-90)	99.4 (99-100)
		Visitect		445	18.7%	73.5 (64-83)	99.7 (99-100)
		Qualpro		415	21.5%	64.0 (54-74)	99.7 (99-100)
		SD Bioline 3.0		415	21.5%	87.6 (81-94)	99.4 (99-100)
Haiti (2006)	Urban STI clinics	Determine	TPPA, TPHA, etc.	761	5.3%	72.5 (59-86)	98.5 (98-99)
		Visitect		516	10.7%	72.7 (61-85)	99.1 (98-100)
		Qualpro		543	7.6%	80.5 (68-93)	97.8 (97-99)
		SD Bioline 3.0		515	5.8%	100 (n=30)	98.3 (97-100)
Peru (2006)	Urban female sex workers	Determine	RPR + TPHA	3586	5.1%	39.3	99.2

1.5 Study Goals

Termed Project CISNE,[3] this study was conducted by researchers at the School of Public Health of *Universidad Peruana Cayetano Heredia* in cooperation with the World Health Organization, the London School of Hygiene and Tropical Medicine, and TDR[4] [33] in order to contribute to RST research in a large-scale, implementable way. Project CISNE introduced SD Bioline Syphilis 3.0 (Standard Diagnostics, Gyeonggi-do, Korea) into maternity services in Lima in order to evaluate the validity and reliability of its performance in primary care settings and in a large urban tertiary care facility. In light of current efforts to incorporate RSTs into MINSA guidelines in Peru, this study is expected to have an immediate in-country impact and provide the basis for further RST studies in Latin America.

[3] Meaning "swan" in Spanish, "CISNE" stands for "Cura Inmediata de la Sifilis Neonatal" (Immediate Cure for Neonatal Syphilis) and metaphorically represents the transformation of a previously at-risk pregnancy/infant into a healthy one with adequate screening and treatment.
[4] TDR, a Special Program for Research and Training in Tropical Diseases, is a scientific collaboration that helps "coordinate, support and influence global efforts to combat a portfolio of major diseases of the poor and disadvantaged." TDR is executed by the World Health Organization (WHO) and sponsored by the United Nations Children's Fund (UNICEF), the United Nations Development Programme (UNDP), the World Bank, and WHO.

2.0 METHODS

Over a fourteen-month period of implementation (Sep 2009 to Nov 2010), Project CISNE introduced the Bioline RST into preexisting health services in two health networks – the urban Maternal Perinatal National Institute (INMP, acronym in Spanish) in Lima and 16 health centers in Ventanilla-Callao[1] – in order to evaluate its impact on screening coverage and treatment efficiency. The diagnostic performance analysis utilized data collected over ten months during this implementation period.

2.1 Study Population

Pregnant women aged 16-55 who attended antenatal care, delivery/postpartum, and abortion services in the intervened health networks were offered the RST.

2.2 Procedures

Project CISNE involved tests at the point of care and in the central laboratory at *Universidad Peruana Cayetano Heredia* (UPCH)**.** Within preexisting care processes, women receiving the rapid syphilis test also provided venous blood samples for laboratory testing and the diagnostic performance analysis. Figure 2 displays the overall process of sample collection, laboratory testing, and data analysis for the diagnostic performance analysis. As shown, a pregnant woman attending an intervened health center consented to participate in the study while receiving her regular health services.

[1] Characterized by industrial areas and *pueblos jovenes* (shantytowns), Ventanilla is the largest district of the province of Callao, located on the northern outskirts of the city of Lima.

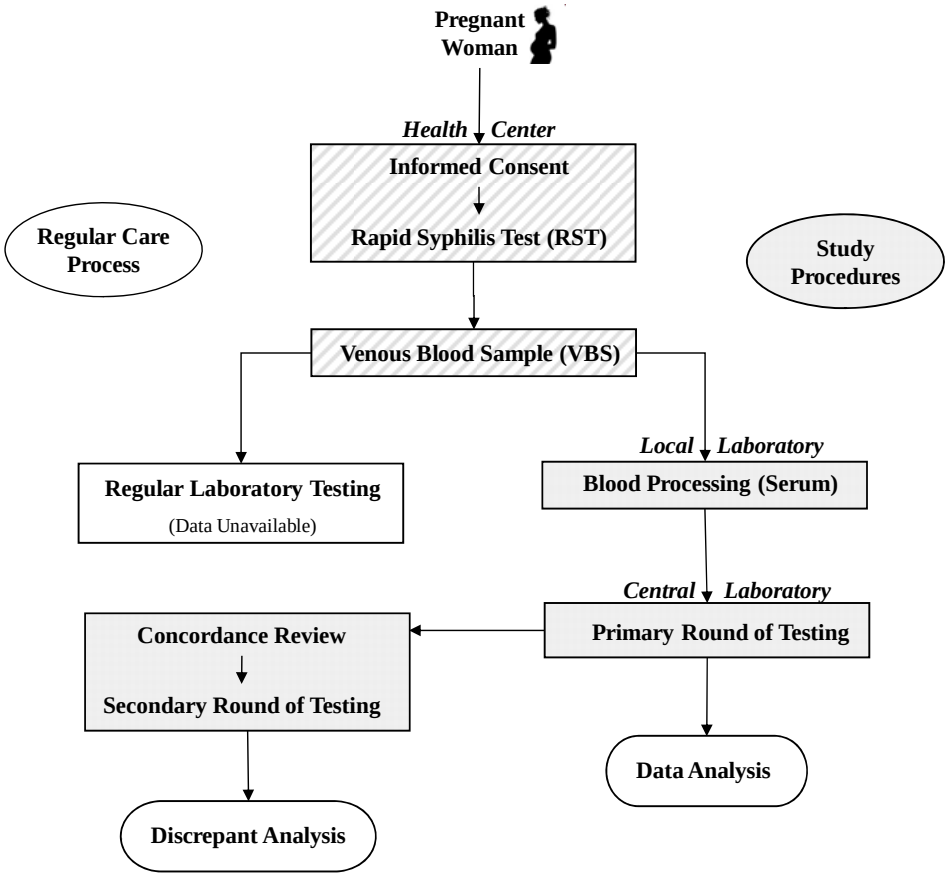

Figure 2: Project CISNE Overall Process

2.2.1 Sample Collection and Storage

Due to practical limitations, not every woman receiving the rapid test provided a venous blood sample (VBS). During sample collection, efforts were made to respect patient comfort and wishes. Therefore, some women were not asked to endure multiple needle insertions or blood draws unnecessarily. For example, women testing negative by RST in the delivery/postpartum service at INMP did not provide venous blood. As part of INMP's single-visit care process, all other women provided venous blood while waiting for their RST results. In Ventanilla-Callao, the regular care

process directed women to the local laboratory to provide venous blood samples for several pregnancy-related tests, including the test for syphilis.

With RST implementation, Ventanilla-Callao health providers decided to immediately draw venous blood from any woman testing positive for syphilis at the point of care, but doing so for all women was logistically impossible. RST-negative women were therefore directed to the laboratory; approximately 27% did not visit the laboratory and were lost to follow-up for the purposes of this analysis. Table 3 summarizes the circumstances under which women attending each health network/service provided fingerstick and venous blood samples in the field.

Table 3: Women Providing Fingerstick and Venous Blood Samples in the Field

Health Network	Health Service	Fingerstick (RST)	RST Result	Venous Blood (Laboratory Tests)
INMP	Antenatal Care (first ANC visit)	All	+	All
			-	All
	Delivery/Postpartum	All	+	All
			-	None
	Abortion Services	All	+	All
			-	All
Ventanilla-Callao	Antenatal Care (first ANC visit)	All	+	All
			-	All women who visited the laboratory (approx. 72%)
	Delivery/Postpartum	All	+	All
			-	All women who visited the laboratory (approx. 73%)
	Abortion Services	All	+	All
			-	All women who visited the laboratory (approx. 89%)

Trained health providers obtained informed consent, administered the Bioline RST, collected venous blood samples for laboratory testing, and recorded results. Samples were labeled with pre-printed patient codes and sent to the local health laboratory within 4 hours of collection for storage at 2-8˚C. To process the samples, on-site laboratory personnel centrifuged the tubes of venous blood into cryovials and

labeled the cryovials with appropriate date, time, and patient code. Cryovials were stored at -20˚C until weekly collection by UPCH central laboratory staff and transported in insulated containers containing dry ice.

2.2.2 Central Laboratory Testing

Upon arrival to the UPCH central laboratory, serum cryovials were again stored at -20˚C until time of testing. Two laboratory technicians blind to the field results processed samples within seven days of arrival. In the primary round of testing, all samples were screened by RPR-Nosticon II (bioMérieux, Boxtel, The Netherlands). RPR-positive samples and a random sample of 700 RPR-negative samples were then tested by Serodia-TPPA (Fujirebio Diagnostics, Malvern, PA, USA) and RST for confirmation, comparison, and quality control. RPR-negative samples beyond the random sample were not tested by TPPA. Laboratory technicians recorded results and periodically delivered updated databases to the laboratory supervisor to use for the performance analysis.

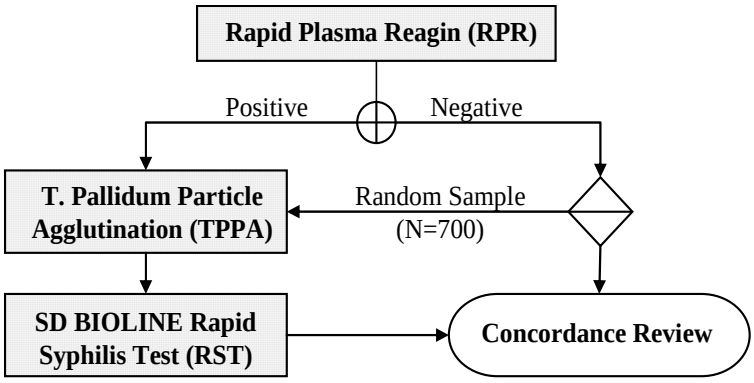

Figure 3: Primary Round of Laboratory Testing

2.2.3 Data Analysis

The analysis of RST performance utilized the field and laboratory test results determined through the process detailed on the preceding page. Two phases of analysis with different definitions of the gold standard enabled both technical and practical (clinically-relevant) comparisons.

The first phase of analysis compared the field RST with a gold standard of TPPA only. This allowed for a direct technical comparison of the two treponemal tests *in theory*. In addition, utilization of the TPPA gold standard aligned the assessment of RST performance with recent WHO measures. The second phase of analysis measured RST against RPR with TPPA confirmation (RPR+TPPA), the method currently used as the gold standard *in practice* in Peru. Along with providing more accurate detection of true-positive and true-negative results, utilization of the RPR+TPPA gold standard served the end goal of implementing rapid syphilis testing without compromising diagnostic quality in the real world.

For each gold standard, the crude sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the field RST were calculated. However, deficiencies of venous blood sample collection were considered in order to yield accurate results. Table 4 summarizes the percentage of RST-positive and RST-negative women providing venous blood samples in each service and consequent adjustments performed in the analysis.

As the table shows, a significant percentage (26.8%) of RST-negative women in Ventanilla-Callao did not contribute venous blood samples. This reality presented two obstacles to analysis: (1) RST-negative women were underrepresented in the total number of laboratory samples used to assess true disease status and RST performance.

(2) Since the subset of women providing a venous blood sample may represent the healthiest, most persistent, or least "at risk" women, the number of RST-negative women testing positive with the gold standard (false negatives) may be lower in the study than is true of the population as a whole. Despite the risk of bias, this analysis assumed that women without a venous blood sample displayed the same percentage of true-positive, true-negative, false-positive, and false-negative results as those with a blood sample.

Table 4: Summary of Analysis

Health Network	Health Service	RST Result	% VBS	Risk/Bias	Action
INMP	Antenatal Care	+	100.0	None	Adjusted the weight of samples by dividing by "% VBS"
		-	98.9		
	Delivery/ Postpartum	+	100.0	Venous blood samples were not collected from RST-negative women, so impossible to calculate sensitivity and specificity	Calculated only PPV and excluded samples from sensitivity and specificity estimates
		-	N/A		
	Abortion Services	+	95.2	Small percentage of RST-negative women refused venous blood sample, so potential upward bias of sensitivity and specificity	Noted risk and adjusted the weight of samples by dividing by "% VBS"
		-	88.6		
Ventanilla-Callao	Antenatal Care	+	100.0	RST-negative women underrepresented in laboratory results and potential upward bias of sensitivity	Adjusted the weight of samples by dividing by "% VBS"
		-	71.5		
	Delivery/ Postpartum	+	100.0	RST-negative women underrepresented in laboratory results and potential upward bias of sensitivity	Adjusted the weight of samples by dividing by "% VBS"
		-	73.0		
	Abortion Services	+	100.0	RST-negative women underrepresented in laboratory results and potential upward bias of sensitivity	Adjusted the weight of samples by dividing by "% VBS"
		-	88.5		

This assumption allowed for a weight adjustment of the data from RST-negative women in Ventanilla-Callao, yielding a more accurate depiction of rapid test performance. In order to limit bias in the adjusted performance results, the individual contribution of samples from each health service (2 networks *x* 3 services in each health network *x* 2 results in each service) was weighted using the reciprocal of the relevant "%VBS" noted in the table.

2.2.4 Discrepant Analysis

In addition to the primary round, the study procedure involved a concordance review and secondary round of laboratory testing for the purposes of quality control. This procedure is often termed *discrepant analysis* and has been shown to upwardly bias sensitivity estimates [34]. Therefore, although part of the original study design, this performance analysis disregards *final* results for its principal findings, instead utilizing results from the primary round of testing. The results and implications of the concordance review are briefly discussed in Section 3.5. Since the concordance review was a key ingredient of the larger study design, it is detailed here.

Upon receiving current databases of field and laboratory results, the laboratory supervisor conducted a concordance review to assess the quality and consistency of results. Whenever a patient's results were discordant, the laboratory supervisor ordered all tests responsible for the discordance repeated for that sample. Table 5 summarizes potential conditions and their implications for retesting. In the secondary round of laboratory testing, laboratory technicians repeated the ordered tests while blind to the original results. After the secondary round of testing, the laboratory supervisor again reviewed the results. If the laboratory test results for a sample did not change, then the *final* RPR+TPPA result was confirmed as shown in Table 5. If the results did change, the sample was retested until achieving a *final* condition.

Table 5: Concordance Review

Central Laboratory Results			Field Result	Laboratory Action	Final Result (RPR+TPPA)
RPR	TPPA	RST	RST		
Positive	Positive	Positive	Positive	FINAL	Positive
Positive	Positive	Negative	Positive	Repeat RST	Positive
Positive	Positive	Positive	Negative	FINAL	Positive
Positive	Positive	Negative	Negative	FINAL	Positive
Positive	Negative	Positive	Positive	Repeat TPPA	Negative
Positive	Negative	Negative	Positive	Repeat all	Negative
Positive	Negative	Positive	Negative	Repeat all	Negative
Positive	Negative	Negative	Negative	Repeat RPR	Negative
Negative	No Result	No Result	Negative	FINAL	Negative
Negative	No Result	No Result	Positive	Repeat all	Negative
Negative	Positive	Positive	Positive	Repeat RPR	Negative
Negative	Positive	Negative	Positive	Repeat all	Negative
Negative	Positive	Positive	Negative	FINAL	Negative
Negative	Positive	Negative	Negative	FINAL	Negative
Negative	Negative	Any Result	Any Result	FINAL	Negative

2.2.5 Ethical Issues

Prior to participating in the study, all pregnant women were provided informational materials about congenital syphilis, syphilis during pregnancy, and rapid syphilis tests. Those wishing to participate signed an informed consent before providing fingerstick or venous blood samples. The signature of a parent/guardian was also required for women aged 16-18 per national policy.

The study was evaluated and approved by the Institutional Review Boards (IRBs) of the Health Directorate of Callao and the institutional ethics committees of INMP, *Universidad Peruana Cayetano Heredia*, and the University of Washington. Duke University approved the performance analysis utilizing a patient de-identified dataset provided by the laboratory supervisor and principal investigator of the project.

3.0 RESULTS

A total of 17,147 rapid syphilis tests were performed in the field (12,656 at INMP and 4,491 in Ventanilla-Callao). Of these, venous blood samples were collected from 11,169 women, including 7,871 (62.2%) women at INMP and 3,298 (73.4%) in Ventanilla-Callao. Field RPR results were unavailable for this analysis. The number of samples collected by health network and health service is shown in Table 6. In the UCPH central laboratory, technicians performed 11,169 RPR, 938 TPPA, and 244 RST tests during the primary round of testing.

Table 6: Number of Samples Collected by Health Network & Health Service

Health Network	Health Service	Number of fingerstick samples collected during field RST	Number of venous blood samples collected for laboratory testing
INMP	Antenatal Care	5,277	5,217
	Delivery/Postpartum	4,436	44
	Abortion Services	2,943	2,610
	All	12,656	7,871
Ventanilla-Callao	Antenatal Care	3,445	2,473
	Delivery/Postpartum	665	487
	Abortion Services	381	338
	All	4,491	3,298
Total	All	17,147	11,169

3.1 Description of Women Screened

The mean age of women screened was 26.7 years (95% CI 26.6-26.8), with no noticeable variations between women attending different health networks. The full age distribution for each health network is shown in Figure 4. The mean gestational age was 30 weeks (or 6.7 months), with women in all antenatal care units averaging 29 weeks, women in INMP's delivery/postpartum unit averaging 33 weeks, and all others (including those in abortion services) departing minimally from the mean.

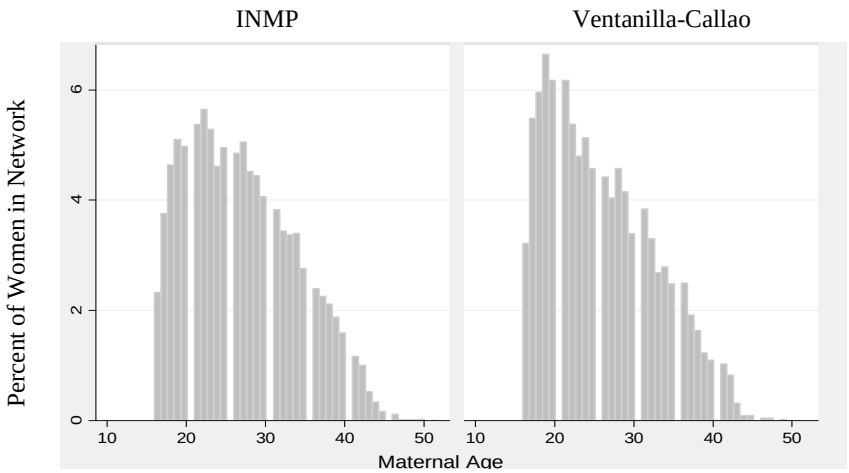

Figure 4: Maternal Age Distribution by Health Network

3.2 Prevalence of Syphilis Among Women Screened

For reasons described, the prevalence of syphilis among women screened varied according to screening method. Crude results are shown in Table 7. According to the RPR+TPPA gold standard, the overall prevalence was 1.05%, with a 0.36% difference between results at Ventanilla-Callao and INMP. Prevalence according to the field RST was lower at 0.90% overall (0.88% INMP, 0.96% Ventanilla-Callao), while prevalence according to only laboratory RPR only was higher at 1.68% (1.41% INMP, 2.33% Ventanilla-Callao).

Table 7: Crude Syphilis Prevalence by Health Network & Screening Method

Health Network	Field RST		Laboratory RPR		RPR+TPPA	
INMP	111/12,656	(0.88%)	111/7,871	(1.41%)	74/7,781	(0.94%)
Ventanilla-Callao	43/4,491	(0.96%)	77/3,298	(2.33%)	43/3,298	(1.30%)
Total	154/17,147	(0.90%)	188/11,169	(1.68%)	117/11,169	(1.05%)

Due to the uneven requisition of venous blood samples, prevalence data according to laboratory RPR and RPR+TPPA were likely inflated. Adjusted values are displayed in Figure 5 below. As perhaps expected from both methods' risks for false-positives, prevalence values found using the nontreponemal RPR and the treponemal RST were both higher than those found using the combined gold standard.

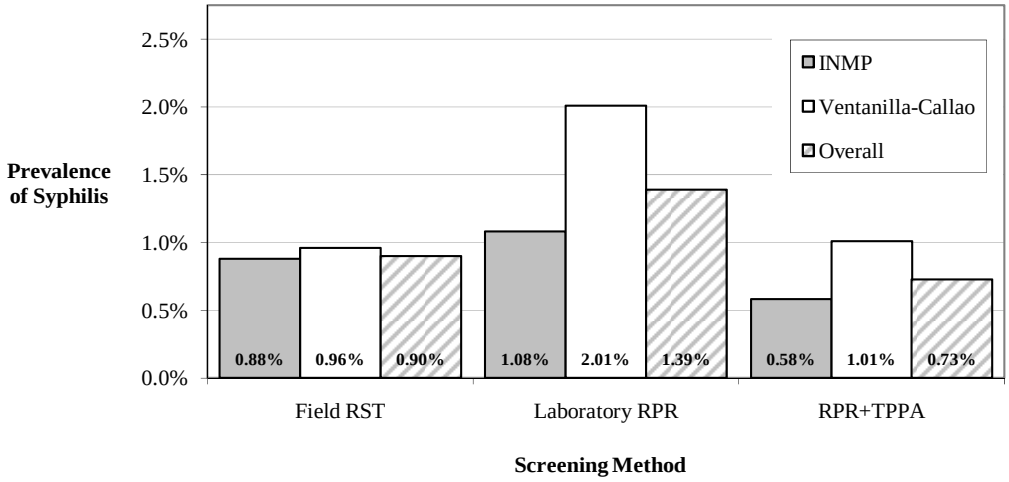

Figure 5: Adjusted Syphilis Prevalence by Health Network & Screening Method

3.3 Performance of the Bioline Rapid Syphilis Test

Overall, the Bioline RST displayed an unadjusted sensitivity of 91.0% (95% CI 86.4-95.0) and specificity of 99.1% (98.1-99.6) compared to TPPA. Compared to the RPR+TPPA gold standard, it displayed an unadjusted sensitivity of 91.5% (84.8-95.8) and specificity of 99.6% (99.4-99.7). In other words, of 200 pregnant women with RPR+TPPA-determined syphilis, the Bioline RST correctly detected approximately 183 as positive cases. Of 200 pregnant women without active syphilis, the Bioline RST correctly detected approximately 199 as negative cases. Full unadjusted results compared to each reference standard are shown in Table 8 and 2x2 tables for each health network are located in Appendix 5.1.

The crude sensitivity was over 12 percentage points lower in Ventanilla-Callao than at INMP according to both reference standards ($p = 0.023$[1]). While the crude specificity displayed less variation between the two health networks, the difference remained statistically significant ($p = 0.034$[2]).

The crude positive predictive value (PPV) and negative predictive value (NPV) could not be calculated for the TPPA reference standard, as the TPPA sample was enriched for cases according to the laboratory testing procedure. The PPV among women attending INMP – with a prevalence of 0.94% – was 64.5% (54.8-73.4), indicating that approximately 129 of every 200 women testing RST positive in the health network would truly have active syphilis. In Ventanilla-Callao – with a prevalence of 1.30% – PPV was 83.7% (69.3-93.2), indicating that approximately 167 of every 200 women testing RST positive would truly have active syphilis. Approximately 71 and 33 of these women, respectively, would be RST false positive.

Although the unadjusted figures give an overall indication of RST performance, the adjusted results are more valid according to the study's sampling limitations.[3] Adjusted performance results compared with both reference standards are presented in Table 9. Due to sampling size robustness and clinical relevance, this analysis focuses on the RPR+TPPA comparison, which yielded an overall sensitivity and specificity of 86.5% (78.8-92.0) and 99.7% (99.6-99.8), respectively.

Even with the adjustment, sensitivity differed greatly between the two health networks. Among women attending INMP (excluding delivery/postpartum), 93.0% of "true positive" (RPR+TPPA-positive) patients were correctly identified as syphilis-positive by the field RST. Among women attending a health center in Ventanilla-Callao, sensitivity was 79.7%. Following from these results, whereas 14 in

[1] Chi-squared test for sensitivity, reference standard: RPR+TPPA, Pearson chi-squared statistic = 5.200
[2] Chi-squared test for specificity, reference standard: RPR+TPPA, Pearson chi-squared statistic = 4.505
[3] This statement assumes that women providing fingerstick samples but not venous blood samples displayed equivalent syphilis prevalences and overall characteristics as those who provided both sample types. If this assumption were false, then the adjusted results would not necessarily be more accurate than the unadjusted.

200 syphilis-positive women attending INMP would receive a false-negative RST result (and generally not receive treatment), approximately 41 attending a Ventanilla-Callao health center would escape detection. The Bioline RST was more sensitive in INMP over Ventanilla-Callao within each health service as well as overall, with a peak sensitivity of 100.0% (88.8-100.0) within antenatal care services.

Specificity was 99.6% and over for each health network/service, indicating a minimal proportional risk of false positives with the RSTs. In this low-prevalence population, however, the overall risk for false-positive results is actually slightly higher than that for false-negative results. To illustrate this reality, one could extrapolate the observed prevalence data to predict 730 cases and 9,270 noncases in a representative population of 10,000 pregnant women. If adjusted, pooled sensitivity and specificity estimates were applied, approximately 30 of the 10,000 women would be false positive (risk = 0.00298) and 22 would be false negative (risk = 0.00218).

Variations in the adjusted sensitivity and specificity among the three health services were significant at INMP. In Ventanilla-Callao, sensitivity differed significantly when comparing antenatal care services against both delivery/postpartum and abortion care units, while no significant variation was seen in specificity estimates. See Table 10 for Pearson chi-squared homogeneity test results (reference standard: RPR+TPPA) and p-values.

Negative predictive value (NPV) was 99.8% to 100.0% for each health network. Positive predictive value (PPV) was considerably lower at 61.2% (51.1-70.1) at INMP, 83.7% (69.3-93.2) in Ventanilla-Callao, and 69.5% (61.3-76.5) overall.

Table 8: Crude RST Performance by Health Network

Health Network	**N**	***n*** (cases)	**% Sensitivity** (95%CI)	**% Specificity** (95%CI)	**% PPV** (95%CI)	**% NPV** (95%CI)
Reference Standard: TPPA						
INMP	649	105	95.2 (89.2-98.4)	98.9 (97.6-99.6)	Not interpretable	
Ventanilla-Callao	288	51	82.4 (69.1-91.6)	99.8 (97.7-100.0)		
All	937	156	91.0 (86.4-95.0)	99.1 (98.1-99.6)		
Reference Standard: RPR+TPPA						
INMP	7,871	74	95.9 (88.6-99.2)	99.5 (99.3-99.6)	64.5 (54.8-73.4)	100.0 (99.9-100.0)
Ventanilla-Callao	3,298	43	83.7 (69.3-93.2)	99.8 (99.6-99.9)	83.7 (69.3-93.2)	99.8 (99.6-99.9)
All	11,169	117	91.5 (84.8-95.8)	99.6 (99.4-99.7)	69.9 (62.0-77.1)	99.9 (99.8-100.0)

Table 9: Adjusted RST Performance by Health Network & Health Service

Health Network	**Health Service**	**N**	***n*** (cases)	**% Sensitivity** (95%CI)	**% Specificity** (95%CI)	**% PPV** (95%CI)	**% NPV** (95%CI)
Reference Standard: TPPA							
INMP	Antenatal Care	401	47	100.0 (92.5- 100.0)†	99.4 (98.0-99.9)	Not interpretable	
	Delivery/ Postpartum	44*	33	Not interpretable			
	Abortion Services	204	25	78.8 (59.3-93.2)	100.0 (98.0-100.0)†		
	All	649	105	92.3 (85.5-96.7)	99.6 (98.7-100.0)		
Ventanilla-Callao	Antenatal Care	196	33	87.7 (71.8-96.6)	99.6 (96.6-100.0)		
	Delivery/ Postpartum	60	8	54.9 (15.7-84.3)	100.0 (93.2-100.0)†		
	Abortion Services	32	10	67.4 (34.8-93.3)	100.0 (84.6-100.0)†		
	All	288	51	78.2 (64.7-88.7)	99.7 (99.4-100.0)		
All	All	937	156	86.4 (80.2-91.5)	99.7 (99.1-100.0)		

*Observations not included in total N or performance measures.
†One-sided (97.5%) confidence interval

Table 9: Adjusted RST Performance by Health Network & Health Service (Continued)

Health Network	Health Service	N	*n* (cases)	% Sensitivity (95%CI)	% Specificity (95%CI)	% PPV (95%CI)	% NPV (95%CI)
Reference Standard: RPR+TPPA							
INMP	Antenatal Care	5,217	31	100.0 (88.8-100.0)†	99.6 (99.4-99.8)	100.0 (93.2-100.0)†	100.0 (99.9-100.0)†
	Delivery/ Postpartum	44*	27	Not interpretable		78.9 (62.7-90.4)	Not interpretable
	Abortion Services	2,610	16	80.1 (54.4-96.0)	99.7 (99.4-99.9)	65.0 (40.8-84.6)	99.9 (99.8-100.0)
	All	7,871	74	93.0 (84.9-97.8)	99.7 (99.6-99.8)	61.2 (51.1-70.1)	100.0 (99.9-100.0)
Ventanilla-Callao	Antenatal Care	2,473	28	90.3 (71.7-97.7)	99.9 (99.7-100.0)	83.9 (66.3-94.5)	99.9 (99.8-100.0)
	Delivery/ Postpartum	487	7	49.3 (14.2-85.9)	99.8 (98.8-100.0)	80.0 (28.4-99.5)	99.4 (98.5-99.9)
	Abortion Services	338	8	72.6 (34.9-96.8)	99.7 (98.3-100.0)	85.7 (42.1-99.6)	99.4 (98.3-100.0)
	All	3,298	43	79.7 (64.0-90.0)	99.8 (99.7-99.9)	83.7 (69.3-93.2)	99.8 (99.6-99.9)
All	All	11,168	117	86.5 (78.8-92.0)	99.7 (99.6-99.8)	69.5 (61.3-76.5)	99.9 (99.8-100.0)

*Observations not included in total N or performance measures.
†One-sided (97.5%) confidence interval

Table 10: Variation Across Health Services

Health Network	Performance Measure	Variation Across...	Pearson Chi^2 Statistic	P-value
INMP	Sensitivity	All three health services	11.335	0.003*
		Antenatal care vs. other services	2.254	0.133
	Specificity	All three health services	1.4×10^3	0.000*
		Antenatal care vs. other services	2.823	0.093
Ventanilla-Callao	Sensitivity	All three health services	5.789	0.055
		Antenatal care vs. other services	4.916	0.027*
	Specificity	All three health services	0.133	0.936
		Antenatal care vs. other services	0.051	0.821

*Statistically significant variation (p level = 0.05)

3.4 Concordance of the Bioline Rapid Syphilis Test

In addition to these standard performance indicators, the reliability of results was important to evaluate in this study. Extraneous to the original study design, we approximated field reliability by assessing the concordance of RST results in the field with RST results in the laboratory. In compliance with the described retesting procedure, laboratory technicians tested 244 samples with the RST during the secondary round of testing. Of these, 230 (93.4%) yielded consistent results in the field and in the laboratory, while 14 (5.7%) yielded discordant results. See Table 11.

Discordant results were scrutinized carefully over the course of the study to ensure that anything less than 100% concordance was due to error within the test cassette itself and not error on the part of field technicians, the efficacy of training, or the quality of provided equipment (i.e. lighting). Any discordant cases deemed correctable were addressed upon result to ensure the highest validity of the field RSTs. For instance, it was discovered early in the study that one health worker did not wear appropriate corrective eyewear and could not accurately discern “borderline” results on the RST cassettes. Once the study provided her with new lenses, she did not

record any inaccurate results.[4] Thus, the concordance analysis provided an additional function of quality control throughout the study period. Of 154 women testing positive in the field, 152 (99.4%) retested RST-positive, 1 (0.7%) retested RST-negative, and 1 (0.7%) was not retested in the laboratory. 107 (69.5%) of these also tested positive by RPR+TPPA.

Table 11: RST Concordance

Concordance	Field RST Result (Whole Blood)	Laboratory RST Result (Serum)	Number of Results	n/N (%)	Action
Concordant	Positive	Positive	152	230/244 (94.3%)	N/A
	Negative	Negative	78		
Discordant	Positive	Negative	1	14/244 (5.7%)	Adjustments in health worker training, eyewear, and field lighting
	Negative	Positive	13		

3.5 Effects of Discrepant Analysis

Following the laboratory supervisor's recommendations, laboratory technicians re-performed 89 RPR, 25 TPPA, and 20 RST tests during the secondary round of testing to yield *final* results for the discrepant analysis. Overall, laboratory prevalence estimates were higher than those found during the primary round of testing (1.22% with RPR+TPPA). Sensitivity estimates, both crude (91.2%, CI 86.4-96.0) and adjusted (86.5%, CI 79.4-93.5), were equivalent to the primary laboratory results. There remained significant variation in performance between health networks. The main difference between the original results and those of the discrepant analysis were the PPV values, as would be expected with a shift in prevalence.

[4] All of this health worker's readings were reviewed for accuracy once the eyewear obstacle was discovered.

4.0 DISCUSSION

This study provides the first extensive field evaluation of the SD Bioline Syphilis 3.0 RST using whole blood in Latin America, as well as the first field study of any RST among pregnant women in Peru. Despite limitations, it displays the performance of the field RST to be reliable, reproducible, as valid as previous studies, and diagnostically apt for implementation in existing maternal care services in Peru.

4.1 Interpretation of Results

Of the various performance measures in this study, the sensitivity of the Bioline RST is the most important statistic of interest, as the medical risks associated with undetected maternal syphilis (false negative result) far outweigh the risk associated with unwarranted treatment (false positive result). Penicillin is relatively innocuous and inexpensive (with allergies rare), while undetected syphilis can lead to the serious health consequences described in Chapter 1.0 of this analysis. In this study, the overall sensitivity of the RST in the field was as expected from previous Bioline studies (both antenatal care and general settings), though lower than published results for Determine Syphilis TP.

In both health networks, the sensitivity of the Bioline RST rivaled that of Determine when limiting the study population to women attending antenatal care services. In the antenatal care unit at INMP, the Bioline RST detected every woman positive for syphilis. Tests performed in the delivery/postpartum units (in Ventanilla-Callao only) or during abortion services were substantially less sensitive, though small sample sizes in these units prevent generalizability of the estimates.

The Bioline RST performed with impressive specificity within each health service and overall; indeed, its figures for specificity represented the highest seen

among published Bioline studies. Despite this result, the low prevalence (and corresponding high proportion of syphilis-negative women) included more false-positive than false-negative results in the sample. The RST yielded 46 false-positive results (0.4% of the sample) versus 10 false-negative results (<0.1% of the sample) – an encouraging balance when working toward the elimination of congenital syphilis.

As the homogeneity test results demonstrate, performance estimates did not vary significantly between women attending antenatal care services and those attending the other services, despite the appearance of low sensitivity. Although sensitivity potentially could be lower in these settings (perhaps due to contamination or more immediate concerns distracting health worker attention), the low sample size in these services prevents conclusions regarding the RST's appropriateness in non-antenatal-care maternity services. Although definitive conclusions cannot be drawn, quality control issues warrant further examination in these settings in future studies.

The 5.3% discordance of the Bioline RST between the field and the laboratory was neither outrageous nor unexpected, especially in light of the different media (whole blood versus serum) used. In general, serum is considered a more accurate medium than whole blood, though inevitably a less appropriate medium for wide-scale implementation. In the laboratory, although the sample size was a fraction of that used to assess the performance of the field RST, the RST displayed a higher unadjusted sensitivity of 99.3% compared to RPR+TPPA. Its unadjusted specificity was lower at 72.2%. Deteriorating test quality (shelf life) of the RST could have also contributed to discordant results found in the later months of the study, along with operator error on the part of health workers reading the RST cassettes (or environmental factors causing the error), as previously noted.

4.2 Limitations

As mentioned, the low sample sizes found in non-antenatal-care services limit the applicability of performance estimates on a stratified level. Researchers tried to minimize this by widening the age-range of eligibility to include women up to 55 years of age (much higher than many pregnancy-related studies), but the limitation still transpired.

In addition, this study's PPV and NPV results are of limited usefulness, as both estimates vary greatly with prevalence. Since syphilis prevalence was low among pregnant women, PPV and NPV varied more erratically than sensitivity or specificity among health networks/services. As can be seen when comparing PPV/NPV results using the primary round of laboratory testing to those using discrepant analysis, even a small change in prevalence (from 1.22% to 0.73%) can shift these estimates substantially. Sensitivity remains a more reliable indicator of performance.

Likewise, the concordance analysis lacks robustness and should be viewed only in support of sensitivity/specificity results. Because the concordance analysis was not built into the original study design, the small sample size resulting from limited RST-retesting was a limitation. If this aspect of the analysis were foreseen, all venous blood samples could have been retested by RST in the laboratory to yield a more complete review.

Essential to note are the assumptions playing into the validity of the adjusted performance results. First and foremost, we assumed that women not contributing a venous blood sample (regardless of health network or service attended) mirrored those who did in demographic characteristics, past history of syphilis and maternal complications, risk behaviors, et cetera. Without knowing the true profile of women who were *not* assessed, this assumption creates a limitation, as it is possible that

women *not* traveling to the laboratory to provide a venous blood sample had non-independent, confounding characteristics.

Another assumption made in this analysis is that the women screened had no past history of syphilis. Due to the treponemal nature of the RST, women with prior treated infections would be classified as syphilis-positive at the time of the test even if active disease were no longer a concern. Upon RST administration, each patient was asked questions regarding past syphilis, genital warts within the past 12 months, and knowledge of partner history of syphilis. However, the data from this *encuesta larga* (long survey) were not available for the performance analysis. In order to complete the picture of diagnostic accuracy in real-world clinical settings, self-reported syphilis history (as would be included in the patient history taken during any standard office visit) is an important consideration.

Also, although equivalent performance of the Bioline RST may be expected in mild climates, the results of this study should not be extended to areas of extreme climate or living conditions, such as in the Peruvian Amazon or similar regions of other nations. As researchers and health representatives have noted, the shelf life and performance of RSTs may decrease in extremely humid or hot environments such as jungle areas [31, 35]. The exact effects are currently unknown.

4.3 Recommendations for Further Research

Further studies into the use of RSTs in maternity care beyond antenatal services are needed to fully determine their usefulness in these settings. It would perhaps be useful to conduct such studies in higher-prevalence areas, as even this extensive study could not capture substantial enough sample sizes. Accordingly, future studies should consider focusing more attention upon potential differences of RST field performance in order to best anticipate problems that may arise with RST

implementation. These could include environmental factors, health worker training, and setting-specific obstacles. For instance, further studies are needed to assess Bioline's performance in conditions beyond the semi-dry, populated capital (i.e. jungle areas) of Peru. In non-antenatal care settings, especially, it is possible that health workers could be distracted by immediate concerns of delivery/emergency situations and therefore not pay enough attention to achieving accurate, punctual RST readings. Albeit speculation, such a tendency could decrease RST performance in field settings; quality control issues likewise should be included in future studies.

Similarly, it would be useful to compare the performance of RPR in field laboratories with the performance of RSTs at the point of care. In order to replace the traditional field method with the new RSTs, the RST's diagnostic performance should be better, or as good as, the field RPR's. With equivalent performance, the added convenience and screening-treatment continuity would make RSTs the better option.

4.4 Implications for RST Implementation

Based on the results of this study, the Bioline RST displays satisfactory performance to allow implementation in health networks like INMP and Ventanilla-Callao with adequate training of health personnel. Aside from the PPV, indicators were particularly strong among women attending antenatal care services. It is therefore recommended that MINSA incorporate RSTs into the normal care routines of antenatal care services in the country, especially in locations that otherwise lack viable screening options. Furthermore, in areas with no laboratory or no capacity for treponemal confirmatory tests like TPPA, RSTs may provide a suitable proxy to use alone or in conjunction with RPR. This proxy method is currently being utilized in Chile [35] and provides an appealing alternative to the difficult, costly laboratory treponemal tests.

Although results are not as impressive among women attending delivery/postpartum or abortion services in this study, there is no reason (apart from sample size) to believe that the RST would not be effective in these settings. Pending further research, it should also be implemented into these services. As the proportion of women testing positive was generally higher among women attending these services – a topic extraneous to this discussion – the goal of reducing complications from maternal syphilis justifies the use of the RST in these situations. Although complications in the current pregnancy cannot be averted by the time a woman reaches the delivery ward or loses her baby, knowledge of syphilis status affects the health of future pregnancies and the risk for syphilis transmission.

Furthermore, lack of visible symptoms should never be equated to lack of disease; as this study indicates, pregnant women can have latent syphilis without their knowledge. Currently, scientific literature has not proven that latent syphilis does not lead to congenital syphilis. For a disease as simple to detect and treat as syphilis, and with new tools available like the Bioline RST, the government should do all it can to prevent the spread and consequences of this disease.

According to the Sexually Transmitted Diseases Diagnostics Initiative (SDI) of WHO, two principal roadblocks to preventing the spread of syphilis are lack of access to prenatal screening and lack of regulatory control (including lack of proof) of the quality of STI diagnostics [19]. The results of this study show that the Bioline RST is a satisfactory tool to increase screening access in the urban and semi-urban population, and likely in rural populations of similar climate. In addition, it contributes in a robust way to the growing scientific evidence in support of RSTs. In Peru, it provides concrete proof of diagnostic quality upon which the government may base its guidelines and provisions.

Increased prevention of congenital syphilis can only be achieved by increased proportion of syphilis-positive women being identified and treated before their pregnancies have progressed too far. Increased proportion is a function of both screening coverage and diagnostic sensitivity. This study shows that the Bioline RST can increase screening coverage with sensitivity comparable to other screening methods. As such, it should be implemented within existing maternal care units immediately.

5.0 APPENDICES

5.1 2x2 Tables, Overall

Overall		Field RST		Total
		Negative	Positive	
Laboratory TPPA	Negative	768	7	775
	Positive	14	142	156
Total		782	149	931

Overall		Field RST		Total
		Negative	Positive	
RPR + TPPA	Negative	11,006	46	11,052
	Positive	10	107	117
Total		11,016	153	11,169

5.1.1 2x2 Tables by Health Network

INMP		Field RST		Total
		Negative	Positive	
Laboratory TPPA	Negative	533	6	539
	Positive	5	100	105
Total		538	106	644

INMP		Field RST		Total
		Negative	Positive	
RPR + TPPA	Negative	7,758	39	7,797
	Positive	3	71	74
Total		7,761	110	7,871

Ventanilla-Callao		Field RST		Total
		Negative	Positive	
Laboratory TPPA	Negative	235	1	236
	Positive	9	42	51
Total		244	43	287

Ventanilla-Callao		Field RST		Total
		Negative	Positive	
RPR + TPPA	Negative	3,248	7	3,255
	Positive	7	36	43
Total		3,255	43	3,298

5.1.2 2x2 Tables by Health Service (INMP)

Antenatal Care		Field RST		Total
		Negative	Positive	
Laboratory TPPA	Negative	349	2	351
	Positive	0	47	47
Total		349	49	398

Antenatal Care		Field RST		Total
		Negative	Positive	
RPR + TPPA	Negative	5,165	21	5,186
	Positive	0	31	31
Total		5,165	52	5,217

Delivery/ Postpartum		Field RST Negative	Positive	Total
Laboratory TPPA	Negative	6	4	10
	Positive	0	33	33
Total		6	37	43

Delivery/ Postpartum		Field RST Negative	Positive	Total
RPR + TPPA	Negative	6	11	17
	Positive	0	27	27
Total		6	38	44

Abortion Services		Field RST Negative	Positive	Total
Laboratory TPPA	Negative	178	0	178
	Positive	5	20	25
Total		183	20	203

Abortion Services		Field RST		Total
		Negative	Positive	
RPR + TPPA	Negative	2,587	7	2,594
	Positive	3	13	16
Total		2,590	20	2,610

5.1.3 2x2 Tables by Health Service (Ventanilla-Callao)

Antenatal Care		Field RST		Total
		Negative	Positive	
Laboratory TPPA	Negative	161	1	162
	Positive	3	31	33
Total		164	31	195

Antenatal Care		Field RST		Total
		Negative	Positive	
RPR + TPPA	Negative	2,440	5	2,445
	Positive	2	26	28
Total		2,442	31	2,473

Delivery/ Postpartum		Field RST Negative	Field RST Positive	Total
Laboratory TPPA	Negative	52	0	52
	Positive	3	5	8
Total		55	5	60

Delivery/ Postpartum		Field RST Negative	Field RST Positive	Total
RPR + TPPA	Negative	479	1	480
	Positive	3	4	7
Total		482	5	487

Abortion Services		Field RST Negative	Field RST Positive	Total
Laboratory TPPA	Negative	22	0	22
	Positive	3	7	10
Total		25	7	32

Abortion Services		Field RST		Total
		Negative	Positive	
RPR + TPPA	Negative	329	1	330
	Positive	2	6	8
Total		331	7	338

5.2 Discrepant Analysis Tables

Table 12: Crude Syphilis Prevalence (Discrepant Analysis)

Health Network	Field RST		Laboratory RPR		RPR+TPPA	
INMP	$^{111}/_{12,656}$	(0.88%)	$^{120}/_{7,871}$	(1.52%)	$^{89}/_{7,781}$	(1.13%)
Ventanilla-Callao	$^{43}/_{4,491}$	(0.96%)	$^{80}/_{3,298}$	(2.43%)	$^{47}/_{3,298}$	(1.43%)
All	$^{154}/_{17,147}$	(0.90%)	$^{200}/_{11,169}$	(1.79%)	$^{136}/_{11,169}$	(1.22%)

Table 13: Crude RST Performance (Discrepant Analysis)

Health Network	**N**	***n*** (cases)	**% Sensitivity** (95%CI)	**% Specificity** (95%CI)	**% PPV** (95%CI)	**% NPV** (95%CI)
Reference Standard: TPPA						
INMP	649	109	95.4 (91.5-99.4)	98.9 (98.0-99.8)	Not interpretable	
Ventanilla-Callao	288	51	82.4 (71.8-92.9)	99.6 (97.7-100.0)		
All	937	160	91.3 (86.9-95.6)	99.1 (98.4-99.8)		
Reference Standard: RPR+TPPA						
INMP	7,871	89	95.5 (91.2-99.8)	99.7 (99.6-99.8)	77.3 (69.4-85.1)	100.0 (99.9-100.0)
Ventanilla-Callao	3,298	47	83.0 (72.1-93.8)	99.9 (99.8-100.0)	90.7 (81.9-99.5)	99.8 (99.6-99.9)
All	11,169	136	91.2 (86.4-96.0)	99.7 (99.6-99.8)	81.1 (74.8-87.3)	99.9 (99.8-100.0)

Table 14: Adjusted RST Performance (Discrepant Analysis)

Health Network	**Health Service**	**N**	***n*** (cases)	**% Sensitivity** (95%CI)	**% Specificity** (95%CI)	**% PPV** (95%CI)	**% NPV** (95%CI)
Reference Standard: TPPA							
INMP	Antenatal Care	401	50	100.0 (92.9-100.0)†	99.4 (98.0-99.9)	Not interpretable	
	Delivery/ Postpartum	44*	34	Not interpretable			
	Abortion Services	204	25	78.8 (59.3-93.2)	100.0 (98.0-100.0)†		
	All	649	109	92.6 (86.0-96.8)	99.6 (98.7-100.0)		
Ventanilla-Callao	Antenatal Care	196	33	87.8 (71.8-96.6)	99.6 (96.6-100.0)		
	Delivery/ Postpartum	60	8	54.9 (19.4-90.4)	100.0 (93.2-100.0)†		
	Abortion Services	32	10	67.4 (37.6-97.2)	100.0 (69.2- 100.0)†		
	All	288	51	78.2 (64.7-88.7)	99.7 (84.6-100.0)		
All	All	937	160	86.7 (80.6-91.7)	99.7 (99.3-100.0)		

*Observations not included in total N or performance measures.
†One-sided (97.5%) confidence interval

Table 14: Adjusted RST Performance (Discrepant Analysis) (Continued)

Health Network	Health Service	N	*n* (cases)	% Sensitivity (95%CI)	% Specificity (95%CI)	% PPV (95%CI)	% NPV (95%CI)
Reference Standard: RPR+TPPA							
INMP	Antenatal Care	5,217	39	100.0 (90.1-100.0)†	99.8 (99.6-99.9)	75.0 (63.2-86.8)	100.0 (99.9-100.0)†
	Delivery/ Postpartum	44*	30	Not interpretable		78.9 (65.4-92.5)	Not interpretable
	Abortion Services	2,610	20	78.9 (60.5-97.1)	99.9 (99.7-100.0)	80.0 (62.5-97.5)	99.9 (99.7-100.0)
	All	7,871	89	92.5 (85.4-99.5)	99.8 (99.7-99.9)	76.4 (66.6-86.2)	99.9 (99.9-100.0)
Ventanilla-Callao	Antenatal Care	2,473	30	90.9 (73.5-97.9)	99.9 (99.8-100.0)	90.3 (74.2-98.0)	99.9 (99.8-100.0)
	Delivery/ Postpartum	487	7	49.3 (11.9-86.8)	99.8 (99.5-100.0)	80.0 (28.4-99.5)	99.4 (98.5-99.8)
	Abortion Services	338	10	67.4 (37.6-97.1)	100.0 (98.9-100.0)†	100.0 (59.0-100.0)†	99.1 (97.7-99.8)
	All	3,298	47	79.1 (66.5-91.7)	99.9 (99.8-100.0)	90.7 (82.0-99.4)	99.8 (99.6-99.9)
All	All	11,168	136	86.5 (79.4-93.5)	99.8 (99.8-99.9)	81.2 (74.7-88.8)	99.9 (99.8-99.9)

*Observations not included in total N or performance measures.
†One-sided (97.5%) confidence interval

6.0 REFERENCES

1. Lyra, P., *Award-winning journalists highlight need for syphilis and HIV prevention among pregnant women*, 2010, Pan American Health Organization.

2. *Reporte Taller PRS*, in *Oportunidades de Aplicación de las pruebas rápidas de sífilis (PRS) en América Latina para la eliminación de la sífilis congénita*2010, Proyecto Cisne, Universidad Peruana Cayetano Heredia: Lima, Peru.

3. Valderrama, J., F. Zacarías, and R. Mazin, *Sífilis materna y sífilis congénita en América Latina: un problema grave de solución sencilla.* Rev Panam Salud Publica/Pan Am J Public Health, 2004. 16(3): p. 211-217.

4. CDC, *Sexually transmitted disease surveillance*, 2008, Atlanta, GA: US Department of Health and Human Services, CDC.

5. *Guidelines for prevention and control of congenital syphilis*, C.f.D.C.a. Prevention, Editor 1988, MMWR.

6. Mullick, S., et al., *Sexually transmitted infections in pregnancy: prevalence, impact on pregnancy outcomes, and approach to treatment in developing countries.* Sexually Transmitted Infections, 2005. 81: p. 294-302.

7. Tucker, J.D., et al., *Accelerating worldwide syphilis screening through rapid testing: a systematic review.* The Lancet Infectious Diseases, 2010. 10(6): p. 381-386.

8. Organización Panamericana de la Salud, U.V.S. *Hoja informativa sobre sífilis congénita*. 2004.

9. Pedersen, B.S., *Economic analysis of maternal screening to prevent congenital syphilis.* Sexually Transmitted Diseases, 1983. 10: p. 167-172.

10. García, P.J., et al., *Congenital syphilis prevention: Implementing rapid tests to improve access to syphilis screening among pregnant women in Peru [Study Protocol]*, 2008, Universidad Peruana Cayetano Heredia, University of Washington, World Health Organization, UNICEF/UNDP/World Bank: Lima, Peru. p. 1-41.

11. Garcia, P.J., et al., *Congenital syphilis prevention: Implementing rapid tests to improve access to syphilis screening among pregnant women in Peru*, 2008. (Study Protocol).

12. González, M.A., *Regional Initiative for the Elimination of Mother-to-Child Transmission of HIV and Congenital Syphilis in Latin America and the Caribbean: Regional Monitoring Strategy*, 2010, Pan American Health Organization: Washington, DC.

13. Lecca, H.G., et al., *Norma Técnica de Salud para la Profilaxis de la Transmisión Madre - Niño del VIH y la Sífilis Congénita*, M.d.S.d. Peru, Editor 2008, MINSA/DGSP.

14. *Action for the Global Elimination of Congenital Syphilis: Rationale and Strategy*, W.H. Organization, Editor 2006: Geneva.

15. Organization, W.H., *Guidelines for the Management of Sexually Transmitted Infections*, 2003, World Health Organization: Geneva. p. 39-45.

16. *RPR-Nosticon Product Manual*.

17. Larsen, S.A., B.M. Steiner, and A.H. Rudolph, *Laboratory Diagnosis and Interpretation of Tests for Syphilis.* Clinical Microbiology Reviews, 1995. 8(1): p. 1-21.

18. Peeling, R.W. and H. Ye, *Diagnostic tools for preventing and managing maternal and congenital syphilis: an overview.* Bulletin of the World Health Organization, 2004. 82(6): p. 439-446.

19. Peeling, R.W., et al., *Rapid tests for sexually transmitted infections (STIs): the way forward.* Sex Transm Infect, 2006. 82 Suppl 5: p. v1-6.

20. Mabey, D., et al., *Diagnostics for the developing world.* Nature Microbiology Reviews, 2004. 2: p. 231-40.

21. Herring, A.J., et al., *A multi-centre evaluation of nine rapid, point-of-care syphilis tests using archived sera.* Sex Transm Infect, 2006. 82 Suppl 5: p. v7-12.

22. Tinajeros, F., et al., *Diagnostic accuracy of a point-of-care syphilis test when used among pregnant women in Bolivia.* Sex Transm Infect, 2006. 82 Suppl 5: p. v17-21.

23. Bronzan, R.N., et al., *On-site rapid antenatal syphilis screening with an immunochromatographic strip improves case detection and treatment in rural South African clinics.* Sex Transm Dis, 2007. 34(7 Suppl): p. S55-60.

24. Hernandez-Trejo, M., et al., *[Maternal and congenital syphilis in two Mexican hospitals: evaluation of a rapid diagnostic test]*. Rev Invest Clin, 2006. 58(2): p. 119-25.

25. Mabey, D., et al., *Prospective, multi-centre clinic-based evaluation of four rapid diagnostic tests for syphilis.* Sex Transm Infect, 2006. 82 Suppl 5: p. v13-6.

26. Montoya, P.J., et al., *Comparison of the diagnostic accuracy of a rapid immunochromatographic test and the rapid plasma reagin test for antenatal syphilis screening in Mozambique.* Bull World Health Organ, 2006. 84(2): p. 97-104.

27. Villazon-Vargas, N., et al., *[Evaluation of a rapid diagnostic test to assess the prevalence of maternal syphilis in Bolivia]*. Rev Med Chil, 2009. 137(4): p. 515-21.

28. West, B., et al., *Performance of the rapid plasma reagin and the rapid syphilis screening tests in the diagnosis of syphilis in field conditions in rural Africa.* Sex Transm Infect, 2002. 78(4): p. 282-5.

29. Benzaken, A.S., et al., *Field evaluation of the performance and testing costs of a rapid point-of-care test for syphilis in a red-light district of Manaus, Brazil.* Sexually Transmitted Infections, 2008. 84: p. 297-302.

30. Nessa, K., et al., *Field evaluation of simple rapid tests in the diagnosis of syphilis.* Int J STD AIDS, 2008. 19(5): p. 316-20.

31. Benzaken, A.S., et al., *Field performance of a rapid point-of-care diagnostic test for antenatal syphilis screening in the Amazon region, Brazil.* International Journal of STD & AIDS, 2010. (Manuscript).

32. Campos, P.E., et al., *Utility of the Determine Syphilis TP rapid test in commercial sex venues in Peru.* Sex Transm Infect, 2006. 82 Suppl 5: p. v22-5.

33. *Innovation for Health: Research that Makes a Difference (TDR Annual Report 2009)*, J.N. Reza, Editor 2010, World Health Organization.

34. Miller, W.C., *Bias in discrepant analysis: when two wrongs don't make a right.* Journal of Clinical Epidemiology, 1998. 51(3): p. 219-31.

35. *Oportunidades de Aplicación de las Pruebas Rápidas de Sífilis (PRS) en América Latina para la Eliminación de la Sífilis Congénita*. in *Oportunidades de Aplicación de las Pruebas Rápidas de Sífilis (PRS) en América Latina para la Eliminación de la Sífilis Congénita Taller*. 2010, Nov. 18-19. Lima, Perú. (Conference Proceedings).