



Accuracy of Rapid Point-of-Care Diagnostic Tests for Hepatitis B Surface Antigen—A Systematic Review and Meta-analysis

Mehnaaz S. Khuroo*, Naira S. Khuroo†, Mohammad S. Khuroo†

*Government Medical College Srinagar and †Digestive Diseases Centre, Dr Khuroo Medical Clinic, Srinagar, Kashmir, Jammu & Kashmir 190010, India

Background: Rapid point-of-care tests provide plausible diagnostic strategy for hepatitis B surface antigen (HBsAg) in low resource areas. However, their utility depends upon their overall performance. Our objective was to meta-analyze the diagnostic accuracy of rapid point-of-care tests for HBsAg. **Methods:** Literature search was done with the help of a metasearch engine Mettā, a query interface for retrieving articles from five leading medical databases. Studies that employed rapid point-of-care tests for detection of HBsAg and compared the results with reference test were included. Two reviewers performed quality assessment of the studies and extracted data for estimating test accuracy. Twenty-seven studies were meta-analyzed and stratified by multiple parameters. **Results:** Twenty-seven studies had evaluated 49 test brands and generated 76 data points. Sensitivity of individual tests varied widely and were heterogeneous (range 43.5%–99.8%); while specificity estimates were more robust and close to 100% (range 90%–100%). Overall pooled sensitivity, specificity, positive likelihood ratio (LR), negative LR and diagnostic odds ratio for all tests were 97.1% (95% CI, 96.1%–97.9%), 99.9% (CI, 99.8%–100%), 118.4 (CI, 84.7–165.5), 0.032 (CI, 0.023–0.045) and 4094.7 (CI, 2504.1–6600.8) respectively. This suggested high pooled accuracy for all studies. We found substantial heterogeneity between studies. Three factors (study location, reference standard and study score) appeared most strongly associated with test estimates and observed heterogeneity. The Determine test showed consistency in performance in studies done across developed and developing countries and the Determine and the BinaxNOW tests had significantly higher estimates than pooled estimates of remaining tests. Tests revealed analytical sensitivity of 4 IU/ml against manufacturer's claim of 0.5 IU/ml; reduced sensitivity with HBsAg mutants and poor performance in seroconversion panels. **Conclusions:** Tests with better analytical sensitivity need to be developed and their feasibility and outcomes in various clinical settings need to be addressed. (J CLIN EXP HEPATOL 2014;4:226–240)

Hepatitis B infection is a global health problem.¹ There are more than 350 million carriers of hepatitis B virus (HBV) in the world today, with addition of 4 million acute infections every year. Around one million die annually from hepatitis B-associated liver dis-

ease including hepatocellular carcinoma. Majority of HBV (over 75%) infections occur in high endemic regions namely Southeast Asia and Africa, with lifetime risk of infection from 60% to 80%. Areas of intermediate risk include parts of southern and Eastern Europe, the Middle East, and the Indian subcontinent with a lifetime risk of infection from 20% to 60%. In the United States, around 1.25 million individuals are chronically infected with HBV with an estimated 78,000 cases of new HBV infections occurring yearly, mostly among sexually active young adults, black and Hispanic people and IV drug users. Within this context, diagnosis of HBV infection, initiation of treatment, and prevention of transmission to susceptible population is a priority.²

Hepatitis B surface antigen (HBsAg) is the serologic hallmark of HBV infection and has been the principal target for laboratory testing to identify active infection by HBV.^{3,4} Additional serologic and molecular tests need to be done depending upon the setting and purpose of HBV testing.⁵ HBsAg is typically detected by a sensitive immunoassays that uses antibody to hepatitis B surface

Keywords: HBsAg, rapid test, performance evaluation, meta-analysis

Received: 3.6.2014; Accepted: 25.7.2014; Available online: 28.8.2014

Address for correspondence. Mohammad S. Khuroo, Director, Digestive Diseases Centre, Dr. Khuroo's Medical Clinic, Srinagar, Kashmir, Jammu & Kashmir 190010, India. Tel.: +91 194 2491190, +91 9906591044 (mobile)
E-mails: khuroo@yahoo.com; mohammad.khuroo@gmail.com

Abbreviations: DOR: diagnostic odds ratio; EIA: enzyme immunoassay; ELISA: enzyme-linked immunosorbent assay; HBsAg: hepatitis B surface antigen; HSROC: hierarchical summary receiver operating characteristic; LR: likelihood ratio; MEIA: micro-particle enzyme immunoassay; PCR: polymerase chain reaction; PRISMA: preferred reporting items for systematic reviews and meta-analysis; QUADAS: quality assessment of diagnostic accuracy of studies; RPOCTs: rapid point-of-care tests; SROC: summary receiver operating characteristic; STARD: standard for reporting diagnostic accuracy studies

<http://dx.doi.org/10.1016/j.jceh.2014.07.008>

antigen (anti-HBs) to capture antigen in the sample. Although immunoassays effectively detect HBsAg, the tests have many limitations in poor resource endemic regions of the world. These include high facility cost, need for sophisticated equipment and trained technicians, continuous supply of electricity, and long turnaround times. Point-of-care testing offers significant advantages which include reduction of facility costs, rapid delivery of results, early diagnosis and rapid initiation of treatment.⁶⁻⁸

Since 1990s, several rapid point-of-care tests (RPOCTs) that primarily use blood samples to test for HBsAg have been developed.⁹ The manufacturers claim high accuracy and utility of their respective tests. However, several vital questions about their use remain unanswered namely accuracy of individual tests, comparative efficacy of the different tests, and applicability of these tests in different scenarios of HBV evaluation.

A recent meta-analysis on accuracy of RPOCTs for HBsAg attempted to address the above mentioned questions.¹⁰ However, this study had many limitations. Two well-conducted major studies done on accuracy of RPOCTs and available by then were not included in the meta-analysis.^{11,12} These 2 studies had evaluated 27 tests against two international sera panels and generated 29 data points. There were errors in data entry in 2×2 tables. Khan et al¹³ while evaluating OneCheck test reported 19 of the 19 negative sera as true negative (specificity of 100%) and not as 19 false positive (specificity 0%) as shown in the study. Authors did not address the issue of inconclusive test results and how to analyze and report these data. We believe complete transparency regarding the handling of inconclusive results in the analysis phase is essential for the reader to understand how key summary statistics have been derived.¹⁴ Analytical sensitivity of the tests based on low titer and seroconversion panels performed in several included studies¹⁵⁻¹⁷ were not evaluated, which affected the conclusions made on the accuracy of the tests under consideration. Authors did not evaluate heterogeneity (differences in reported estimates among studies) and its potential sources, an important component in meta-analysis studies.¹⁸ We believe recommendations on use of RPOCTs can have far reaching effects on the healthcare in developing countries. For example recommending tests with low analytical sensitivity in blood banks can pose a serious threat to recipients as infected otherwise healthy donors often have low titer HBsAg viremia.¹⁹ Keeping the above in consideration, we conducted another systematic review and meta-analysis of studies pertaining to diagnostic accuracy and applicability of RPOCTs for HBsAg.

METHODS

Protocol

We sought to conduct a systematic review and meta-analysis on studies which evaluated accuracy of RPOCTs for HBsAg.

A protocol was written that specified several aspects of the meta-analysis (available at our website www.drkhuroo.com; Appendix Document 1) as per PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analysis) guidelines.²⁰ Two independent reviewers made literature search, performed quality assessment of the included studies and extracted data for estimating test accuracy.²¹ Any discrepancies were referred to third reviewer.

Acquisition of Data

Literature search was done with the help of a metasearch engine "Mettā" which can be accessed at <http://mengs1.cs.binghamton.edu/metta/search.action>. Mettā is a query interface that helps systematic reviewers in retrieving, filtering and assessing articles from five leading medical databases: PubMed, EMBASE, CINAHL, PsycINFO, and the Cochrane Central Register of Controlled Trials.²² De-duplication of records returned from multiple databases is carried out in a prioritized fashion that favors retaining citations returned from PubMed. Upto several thousand retrieved records can be exported within a few minutes and the process of writing systematic reviews is accelerated. The articles can be exported in XML format for further processing by a series of other computer-based tools, or exported in BibTeX format for users who employ reference manager software.

MeSH terms used for key and text word searching included "Hepatitis B" OR "Hepatitis B Surface Antigens" AND "Point-of-Care Systems" OR "rapid test" OR "diagnostics" AND "Sensitivity and Specificity" OR "diagnostic accuracy" OR "validity". We also searched bibliographies and reference lists of eligible papers and related reviews, consulted experts in the field, and contacted several authors from the included studies to locate additional studies. Titles and abstracts of all the above articles identified in the primary search were evaluated and a list of potential eligible studies identified. These studies were considered for full text review. Studies which fulfilled the criteria for selection were included in the systematic review and meta-analysis.

Criteria for Study Inclusion

Studies were included that employed RPOCTs for detection of HBsAg (Index test) and compared the results with a reference standard and reported results to recreate the 2×2 diagnostic table for estimating test accuracy. Index test was defined as any commercially available assay that identified HBsAg at or near the site of patient care. The test had to have quick turnaround time; ease of sampling, performance and reading results and no requirement of cold chain and specialized equipment. The test results are available to the patient, physician and care team within minutes, which allows for clinical management decisions to be made in the same clinical encounter.⁸ Acceptable reference standards included enzyme immunoassay (EIA),

enzyme-linked immunosorbent assay (ELISA) or micro-particle enzyme immunoassay (MEIA) for detection of HBsAg. Additional tests (confirmatory neutralization test for HBsAg positive sera, HBsAg quantification and detection of HBV DNA) were considered to make the reference standard as perfect. We included studies conducted in adults (age > 18 years), published both as abstracts and full-text articles, using all study-designs, conducted in any study settings (laboratory or field-based) and regardless of sample size, study location, language of publication, and country of origin of test. We excluded studies which dealt with accuracy of laboratory-based tests; studies with data unable to recreate 2×2 diagnostic table; reports from the manufacturers and package inserts which are subjected to overt conflict of interest, and duplicate reports.

Data Extraction

Each study was subjected to following search: study author, year of publication, location of study, index test (one or more), reference standard, study design, source of sera, sample size, characteristics of the population employed for sera collection, cross reactive sera included in panel and analytical sera included for evaluating test sensitivity. Detailed information about the index test were extracted from the studies which included: name of the test, country of origin and name of the manufacturer, time taken to read results, specimen (serum, plasma or blood) needed for test, volume of the sample (μ l) needed to test, storage conditions for maintaining test kit, special equipment if any needed to perform the test and shelf life of the test kit. For purposes of data synthesis we extracted raw cell numbers namely true positives, false negatives, false positive and true negatives for each test run.

Quality Assessment

We performed quality assessment of the studies using QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) tool²³ and the STARD (Standards for the Reporting of Diagnostic Accuracy Studies) checklists.²⁴ QUADAS-2 sheet was completed by following stepwise guidelines to judge risk of bias (4 domains) and concerns about applicability (3 domains) for each study. STARD checklist consisted of 25 questions and each question was weighted equally (yes = 1, No = 0) and total score for each study calculated.

Statistical Analysis and Data Synthesis

Data was extracted to construct 2×2 tables (reference test results vs. index test results). We defined HBsAg positive as those with disease and HBsAg negative as those without disease based on the reference test results. Index test results were reported as true positive, false positive, false negative and true negative. To maintain transparency of the test results reporting, we searched for inconclusive index test re-

sults. We believe inconclusive test results should be analyzed depending on how these patients would be treated in the clinical context. For the present meta-analysis all valid inconclusive index test results were clubbed with either false negative results (sera which were HBsAg positive by reference test and inconclusive by index test) or false positive results (sera which were HBsAg negative by reference test and inconclusive by index test) (Appendix Document 2).¹⁴

For estimates of accuracy we calculated sensitivity & specificity, positive & negative likelihood ratio (LR) and diagnostic odds ratio (DOR) along with 95 percent confidence intervals (CIs) (see Appendix Document 2). Sensitivity of a test is defined as the probability that the index test result will be positive in a case with disease. Specificity of a test is defined as the probability that the index test result will be negative in a non-diseased case. Both Sensitivity, & Specificity were expressed as percentages. LR is used to update the pre-test probability of disease using Bayes' theorem, once the test result is known. The updated probability is referred to as the post-test probability. For a test that is informative, the post-test probability should be higher than the pre-test probability if the test result is positive, whereas the post-test probability should be lower than the pre-test probability if the test result is negative. Positive LR of >10 and negative LR of <0.1 gives conclusive evidence of test accuracy. DOR, commonly considered a global measure of test performance, was estimated by the ratio of the odds of a positive test result in a diseased subjects to odds of a positive test result in a non-diseased subjects. DOR describes how many times higher the odds are of obtaining a test positive result in a diseased rather than a non-diseased person. DORs summarizes test accuracy in a single number and makes it easy to use this in subgroup/meta-regression model to derive statistical values.

We pooled test estimates using a bivariate random-effects regression model.²⁵ The bivariate model takes in to consideration the potential tradeoff between sensitivity and specificity by explicitly incorporating this negative correlation in the analysis. The model was used to draw hierarchical summary-operating characteristic (HSROC) curves. The curve of a test is the graph of the values of sensitivity and specificity that are obtained by varying the positivity threshold across all possible values. The graph plots sensitivity (true-positive rate) against 1-specificity (false-positive rate). The shape of the curve between these two fixed points depends on the discriminatory ability of the test. The closer the curve is to the upper left hand corner of the curve plot, the better the overall accuracy of the test.

Heterogeneity in meta-analysis refers to the degree of variability in the results across studies.²⁶ Diagnostic accuracy studies are expected to show considerable heterogeneity and models used are by default random-effects models, taking in to account study variation as well as chance variation. To detect heterogeneity test estimates were graphically

displayed using forest plots and overlap in the confidence intervals of individual studies compared. Poor overlap suggested statistical heterogeneity. *Q* test for heterogeneity with *P* value was calculated. A large *Q* value with *P* < 0.10 (rather than conventional 0.05) suggested heterogeneity. Finally the inconsistency index (*I*²) test was calculated. We regarded values of ≤25%, >25–50%, >50–75% and >75% as low, moderate, substantial and considerable statistical heterogeneity, respectively (See Appendix Document 2).

To investigate the sources of heterogeneity the following variables were selected a priori as potential sources of heterogeneity: year of publication (prior vs. after 2005), location of study (developed vs. developing countries), reference standard (EIA alone vs. EIA with neutralization test or nucleic acid test), study design (cross sectional vs. case control), source of sera (blood banks vs. hospitals/clinics), cross reactive sera (included vs. not included in sera panel) and study quality (high >15 vs. low ≤15 STARD score). Summary estimates for each covariate were generated, along with their 95% CIs and compared in the meta-regression model. A *P* value below 0.05 for DORs was defined as significant difference among the levels of particular covariate. Because the effects of some of the covariates may influence each other, multivar-

iate meta-regression was also done to take in to account the possible interrelations among the variables.²⁷

In order to compare relative efficacy of tests, we appreciated the studies had included different test brands and assays were conducted in different study populations and under varied laboratory conditions. Consequent to this, the accuracy of different assays varied markedly within and between different tests. To overcome this, the studies were stratified in to 2 subgroups based on location where studies had been performed (developed vs. developing countries). Pooled estimates of tests that had acquired ≥3 data points (Determine, Hepacard, BinaxNOW, Genedia, and SD) were calculated across and within the 2 subgroups and compared in the meta-regression model.

We did all statistical analysis in Meta-Analyst (Tufts Medical Centre, Boston, MA).²⁸

RESULTS

Literature Search

We identified 29 articles eligible for inclusion (Figure 1).^{11-13,15-17,29-51} Data from 2 studies had duplicate publications. Official report of International Consortium for Blood Safety 2007 (ICBS,¹²) was

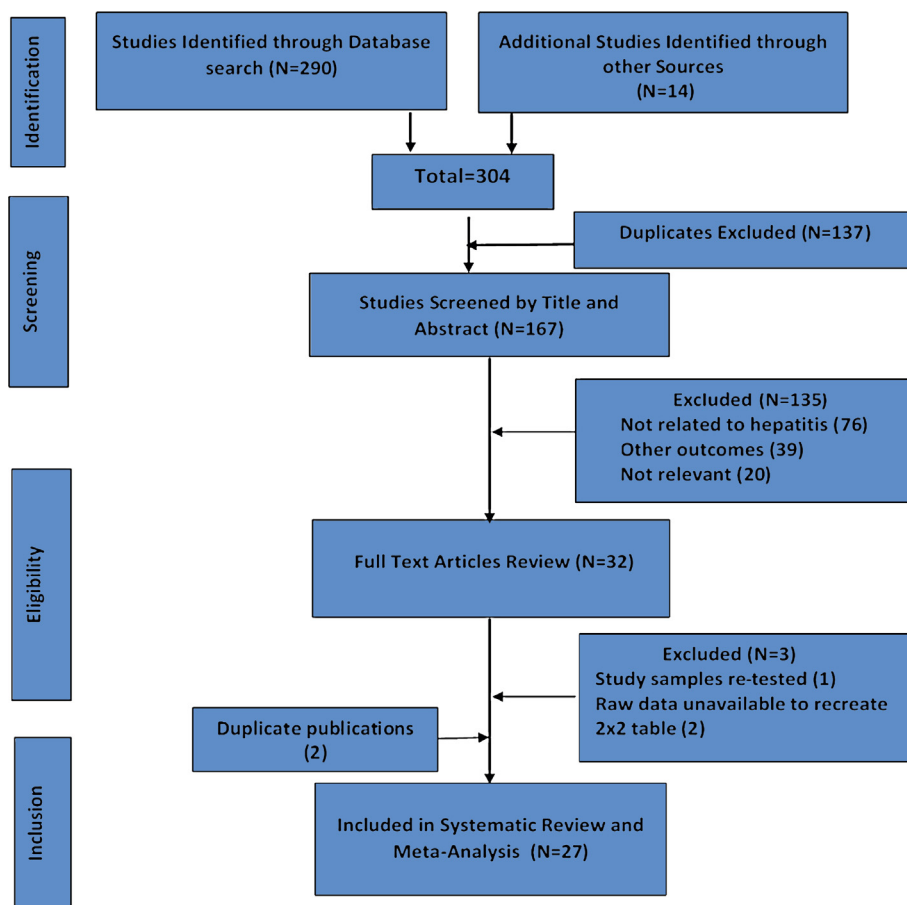


Figure 1 Flow diagram of study selection.

Table 1 Characteristics of the 27 Studies^a.

Author (Reference)	Year	Location	Index test	Reference standard	Study design	Source	Size	Patient population	CRS ^b	Analytical analysis
Nakata ²⁹	1990	NK	Hybritch	EIA	Cross sectional	Prison inmates	300	NK	No	No
Sato ¹⁵	1996	Japan	Dainascreen	EIA	Case control	Hospital samples	462	NK	NK	Yes; Dainabot panel
Mvere ³⁰	1996	Zimbabwe	SimpliRed, Dipstick (PATH)	EIA	Cross sectional	Blood Bank	206	NK	No	No
Abraham ³¹	1998	India	Virucheck, QuickChaser	ELISA	Cross sectional	NK	400	NK	No	No
Oh ³²	1999	Korea	Genedia	EIA	Case control	Blood donors	250	NK	NK	No
Palmer ³³	2000	Honduras, Dominican Republic, Trinidad, Jamaica	Determine	EIA	Case control	NK	298	NK	No	No
Kaur ³⁴	2000	India	Hepacard	EIA	Cross sectional	Hospital samples	2754	NK	Nil	No
Lien ³⁵	2000	Vietnam	Dainascreen, Determine, Serodia	EIA	Case control	High risk volunteers	328	NK	Yes, 128 sera	No
WHO ¹¹	2001	Global	Ten tests ^g	EIA ^c	Serum panel study (Case control)	Africa, Asia, Europe, Latin America	277	Well described	No	Low titer panel, Seroconversion panel
Raj ³⁶	2001	India	Hepacard	EIA/MEIA	Cross sectional	Hospital samples	1000	NK	No	No
Clement ¹⁶	2002	Belgium	BinaxNOW	MEIA	Cross sectional	Hospital samples	942	Chronic hepatitis	No	in-house standards sera panel
Lau ³⁷	2003	USA	BinaxNOW	EIA	Cross sectional	Criminals, Hepatology clinic, Health fair,	2463	NK	Fresh & frozen sera	No
Whang ³⁸	2005	Korea	Daewoong, Genedia	MEIA	Case control	NK	200	NK	No	No
Cha ³⁹	2006	Korea	Asan, Genedia, SD	MEIA	Case Control	NK	40	NK	No	No
Akanmu ⁴⁰	2006	Nigeria	Binax Now	ELISA	Cross sectional	Donors	101	HBV carriers	No	No
Ansari ⁴¹	2007	Iran	6 tests ^h	EIA/PCR	Case control	Hospital Patients	240	NK	No	No

ICBS ¹²	2007	Global	19 tests ⁱ	ICBS panel testing ^f	Serum panel study (Case control)	ICBS master panel ^f	346	NK	No	ICBS Quantitative panel
Lin ¹⁷	2008	China, Guinea	Determine, DRW	EIA	Case Control	Blood donors (Stanford & Ghana)	1250	NK	No	WHO sera, Seroconversion panel
Randrianirina ⁴²	2008	Madagascar	Cypress, Determine, Hexagon, Virucheck	EIA	Case control	NK	200	NK	No	No
Torane ⁴³	2008	India	Hepacard	ELISA	Case control	Blood bank	60	NK	No	No
Ola ⁴⁵	2009	Nigeria	GWHB	ELISA	Cross sectional	Medical clinic	25	NK	No	No
Davies ⁴⁶	2010	Malawi ^e	Determine	EIA ^c	Cross sectional	HIV clinic	75	NK	HIV	No
Khan ¹³	2010	Pakistan	OneCheck, Accurate	ELISA	Case control	NK	57	NK	No	No
Maity ⁴⁸	2012	India	Hepacard, Span, SD	ELISA ^c	Case control	NK	100	NK	No	No
Bottero ⁴⁹	2013	France	Determine, QuickProfile VIKIA	ELISA ^d	Cross sectional	Health care centers	3956	NK	No	No
Franzeck ⁵⁰	2013	Tanzania	Determine	EIA	Cross sectional	HIV Clinic	272	Chronic carriers	HIV	No
Chemeera ⁵¹	2013	Sri Lanka	Cortez Onsite	EIA	Cross sectional	Hospital samples	50	NK	No	No

NK = Not known or not specified in the report, EIA = Enzyme immunoassay, MEIA = Micro-particle enzyme immunoassay, ELISA = Enzyme-Linked Immunosorbent Assay, PCR = Polymerase Chain Reaction.

^a2 studies^{44,47} were excluded for duplicate publications.

^bCRS = Cross reactive sera included in serum panel.

^cNeutralization test for all positive sera performed.

^dAll positive sera had HBsAg quantification and HBV DNA quantification.

^eSera collected in Malawi and tested in UK.

^fICBS clinical master panel sera were collected in blood banks from diverse regions of world and have been extensively analyzed by serology, HBV DNA, HBV genotyping, HBV subtyping and sequencing.

^gWHO 10 tests: 1. Advanced, 2. Determine, 3. DoubleCheck, 4. Genedia, 5. HEP B STAT-PAK, 6. Hepacard, 7. Immunocomb II, 8. Rapid, 9. Serodia, 10. Unigold.

^hAnsari 6 tests: 1. Acon; 2. Atlas; 3. Intec, 4. Blue Cross, 5. DIMA, 6. Cortez.

ⁱICBS 19 tests: 1. Acon, 2. Assure, 3. Boline, 4. Cassette, 5. Cypress, 6. Determine, 7. Equipar, 8. HBsAg (WB), 9. Hepacard, 10. Hep-Alert-B, 11. HepaScan, 12. Hep-Check-1 Miniclip, 13. Hep-Check-1 Strip, 14. i + LAB, 15. Line test, 16. Newmarket, 17. QuickChaser, 18. QuickTest, 19. SD.

published as a full-text paper in a peer-reviewed journal.⁴⁷ Sera from another study were collected and tested in Malawi⁴⁴ and had shown conflicting results. The same sera were retested in UK and results published as a full-text paper in a peer-reviewed journal.⁴⁷ Consequently to this we excluded 2 studies from further analysis.^{44,47} Of the 27 studies, 24 were published as full-length manuscripts in peer-reviewed journals, two studies were published as official reports of WHO & ICBS^{11,12} and two studies were reported as conference abstracts.^{29,33} There was unanimity between the authors about the selection of relevant articles.

Study Characteristics

Table 1 shows the study characteristics. Of the 27 studies, 20 were conducted in developing countries and 14 were case-control. Source of sera were from blood banks in 7 studies, hospitals and medical/surgical clinics in 13 studies, and from unknown source in 7 studies. Four studies had included cross reactive sera in sera panels.^{35,37,46,50} Four studies had described patient's characteristics.^{11,16,40,50} Sample size varied from 25 to 3956. For defining analytical sensitivity of the tests, 5 studies employed low titer sera panels^{11,12,15-17} and 1 included seroconversion sera panels.¹¹ None of the studies had included sera from general population, patients with acute hepatitis syndrome, fulminant hepatic failure or HBV infected patients on antiviral therapy.

Appendix Table 1 lists the index test characteristics. Twenty-seven studies had used 49 test brands and generated 76 data points. A single data point was generated for 37 tests; 2 data points for 7 test brands (Dianascree, ViruCheck, QuickChaser, Serodia, Cortez, Cypress & Acon), 3 data points for 2 test brands (BinaxNow, SD); 4, 6 and 9 data points for one test brand each (Genedia, Hepacard and Determine respectively). Reference standard in all studies was detection of HBsAg either by EIA, MEIA or ELISA. Neutralization test for HBsAg (CDC guideline for HBsAg testing by EIA), HBsAg quantification or molecular testing for HBV were employed in only 6 studies.^{11,12,41,46,48,49} All the studies administered the same reference test to all patients, thus avoiding partial or differential verification bias.

Study Quality

QUADAS-2 sheet shows major risk of bias in patient selection and index test (Table 2). Bias in patient selection resulted from: i) case-control study design, ii) poor description of patient's selection and clinical scenario. Bias in index test was mostly related to lack of reporting blinding while reading test results. The quality of study reporting ranged from poor to good (STARD scores from 8 to 23 of 25), with a number of items missing from reporting of diagnostic accuracy.

Conflict of Interest

Two studies reported a financial relationship with or received funding from the industry,^{17,35} 3 explicitly declared no conflict of interest,^{37,42,46} and 2 were independent evaluation by International Organizations.^{11,12} Most remaining studies omitted disclosure of conflict of interest.

Test Accuracy

Figure 2 reports on the estimates of sensitivity & specificity with 95% confidence intervals and related Forest plots stratified by type of test brand. Sensitivity of individual tests varied widely and were heterogeneous (range 43.5%–99.8%); while specificity estimates were more robust and close to 100% (range 90%–100%). Overall pooled sensitivity, specificity, positive LR, negative LR and DORs for all tests were 97.1% (95% CI, 96.1%–97.9%), 99.9% (CI, 99.8%–100%), 118.4 (CI, 84.7–165.5), 0.032 (CI, 0.023–0.045) and 4094.7 (CI, 2504.1–6600.8) respectively. This suggested high pooled accuracy for all studies. The ROC curve also indicated high sensitivity with a high specificity as the curve approached the upper left hand corner of the graph where sensitivity is 1 and specificity is 1 (Appendix Figure 1).

We found substantial heterogeneity between studies when calculating the pooled sensitivity ($I^2 = 83\%$), specificity ($I^2 = 83\%$), positive LR ($I^2 = 64\%$), negative LR ($I^2 = 66\%$) and DORs ($I^2 = 66\%$). Multiple covariates were used in subgroup analysis and tested statistically in meta-regression model to find out the reason/s of heterogeneity (Table 3). Three factors (location of study, reference standard and study score) appeared most strongly associated with test estimates and observed heterogeneity. In contrast study year (studies before & after 1995), study design (case control vs. cross sectional studies), source of sera (blood banks vs. hospital/clinics) and cross reactive sera (presence or absence of cross reactive sera in panels) had no significant effect on test performance and heterogeneity. Studies conducted in developed and developing countries both revealed high pooled accuracy values; however DORs were 4.5 times higher in studies conducted in the developed than in developing countries (DORs 9812.7 [CI, 6467.3–14888.7] vs. 1467.7 [CI, 762.8–2824.1] respectively; $P \leq 0.001$) (Figure 3). Studies conducted in developed countries had minimal heterogeneity ($I^2 = 0\%$), while those in developing countries showed substantial heterogeneity ($I^2 = 69\%$). Studies which used additional neutralization test with or without nucleic acid testing as a reference standard had higher DORs than studies which used enzyme immunoassays alone (DORs 7179.6 [CI, 4901.5–10497.3] vs. 1663.7 [CI, 745.3–3714.0] respectively; $P < 0.001$) and were less heterogeneous ($I^2 = 0\%$ vs. 76% respectively) (Appendix Figure 2). On a similar note, studies with STARD score of >15 had higher DORs than those with STARD score ≤ 15 (DORs 10414.9 [6464.8–16778.6] vs. 1838.4 [915.8–3690.2] respectively; $P < 0.001$)

Table 2 QUADAS-2 Assessments and STARD Scores.

Author (Reference)	QUADAS-2 sheet						STARD	
	Patient selection		Index test		Reference standard		Flow & timing	Score
	Risk of bias	Concern about applicability	Risk of bias	Concern about applicability	Risk of bias	Concern about applicability	Risk of bias	Total 25
Nakata ²⁹	Published as an abstract; data insufficient for evaluation							
Sato ¹⁵	UC	UC	HR	LR	LR	LR	LR	10
Mvere ³⁰	LR	LR	LR	LR	LR	LR	LR	13
Abraham ³¹	HR	LR	LR	LR	LR	LR	LR	9
Oh ³²	UC	LR	HR	LR	LR	LR	LR	9
Palmer ³³	Published as an abstract; data insufficient for evaluation							
Kaur ³⁴	LR	LR	HR	LR	LR	LR	LR	11
Lien ³⁵	LR	LR	HR	LR	LR	LR	LR	19
WHO ¹¹	LR	LR	LR	LR	LR	LR	LR	24
Raj ³⁶	HR	LR	HR	LR	LR	LR	LR	9
Clement ¹⁶	LR	LR	HR	LR	LR	LR	LR	17
Lau ³⁷	LR	LR	HR	LR	LR	LR	LR	10
Whang ³⁸	HR	LR	HR	LR	LR	LR	LR	10
Cha ³⁹	HR	LR	HR	LR	LR	LR	LR	10
Akanmu ⁴⁰	HR	LR	HR	LR	LR	LR	LR	15
Ansari ⁴¹	HR	LR	HR	LR	LR	LR	LR	12
ICBS ¹²	LR	LR	LR	LR	LR	LR	LR	23
Lin ¹⁷	UC	LR	HR	LR	LR	LR	LR	11
Randrianirina ⁴²	HR	LR	HR	LR	LR	LR	LR	9
Torane ⁴³	HR	LR	HR	LR	LR	LR	LR	8
Ola ⁴⁵	UC	LR	HR	LR	LR	LR	LR	9
Davies ⁴⁶	HR	LR	LR	LR	LR	LR	LR	18
Khan ¹³	HR	LR	LR	LR	LR	LR	LR	14
Maity ⁴⁸	HR	LR	HR	LR	LR	LR	LR	14
Bottero ⁴⁹	LR	LR	HR	LR	LR	LR	LR	13
Franzeck ⁵⁰	LR	LR	LR	LR	LR	LR	LR	20
Chemeera ⁵¹	LR	LR	HR	LR	LR	LR	LR	14

LR = low risk, HR = high risk, UC = unclear.

and were less heterogeneous [$I^2 = 0\%$ vs. 74%] respectively) (Appendix Figure 3). We performed multivariate meta-regression to take in to account the interrelations among these 3 variables associated with test estimates and heterogeneity. Study location was the only variable which was significantly associated with test estimates and heterogeneity (Appendix Table 2).

Comparative accuracy of test brands was a complex issue. Five test brands with sufficient data points (≥ 3) (Determine, Hepacard, Genedia, BinaxNow, and SD) showed grossly heterogeneous pooled estimates ($I^2 = 59\%$, 75% , 42% , 66% and 73% respectively) (Appendix Figure 4). Next we pooled estimates of these 5 test brands

separately for developed and developing countries. Pooled estimates of all test brands showed minimal heterogeneity for studies done in developed countries (I^2 for Determine 0% , Hepacard 0% and BinaxNow 0% respectively) (Appendix Figure 5) and substantial for studies done in developing countries (I^2 for Determine 70% , Hepacard 80% , Genedia 43% and SD 74% respectively) (Appendix Figure 6). Accuracy estimates of all the 5 test brands were 2.3–4.3 times higher in studies done in developed vs. developing countries. Determine test estimates were consistent and revealed high sensitivity ($98.0 [95.1-99.2]$ vs. $96.7 [95.1-97.8]$ respectively) and specificity ($98.9 [92.1-99.9]$ vs. $99.6 [97.0-99.9]$ respectively) across developed

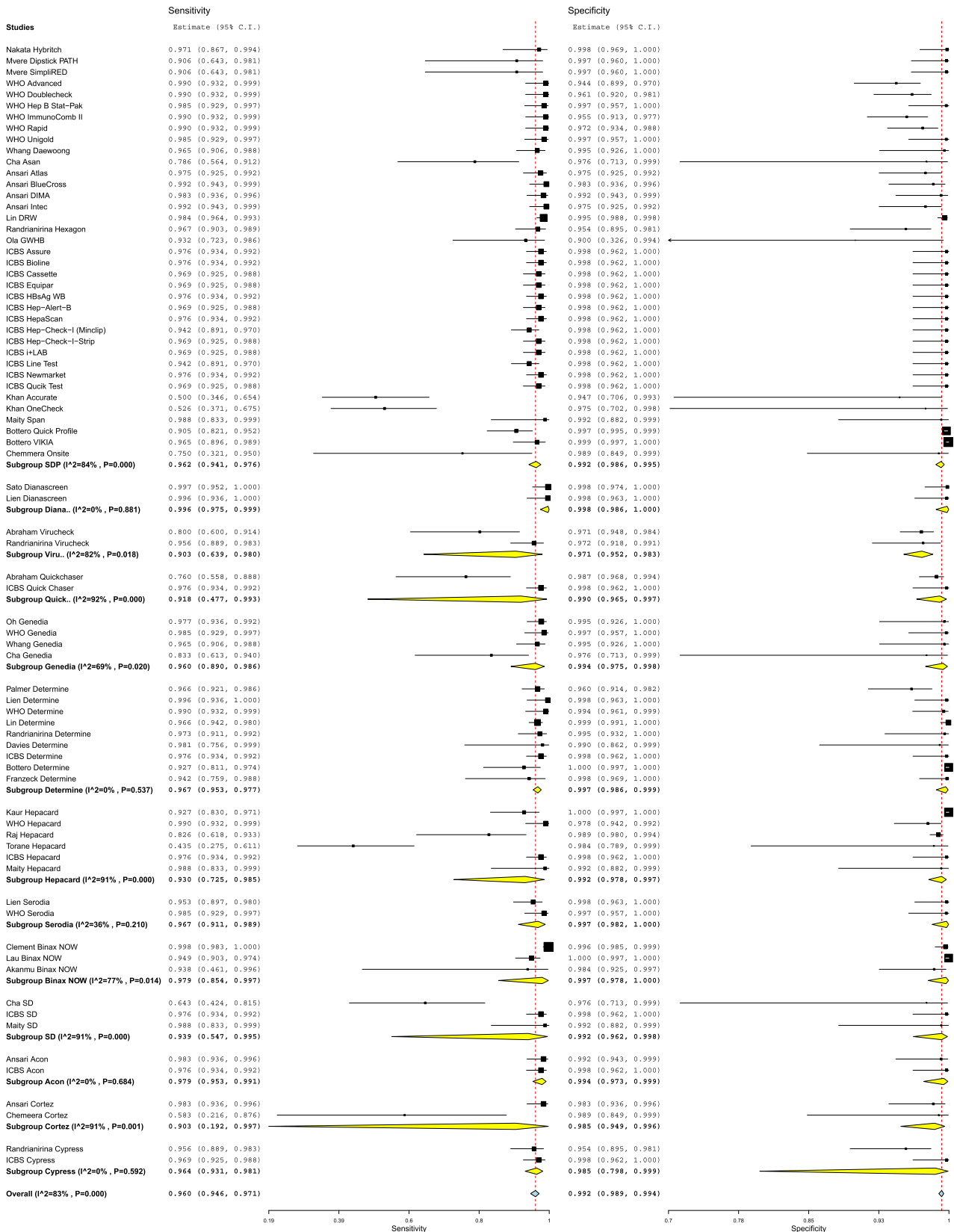


Figure 2 Forest plot of the sensitivity and specificity estimates and 95% confidence intervals (CIs) of 27 studies stratified by test brands. 37 test brands had generated single data point each (subgroup SDP) while 12 test brands had generated ≥ 2 data points each. Estimates of sensitivity and specificity from each study are shown as solid squares. Solid lines represent the 95% CIs. Squares are proportional to the weights based on the random effect model. Pooled estimates and 95% CIs is denoted by the diamond at the bottom. I^2 and p values represents heterogeneity of studies.

Table 3 Accuracy Estimates from Subgroup Analysis.

Variable	Subgroups (no of studies/data points)	Pooled sensitivity (95% CI), %	Pooled specificity (95% CI), %	Positive LR (95% CI)	Negative LR (95% CI)	DOR (95% CI)	$I^2_{\%}$	p Value ^a
All studies	27 studies/76 data points	97.1 (96.1–97.9)	99.9 (99.8–100)	118.4 (84.7–165.5)	0.032 (0.023–0.045)	4094.7 (2504.1–6600.8)	66%	–
Abstracts excluded	25 full text studies/74 data points	97.1 (96.0–97.9)	99.9 (99.8–100)	126.6 (90.2–177.8)	0.03 (0.02–0.04)	4327.2 (2753.2–6800.9)	66%	–
Year	Prior 2005 (12/25)	97.1 (95.2–98.2)	99.0 (98.4–99.4)	93.4 (56.8–153.6)	0.023 (0.012–0.044)	5931.9 (2393.4–14703.4)	74%	Reference
	After 2005 (15/51)	95.4 (91.1–97.5)	99.3 (98.9–99.5)	129.1 (84.4–196.1)	0.038 (0.029–0.050)	3501.5 (2016.6–6079.2)	59%	0.39
Location of Test ^b	Developing Countries (20/39)	94.3 (89.8–96.4)	99.1 (98.4–99.5)	72.2 (48.4–109.3)	0.049 (0.032–0.077)	1467.7 (762.8–2824.1)	69%	Reference
	Developed Countries (6/36)	97.0 (96.2–97.7)	99.4 (99.0–99.6)	193.5 (115.8–323.6)	0.018 (0.012–0.036)	9812.7 (6467.3–14888.7)	0%	<0.001
Reference Test	EIA/ELISA/MEIA alone (21/34)	94.1 (90.6–96.3)	98.9 (98.4–99.3)	82.9 (55.5–123.7)	0.060 (0.037–0.098)	1663.7 (745.3–3714.0)	76%	Reference
	Neutralization/NAT test (6/42)	97.0 (96.4–97.5)	99.5 (99.1–99.7)	185.5 (104.5–329.2)	0.018 (0.014–0.023)	7179.6 (4901.5–10497.3)	0%	<0.001
Study design	Cross sectional (13/18)	95.4 (93.3–96.8)	99.3 (98.8–99.8)	129.9 (84.9–196.7)	0.078 (0.043–0.140)	3111.7 (984.6–9833.7)	81%	Reference
	Case control (14/58)	97.1 (95.2–98.2)	99.0 (98.4–99.4)	93.9 (56.4–153.2)	0.021 (0.016–0.026)	4373.0 (2656.9–7197.7)	54%	0.63
Source of sera ^c	Hospital/Clinics (13/23)	96.2 (94.0–97.6)	99.5 (99.1–99.7)	188.0 (103.1–344.0)	0.027 (0.014–0.60)	7052.4 (3063.5–16238.3)	64%	Reference
	Blood Banks (7/36)	97.1 (95.9–97.9)	99.4 (99.0–99.6)	157.2 (94.9–260.4)	0.015 (0.012–0.020)	9974.1 (5115.8–12429.4)	0%	0.36
CRS \ddagger	No CRS \ddagger (23/70)	95.9 (94.4–97.1)	99.1 (98.8–99.4)	107.6 (76.6–151.2)	0.033 (0.024–0.047)	3619.5 (2217–5908.4)	66%	Reference
	HIV Clinics with CRS \ddagger (4/6)	96.2 (93.1–97.9)	99.8 (99.4–99.9)	493.6 (1594.7–1527.9)	0.019 (0.006–0.059)	24709.4 (6264.0–97467.0)	0%	0.18
Score	≤ 15 (19/39)	93.7 (89.9–96.2)	99.0 (98.5–99.4)	91.5 (58.6–142.8)	0.053 (0.035–0.081)	1838.4 (915.8–3690.2)	74%	Reference
	> 15 (6/35)	97.0 (96.5–97.5)	99.5 (99.1–99.7)	185.2 (107.8–318.3)	0.012 (0.009–0.016)	10414.9 (6464.8–16778.6)	0%	<0.001

LR = Likelihood ratio, DOR = Diagnostic Odds Ratio, $I^2_{\%}$ = inconsistency index test for heterogeneity (<25 = low, >25–50% = moderate; >50–75% = substantial & >75% = considerable statistical heterogeneity), CRS \ddagger = cross reactive sera.

^a p Value determined from meta-regression model.

^bLocation of one study (Reference 29) not known.

^cSource of sera of 7 studies (Ref 20, 22, 29, 30, 34, 40, 41) not known.

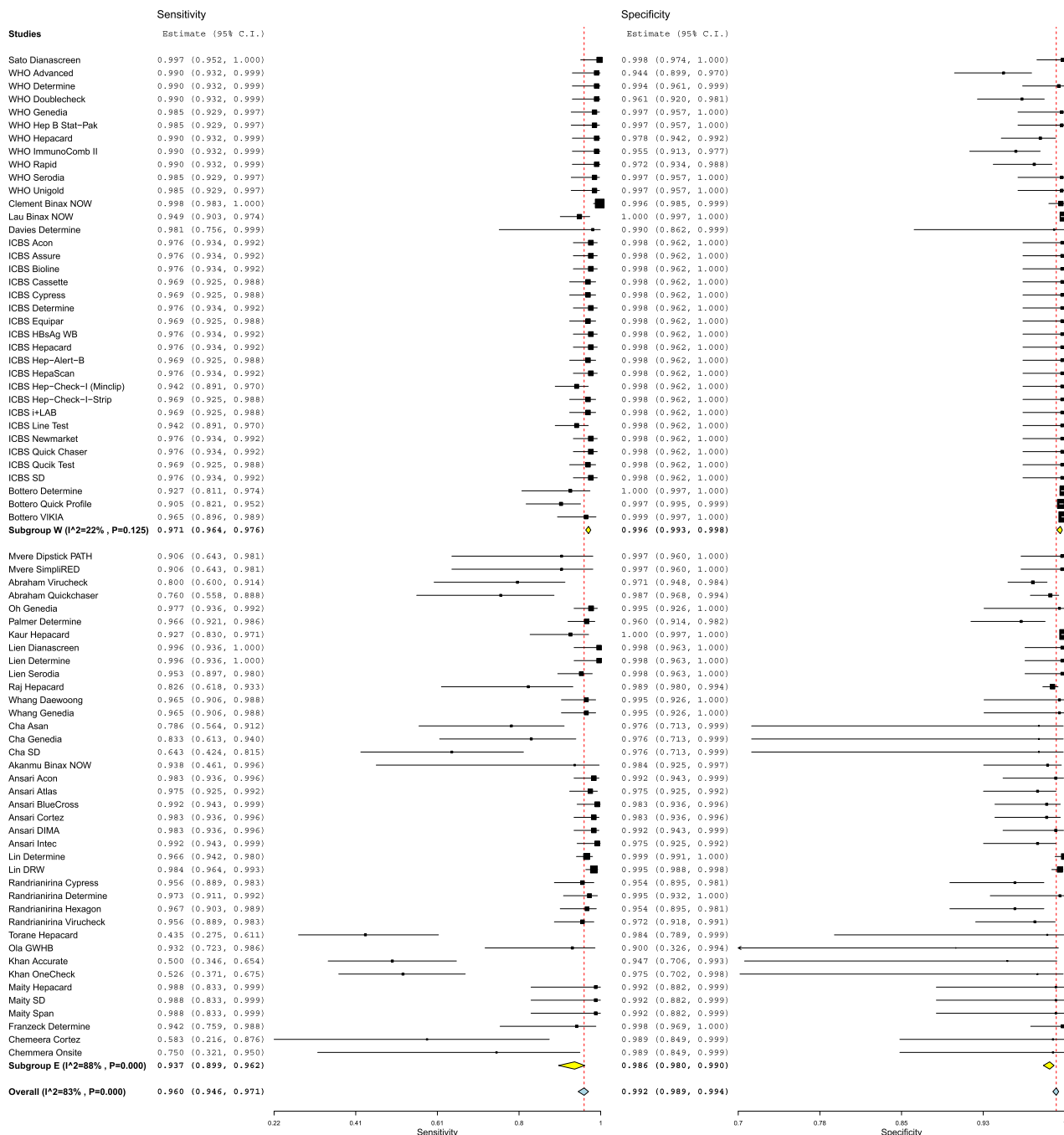


Figure 3 Forest plot of the sensitivity and specificity estimates and 95% confidence intervals (CIs) stratified by study location. 7 studies generating 36 data points were conducted in developed countries (Subgroup W) and 19 studies generating 39 data points were conducted in developing countries (Subgroup E). Estimates of sensitivity and specificity from each study are shown as solid squares. Solid lines represent the 95% CIs. Squares are proportional to the weights based on the random effect model. Pooled estimates and 95% CIs is denoted by the diamond at the bottom. I^2 and p values represents heterogeneity of studies.

and developing countries. Hepacard test showed inconsistency in pooled sensitivity (98.0 [CI, 95.1–99.2] vs. 85.7 [CI, 51.6–97.1] respectively) in studies between developed countries and developing countries. In contrast pooled specificity of Hepacard was consistent and high in studies done both developed vs. developing countries (98.9 [CI,

92.1–99.9] vs. 99.5 [CI, 97.2–99.9] respectively) (Appendix Table 3).

Next we compared the DOR's of these 5 tests and compared these with pooled DOR's of remaining 37 tests with single data points (reference tests) in the meta-regression model. DORs of Determine test performances

were significantly higher in studies done in developed (DORs 19971.5 [CI, 4263.2–93550.1] vs. 8169.7 [CI, 5155.7–12945.7]; $P = 0.01$) as well as in developing countries (8505.5 [993–72849.4] vs. 1116.9 [505.3–2468.8]; $P = 0.03$). BinxNow also showed significantly higher test estimates in studies done in developed countries (DORs 97845.1 [CI, 155549.6–615683.0] vs. 8169.7 [CI, 5155.7–12945.7]; $P = 0.007$). DOR's of SD, Hepacard, and Genedia did not differ from the DOR's of the reference tests in either developed or developing countries (Appendix Table 3).

Test brands with one or 2 data points could not be evaluated in the meta-analysis statistical model. Of these 42 such test brands, 7 tests had acquired extremely poor sensitivity estimates: ViruCheck (79.0 [CI, 71.7–83.2]), Quick-Chaser (77.0 [CI, 71.7–83.2]), Asan (78.6 [CI, 56.4–91.2]), Accurate (50.0 [CI, 34.6–65.4]), OneCheck (52.6 [CI, 37.1–67.5]), Cortez (58.3 [CI, 21.6–87.6]), and Onsite (75.0 [CI, 32.1–95.0]). In contrast, the specificity of these tests was consistency high (97.0 [CI, 96.5–99.9], 99.0 [CI, 96.5–99.9], 97.6 [CI, 71.3–99.9], 94.7 [CI, 70.6–99.3], 97.5 [CI, 70.2–99.8], 98.9 [CI, 84.9–99.9], 98.9 [CI, 84.9–99.9] respectively).

Analytical Analysis

Five studies used panels with low titer HBsAg sera and one study used seroconversion panel to evaluate analytical sensitivity of index tests. Sato et al¹⁵ used Dainabot panel with six concentrations. Index test sensitivity was 3.1 ng/ml (HBs/ad) and 6.3 ng/ml (HBs/ay). Nine of the 10 tests failed to detect WHO HBsAg low (0.1–0.8 IU/ml) titer sera. WHO¹¹ five seroconversion panels with 43 sera (23 positive by reference test) reacted poorly by index tests. Clement et al¹⁶ used in-house standards with HBsAg concentrations and found index test lowest detection limit of 9.5 IU/ml as against 0.2–0.7 ng/ml by reference test. Lin et al¹⁷ used 2 analytical panels (In-house sensitivity panel & WHO HBsAg reference panel) and claimed DRW test detected HBsAg with 0.2–0.5 IU/ml irrespective of the serotype. ICBS¹² HBsAg quantitative panel consisted of 146 sera with 8 HBV genotypes and 14 subtypes. Of the 19 RPOCTs, most could only detect undiluted sera with HBsAg levels of >4 IU/ml. 17 test kits showed genotype (D/ayw3, E/ayw4, F/adw4 and by S gene mutants) dependent sensitivity reduction. Analytical sensitivities for HBsAg of >1 IU/ml significantly reduce the length of the HBsAg positive period which renders them less reliable for detecting HBsAg in asymptomatic HBV infections.

DISCUSSION

The present meta-analysis had six strengths when compared to 2 earlier studies^{10,52}: i) global and complete literature search, with strict inclusion criteria, ii) defined

strategy to include and analyze inconclusive results, iii) use of Bivariate statistical model, iv) statistical comparison in meta-regression model of summary estimates of diagnostic accuracy within subgroups to make relevant conclusions, v) evaluation of heterogeneity and its potential sources in the meta-regression model, vi) assessing analytical sensitivity of tests. We do not agree with use of Bayesian hierarchical summary receiver operating characteristic (HSROC) model on the assumption that IgM antibodies to HBcAg were not included in reference standard.⁵³ The meta-analysis was related to detection of HBsAg, which is the hallmark of HBV infection. Further serological and molecular tests follow detection of HBsAg.³ In addition use of Bivariate model was appropriate as: i) reference standards were enzyme immunoassays which employed consistent standard positive thresholds across the studies as per manufacturer's guidelines, ii) Index test yielded positive and negative results as consistent cut-offs on the device (appearance of a colored line or dot) across the studies, iii) all studies administered the same reference standard to all patients, thereby avoiding partial or differential bias.²⁵ HSROC model widens the confidence intervals due assumed conservative ranges of sensitivity and specificity.⁵³ Interestingly, the same authors used bivariate model in similar meta-analysis for HCV.⁵⁴

Our meta-analysis showed high pooled accuracy for RPOCTs for HBsAg. However, sensitivity of individual tests showed wide range from 43.5 to 99.8%. While evaluating the high accuracy of tests, we should be cautious as our meta-analysis was subject to the detection, spectrum, and sampling bias of the original studies.¹⁸ Case-control design employed in studies is apt to overestimate accuracy, with a potential of spectrum bias. In addition, only 4 of the 27 studies explicitly mentioned blinded reading of index test results, suggesting possible detection bias in the remaining studies. This could artificially inflate sensitivity and specificity estimates of the index test.

In the present meta-analysis all tests performed better in studies conducted in developed versus developing countries. It is well known that performance characteristic of any test varies markedly with prevalence of the condition in that population. Studies conducted in developed countries showed minimal heterogeneity while those in developing countries showed substantial heterogeneity. These observations need to be kept in mind as all these tests are meant to be employed in low resource endemic regions of the World. Satisfying was the observation that cross reactive sera (for example HIV, HCV, Tuberculosis etc.) did not influence the diagnostic accuracy of the tests. Such co-infections are highly prevalent in endemic regions and this findings is significant while evaluating scope of these test usage.

We could evaluate the performances of 5 tests with ≥ 3 data points in our statistical model. Determine and BinxNow tests showed significantly higher test estimates than

the reference standard. In addition Determine test estimates were consistent across studies performed in the developed and developing countries. However, this finding should not underestimate tests which could not be evaluated in the statistical model. We believe that the high accuracy of some tests may have been obscured. For example, Lin et al¹⁷ made head-to-head comparison of DRW test and Determine and found DRW did better in overall accuracy, analytical and seroconversion panels.

We believe evaluation of tests by analytical and seroconversion panels gave concerning results. Analytical sensitivity of tests was around 4 IU/ml, which is much higher (<0.5 IU/ml) than claimed by most manufacturers.⁹ Tests performed poorly in seroconversion panels and showed sensitivity reduction with several HBV mutants. Thus these tests are likely to do poorly in situations with low concentrations of HBsAg prevalent in healthy blood donors, general populations, patients recovering from acute HBV infection and those on antiviral therapy.¹⁹

Limitations

While the included studies in the present meta-analysis addressed accuracy of RPOCTs for HBsAg, there was little information available on the evaluation of these tests on HBV-related patient centered outcomes. The most important of these is the feasibility and outcomes of these tests in various setting where HBsAg testing is routinely done. Studies need to be done how these tests shall perform in following situations: i) seroepidemiologic studies in low and high endemic zones, ii) screening of blood and blood products for HBsAg and prevention of transfusion transmitted HBV infection, iii) diagnosis and follow up of acute hepatitis and fulminant hepatic failure, iv) diagnosis of HBV-related chronic hepatitis and cirrhosis and evaluation of antiviral therapy. These outcomes are key to shaping global policy in diagnostics as it becomes more patient centered in the future. In addition we need to do studies which shall address the variability in accuracy with variable prevalence in different population and settings. Also impact of HBV e-mutants which constitute upto 30% of chronic infections in tropical regions need to be defined.⁵⁵

CONCLUSIONS

RPOCTs have major benefits to offer in screening and control of HBV infection especially in far flung low resource endemic zones of the world. The tests are cheap, give rapid diagnosis, are convenient and easy to store, need small blood samples for testing, need minimal training of professionals for testing and need no or minimal expensive equipment. However, their global uptake needs to be addressed keeping in view the overall performance of the tests. With present low level of test accuracy especially sensitivity, poor performance in developing countries,

poor analytical sensitivity and possible influence of HBV mutants, these tests shall significantly underdiagnose HBV infections. This is of major importance in blood banks. Many potential donors in blood banks in tropical countries are likely to have low HBsAg titers and tests are unlikely to diagnose such infections, putting huge health risks to recipients. We believe tests with better analytical sensitivity need to be developed and their feasibility and outcomes in various clinical settings need to be addressed.

FUNDING SOURCE

Our study was funded by the “Dr. Khuroo’s Medical Trust”, a non-profit organization to promote research and help poor patients for medical treatment. The funding source had no role in the conception, design or conduct of this review.

CONFLICTS OF INTEREST

All authors have none to declare.

REFERENCES

1. Lee WM. Hepatitis B virus infection. *N Engl J Med*. 1997;337:1733–1745.
2. Alter MJ. Epidemiology and prevention of hepatitis B. *Semin Liver Dis*. 2003;23:39–46.
3. Krajden M, McNabb G, Petric M. The laboratory diagnosis of hepatitis B virus. *Can J Infect Dis Med Microbiol*. 2005;16(2):65–72.
4. Dufour DR. Hepatitis B surface antigen (HBsAg) assays—are they good enough for their current uses? *Clin Chem*. 2006;52:1457–1459.
5. Bayliss J, Nguyen T, Lesmana CR, Bowden S, Revill P. Advances in the molecular diagnosis of hepatitis B infection: providing insight into the next generation of disease. *Semin Liver Dis*. 2013;33(2):113–121.
6. John AS, Price CP. Economic evidence and point-of-care testing. *Clin Biochem Rev*. 2013;34:61–74.
7. Centers for Disease Control and Prevention. *Update: HIV Counseling and Testing Using Rapid Tests—United States: Morbidity Mortality Weekly Report*. vol. 47. 1995:211–215.
8. Mabey D, Peeling RW, Perkins MD. Rapid and simple point of care diagnostics for STIs (Editorial). *Sex Transm Infect*. 2001;77:397–398.
9. *RDT Info: Current Information on Rapid Diagnostic Tests*. Rapid tests for Hepatitis B. Diagnostic overview. www.rapid-diagnostics.org/rti-hepb-diag.htm.
10. Shivkumar S, Peeling R, Jafari Y, Joseph L, Pai NP. Rapid point-of-care first-line screening tests for hepatitis B infection: a meta-analysis of diagnostic accuracy (1980-2010). *Am J Gastroenterol*. 2012;107(9):1306–1313.
11. World Health Organization. *Hepatitis B Surface Antigen Assays: Operational Characteristics (Phase I)*. http://www.who.int/diagnostics_laboratory/evaluations/en/hep_B_rep1.pdf. (Updated on July 19 2007).
12. International Consortium for Blood Safety (ICBS), New York, USA & ICBS-Europe, Düsseldorf, Germany: *HBV Test Kit Evaluation Results*; <http://www.icbs-web.org/page31.html>.
13. Khan J, Lone D, Hameed A, et al. Evaluation of the performance of two rapid immuno chromatographic tests for detection of hepatitis

- B surface antigen and anti HCV antibodies using ELISA tested samples. *AKEMU*. 2010;16(1S1):84–87.
14. Shinkins B, Thompson M, Mallett S, Perera R. Diagnostic accuracy studies: how to report and analyse inconclusive test results. *BMJ*. 2013;346:f2778.
 15. Sato K, Ichiyama S, Iinuma Y, Nada T, Shimokata K, Nakashima N. Evaluation of immunochromatographic assay systems for rapid detection of hepatitis B surface antigen and antibody, Dainascreen HBsAg and Dainascreen Ausab. *J Clin Microbiol*. 1996;34:1420–1422.
 16. Clement F, Dewint P, Leroux-Roels G. Evaluation of a new rapid test for the combined detection of hepatitis B virus surface antigen and hepatitis B virus e antigen. *J Clin Microbiol*. 2002;40(12):4603–4606.
 17. Lin YH, Wang Y, Loua A, et al. Evaluation of a new hepatitis B virus surface antigen rapid test with improved sensitivity. *J Clin Microbiol*. 2008;46:3319–3324.
 18. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21:1539–1558.
 19. Hadziyannis E. Quantification of HBsAg in serum: characteristics of the assays. *OA Hepatol*. 2013;1:1–6.
 20. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.
 21. Khuroo MS, Khuroo M, Farahat KLC, Kagevi I. Treatment with proton pump inhibitors in acute non-variceal upper gastrointestinal bleeding: a meta-analysis. *J Gastroenterol Hepatol*. 2005;20:11–25.
 22. Smalheiser NR, Lin C, Jia L, et al. Design and implementation of metta, a metasearch engine for biomedical literature retrieval intended for systematic reviewers. *Health Inf Sci Syst*. 2014;2:1. <http://dx.doi.org/10.1186/2047-2501-2-1>.
 23. Whiting PF, Rutjes AW, Westwood ME, et al, QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529–536.
 24. Bossuyt PMM, Reitsma JB, Bruns DE, et al, Standards for Reporting of Diagnostic Accuracy. Toward complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Ann Intern Med*. 2003;138:40–44.
 25. Reistma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. 2005;58:982–990.
 26. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7:177–188.
 27. Vamvakas EC. Applications of meta-analysis in pathology practice. *Am J Clin Pathol*. 2001;116(suppl 1):S47–S64.
 28. Wallace CB, Schmid CH, Lau J, Trikalinos TA. Meta-Analyst: software for meta-analysis of binary, continuous and diagnostic data. *BMC Med Res Methodol*. 2009;9:80. <http://dx.doi.org/10.1186/147-2288-9-80>.
 29. Nakata JM, Johnson JM. Evaluation of rapid immunoenrichment assay for HBsAg in a prison inmate population. In: *Annual Meeting of the American Society for Microbiology, Anaheim*. 1990.
 30. Mvere D, Constantine NT, Kalaswue E, Tobaiwa O, Dambire S, Corcoran P. Rapid and simple hepatitis assays: encouraging results from a blood donor population in Zimbabwe. *Bull World Health Organ*. 1996;74:19–24.
 31. Abraham P, Sujatha R, Raghuraman S, Subramaniam T, Sridharan G. Evaluation of two immunochromatographic assays in relation to RAPID screening of HbsAg. *Ind J Med Microbiol*. 1998;16:23–25.
 32. Oh J, Kim TY, Yoon HJ, Min HS, Lee HR, Choi TY. Evaluation of Genedia HBsAg rapid and Genedia anti-HBs rapid for the screening of HBsAg and anti-HBs. *Korean J Clin Pathol*. 1999;19:114–117.
 33. Palmer C, Cuadrado R, Koenig E. Multi-center evaluation of the Determine™ rapid test for the diagnosis of HIV, hepatitis B surface antigen and syphilis. In: *Interscience Conference on Antimicrobial Agents and Chemotherapy, Ft Lauderdale*. 1999.
 34. Kaur H, Dhanao J, Oberoi A. Evaluation of rapid kits for detection of HIV, HBsAg and HCV infections. *Ind J Med Sci*. 2000;54:432–434.
 35. Lien TX, Tien NTK, Chanpong GF, et al. Evaluation of rapid diagnostic tests for the detection of human immunodeficiency virus types 1 and 2, hepatitis B surface antigen and syphilis in Ho Chi Minh City, Vietnam. *Am J Trop Med Hyg*. 2000;62(2):301–309.
 36. Raj AA, Subramaniam T, Raghuraman S, Abraham P. Evaluation of an indigenously manufactured rapid immunochromatographic test for detection of HBsAg. *Ind J Pathol Microbiol*. 2001;44:413–414.
 37. Lau DT, Ma H, Lemon SM, et al. A rapid immunochromatographic assay for hepatitis B virus screening. *J Viral Hepat*. 2003;10:331–334.
 38. Whang DH, Um T. Comparison of immunochromatography assays and quantitative immunoassays for detecting HBsAg and anti-HBs. *Korean J Lab Med*. 2005;25:186–191.
 39. Cha YJ, Yang JS, Chae SL. Evaluation of indigenously manufactured immunochromatographic assay systems for rapid detection of hepatitis B surface antigen and antibody. *Korean J Lab Med*. 2006;26:52–57.
 40. Akanmu AS, Esan OA, Adewuyi JO, et al. Evaluation of a rapid test kit for detection of HBsAg/eAg in whole blood: a possible method for pre-donation testing. *Afr J Med Med Sci*. 2006;35:5–8.
 41. Ansari M, Omrani M-D, Movahedi V. Comparative evaluation of immunochromatographic rapid diagnostic tests (Strip and Device) and PCR methods for detection of human hepatitis B surface antigens. *Hepat Mon*. 2007;7:87–91.
 42. Randrianirina F, Carod JF, Ratsima E, Chretien JB, Richard V, Talarmin A. Evaluation of the performance of four rapid tests for detection of hepatitis B surface antigen in Antananarivo, Madagascar. *J Virol Methods*. 2008;151:294–297.
 43. Torane VP, Shastri JS. Comparison of ELISA and rapid screening tests for the diagnosis of HIV, hepatitis B and hepatitis C among healthy blood donors in a tertiary care hospital in Mumbai. *Ind J Med Microbiol*. 2008;26:284–285.
 44. Nyirenda M, Beadsworth MJB, Stephany P, et al. Prevalence of infection with hepatitis B and C virus and co-infection with HIV in medical inpatients in Malawi. *J Infect*. 2008;57:72–77.
 45. Ola SO, Otegbayo JA, Yakubu A, Oaje A, Odaibo GN, Shokunbi WI. Pitfalls in diagnosis of hepatitis B virus infection among adults Nigerians. *Niger J Clin Pract*. 2009;12(4):350–354.
 46. Davies J, van Oosterhout JJ, Nyirenda M, et al. Reliability of rapid testing for hepatitis B in a region of high HIV endemicity. *Trans R Soc Trop Med Hyg*. 2010;104(2):162–164.
 47. Scheiblaue H, El-Nageh M, Diaz S, et al. Performance evaluation of 70 hepatitis B virus (HBV) surface antigen (HBsAg) assays from around the world by a geographically diverse panel with an array of HBV genotypes and HBsAg subtypes. *Vox Sang*. 2010;98:403–414.
 48. Maity S, Nandi S, Biswas S, Sadhukhan SK, Saha MK. Performance and diagnostic usefulness of commercially available enzyme linked immunosorbent assay and rapid kits for detection of HIV, HBV and HCV in India. *Viral J*. 2012;9:290.
 49. Bottero J, Boyd A, Gozlan J, et al. Performance of rapid tests for detection of HBsAg and anti-HBsAb in a large cohort, France. *J Hepatol*. 2013;58(3):473–478.
 50. Franzeck FC, Ngwale R, Msongole B, et al. Viral hepatitis and rapid diagnostic test based screening for HBsAg in HIV-infected patients in rural Tanzania. *PLoS One*. 2013;8(3):e58468.
 51. Chameera EWS, Noordeen F, Pandithasundara H, Abeykoon AMSB. Diagnostic efficacy of rapid assays used for the detection of

- hepatitis B virus surface antigen. *Sri Lankan J Infect Dis*. 2013;3(2):21–27.
52. Hwang SH, Oh HB, Choi SE, et al. Meta-Analysis for the pooled sensitivity and specificity of hepatitis B surface antigen rapid tests. *Korean J Lab Med*. 2008;28:160–168.
 53. Dendukuri N, Schiller I, Joseph L, Pai M. Bayesian meta-analysis of the accuracy of a test for tuberculous pleuritis in the absence of a gold standard reference. *Biometrics*. 2012;68(4):1285–1293.
 54. Shivkumar S, Peeling R, Jafari Y, Joseph L, Pai NP. Accuracy of rapid and Point-of-Care screening tests for hepatitis C. A systematic review and meta-analysis. *Ann Intern Med*. 2012;157:558–566.
 55. Funk ML, Rosenberg DM, Lok AS. World-wide epidemiology of HBsAg-negative chronic hepatitis B and associated precore and core promotor variants. *J Viral Hepat*. 2002;9(1):52–61.