JOURNAL OF MEDICAL MICROBIOLOGY

RESEARCH ARTICLE

Osbak *et al.*, *Journal of Medical Microbiology* 2017;66:1130–1139
DOI 10.1099/jmm.0.000559

MICROBIOLOGY SOCIETY

Evaluation of an automated quantitative latex immunoturbidimetric non-treponemal assay for diagnosis and follow-up of syphilis: a prospective cohort study

Kara Osbak,[1,*] Said Abdellati,[1] Achilleas Tsoumanis,[1] Marjan Van Esbroeck,[1] Luc Kestens,[2,3] Tania Crucitti[1,*†] and Chris Kenyon[1,4†]

Abstract

Purpose. We evaluated the Sekure rapid plasma reagin (RPR-S) (Sekisui Diagnostics) automated quantitative latex immunoturbidimetric assay performed on the SK500 clinical chemistry system for clinical appropriateness.

Methodology. Syphilis-infected individuals and controls were recruited into a prospective cohort study conducted at a sexually transmitted infection clinic in Antwerp, Belgium. Sera collected at diagnosis (baseline) and at 3, 6, 9 and 12 months post-treatment were tested with RPR-S and Macro-Vue RPR card (RPR-C) (Becton Dickinson) assays; RPR-C was considered the reference test. IgG/IgM enzyme immunoassay and *Treponema pallidum polA* serum PCR results were consulted by discordancy at baseline. Categorical analyses were performed and correlations were assessed with (non)-linear regression. Post-treatment longitudinal serological evolution was evaluated.

Results. A total of 463 samples from 120 new syphilis cases from a variety of clinical stages and 30 syphilis-negative controls were tested. Initially, there was a weak correlation between quantitative RPR-C/S (r=0.15). In 70 samples there was a strong suspicion of hook effect. Of these, 57/70 sera were retested with an extra dilution step, resulting in an average 12-fold increase in quantitative RPR-S results. After the extra dilution, the overall qualitative RPR-C/S agreement was 78.89 %, (κ–coefficient: 0.484). Of the 92 discordant samples, 9 were from the baseline visit (RPR-C titre: 1–8), which could have led to possible missed diagnoses using the RPR-S.

Conclusions. The sensitivity and accuracy of the RPR-S test requires improvement before it can be used to diagnose syphilis and evaluate treatment efficacy in clinical practice.

INTRODUCTION

Syphilis, a multi-stage chronic disease transmitted sexually or vertically through the placenta, remains a substantial public health burden, with more than 8 million new infections per year worldwide [1]. The past 15 years has witnessed a re-emergence of syphilis in Belgium. An increasing proportion of cases are reinfections [2, 3], which present asymptomatically more frequently [3], making diagnosis dependent on serological test results.

Serological tests to diagnose syphilis fall into two main categories: (1) treponemal tests (TTs), which measure antibodies directed against *Treponema pallidum* ssp. *pallidum* (henceforth referred to as *T. pallidum*), the aetiological agent of syphilis, using recombinant or wild-type antigen–antibody reactions; and (2) non-treponemal tests (NTTs) [4], which measure nonspecific antibodies directed against cardiolipin that are released from damaged human cells and may be present in *T. pallidum's* cell wall [5, 6]. There are two main diagnostic algorithms for syphilis screening: a 'traditional algorithm', which screens with a NTT and confirms with a TT [7], and a 'reverse algorithm' that uses a TT for screening and a quantitative NTT [4] or a

Received 22 May 2017; Accepted 11 July 2017
Author affiliations: [1]Department of Clinical Sciences, Institute of Tropical Medicine, Nationalestraat 155, 2000 Antwerp, Belgium; [2]Department of Biomedical Sciences, Institute of Tropical Medicine, Antwerp, Belgium; [3]Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium; [4]Department of Medicine, University of Cape Town, Cape Town, South Africa.
***Correspondence:** Kara Osbak, kosbak@hotmail.com *or* kosbak@itg.be; Tania Crucitti, tcrucitti@itg.be
Keywords: diagnostics; *Treponema pallidum*; syphilis; non-treponemal test; hook effect; sexually transmitted infections; screening.
Abbreviations: ITM, Institute Tropical Medicine Antwerp; LTIA, latex turbidimetric immunoassay; NTT, non-treponemal test; RPR-C, Rapid Plasma Reagin Card test; RPR-S, Sekure Rapid Plasma Reagin Test; R.U., RPR units; STI, sexually transmitted infection; TPPA, *Treponema pallidum* particle agglutination test; TT, treponemal test; VDRL, venereal disease research laboratory test.
†These authors contributed equally to this work.
Two supplementary tables and three supplementary figures are available with the online Supplementary Material.

second TT assay (followed by an NTT assay to measure infection activity) for confirmation. The latter has been promoted in low-prevalence settings with the advent of automated low-cost *T. pallidum* chemiluminescent/luminescent immunoassays [4]. Only NTTs, such as the rapid plasma reagin card test (RPR-C) or the Venereal Disease Research Laboratory (VDRL) test, can be used for post-treatment follow-up to determine *T. pallidum* eradication and for the diagnosis of reinfection [8]. A fourfold decrease in RPR-C titre, which is considered to be a clinically significant difference between two NTT results, is necessary to qualify a response as adequate 6 to 12 months post-treatment. A recent systematic review estimated that the median proportion of adequately treated patients classified as 'serofast' or 'non-responders' was 20.5 % at 6 months and 11.2 % at 12 months post-treatment for all disease stages [9]. To diagnose a syphilis reinfection, a fourfold NTT titre increase, using the same test, should be demonstrated. However, increases in RPR-C titres can be secondary to a large variety of (auto)-inflammatory conditions, such as HIV and hepatitis [8, 10–12], leading to false positive diagnoses. Inaccurately low or false negative RPR assay results have been attributed to the hook or prozone effect during syphilis testing [13–15], an undesirable phenomenon observed in immunoassays when an unbalanced high antibody-to-antigen ratio interferes with the antigen–antibody reaction, which is typically alleviated by sample dilution.

In most laboratories, NTTs are performed manually and are thus subject to variation in inter-individual interpretation and reagent batch variation [16, 17]. Automation of NTTs could contribute in various ways to improve syphilis diagnosis, for example, by increasing test accuracy with the use of an internal standardized calibrant, higher throughput capability and reporting of more accurate continuous values instead of titre values. In Korean and Japanese laboratories, automated quantitative latex turbidimetric immunoassays (LTIA) for syphilis diagnostics have been widely implemented [18]. However, large-scale comparative studies, in particular in high sexually transmitted infection (STI) prevalence settings, are lacking. Currently, two automated RPR tests are commercially available, the Sekure RPR LTIA (RPR-S), also marketed as the 'Mediace RPR assay' (Sekisui Chemical Co. Ltd., Osaka, Japan) and the HISens Auto RPR LTIA (HBI Co., Ltd., Korea). The RPR-S test utilizes latex particles coated with purified cardiolipin-lecithin antigens extracted from bovine heart and measures agglutination secondary to antibody binding. In previous studies, the Mediace RPR assay, performed on other analytical platforms, was shown to have a high sensitivity (100 %) for early stage disease [19], but low qualitative correlation (83.8 %) with the traditional RPR-C test, low sensitivity for latent stage syphilis (55.6 %) [19, 20] and problems with prozone effect at high RPR-C titre values [20] have been reported.

In this study, we conducted a clinical laboratory evaluation to investigate the clinical appropriateness of the Sekure RPR assay for syphilis diagnosis and post-treatment serological follow-up. The SK500 clinical chemistry system was used to perform RPR-S testing on sera from 120 syphilis-infected patients and 30 controls collected during the course of a prospective cohort study.

METHODS

Study participants

This study was conducted as a sub-study of the '*Treponema pallidum*-specific proteomic changes in patients with incident syphilis infection (SeTPAT)' study (ClinicalTrials.gov registration number: NCT02059525) investigating novel biomarkers for syphilis infection. Potentially eligible study participants over the age of 17 in whom a new syphilis diagnosis was made were screened and recruited between January 2014 and August 2015 at the STI clinic of the ITM in Antwerp, Belgium. Additionally, HIV-infected controls with negative syphilis serological tests [RPR-C and *T. pallidum* particle agglutination test (TPPA) negative] were included at the same time as the study recruitment. The exclusion criteria were the use of beta-lactam, doxycycline or macrolide antibiotics during the 28 days preceding enrolment. Syphilis diagnosis and disease staging was performed according to the Centers for Disease Control guidelines [21]. Stage-appropriate treatment was administered according to European guidelines [22] in the form of intramuscular injection with benzathine penicillin G, or oral doxycycline in the case of penicillin allergy or penicillin unavailability. All syphilis-infected participants were followed up at 3, 6, 9 and 12 months post-treatment and in cases where syphilis reinfection was suspected. The study physician systematically recorded clinical and laboratory details in a standardized fashion during each study visit.

After screening 167 potential study participants, a total of 150 individuals were included in the study, 120 with an active episode of syphilis (RPR-C- and TPPA-positive) and 30 controls (Fig. 1, Table 1), with the exception of one case of secondary syphilis with a negative RPR-C but a positive TPPA and serum *T. pallidum* PCR at baseline. The median age of the syphilis-infected participants was 40 years (IQR 31.5–48), while for the controls it was 37.5 years (IQR 32–45). Only one syphilis-infected participant was a woman. The diagnosed syphilis stages were: *N*=23 primary, *N*=50 secondary, *N*=31 early latent and *N*=16 late latent stage disease. Of the syphilis-infected individuals, 103/120 (85.8 %) were HIV-infected, of which 91/103 (88.4 %) were taking antiretroviral therapy. All except one self-reported as being men who have sex with men (MSM), 43 % reported having had more than 10 sexual partners during the last year and almost two-thirds had a previous history of syphilis infection, indicating that this is generally a high-risk population. The median RPR-C titre at the time of syphilis diagnosis was 64 (IQR 32–128). During follow-up, participants provided a total of 86 3-month samples, 108 6-month samples, 47 9-month samples and 68 12-month samples. Six 'reinfection' samples were provided for analysis (Fig. 1). These were

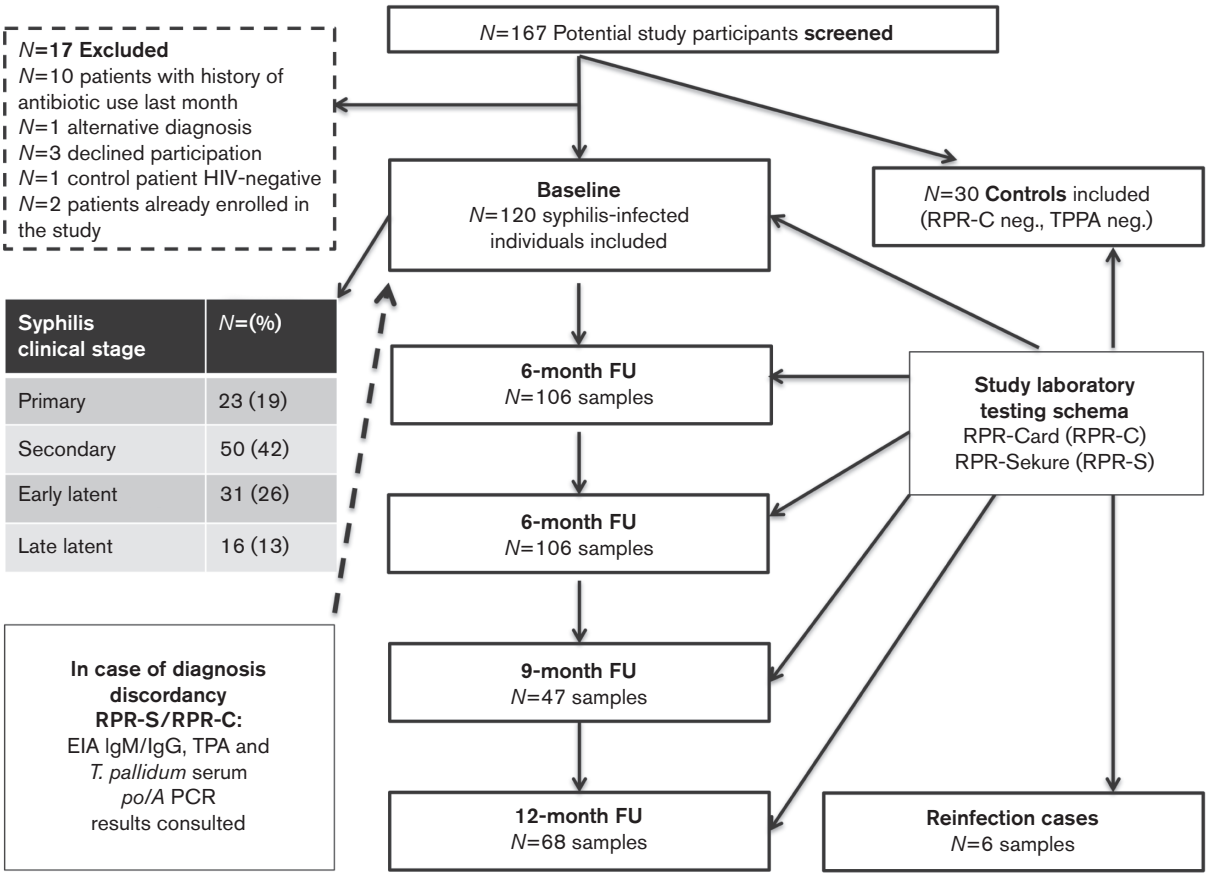

Fig. 1. Prospective cohort sub-study overview and laboratory testing workflow description. A total of 150 study participants, 120 with a new syphilis diagnosis and 30 syphilis negative controls, were recruited at the ITM STI clinic in Antwerp, Belgium. FU, follow-up.

Table 1. Clinical cohort characteristics of participants in this study

Characteristics	Syphilis-infected cases *N*=120	Controls *N*=30
	N=; (%/IQR)	*N*=; (%/IQR)
Gender (male)	119 (99)	30 (100)
MSM	117 (98)	24 (80)
HIV-infected	103 (86)	30 (100)
Taking ART	91 (88)	24 (80)
Benzathine penicillin G treatment	117 (98)	NA
RPR-C titre at baseline	64 (32–128)	0
Age	40 (31.5–48)	37.5 (32–45)
Number of previous syphilis episodes		
0	48 (40)	30 (100)
1	41 (34)	0
2	17 (14)	0
3	9 (8)	0
≥4	5 (4)	0

ART, antiretroviral therapy; NA, not applicable.

collected in the case of syphilis reinfection between regular study visits, or after the 12-month sub-study period.

Routine clinical serological testing

Blood was drawn into serum gel tubes (Sarstedt Monovette, Nümbrecht, Germany) and separated at 2000 ***g*** for 10 min at ambient temperature. Serum samples were divided and subsequently either (1) refrigerated at 4 °C until routine clinical syphilis serological testing (as detailed below) within 4 days, or (2) aliquoted into 2 ml cryovials and stored at −80 °C within 3 h for later evaluative testing. All fresh sera were analysed with Macro-Vue RPR Card (RPR-C) (Becton Dickinson, Sparks, MD, USA) and a TPA assay (Ortho-Clinical Diagnostics, Rochester, NY, USA) according to the manufacturer's instructions. If the sera tested RPR-C-positive, the titre was determined to an endpoint. Since these tests were performed during regular clinical routine testing, different experienced laboratory technicians performed the tests during the course of the study. All baseline fresh sera were also tested with an in-house *T. pallidum* PCR directed against *polA* [23] and a SERODIA-TPPA (Fujirebio Inc., Tokyo, Japan) assay.

Implementation of the SK500 clinical chemistry system for RPR-S testing

In collaboration with Sekisui Diagnostics, LLC, a SK500 analytical system was installed in the ITM STI reference laboratory. Two laboratory technicians undertook practical on-site training for two days, including trial-run analysis of non-study-related reference samples. Evaluative testing of the RPR-S assay took place after the samples of interest had been collected, a maximum of 974 days previously, and samples were stored at −20 °C until analysis, which is outside of the RPR-S test insert stability claim. Samples were subsequently thawed, batched, analysed and stored at 4 °C until all sub-study testing was completed. Sera testing was conducted on all available study samples from baseline, 3, 6, 9 and 12 months post-treatment visits, and on samples from serologically (RPR-C) confirmed cases of reinfection, following the manufacturer's instructions (version KI486616.EN.02, June 2016). Sera were thawed at ambient temperature and centrifuged at 15 000 ***g*** for 10 min prior to testing. For the RPR-S testing, 180 µl of reagent 1 and 20 µl of serum were added to the assay tube and incubated for 5 min, and then 60 µl of reagent 2 was added and incubated for 5 min. The absorbance at 700 nm was measured at regular intervals up to 7.5 min. The results are expressed in RPR units (RU), whereby 1 RU is equivalent to a onefold RPR-C titre according to the manufacturer. A measurement of 1 RU or higher indicates that the sample is antibody-positive and measurements of 2 RU and 4 RU are equivalent to RPR-C dilutions of 1/2 and 1/4, respectively. During the first-run analysis of all samples, the analyser was programmed to automatically subject samples with an initial RU score of ≥8.0 to a 10× dilution step with 0.9 % NaCl solution before retesting. In addition, the analyser was programmed to detect the prozone phenomenon.

Intermittent statistical analysis was performed after the first-run RPR-S testing of all samples. The results showed possible hook effects, after which the analyser was programmed to automatically dilute (10×) samples with an initial RU score of ≥4.0 and retest.

All RPR-S testing was performed in single-run analyses conducted by the same study technician, with the exception of the second analysis of the samples that showed a possible hook effect, which was performed by the company technicians at the ITM. Industry assistants were available for troubleshooting consultations through mail correspondence, by telephone and through visits to the facility during the testing period.

Quality control

The ITM laboratories in Antwerp where these samples were analysed are ISO15189 accredited. The RPR-S reagent kit and respective controls used in this study were from the same lot number. Before each batch analysis, the RPR-S assay was calibrated with the RPR calibrator set and the control samples provided; the control sample results were compared to the sample validation range provided by the manufacturer. The reagents used for the RPR-C, TPPA and TPA tests were from different lot numbers due to the prolonged testing period and the nature of the routine clinical testing setting. All testing reagents were stored in the same conditions, according to the manufacturer's instructions. The reanalysis of the suspect hook effect samples was first re-examined without the extra dilution step to assess possible testing variation after prolonged (defined as longer than 2 weeks) storage at 4 °C. Details of the freeze-thaw cycle frequency were recorded for each sample.

Operator feedback

After the two-day on-site training provided by Sekisui Diagnostics technical experts, the study laboratory technicians were able to perform the testing, at first under supervision and with coaching, and then after one week independently, without major problems. Overall, the technicians found the SK500 system easy to use. However, a comprehensive understanding of the analyser is required to run the samples independently, and one must diligently focus on the test progress and reaction dynamics during each run to ensure that the testing runs smoothly and efficiently.

Statistical analysis

Henceforth, the 'first-run' analysis refers to the initial data set with an RU cutoff of 8.0, triggering 10× dilution. The 'second-run' analysis refers to the same data as those used in the first run, except for the 57 suspected-hook-effect samples that were re-run at an RU 10× dilution cutoff of 4.0. For the comparative categorical and continuous data analysis, the results were substantively equivalent when analysed per visit, or when all of the visits were pooled. For the longitudinal serological evaluation of the RPR-S/C results, the baseline, the 6-month, 12-month and reinfection time points were considered individually.

The RPR-C test was the reference test. Categorical analysis involved the construction of two-by-two contingency tables (positive/negative) for the RPR-S/C result comparisons. The percentage agreement and Cohen's kappa (κ)-coefficient value [24] were calculated to estimate the agreement between both tests. The sensitivity, specificity, and positive (PPV) and negative predictive values (NPV) with 95 % confidence intervals (CIs) were calculated using the baseline visit samples. The results of the enzyme immuno assays (EIAs) and PCR testing were consulted in the case of discordance between the RPR-S/C results at the baseline visit. The PCR analyses were performed in the context of the main study and the EIAs (Anti-*Treponema pallidum* ELISA IgM/IgG; Euroimmun, Lübeck, Germany; recomWell Treponema ELISA IgM/IgG; Mikrogen, Neuried, Germany) were performed in the context of another sub-study evaluating two different IgG and IgM EIAs.

Continuous variables are expressed with the median value and interquartile range (IQR). Histograms were plotted and Shapiro–Wilk testing [25] was conducted to assess data normality. (Non)-linear regression was performed after scatterplots were constructed to evaluate the relationship of the

transformed results. The initial first-run RPR-S/C results were assessed using Pearson's correlation analysis (r). Non-normal continuous variables from the second-run dataset were transformed in order to achieve more linear relationships for correlation analyses; the linear agreement was assessed by Passing and Bablok regression analysis [26] within the linear range test values.

To assess the potential effects of prolonged storage at 4 °C, sample means were compared with a Wilcoxon- signed rank test [27]. The serofast and seroreversion classification frequency between RPR-S/C testing were evaluated with a chi-squared (χ^2) test [28] and κ-coefficient value calculations.

All analyses were performed in Stata 12.1 (StataCorp LP, College Station, TX, USA) and MedCalc version 17.1 (MedCalc Software, Mariakerke, Belgium). A P-value of less than 0.05 was considered statistically significant.

Serofast, seroreversion and syphilis reinfection definitions

We defined the 'serofast' state [21, 29] as the failure of the RPR-C titre to decline fourfold between the baseline and the 6- (6-month serofast) and 12-month (12-month serofast) visits, respectively. Seroreversion was defined as RPR-C titre conversion to negative at the 6- (6-month seroreversion) or 12-month (12-month seroreversion) time points. Reinfection was defined as an episode in which the RPR titre increased ≥fourfold after a previous syphilis episode demonstrated a serological response to therapy (≥fourfold decrease in RPR-C titre). We assumed a 1 : 1 relationship between the RPR-C/S values, as dictated by the manufacturer's testing specifications. As a result we extrapolated the definitions mentioned above to the RPR-S testing. Thus, a fourfold change in RPR-S result was calculated by dividing (decrease) or multiplying (increase) the RU value by four.

RESULTS

Summary of the quantitative RPR-S serological test evaluations

Serological testing of 463 sera with the RPR-S kit was performed over the course of a two-month period in late 2016 (first-run analysis). Once all of the RPR-S testing was complete, an intermittent quantitative statistical analysis was performed to examine the test correlations with the RPR-C results. The initial RPR-S/C correlation was r=0.15, whereby 70 samples had a strong suspicion for hook effect (low RPR-S RU/high-RPR-C titre) as identified by scatterplot analyses, corresponding to a median RPR-C titre of 128 (IQR 128–256) and an RPR-S median RU value of 6.6 (IQR 5.9–7.2). Fifty-seven of these samples were available for retesting, which was performed a median of 40 days (IQR 27–45) after thawing and continuous storage at 4 °C. The 57 samples were first reanalysed without 10× dilution as a control check; these results did not differ significantly from those of the original first-run testing (P=0.22). After implementation of a sample dilution step (10×) at an RPR cutoff of 4.0 RU, the RPR-S values for the 57 samples increased significantly (P<0.001) (Table 2, Fig. 2b), resulting in a median RU value of 90.2 (IQR 81.2–96), equalling an average 12-fold increase compared to the first-run. Only one RPR-S result did not increase after retesting, and this corresponded to a RPR-C of 32 and a RPR-S of 4.7/4.3 for the first-/second-run results, respectively. A total of 450 samples were included in the second-run dataset, including 393 samples from the first-run dataset (no suspicion of prozone or hook effect) and the 57 samples subjected to the 4.0 RU cutoff 10× dilution, and excluding the 13 potential hook effect samples, which were not available for retesting (Fig. S1, available with the online Supplementary Material). All samples with an RPR-C titre of 1 (N=29) tested RPR-S-negative, whilst

Table 2. RPR-Sekure RU result breakdown per RPR-C titre for the first- and second-run analyses

RPR-C titre	RPR-Sekure first-run analysis					RPR-Sekure second-run analysis				
	Neg.	Pos.	Total	Median RU	IQR	Neg.	Pos.	Total	Median RU	IQR
0	71	1	72	0	0–0	71	1	72	0	0–0
1	29	0	29	0	0–0.2	29	0	29	0	0–0.1
2	38	5	43	0.1	0–0.4	38	5	43	0.1	0–0.4
4	22	21	43	0.9	0.5–1.5	22	21	43	0.9	0.5–1.5
8	5	42	47	2.3	1.6–3.8	5	42	47	2.3	1.6–3.8
16	0	57	57	6.9	4.2–19.4	0	57	57	6.9	4.1–19.4
32	0	55	55	28.9	23.6–32.9	0	54	54	29.7	24.2–33.1
64	0	48	48	40.5	22.4–65.8	0	45	45	62.7	39.5–81.9
128	0	35	35	6.9	6.5–55.7	0	30	30	92.9	79–96
256	0	21	21	6.8	6.1–7.6	0	19	19	90.2	81.9–98.8
512	0	9	9	5.9	5.8–6	0	8	8	78.9	74.3–84.05
1024	0	4	4	5.6	5.5–6.05	0	3	3	68.8	57.4–83.2
Total	165	298	463			165	285	450		

IQR, interquartile range; RU, RPR units; RPR-C, manual RPR-card testing.

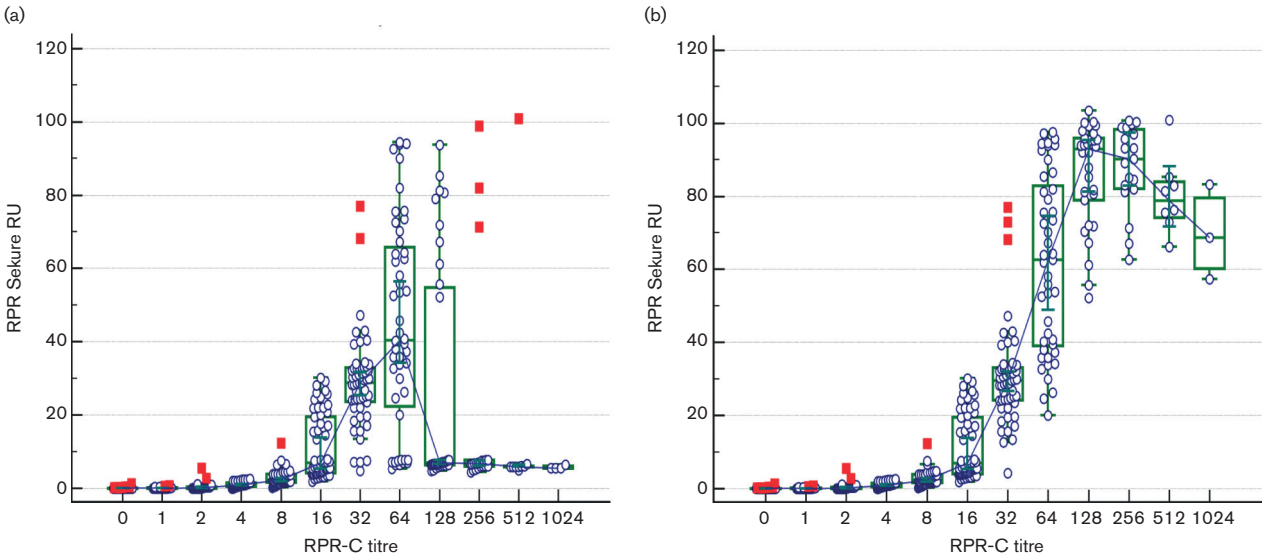

Fig. 2. Box and whisker plots depicting RPR-S RU results per RPR-C titre for (a) the first-run and (b) the second-run analyses. RPR-S results are shown as median and interquartile range (box), with 10th and 90th percentiles (whiskers). Red squares represent outliers.

11.6 % (5/43), 48.8 % (21/43) and 98.4 % (42/47) of RRP-C titres 2, 4 and 8, respectively, tested RPS-S-positive. A breakdown of the results per RPR-C titre level and schematic graphical overview are provided in Fig. 2(a, b) and Table 2.

(Non-linear) correlation analysis RPR-S performance for second-run analysis

There was a non-normal data distribution (RPR-S: W=0.77, P<0.001; RPR-C: W=0.45, P<0.001), thus the RPR-S second-run and RPR-C values were $\log_e(x)$-transformed. Fig. S2 is a graphical depiction of the sigmoidal relationship that was fitted with the regression equation. When considering the RPR-C titre ranges 4–64 (N=246 samples) for the Passing–Bablok linear regression analysis, intercept A was −2.27 (95 % CI: −2.54–2.04), representing the systemic differences, and the proportional difference (slope B) was 1.60 (95 % CI: 1.52–1.70), yielding the equation y=−2.27+1.60x (Figs S2 and S3). The RPR-S/C linear correlation range was estimated to be between RPR-C titres 8–32, corresponding to a median RPR-S of 2.3–29.7 (IQR 1.6–32.9) RU.

Categorical assessment of the RPR-S test

With respect to the RPR-S first- and second-run analyses, in both runs 94/463 and 94/450 samples were classified as false negative when compared with RPR-C testing, respectively (Table 3). There was 78.48 % agreement for the first-run and 78.89 % agreement for the second-run, together with moderate [30] κ-coefficients of 0.488 and 0.484, respectively. The quantitative agreement ±1 RPR-C dilution (margin of error) corresponding to a ±2-fold RPR-S range was 42.98 % for the first-run analyses and 50.44 % for the second-run analyses. When considering the RPR-C titre breakdown, the false negative RPR-S results occurred in the low RPR-C titre ranges 1–8 (Table 2) for both runs. Only one RPR-S sample was false positive, with a weakly positive value of 1.3. The overall sensitivity was 93.3 % for the first-run (baseline visit; N=150) and 92.5 % for the second-run (baseline visit; N=138), and the specificity was 100 % for both runs (Table 4).

Table 3. Cross-tabulation of the RPR-S (first and second runs) with the results of the RPR-C testing

		RPR-C		
		Positive	**Negative**	**Total**
First-run analysis, all				
RPR-S	Positive	297	1*	298
	Negative	94†	71	165
	Total	391	72	463
Second-run analysis, all				
RPR-S	Positive	284	1*	285
	Negative	94†	71	165
	Total	378	72	450
First-run analysis, baseline				
RPR-S	Positive	111	0	111
	Negative	8	31	39
	Total	119	31	150
Second-run analysis, baseline				
RPR-S	Positive	99	0	99
	Negative	8	31	39
	Total	107	31	138

*RPR-S value was weakly positive: 1.3.

†RPR-S tested negative values corresponded to RPR-C titre ranges 1–8 (see Table 2).

Table 4. Categorical analysis of RPR-S test *versus* reference test (RPR-C)

	RPR-C *versus* RPR-S First-run *N*=463 (95 % CI)	RPR-C *versus* RPR-S Second-run *N*=450 (95 % CI)
κ-coefficient	0.488 (0.409–0.568)	0.484 (0.405–0.564)
% Agreement, qualitative*	79.48 %	78.89 %
% Agreement, quantitative†	42.98 %	50.44 %
	N=150 (95 % CI)	*N*=138 (95 % CI)
Sensitivity‡	93.3 % (87.2–97.1)	92.5 % (85.8–96.7)
Specificity‡	100 % (88.8–100)	100 % (88.8–100)
PPV‡	100 % (96.7–100)	100 % (96.3–100)
NPV‡	79.5 % (63.5–91.8)	79.5 % (63.5–90.7)

PPV, Positive predictive value; NPV, negative predictive value; CI, confidence interval.

*Qualitative agreement, positive/negative.

†Quantitative ± twofold RPR-C/S agreement equals ± one dilution RPR-C/ ± twofold RPR-S RU result.

‡Calculated using baseline samples only.

Potential clinical misclassifications at baseline according to RPR-S testing

Of the 120 syphilis-infected individuals included in this study, nine would have been considered syphilis-negative according to the RPR-S testing at baseline [7]. Additional laboratory information, such as TPA, EIAs IgM/IgG and serum *polA* PCR results were consulted in order to serologically characterize these nine cases (full details are provided in Table S1).

All nine individuals were HIV-infected males and two individuals had a previous history of syphilis, with positive RPR-C and TPPA tests. All except one had negative RPR-C results available from before the study period. Two individuals presented with primary (ulcers), one with secondary (maculopapular rash), four with early latent and two with late latent disease stages.

All had positive RPR-C titres at baseline in the low range (2–4), except for one secondary syphilis case that was RPR-C-negative, serum PCR-positive, EIA IgM-positive at baseline, and without a history of syphilis. His RPR-C became positive at his 9-month visit.

Eight of the nine were classified as definite new syphilis infections, based on seroconversion of RPR-C results within the prior 12 months. In addition, three were PCR-positive, three exhibited convincing clinical signs of syphilis and five were IgM-positive. The ninth individual classified with probable syphilis presented with recent high-risk behaviour, no history of syphilis symptoms or treatment, a new diagnosis of HIV, an RPR-C of 2, a TPA of 151, and positive IgG but negative IgM and PCR tests.

Six and 12-month post-treatment serofast and seroreversion cases according to RPR-C/S testing

The details of the 12-month RPR-S/C follow-up for the syphilis positive participants with available longitudinal data are reported in Table S2. The serological follow-up results at the 6-month and 12-month time points were available for 96/120 and 63/120 individuals, respectively. Samples were excluded from the longitudinal analysis if they were taken after an episode of reinfection during the course of the study (*N*=6), and if they had negative RPR-C (*N*=1) or RPR-S (*N*=9) results at baseline. All analyses were performed with the second-run analysis.

According to the RPR-S testing, there were significantly more 6-month serofast cases (22.6 *versus* 10.2 %, *P*=0.001); at 12 months there was no significant difference between the two testing methods (14.8 *versus* 5.6 %, *P*=0.353). There were significantly more seroreversions at both the 6-month (42.1 *versus* 9.1 %, *P*<0.001) and 12-month (46.3 *versus* 20.4 %, *P*=0.001) time points according to the RPR-S testing. There was general moderate to low percentage agreement and minimal agreement according to the κ–coefficient calculations for all time points and classifications (serofast/seroreversion) (Table 5).

Reinfections during the course of the study

A total of nine reinfection cases could be evaluated with both the RPR-C and RPR-S tests (second-run analysis) during the total course of the study follow-up. Seven of nine were asymptomatic and serologically confirmed (RPR-C). Only 6/9 (67 %) cases would have been classified as reinfections according to the RRP-S test (Table S2).

DISCUSSION

The objective of this study was to assess the clinical applicability of the Sekure RPR test performed on the SK500 clinical chemistry system, specifically for the applications of syphilis diagnosis and post-treatment serological follow-up. A novel high-throughput, sensitive and precise NTT would be a welcome improvement for syphilis diagnostics. We found two main problems with the RPR-S test. Firstly, the hook effect and/or ineffectual sample dilution presumably affected samples with high antibody levels, corresponding to RPR-C titre $\geq$128, leading to lower than expected RPR-S results. This problem was partially rectified when samples were rerun at an RU cut-off of 4.0 to trigger a 10× dilution step. Secondly, the RPR-S test was not sufficiently sensitive at RPR-C titres $\leq$8 – missing nine cases of syphilis at the baseline visit. Consequently, we estimated the limited linear RPR-C/S correlation range to be between RPR 8–32, based on the second-run analysis. This lack of RPR-S test sensitivity and accuracy would have negative clinical outcome consequences, including potential misdiagnoses and misclassified treatment outcomes, such as an overestimation of serofast status and resulting antibiotic overtreatment.

Our results are concordant with a previous study investigating the Mediace RPR LTIA assay, which found a poor correlation

Table 5. Contingency tables for the 6- and 12-month serofast and seroreversion cases according to the RPR-S and RPR-C testing

		RPR-C				
		No *N*=(%)	**Yes *N*=(%)**	**Total *N*=(%)**	**Agreement (%)**	**κ–coefficient (95 % CI)**
Serofast 6 months						
RPR-S	No	65	3	68 (77.3)	80.68 %	0.318 (0.081–0.554)
	Yes	14	6	20 (22.6)		
	Total	79 (89.8)	9 (10.2)	88		
Serofast 12 months						
RPR-S	No	44	2	46 (85.2)	83.33 %	0.110 (−0.202–0.421)
	Yes	7	1	8 (14.8)		
	Total	51 (94.4)	3 (5.6)	54		
Seroreversion 6 months						
RPR-S	No	51	0	51 (58.0)	67.05 %	0.242 (0.095–0.390)
	Yes	29	8	37 (42.1)		
	Total	80 (90.9)	8 (9.1)	88		
Seroreversion 12 months						
RPR-S	No	28	1	29 (53.7)	70.37 %	0.380 (0.167–0.593)
	Yes	15	10	25 (46.3)		
	Total	43 (79.6)	11 (20.4)	54		

between the RPR-S and RPR-C testing and evidence of hook effects at high RPR-C titres [31]. Sample dilution presumably (partially) corrected the hook effect, as demonstrated in the present study, where 56/57 attained higher values closer to the expected range after dilution. The maximum RPR-S value in our study was 103.5 (corresponding to a RPR-C titre of 128), which is lower than the previously reported maximum result of 384 RU [32], although it was not specified whether the samples were diluted and on which analytical platform the tests were performed in the earlier report. These findings could be due in part to the low maximum value of the calibration curve (8.0 RU), whereby 10× dilution only results in a maximum measurement of 80 RU within the linear range. Per procedure, the extra dilution step did not affect the qualitative analysis and agreement. Notably, few previous studies of the RPR-Mediace test have consistently evaluated samples with RPR-C titre values of higher than 128 [31–33]. This oversight is important, since many syphilis patients present with high RPR titres at diagnosis, as seen in this study, where 44/120 (36.7 %) of individuals presented with RPR titres ≥128 at baseline.

A limited RPR-S test sensitivity was observed at the lower antibody level range, corresponding to an RPR titre ≤8. The RPR-S had an overall good sensitivity of 93.3 % and a specificity of 100 %, although these values must be interpreted with caution, since this cohort biomarker discovery study was not designed to adequately address the performance of the test as a population-screening tool. The attained sensitivity is higher than that in a previous study of 101 serum samples, reporting a RPR-S sensitivity of 60.5 % compared to the RPR-C test [34]. Further, if the low-titre false negative results from the other study visits were taken into account and extrapolated to theoretical disease events, the clinical consequences would be important.

Regarding clinical diagnosis discrepancies between RPR-S/C testing, 9/120 (7.5 %) syphilis diagnoses would have been missed at baseline and 3/9 (33 %) cases of reinfection would have been missed if only the RPR-S results were taken into account. It is important to note that these reinfection cases could have been misclassified, since syphilitic reactivation due to subcurative treatment and treponemal antibiotic resistance [35, 36] can also cause significant increases in NTT results, although clinically significant high-level resistance has only been reported for macrolides. Specifically, 2/9 (22 %) of these missed cases were from the primary stage, in which antibody levels are often low or insufficiently high for detection in serological tests, as reflected by the low TPA/TPPA results for these cases, while one case was deemed 'probable', since we could not rule out the possibility that the syphilis infection could have been indirectly treated by antibiotics long before diagnosis.

The appropriateness of RPR-S testing for post-treatment syphilis follow-up is a concern, due to the aforementioned suboptimal performance at high antibody titres and low quantitative agreement between the tests (43 % first-run and 50 % second-run), which mean that it might not be possible to deduce a correct delta RPR change after treatment. This is reflected in the significantly higher amount of serofast 6-month cases according to the RPR-S testing, which when translated into clinical practice could result in unnecessary antibiotic treatment. Nevertheless, pending RPR-S optimization, the definition of successful treatment (serofast/non-responders) could change with the introduction of a test providing continuous values instead of titres.

There are a number of caveats related to biological, technical and analytical biases, which may have influenced the results, thereby limiting generalizability. For example, we did not control for the possible effects of HIV and the

participants were almost all MSM. The RPR-C test was used as the reference test, but it is by no means a perfect test [16, 17], and there is a possibility that some RPR-C results could have been interpreted incorrectly beyond the error margin of ±1 RPR-C dilution. Some samples were subjected to sub-optimal testing conditions, namely storage at 4 °C for longer than 2 weeks. Statistical testing, however, revealed that this did not produce a significant difference in the RPR-S results. RPR-C testing was performed on fresh sera as opposed to thawed samples for the batched RPR-S testing; ideally the RPR-S and -C testing would have been run in parallel to show the effect of the extended frozen storage period on the samples. Advanced statistical modelling was not performed to assess the possible role of biological bias, since the samples originated from the same individuals at multiple testing time points. Finally, despite efforts to transform the data, we could not achieve a sufficiently linear distribution (all RPR-C ranges) to warrant a Bland–Altman analysis, which is typically recommended for the comparison of two methods measuring the concentration of an analyte [37].

In this study, we implemented an automatic dilution of samples with an initial RU value of 8.0 as set by the company. Dilution at a cut-off of 4.0 RU would likely alleviate the hook effects, since the first-run analysis samples with suspected hook effect had a median initial value of 6.6 (IQR 5.9–7.2). Moreover, further dilution steps (higher than 10×) are likely warranted and adjusting the calibration curve to measure higher than 8.0 RU and quantitate higher than 4.0 RU could improve test performance. Further studies could also investigate whether lowering the RPR-S test cut-off level to less than 1.0 RU, which is closer to the limit of detection of 0.2 RU, would increase sensitivity, although this might increase the number of false positive results secondary to noise or cross-reactions at lower cut-off levels. For example, a study by Noh *et al.* implemented a 0.5 RU cut-off level and 45.4 % of positive results were rated as false positive compared to RPR-C/TPLA/EIA testing [19]. This is the first study investigating the testing efficacy of the RPR-S reagent on the SK500 clinical chemistry system. Further optimization of the platform settings may improve test performance.

Conclusions

This study exemplifies how an industry-independent evaluation of a diagnostic test can be beneficial for improving laboratory technology through independent testing of well-characterized clinical samples. Although we identified a number of important problems with the RPR-S test, our improved results with a dilution step at a lower cut-off value suggests a remedy for the suboptimal performance. Further test optimization is required by the manufacturer.

Funding information
This work was part of project ID 757003, funded by the Flemish Government-Department of Economy, Science and Innovation, through a grant to C.K.

Acknowledgements
We would like to thank the ITM laboratory teams for their involvement in the study and the authors gratefully acknowledge the individuals who participated in this study.

Conflicts of interest
The authors declare that there are no conflicts of interest.

Ethical statement
The Institutional Review Board of the Institute of Tropical Medicine (ITM) and the Ethics Committee of University Hospital Antwerp approved this study (13/44/426). Written informed consent was obtained from all participants upon study inclusion.

References

1. Newman L, Rowley J, vander Hoorn S, Wijesooriya NS, Unemo M *et al.* Global estimates of the prevalence and incidence of four curable sexually transmitted infections in 2012 based on systematic review and global reporting. *PLoS One* 2015;10:e0143304.
2. Apers L, Crucitti T, Verbrugge R, Vandenbruaene M. Sexually transmitted infections: what's new? *Acta Clin Belg* 2012;67:154–159.
3. Kenyon C, Lynen L, Florence E, Caluwaerts S, Vandenbruaene M *et al.* Syphilis reinfections pose problems for syphilis diagnosis in Antwerp, Belgium - 1992 to 2012. *Euro Surveill* 2014;19:20958.
4. Seña AC, White BL, Sparling PF. Novel *Treponema pallidum* serologic tests: a paradigm shift in syphilis screening for the 21st century. *Clin Infect Dis* 2010;51:700–708.
5. Matthews HM, Yang TK, Jenkin HM. Unique lipid composition of *Treponema pallidum* (Nichols virulent strain). *Infect Immun* 1979; 24:713–719.
6. Belisle JT, Brandt ME, Radolf JD, Norgard MV. Fatty acids of *Treponema pallidum* and *Borrelia burgdorferi* lipoproteins. *J Bacteriol* 1994;176:2151–2157.
7. Centers for Disease Control and Prevention (CDC). Discordant results from reverse sequence syphilis screening–five laboratories, United States, 2006–2010. *MMWR Morb Mortal Wkly Rep* 2011;60:133–137.
8. Lafond RE, Lukehart SA. Biological basis for syphilis. *Clin Microbiol Rev* 2006;19:29–49.
9. Seña AC, Zhang XH, Li T, Zheng HP, Yang B *et al.* A systematic review of syphilis serological treatment outcomes in HIV-infected and HIV-uninfected persons: rethinking the significance of serological non-responsiveness and the serofast state after therapy. *BMC Infect Dis* 2015;15:479.
10. Nandwani R, Evans DT. Are you sure it's syphilis? A review of false positive serology. *Int J STD AIDS* 1995;6:241–248.
11. Hernández-Aguado I, Bolumar F, Moreno R, Pardo FJ, Torres N *et al.* False-positive tests for syphilis associated with human immunodeficiency virus and hepatitis B virus infection among intravenous drug abusers. Valencian study group on HIV epidemiology. *Eur J Clin Microbiol Infect Dis* 1998;17:784–787.
12. Liu F, Liu LL, Guo XJ, Xi Y, Lin LR *et al.* Characterization of the classical biological false-positive reaction in the serological test for syphilis in the modern era. *Int Immunopharmacol* 2014;20:331–336.
13. Liu LL, Lin LR, Tong ML, Zhang HL, Huang SJ *et al.* Incidence and risk factors for the prozone phenomenon in serologic testing for syphilis in a large cohort. *Clin Infect Dis* 2014;59:384–389.
14. Haslett P, Laverty M. The prozone phenomenon in syphilis associated with HIV infection. *Arch Intern Med* 1994;154:1643–1644.
15. Jurado RL, Campbell J, Martin PD. Prozone phenomenon in secondary syphilis. *Arch Intern Med* 1993;153:2496.
16. Gupta SM, Bala M, Muralidhar S, Ray K. Evaluation of test results of microbiology laboratories of North India for standard tests for syphilis under an external quality assurance scheme. *Eur J Clin Microbiol Infect Dis* 2009;28:461–468.

17. **Müller I, Brade V, Hagedorn HJ, Straube E, Schörner C *et al.*** Is serological testing a reliable tool in laboratory diagnosis of syphilis? Meta-analysis of eight external quality control surveys performed by the german infection serology proficiency testing program. *J Clin Microbiol* 2006;44:1335–1341.
18. **Huh HJ, Chung JW, Park SY, Chae SL.** Comparison of automated treponemal and nontreponemal test algorithms as first-line syphilis screening assays. *Ann Lab Med* 2016;36:23–27.
19. **Noh J, Ko HH, Yun Y, Choi YS, Lee SG *et al.*** Evaluation of performance and false positivity of Mediace RPR test that uses a chemistry autoanalyzer. *Korean J Lab Med* 2008;28:312–318.
20. **Kim YS, Lee J, Lee HK, Kim H, Kwon HJ *et al.*** Comparison of quantitative results among two automated rapid plasma reagin (RPR) assays and a manual RPR test. *Korean J Lab Med* 2009;29:331–337.
21. **Workowski KA, Berman S, Centers for Disease Control and Prevention (CDC).** Sexually transmitted diseases treatment guidelines, 2010. *MMWR Recomm Reports* 2010;59:1–110.
22. **French P, Gomberg M, Janier M, Schmidt B, van Voorst Vader P *et al.*** IUSTI: 2008 European guidelines on the management of syphilis. *Int J STD AIDS* 2009;20:300–309.
23. **Liu H, Rodes B, Chen CY, Steiner B.** New tests for syphilis: rational design of a PCR method for detection of *Treponema pallidum* in clinical specimens using unique regions of the DNA polymerase I gene. *J Clin Microbiol* 2001;39:1941–1946.
24. **Cohen J.** A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.
25. **Shapiro SS, Wilk MB.** An analysis of variance test for normality (complete samples). *Biometrika* 1965;52:591–611.
26. **Passing H, Bablok W.** A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, part I. *J Clin Chem Clin Biochem* 1983;21:709–720.
27. **Wilcoxon F.** Individual comparisons by ranking methods. *Biometrics Bulletin* 1945;1:80–83.
28. **Yates F.** Contingency tables involving small numbers and the $\chi 2$ test. *J R Stat Soc* 1934;1:217–235.
29. **Rolfs RT, Joesoef MR, Hendershot EF, Rompalo AM, Augenbraun MH *et al.*** A randomized trial of enhanced therapy for early syphilis in patients with and without human immunodeficiency virus infection. The syphilis and HIV study group. *N Engl J Med* 1997;337:307–314.
30. **Mchugh ML.** Interrater reliability: the kappa statistic. *Biochem Med* 2012;22:276–282.
31. **Kim YS, Lee J, Lee HK, Kim H, Kwon HJ *et al.*** Comparison of quantitative results among two automated rapid plasma reagin (RPR) assays and a manual RPR test. *Korean J Lab Med* 2009;29:331–337.
32. **Osato K, Nagao T, Inuzumi K, Araki H, Kawai K.** Clinical evaluation of latex agglutination test kits for detecting anti-syphilitic lipoidal antibodies and anti-treponemal antibodies. *Japanese J Sex Transm Dis* 2002;13:124–130.
33. **Onoe T, Honda M, Matsuo K, Sasaki H, Sawamura M *et al.*** Examination of the correlation between the manual and automated serological testing methods for syphilis. *J Dermatol* 2012;39:355–361.
34. **Kinjyo T, Nago T, Sakiyama K, Oshiro M, Nagamine T *et al.*** Laboratory -based evaluation of latex agglutination turbidimetric assay by Mediace RPR on P module of Hitachi Auto analyzer 7600 to quantitatively determine serum RPR antibody. *Japanese J Clin Lab Autom* 2005;30:257–262.
35. **Stamm LV.** Syphilis: antibiotic treatment and resistance. *Epidemiol Infect* 2015;143:1567–1574.
36. **Ghinsberg RC, Nitzan Y.** Is syphilis an incurable disease? *Med Hypotheses* 1992;39:35–40.
37. **Bland JM, Altman DG.** Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–310.