Clinical trials and observations

Diagnostic value of immunoassays for heparin-induced thrombocytopenia: a systematic review and meta-analysis.

Michael Nagler\textsuperscript{1,3}; Lucas M. Bachmann\textsuperscript{2}; Hugo ten Cate\textsuperscript{1}; Arina ten Cate-Hoek\textsuperscript{1}

Affiliations:
\textsuperscript{1} Laboratory of Clinical Thrombosis and Haemostasis, and Cardiovascular Research Institute, Maastricht University Medical Center, 6229 ER Maastricht, The Netherlands
\textsuperscript{2} medignition Inc., 8036 Zurich, Switzerland
\textsuperscript{3} Division of Haematology and Central Haematology Laboratory, Luzerner Kantonsspital, CH-6000 Lucerne 16, Switzerland

Correspondence to: michael.nagler@insel.ch; Department of Haematology and Central Haematology Laboratory, Inselspital University Hospital, 3010 Bern, Switzerland; phone +41 31 632 2405; fax +41 31 632 5193

Running head: Immunoassays for heparin-induced thrombocytopenia

Text word count: 4039
Abstract word count: 246
Figures: 5
Tables: 3
References: 83
Key points

- Performance of immunoassays used to diagnose HIT varies substantially with regard to the specific characteristics of the tests.
- High sensitivity (>95%) in combination with a high specificity (>90%) was found in only five tests.

Abstract

Immunoassays are essential in the work-up of patients with suspected heparin-induced thrombocytopenia (HIT). However, the diagnostic accuracy is uncertain with regard to different classes of assays, antibody specificities, thresholds, test variations, and manufacturers. We aimed to assess diagnostic accuracy measures of available immunoassays and to explore sources of heterogeneity. We performed comprehensive literature searches and applied strict inclusion criteria. Finally, 49 publications comprising 128 test evaluations in 15199 patients were included in the analysis. Methodological quality according to QUADAS2 was moderate. Diagnostic accuracy measures were calculated with the unified model (comprising a bivariate random-effects model and a hierarchical summary receiver operating characteristics model). Important differences were observed between classes of immunoassays, type of antibody specificity, thresholds, application of confirmation step, and manufacturers. Combination of high sensitivity (>95%) and a high specificity (>90%) was found in five tests only: polyspecific ELISA with intermediate threshold (GTI, Asserachrom), particle gel immunoassay, lateral flow immunoassay, polyspecific chemiluminescent immunoassay (CLIA) with a high threshold, and IgG-specific CLIA with low threshold. Borderline results (sensitivity, 99.6%; specificity, 89.9%) were observed for IgG-specific GTI-ELISA with low threshold. Diagnostic accuracy appears to be inadequate in tests with high thresholds (ELISA; IgG-specific CLIA), combination of IgG-specificity and intermediate thresholds (ELISA, CLIA), high-dose heparin confirmation step (ELISA), and particle immunofiltration assay. When making treatment decisions, clinicians should be aware of diagnostic characteristics of the tests used and are recommended to estimate post-test probabilities according to likelihood ratios as well as pre-test probabilities using clinical scoring tools.
Introduction

In clinical practice, immunoassays are pivotal for the work-up of patients with suspected heparin-induced thrombocytopenia (HIT)\textsuperscript{1-4}. HIT is a life-threatening complication of heparin therapy, which affects a significant number of patients\textsuperscript{5}. It is associated with a high morbidity and mortality due to a massive pro-coagulant state, with a high incidence of extensive venous and arterial thrombosis, limb loss and even death\textsuperscript{6,7}. Suspicion of HIT requires an immediate diagnostic workup to prevent severe complications\textsuperscript{1,8,9}. Still, diagnosis of HIT is challenging\textsuperscript{1,10}. Functional assays, such as the serotonin release assay (SRA) or the heparin-induced platelet activation assay (HIPA) are accepted as gold standard, but are rarely available in a timely manner\textsuperscript{2,3,11}. Clinical assessment tools such as the 4T’s score can exclude HIT in many patients if conducted by experienced observers\textsuperscript{12}. However, the positive predictive value is low\textsuperscript{12}, they are subject to a relevant inter-observer variability\textsuperscript{13}, and they are not adequately evaluated in all settings\textsuperscript{14}. To make a treatment decision at bedside, physicians are recommended by current guidelines and recent reviews to order an HIT immunoassay in all patients with an intermediate or high risk 4T’s score\textsuperscript{1,8,10-12,15}.

Many tests and several classes of assays have been developed by a number of manufacturers. They recognize different classes of antibodies, utilize varying thresholds and may apply test variations such as the high-dose heparin confirmation step. All tests are based on detecting antibodies targeting complexes of platelet factor 4 (PF4) bound to heparin (or other polyanions respectively)\textsuperscript{16,17}. A solid-phase enzyme-linked immunosorbent assay (ELISA) using a low optical-density threshold (OD) and targeting IgG-, IgM and IgA-antibodies was already developed in the mid 90ies\textsuperscript{16,18}. Later on, other tests were introduced: IgG-specific assays\textsuperscript{19}, high- and intermediate OD thresholds\textsuperscript{19,20}, as well as the high-dose heparin confirmation step\textsuperscript{21}. Other classes of assays have been developed to enable short turnaround times and a 24-hour-service: particle gel immunoassay (PaGIA), particle immunofiltration assay (PIFA), lateral flow immunoassay, latex agglutination assay and chemiluminescent immunoassay (CLIA). Different antibody specificities are available for the CLIA and a number of thresholds have been proposed for CLIA and PaGIA. The corresponding diagnostic accuracy studies are hard to interpret because of varying results, imprecise estimates and sometimes conflicting data. In clinical practice, it is difficult to choose the right assay because it remains unclear, if and how the
mentioned test characteristics influence the diagnostic value. Several authors raised concerns that the diagnostic value may differ relevantly among the individual tests\textsuperscript{2,22-24}. As a result of this uncertainty, the application of assays varies greatly among laboratories\textsuperscript{25}.

The aim of the present investigation was to assess the diagnostic accuracy with regard to different classes of assays, antibody specificities, thresholds, manufacturers, application of high-dose heparin confirmation step, and material used. In addition, we assessed whether different reference standards used in diagnostic accuracy studies have an impact on the diagnostic accuracy. We retrieved all available data, applied strict inclusion criteria, systematically assessed methodological quality, and pooled diagnostic accuracy measures if possible.

**Methods**

Before start of the investigation, we developed a research protocol according to recommendations of the Cochrane Diagnostic Test Accuracy Working Group\textsuperscript{26}.

*Study identification*

A search strategy was developed for MEDLINE, EMBASE and Cochrane Collaboration databases to identify diagnostic accuracy studies of immunoassays for diagnosis of HIT (see supplemental data). Search strategy was refined using keywords of references found in a pilot search, and after manual review of reference lists. Search strategy was tested in 12 index publications (100 percent sensitivity). Literature search was supplemented by manual review of reference lists, including the ACCP antithrombotic practice guideline (ninth edition)\textsuperscript{2}. No restrictions or filters with regard to language, publication date and age range were applied. Last search run was conducted on 2\textsuperscript{nd} of November 2014. Records were screened by two investigators (MN, ATC); all hits were assessed in full-text for eligibility.

*Study eligibility*

The following inclusion criteria were applied: (1) evaluation of a commercially available or fully described immunoassay to detect Heparin/PF4 antibodies, (2) application of the index test to patients with suspected HIT, (3) application of a reference standard that is established or fully described, (4) numbers of true positives, false positives, true negatives and false negatives are reported or can be calculated. We considered a reference standard test to be *established* if it is evaluated in reasonably designed studies and implemented in routine practice. Studies were
excluded in case of (a) application of the index test to patients without any suspicion for HIT (screening test), (b) evaluation of tests that are not immunoassays, (c) evaluation of confirmatory tests after application of a previous laboratory test, (d) reference standard not clearly defined, (e) studies that compare diagnostic accuracy of one immunoassay with another one without using an established reference standard. Two reviewers assessed eligibility and consensus was achieved by discussion (MN, AtC; raw agreement 0.91; Cohen’s kappa 0.71).

Data extraction

Included studies were reviewed in duplicate and the following data were extracted: author, year of publication, study design, type of patient selection, characteristics of study population including cohort criteria, inclusion procedure, setting, characteristics of index test including antibody specificity, thresholds, material used, and use of high-dose heparin confirmation test, application of index test, characteristics of reference standard, application of reference standard, sample size, disease prevalence, numbers of true positives, false positives, true negatives and false negatives (reported or calculated). Test results were categorized according to the “intention-to-diagnose” principle to avoid biased overestimation of diagnostic accuracy: inconclusive results of the index test were classified as negative if the reference standard was positive, and rated as positive if the reference standard was negative. Observations were excluded from analysis if the reference standard revealed inconclusive results. For example, in one study an inconclusive PIFA result was classified as negative because SRA was positive, another inconclusive PIFA result was counted as positive because SRA was negative, and two observations were excluded because of a borderline SRA result in combination with a negative ELISA and a low 4T score. Two-by-two tables were created for every evaluation study. Sensitivity, specificity, positive likelihood ratio, negative likelihood ratio and pre-test probability (prevalence) were calculated. To account for differences in the scale of the threshold among different assays (ELISA assays in particular), we categorized thresholds into low, intermediate and high. For ELISA, optical densities (OD) below or equal to 0.7 were classified as low, OD between 0.8 and 1.4 as intermediate, and OD above 1.4 as high. A positive/negative PaGIA result was ranked as low, and a titer of 2 or 3 as intermediate. LFI results were categorized as low (no studies using titration available). For CLIA, a threshold of 1.0 U/ml was classified as low, between 1.0 and 2.8 U/ml as intermediate, and above 2.8 U/ml as high. Threshold of latex agglutination assay (1.0 U/ml) was classified as low.
Assessment of methodological quality

The methodological quality of the individual studies was assessed in terms of risk of bias and concerns regarding applicability using the revised tool for quality assessment of diagnostic accuracy studies (QUADAS-2). Details of this validated and widely accepted instrument are published elsewhere. Shortly, methodological quality is assessed in four domains (patient selection, index test, reference standard, flow and timing) by the use of signaling questions. According to the published guidelines, we adjusted the signaling questions to our particular research questions and developed detailed decision criteria. In a pilot study, the adapted QUADAS-2 tool was applied to a subset of 7 studies by two investigators (MN, ATC) with a good agreement. Signaling questions and decision criteria were refined after these pilot studies to further improve reliability. Application of adapted QUADAS-2 tool to all included studies was done in duplicate and disagreement was solved by discussion between two investigators.

Statistical analysis

To assess the diagnostic accuracy of different classes of assays, we grouped observations by class, antibody specificity, threshold, and the use of the high-dose heparin confirmation step. In addition, we considered differences in ELISA assay designs by grouping according to manufacturer, antibody specificity, threshold, and use of the high-dose heparin confirmation step. To observe the impact of different reference standards used in evaluation studies, we conducted a stratified analysis in the assay class with the most available studies and performed a meta-regression analysis (see below). To maintain independency of observations, studies which were conducted in the same study population were included only once in each group. In these cases, we selected the study to be included according to the following priorities: 1.) largest number of observations, 2.) type of manufacturer that was included less often within the particular category. As heterogeneity is expected to be high in systematic reviews of diagnostic accuracy studies, we followed current recommendations and did not calculate an I^2 statistic. Diagnostic accuracy measures were pooled with the use of the unified model, comprising a bivariate random-effects model and a hierarchical summary receiver operating characteristics model. However, it was not possible to fit these models in every category because at least four studies are required to perform the analysis. To give an approximation of the diagnostic accuracy in these tests, we fitted a fixed effect model according to Mantel and Haenszel.
Potential sources of heterogeneity and bias

A number of measures were carried out to explore sources of heterogeneity and to recognize potential sources of bias. First, results of studies with a higher methodological quality were compared with estimates obtained from all studies. We selected the criterion of cohort studies (in contrast to case-control studies) because it is considered to be of highest impact on diagnostic accuracy study results. Second, we compared measures between tests within the same study population using the same reference standard. Third, estimates obtained with accepted gold standards (SRA/HIPA) were compared with estimates from all studies. Finally, we conducted a random-effects meta-regression using diagnostic odds ratio (DOR) as outcome variables and important characteristics as predictors (type of reference standard [SRA/HIPA vs. others], material used [plasma vs. serum], setting [surgical vs. mixed population; ICU patients vs. mixed population]).

All statistical analyses were conducted using Stata® 13.1 (StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP), figures of methodological quality were generated using Review Manager (Computer program. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014).

Results

Study identification and selection

Literature search yielded 2716 records, including 2 publications identified by manual review (Figure 1). After removal of duplicates, title and abstract were screened in 2261 records. We selected 258 publications for full-text review. Of these, 37 were excluded because of duplicate publication (conference abstract as well as original article), 78 studies did not evaluate an immunoassay, 24 investigations were not conducted in patients with suspected HIT, 39 articles because of insufficient data for analysis, and the reference standard was not clearly defined in 31 publications. Finally, we included 49 studies comprising 128 test evaluations in 15199 patients.

Study characteristics

Detailed characteristics of all 128 diagnostic accuracy studies are given in Table S1 of supplemental data. Six different classes of assays were evaluated: ELISA (79 evaluation studies)
immunoassays for heparin-induced thrombocytopenia, PaGIA (14), particle immunofiltration assay (1), lateral flow immunoassay (9), latex agglutination assay (1), and CLIA (24). Available assays, antibody specificities, thresholds, test variations, and manufacturers are illustrated in Table 1. Often, more than one evaluation study was reported in a publication (range 1 to 13), and more than one assay was evaluated in the same study population (range 1 to 11). Low thresholds were studied in 96 cases, intermediate in 21, and high in 11. Assays targeting polyspecific antibodies were used in 78 studies, and IgG specific antibodies in 49 investigations (unclear in one study). A high-dose heparin confirmation step (ELISA) was studied in 9 cases. The following reference standard tests have been used: SRA (43 evaluation studies), HIPA (42), combination of HIPA/SRA with clinical criteria (10), flow cytometry (12), heparin-induced platelet aggregation test (PAT; 4), and others (17).

Methodological quality

A summary of the methodological quality is shown in Figure 2, the quality of the individual studies is reported in Figure S1 of the supplemental data. A low risk of bias in all four domains was observed in 2 out of 128 evaluation studies only (1.6%). The median number of low-risk ratings was 1 (25%). A high risk of bias was assessed in 125 cases (24.4%). Most high-risk ratings were assigned in the domain patient selection (57) and the fewest in index test (11). Most low-risk ratings were found in the reference standard domain (90). In contrast, low concerns regarding applicability were assessed in 276 cases (71.9%). Methodological criteria that were frequently inadequately addressed were (1) prospectively enrolled patients, (2) specified cohort criteria, (3) numbers and reasons for excluded patients, (4) interpretation of reference standard without knowledge of the index test, and (5) clearly described sequence of testing.

Diagnostic accuracy of immunoassay classes - HSROC and bivariate model

We were able to calculate diagnostic accuracy measures according to the unified model for nine classes of assays: polyspecific ELISA with low and intermediate threshold, IgG-specific ELISA with low and intermediate threshold, PaGIA, lateral-flow immunoassay, polyspecific CLIA with low and high threshold, and IgG-specific CLIA with low threshold. Detailed results are reported.
in Table 2; figure 3 illustrates diagnostic accuracy as HSROC curve. Among the categories that could be analyzed with the unified model, sensitivity was above 95% for polyspecific ELISA with low and intermediate threshold, IgG-specific ELISA with low threshold, PaGIA, lateral flow immunoassay, polyspecific CLIA with low and high threshold, as well as IgG-specific CLIA with low threshold. Specificity was above 90% in polyspecific-ELISA with intermediate threshold, IgG-specific ELISA with intermediate threshold, PaGIA, lateral flow immunoassay, polyspecific CLIA with high threshold, and IgG-specific CLIA with low threshold. Both criteria were fulfilled in five assay classes (poly-ELISA with intermediate threshold, PaGIA, lateral flow immunoassay, polyspecific CLIA with high threshold, and IgG-specific CLIA with low threshold). A forest plot was generated to compare diagnostic accuracy measures (Figure 4). To illustrate the diagnostic value of the tests at different pre-test probabilities (prevalence or probability according to clinical scoring tools), we show post-test probabilities of positive and negative test results in Table 2 and in Figure 5 (assuming constant likelihood ratios).

Approximate diagnostic accuracy of immunoassay classes – fixed effects model

Because of a limited number of studies, we were unable to calculate diagnostic accuracy measures using the unified model in 9 tests: polyspecific ELISA and IgG-specific ELISA with high threshold, polyspecific ELISA and IgG-specific ELISA with high-dose heparin confirmation step (low and intermediate threshold), PaGIA with intermediate threshold, polyspecific CLIA with intermediate threshold, and IgG-specific CLIA with intermediate and high threshold. To give an impression of the diagnostic accuracy in these tests, we calculated diagnostic accuracy measures using a fixed effects model (if possible) and report the results in Table 2. Sensitivity was below 95% in all of the above-mentioned tests except poly-ELISA with low threshold and high heparin-dose confirmation step.

Diagnostic accuracy of ELISA assays by manufacturer

We were able to calculate diagnostic accuracy measures for the individual ELISA manufacturers by means of the unified model in 8 assays and by the use of a fixed effects model in two assays. No pooling was possible in four assays with a unique evaluation study. Results are reported in Table 3. In most of the assays, diagnostic accuracy corresponds to the general characteristic of the class (sensitivity above or below 95%, specificity above or below 90%). However, some assays appear to be inferior (e.g. IgG-specific Asserachrom assay with a low threshold).
Diagnostic accuracy was superior to corresponding class in case of IgG-specific GTI-ELISA with low threshold (sensitivity 99.6%, 95%CI 22.7 to 100.0; specificity 89.9%, 95%CI 86.2 to 92.6).

Potential sources of heterogeneity and bias

For the first sensitivity analysis, calculations were repeated in studies with cohort designs only (lower risk of bias). Only minor differences were observed (range 0.0 to 0.5% for sensitivity, and 0.0 to 3.6% for specificity), the results are shown in Table S2 of the supplementary data. In the second sensitivity analysis, data of studies that used SRA/HIPA as reference standard were pooled only (Table S5 of the supplementary data) and compared with results from all studies. Differences ranged between 0.2 and 1.7% for sensitivity, and 0.7 and 1.5% for specificity. Third, we compared data of different tests obtained within the same study population using the same reference standard. Table S4 of the supplemental data shows the results of the two populations with the most tests evaluated. The results of the individual studies and the differences between different categories (thresholds, antibody specificity, class of assays) correspond very well with the pooled data. Table S3 of the supplemental data shows results of a meta-regression analysis exploring the impact of the reference standard used, setting, as well as specimen material on diagnostic accuracy. No significant results were observed for any of the above-mentioned variables.

In general, neither sensitivity analyses nor meta-regression recognized any characteristic of study design as possible source of heterogeneity and bias.

Discussion

Key findings

Important differences exist between classes of immunoassays, type of antibodies, thresholds, and application of confirmation step, and only five tests appear optimal for clinical use. Relevant variations were found even between manufacturers of ELISA assays. In general, methodological quality of primary studies was moderate. Nevertheless, the results were consistent and no differences were found in sensitivity analyses.
**Strengths and limitations**

The present investigation is the first systematic review and meta-analysis focusing on immunoassays for diagnosis of HIT. We conducted a comprehensive literature search to retrieve the published evidence, applied strict inclusion criteria, assessed the methodological quality of the studies systematically, and used an appropriate meta-analytic technique to pool the existing data. Thus, we were able to compare diagnostic accuracy measures between classes of assays, manufacturers, antibody specificities, thresholds and the application of high-dose heparin confirmation step.

Our study has several limitations. First, methodological quality of primary studies was only moderate in general. Even though we cannot fully exclude that this might have influenced our results, we estimate the risk to be low: (1) applying strict inclusion criteria, we excluded 94 publications because of low quality, (2) sensitivity analysis did not show relevant differences if compared with high-quality publications only, and (3) the results were consistent across subgroups. Second, the number of available studies was limited in certain categories of tests, restricting the analysis to a fixed effects meta-analytic model only (Table 2 and 3). The results of this analysis must be interpreted with caution, because it is not accepted as a valid statistical technique in this situation. The number of studies and cases is even more limited in certain categories. For example, the favorable sensitivity of poly-ELISA with low threshold and high heparin-dose confirmation step largely depends on one study in which the only one HIT case was detected (corresponding to a sensitivity of 100%) 67. The number of studies and patients is limited even for the tests most investigated, leading to a restricted precision with wide confidence intervals. For example, expected differences between polyspecific and IgG-specific ELISA in terms of sensitivity and specificity could not be reproduced in every subgroup (Table 2). And third, a number of different reference standard tests have been used in evaluation studies and we cannot fully preclude that this might have influenced the results of our investigation. However, we expect the risk to be low for several reasons: (1) a sensitivity analysis comparing diagnostic accuracy measures obtained with one of the accepted gold standards SRA/HIPA only and all results did not find any difference, (2) meta-regression analysis revealed no difference in diagnostic odds ratio if SRA/HIPA were compared to other defined reference standard tests, (3) we applied strict inclusion criteria requiring a clear definition of reference standard tests, and (4)
in the majority of studies (95/128) either SRA/HIPA or a combination of SRA/HIPA with clinical criteria was used.

One may argue that a systematic review of diagnostic tests for HIT should focus on sensitivities and specificities. Indeed, sensitivity and specificity are well-known (even though often misinterpreted) measures of diagnostic accuracy, and reporting them is essential. However, the important question in clinical practice is, how a particular test result predicts disease and sensitivities/specificities cannot answer this question. Predictive values are often used, but these measures depend on the prevalence of the disease and generalization beyond a study population is difficult. Likelihood ratios are powerful diagnostic accuracy measures because they are assumed to be independent from prevalence, they express how many times more (or less) likely a test result is in patients with the disease in contrast to patients without the disease, and they can be used to calculate post-test probabilities. In addition, differences in likelihood ratios between tests can be seen clearly in a forest plot. In contrast, it is hard to read a forest plot with estimates of sensitivities or specificities along with confidence intervals when most data are between 90 and 100% (ceiling effect). Following these considerations, we decided to report all sensitivities/specificities in a table and likelihood ratios as a forest plot.

**Implications for clinical practice**

With the present investigation, we provide aggregated diagnostic accuracy measures for all frequently applied immunoassays used to diagnose HIT. Our results will raise the clinicians’ awareness for the characteristics of tests they are confronted with. Using the post-test probabilities provided in Table 2 and Figure 5, clinicians may be enabled to estimate the probability of HIT in individual patients. If used in combination with clinical scoring systems (which assess the pre-test probability), they might improve care in patients with suspected HIT substantially. In addition, laboratory managers can use our results to compare test characteristics directly, and base decisions about which test to implement on a sound scientific background.

**Implications for future research**

No antibody test is both 100% sensitive and specific and such a test is not expected in the near future. Functional assays are rarely available in a timely fashion. Clinical scoring systems such as the 4T’s score could help identifying patients with a low pre-test probability and might reduce the number of “overtreated” patients. However, there are several unsolved issues associated
with the 4T’s score \textsuperscript{12-14}. Thus, we will always miss a fraction of patients with HIT, or treat a number of non-HIT patients. Both situations are dangerous. A promising concept to overcome these limitations might be a subsequent or parallel determination of different tests and estimation of post-test probabilities using nomograms. We suggest researchers to focus on diagnostic algorithms in order to improve diagnostic procedures in this vulnerable group of patients.

Conclusions

Important differences exist between immunoassays for diagnosis of HIT, in particular with regard to different classes of assays, antibody specificities, manufacturers, thresholds and the application of the high-dose heparin confirmation step. Only five tests have a high sensitivity combined with a high specificity: polyspecific ELISA with an intermediate threshold (GTI, Asserachrom), PaGIA, lateral flow immunoassay, polyspecific CLIA with a high threshold, and IgG-specific CLIA with a low threshold. Diagnostic accuracy appears to be inadequate in tests with high thresholds (ELISA, IgG-specific CLIA), combination of IgG-specificity and intermediate thresholds (ELISA, CLIA), high-dose heparin confirmation step (ELISA), particle immunofiltration assay, and IgG-specific Asserachrom assay at low threshold. Clinicians should be aware of the pre-test probability as well as the diagnostic characteristics of the tests they are confronted with. Investigations focused on diagnostic algorithms combining different tests might further improve care in patients with suspected HIT.

Acknowledgements

This study was supported by Research Fund Haematology Luzerner Kantonsspital. We thank Prof. Greinacher, Greifswald (Germany) for his valuable comments on an early version of this manuscript.

Authorship contributions

MN developed the protocol, conducted the literature search, assessed methodological quality, performed data extraction, conducted the statistical analysis, and wrote the manuscript. LMB developed the analysis plan, and reviewed the analysis. HTC reviewed the protocol and the manuscript. ATC reviewed the protocol, assessed methodological quality, performed data
extraction, and reviewed the manuscript. All authors approved the final version of the manuscript.

**Disclosure of Conflict of Interest**

MN has received research grants or lecture fees from Pfizer, Bayer, CSL Behring, Sanofi-aventis, Roche Diagnostics, and Axonlab. HtC is a Fellow of the Gutenberg Forschungscolleg, Gutenberg University, Mainz, Germany.
References


### Tables

**Table 1:** Immunoassays for diagnosis of HIT – available classes of assays, antibody specificities, thresholds, test variations, and manufacturers

<table>
<thead>
<tr>
<th>Classes of assays</th>
<th>Antibody specificity</th>
<th>Threshold</th>
<th>Test variation</th>
<th>Manufacturer (name of tests)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme-linked immunosorbent assay (ELISA)</td>
<td>Polyspecific</td>
<td>Low*</td>
<td>High-heparin dose confirmation step</td>
<td>In-house assays</td>
</tr>
<tr>
<td></td>
<td>IgG specific</td>
<td>Intermediate+</td>
<td></td>
<td>Genetic testing institute [GTI] Diagnostics, Waukesha, WI, USA (GTI-PF4; HAT; PF4-Enhanced; GTI-IgG)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High#</td>
<td></td>
<td>Hyphen-BioMed, Neuville-Sur-Oise, France (Zymutest HIA IgGAM; Zymutest HIA IgG)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Diagnostica Stago, Asnières-sur-Seine, France (Asserachrom HPIA)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gen-Probe-Waukesha, Waukesha, WI, USA (Gen-Probe PF4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Technoclone GmbH, Vienna, Austria (Technozym)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Diamed, Cressier sur Morat, Switzerland (ID-H/PF4 PaGIA)</td>
</tr>
<tr>
<td>Particle gel immunoassay (PaGIA)</td>
<td>Polyspecific</td>
<td>Low°</td>
<td></td>
<td>Akers Biosciences Inc, Thorofare, NJ, USA (HealthTEST)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermediate§</td>
<td></td>
<td>Diagnostica Stago, Asnières-sur-Seine, France (STic EXPERT HIT)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Milenia Biotec, Giessen, Germany (Milenia QuickLine HIT)</td>
</tr>
<tr>
<td>Particle immunofiltration assay</td>
<td>Polyspecific</td>
<td>Positive/negative</td>
<td></td>
<td>Instrumentation Laboratory, Bedford, MA, USA (HemosIL AcuStar HIT-Ab; HemosIL AcuStar HIT-IgG)</td>
</tr>
<tr>
<td>Lateral flow immunoassay</td>
<td>IgG specific</td>
<td>Positive/negative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemiluminescent immunoassay (CLIA)</td>
<td>Polyspecific</td>
<td>Low†</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IgG specific</td>
<td>Intermediate‖</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>High‡</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latex agglutination assay</td>
<td>Polyspecific</td>
<td>Low†</td>
<td></td>
<td>Instrumentation Laboratory, Bedford, MA, USA (HemosIL HIT-Ab)</td>
</tr>
</tbody>
</table>

* below or equal to OD 0.7; + between OD 0.8 and 1.4; # above OD 1.4; ° positive/negative; § titer 2 to 3; † 1.0 U/ml; ‖ between 1.0 and 2.8 U/ml; ‡ above 2.8 U/ml; ** above 3.85 U/ml; & technically identical with GTI assay
### Table 2: Pooled diagnostic accuracy measures of different classes of immunoassays for diagnosis of HIT

<table>
<thead>
<tr>
<th>Type of test</th>
<th>Number of studies; patients*</th>
<th>Statistical model</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Likelihood ratio</th>
<th>Post-test probability of a positive test result</th>
<th>Post-test probability of a negative test result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(percentages)</td>
<td></td>
<td></td>
<td>1% Prevalence</td>
<td>7% Prevalence</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Positive (95% CI)</td>
<td>Negative (95% CI)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Polyspecific ELISA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low threshold19, 20, 28, 39, 41, 43, 44, 46, 47, 48, 54, 59, 61, 63, 64, 66, 68, 69, 71, 74, 75, 77-79</td>
<td>31; 8933</td>
<td>Unified*</td>
<td>96.7 (89.7, 99.0)</td>
<td>86.8 (82.0, 90.5)</td>
<td>7.3 (5.4, 10.0)</td>
<td>0.04 (0.01, 0.12)</td>
<td>6.9 (5.1, 9.2)</td>
</tr>
<tr>
<td>Intermediate threshold19, 47, 61, 65, 77</td>
<td>5; 2334</td>
<td>Unified*</td>
<td>98.4 (90.8, 99.7)</td>
<td>94.9 (90.5, 97.3)</td>
<td>19.3 (10.4, 36.0)</td>
<td>0.02 (0.00, 0.1)</td>
<td>16.3 (9.5, 26.7)</td>
</tr>
<tr>
<td>High threshold20, 29, 39, 47</td>
<td>3; 1014</td>
<td>Fixed effects+</td>
<td>15.0 (14.5, 15.5)</td>
<td>100 (99.3, 100)</td>
<td>73.4 (28.2, 190.9)</td>
<td>0.3 (0.2, 0.5)</td>
<td>42.6 (22.2, 65.9)</td>
</tr>
<tr>
<td><strong>IgG-specific ELISA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low threshold19, 20, 44, 55, 63, 66, 69-72, 77</td>
<td>12; 3116</td>
<td>Unified*</td>
<td>98.3 (95.1, 99.4)</td>
<td>85.4 (78.2, 90.6)</td>
<td>6.7 (4.5, 10.2)</td>
<td>0.02 (0.01, 0.05)</td>
<td>6.3 (4.3, 9.3)</td>
</tr>
<tr>
<td>Intermediate threshold19, 20, 77</td>
<td>4; 2545</td>
<td>Unified*</td>
<td>91.2 (86.2, 94.5)</td>
<td>93.5 (89.1, 96.2)</td>
<td>14.1 (8.1, 24.5)</td>
<td>0.09 (0.05, 0.15)</td>
<td>12.5 (7.6, 19.8)</td>
</tr>
<tr>
<td>High threshold19</td>
<td>2; 1958</td>
<td>Fixed effects+</td>
<td>60.9 (59.7, 62.1)</td>
<td>99.4 (97.6, 100)</td>
<td>97.0 (53.0, 177.6)</td>
<td>0.4 (0.3, 0.5)</td>
<td>49.5 (34.9, 64.2)</td>
</tr>
<tr>
<td><strong>PaGIA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low threshold20, 36, 42, 45, 46, 53, 58, 64, 66, 70, 72, 75, 79</td>
<td>12; 2205</td>
<td>Unified*</td>
<td>96.5 (89.8, 98.9)</td>
<td>93.7 (83.1, 97.8)</td>
<td>15.3 (5.5, 42.3)</td>
<td>0.04 (0.01, 0.11)</td>
<td>13.4 (5.3, 29.9)</td>
</tr>
<tr>
<td>Intermediate threshold56</td>
<td>1; 1291</td>
<td>None°</td>
<td>98.9</td>
<td>95.9</td>
<td>24.1</td>
<td>0.01</td>
<td>19.6</td>
</tr>
<tr>
<td>Lateral flow immunosassay54, 56, 58, 66, 70, 72</td>
<td>7; 1163</td>
<td>Unified*</td>
<td>98.4 (85.3, 99.9)</td>
<td>90.3 (84.4, 94.1)</td>
<td>10.1 (6.2, 16.5)</td>
<td>0.02 (0.00, 0.18)</td>
<td>9.3 (5.9, 14.3)</td>
</tr>
<tr>
<td><strong>Particle immunofiltration assay</strong></td>
<td>1; 88</td>
<td>None°</td>
<td>0.0</td>
<td>70.1</td>
<td>2.3</td>
<td>0.5</td>
<td>2.3</td>
</tr>
<tr>
<td><strong>Latex agglutination assay</strong></td>
<td>1; 119</td>
<td>None°</td>
<td>100.0</td>
<td>84.3</td>
<td>3.7</td>
<td>0.0</td>
<td>3.6</td>
</tr>
</tbody>
</table>
### Polyspecific chemiluminescent immunoassay

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Low: 1008</th>
<th>Unified*</th>
<th>98.9 (92.7, 99.8)</th>
<th>85.6 (79.3, 90.3)</th>
<th>6.9 (4.7, 10.0)</th>
<th>0.01 (0.00, 0.09)</th>
<th>6.5 (4.5, 9.2)</th>
<th>34.2 (26.1, 42.9)</th>
<th>54.9 (45.4, 63.8)</th>
<th>0.0 (0.0, 0.1)</th>
<th>0.1 (0.0, 0.7)</th>
<th>0.2 (0.0, 1.6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intermediate: 448</td>
<td>Fixed effects+</td>
<td>97.9 (94.6, 100.0)</td>
<td>93.1 (90.4, 95.8)</td>
<td>13.5 (9.5, 18.9)</td>
<td>0.0 (0.0, 0.1)</td>
<td>12.0 (8.8, 16.0)</td>
<td>50.4 (41.7, 58.7)</td>
<td>70.4 (62.6, 76.9)</td>
<td>0.0 (0.0, 0.1)</td>
<td>0.0 (0.0, 0.7)</td>
<td>0.0 (0.0, 1.7)</td>
</tr>
<tr>
<td></td>
<td>High: 755</td>
<td>Unified*</td>
<td>98.3 (69.5, 99.9)</td>
<td>97.5 (94.4, 98.9)</td>
<td>39.5 (17.5, 89.2)</td>
<td>0.0 (0.0, 0.40)</td>
<td>28.5 (15.0, 47.6)</td>
<td>74.8 (56.8, 87.1)</td>
<td>98.9 (75.5, 94.1)</td>
<td>0.0 (0.0, 0.4)</td>
<td>0.1 (0.0, 0.9)</td>
<td>0.2 (0.0, 6.6)</td>
</tr>
</tbody>
</table>

### IgG-specific chemiluminescent immunoassay

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Low: 5741</th>
<th>Unified*</th>
<th>98.8 (69.2, 100.0)</th>
<th>94.6 (90.7, 96.9)</th>
<th>18.3 (10.6, 31.5)</th>
<th>0.01 (0.00, 0.40)</th>
<th>15.6 (9.7, 24.1)</th>
<th>57.9 (44.4, 70.3)</th>
<th>76.3 (65.2, 84.8)</th>
<th>0.0 (0.0, 0.4)</th>
<th>0.1 (0.0, 0.7)</th>
<th>0.2 (0.0, 7.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intermediate: 448</td>
<td>Fixed effects+</td>
<td>78.6 (75.9, 81.2)</td>
<td>98.7 (94.6, 100)</td>
<td>42.3 (20.1, 88.7)</td>
<td>0.2 (0.1, 0.3)</td>
<td>29.9 (16.9, 47.3)</td>
<td>76.1 (60.2, 87.0)</td>
<td>88.2 (78.0, 94.0)</td>
<td>0.2 (0.1, 0.3)</td>
<td>1.5 (0.7, 2.2)</td>
<td>3.4 (1.7, 5.0)</td>
</tr>
<tr>
<td></td>
<td>High: 552</td>
<td>Fixed effects+</td>
<td>74.2 (71.9, 76.5)</td>
<td>99.1 (95.4, 100)</td>
<td>47.8 (23.2, 98.7)</td>
<td>0.2 (0.1, 0.4)</td>
<td>32.6 (19.0, 49.9)</td>
<td>78.3 (63.6, 88.1)</td>
<td>89.4 (80.4, 94.6)</td>
<td>0.2 (0.1, 0.4)</td>
<td>1.5 (0.7, 2.9)</td>
<td>3.4 (1.7, 6.6)</td>
</tr>
</tbody>
</table>

### High-dose heparin confirmation step

<table>
<thead>
<tr>
<th>Assay Type</th>
<th>Polyspecific ELISA, low threshold: 2; 116</th>
<th>Fixed effects+</th>
<th>99.4# (94.7, 100)</th>
<th>85.8 (83.1, 88.5)</th>
<th>2.7 (1.6, 4.7)</th>
<th>0.3 (0.1, 1.3)</th>
<th>2.7 (1.6, 4.5)</th>
<th>16.9 (10.7, 26.1)</th>
<th>32.3 (22.0, 45.3)</th>
<th>0.3 (0.1, 1.3)</th>
<th>2.2 (0.7, 8.9)</th>
<th>5.0 (1.7, 18.7)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IgG-specific ELISA, low threshold: 3; 1459</td>
<td>Fixed effects+</td>
<td>88.3 (87.0, 89.8)</td>
<td>90.7 (89.2, 92.1)</td>
<td>9.3 (7.8, 11.1)</td>
<td>0.1 (0.1, 0.2)</td>
<td>8.6 (7.3, 10.1)</td>
<td>41.2 (37.0, 45.5)</td>
<td>62.2 (57.9, 66.2)</td>
<td>0.1 (0.1, 0.2)</td>
<td>0.7 (0.7, 1.5)</td>
<td>1.7 (1.7, 3.4)</td>
</tr>
<tr>
<td></td>
<td>IgG-specific ELISA, intermediate threshold: 2; 1000</td>
<td>Fixed effects+</td>
<td>69.7 (68.2, 71.3)</td>
<td>96.2 (94.7, 97.8)</td>
<td>18.2 (12.8, 25.8)</td>
<td>0.3 (0.2, 0.4)</td>
<td>15.5 (11.4, 20.7)</td>
<td>57.8 (49.1, 66.0)</td>
<td>76.3 (69.3, 82.0)</td>
<td>0.3 (0.2, 0.4)</td>
<td>2.2 (1.5, 2.9)</td>
<td>5.0 (3.4, 6.6)</td>
</tr>
</tbody>
</table>

*Studies which were conducted in the same study population were included only once to maintain independence of observations; * Unified model 31, comprising a bivariate random-effects model 32 and a hierarchical summary receiver operating characteristics model 33; + in tests with < 4 studies available, summary estimates were calculated using a fixed-effects model (Mantel-Haenszel method) - however, the results of this analysis must be interpreted with caution, because it is not accepted as a valid statistical technique in this situation; ° only one study available; # result of fixed effects model appears invalid because the favorable sensitivity depends on one study in which the only one HIT case was detected 67
Table 3: Diagnostic accuracy of different ELISA assays by manufacturer

<table>
<thead>
<tr>
<th>Assay</th>
<th>Number of studies</th>
<th>Number of patients</th>
<th>Statistical model</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Positive likelihood ratio (95% CI)</th>
<th>Negative likelihood ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polyspecific ELISA, low threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled estimates</td>
<td>31*</td>
<td>8933*</td>
<td>Unified</td>
<td>96.7 (89.7, 99.0)</td>
<td>86.8 (82.0, 90.5)</td>
<td>7.3 (5.4, 10.0)</td>
<td>0.04 (0.01, 0.12)</td>
</tr>
<tr>
<td>GTI</td>
<td>13</td>
<td>3804</td>
<td>Unified</td>
<td>99.9 (90.9, 100.0)</td>
<td>87.4 (79.2, 92.7)</td>
<td>7.9 (4.7, 13.4)</td>
<td>0.001 (&lt;0.001, 0.11)</td>
</tr>
<tr>
<td>ZymuTest</td>
<td>3</td>
<td>295</td>
<td>Fixed effects+</td>
<td>99.5 (94.8, 100.0)</td>
<td>77.6 (72.9, 82.3)</td>
<td>4.1 (3.1, 5.5)</td>
<td>0.05 (0.01, 0.12)</td>
</tr>
<tr>
<td>Asserachrom</td>
<td>15</td>
<td>2621</td>
<td>Unified</td>
<td>92.7 (73.6, 98.3)</td>
<td>87.3 (79.9, 92.3)</td>
<td>7.3 (4.6, 11.7)</td>
<td>0.08 (0.02, 0.34)</td>
</tr>
<tr>
<td>In-house assay</td>
<td>5</td>
<td>2673</td>
<td>Unified</td>
<td>93.2 (76.3, 98.3)</td>
<td>92.6 (90.1, 94.0)</td>
<td>12.6 (9.7, 16.3)</td>
<td>0.07 (0.02, 0.28)</td>
</tr>
<tr>
<td>Polyspecific ELISA, intermediate threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled estimates*</td>
<td>5*</td>
<td>2334*</td>
<td>Unified</td>
<td>98.6 (95.2, 99.6)</td>
<td>94.8 (92.2, 96.5)</td>
<td>18.8 (12.7, 27.9)</td>
<td>0.01 (0.004, 0.05)</td>
</tr>
<tr>
<td>GTI</td>
<td>5</td>
<td>2646</td>
<td>Unified</td>
<td>97.4 (92.4, 99.1)</td>
<td>95.9 (94.3, 97.0)</td>
<td>23.6 (17.2, 32.4)</td>
<td>0.03 (0.009, 0.08)</td>
</tr>
<tr>
<td>ZymuTest</td>
<td>1</td>
<td>87</td>
<td>-</td>
<td>100.0</td>
<td>82.5</td>
<td>5.7</td>
<td>0.00</td>
</tr>
<tr>
<td>Asserachrom</td>
<td>1</td>
<td>1291</td>
<td>-</td>
<td>99.6</td>
<td>93.6</td>
<td>15.6</td>
<td>0.01</td>
</tr>
<tr>
<td>IgG-specific ELISA, low threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled estimates*</td>
<td>12*</td>
<td>3116*</td>
<td>Unified</td>
<td>98.3 (95.1, 99.4)</td>
<td>85.4 (78.2, 90.6)</td>
<td>6.7 (4.5, 10.2)</td>
<td>0.02 (0.01, 0.05)</td>
</tr>
<tr>
<td>GTI</td>
<td>4</td>
<td>639</td>
<td>Unified</td>
<td>99.6 (22.7, 100.0)</td>
<td>89.9 (86.2, 92.6)</td>
<td>9.8 (7.2, 13.4)</td>
<td>0.004 (&lt;0.001, 3.7)</td>
</tr>
<tr>
<td>ZymuTest</td>
<td>5</td>
<td>829</td>
<td>Unified</td>
<td>99.2 (86.4, 100.0)</td>
<td>85.8 (77.1, 91.5)</td>
<td>7.0 (4.3, 11.4)</td>
<td>0.008 (&lt;0.001, 0.18)</td>
</tr>
<tr>
<td>Asserachrom</td>
<td>2</td>
<td>160</td>
<td>Fixed effects+</td>
<td>72.0 (68.4, 75.5)</td>
<td>93.8 (90.3, 97.4)</td>
<td>3.6 (2.04, 5.6)</td>
<td>0.27 (0.10, 0.72)</td>
</tr>
<tr>
<td>Technozym</td>
<td>1</td>
<td>41</td>
<td>-</td>
<td>100.0</td>
<td>0.92</td>
<td>12.7</td>
<td>0.00</td>
</tr>
<tr>
<td>In-house assay</td>
<td>4</td>
<td>2958</td>
<td>Unified</td>
<td>99.2 (96.4, 99.8)</td>
<td>83.5 (71.5, 91.0)</td>
<td>6.0 (3.4, 10.7)</td>
<td>0.009 (0.002, 0.04)</td>
</tr>
<tr>
<td>IgG-specific ELISA, intermediate threshold</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled estimates*</td>
<td>4*</td>
<td>2545*</td>
<td>Unified</td>
<td>91.2 (86.3, 94.5)</td>
<td>93.5 (89.1, 96.2)</td>
<td>14.1 (8.1, 24.5)</td>
<td>0.09 (0.06, 0.15)</td>
</tr>
<tr>
<td>ZymuTest</td>
<td>1</td>
<td>87</td>
<td>-</td>
<td>100.0</td>
<td>82.5</td>
<td>5.7</td>
<td>0.00</td>
</tr>
<tr>
<td>In-house assay</td>
<td>5</td>
<td>4416</td>
<td>Unified</td>
<td>89.8 (82.7, 94.2)</td>
<td>96.0 (94.5, 97.1)</td>
<td>22.5 (16.3, 31.2)</td>
<td>0.10 (0.06, 0.18)</td>
</tr>
</tbody>
</table>

* Studies which were conducted in the same study population were included only once to maintain independency of observations; + in tests with < 4 studies available, summary estimates were calculated using a fixed-effects model (Mantel-Haenszel method; to be interpreted with caution); § including technically identical Gen-Probe assay; ° application of pooled ‘in-house assay’ data to individual tests must be done with caution because these tests are constructed differently (every assay must be evaluated individually)
Figure Legends

**Figure 1:** Selection process of primary studies. PRISMA flow chart is shown.

**Figure 2:** Summary of methodological quality. Methodological quality of studies investigating diagnostic accuracy of immunoassays for diagnosis of HIT was assessed using the *Quality Assessment of Diagnostic Accuracy Studies* tool, version 2 (QUADAS-2)\(^\text{29}\). Detailed results for individual studies are shown in Figure S1 of supplemental data.

**Figure 3:** Sensitivity and specificity of immunoassay classes for diagnosis of HIT. Results of a hierarchical summary receiver operating characteristics (HSROC) model are reported showing the summary point (red quadrat), a 95% confidence region (dashed yellow line) and 95% prediction region (dashed green line), illustrating a confidence region for the true sensitivity and specificity in a possible future study. Results of primary studies are shown as circles, the sample size of the studies is represented by the size of the circle.

**Figure 4:** Diagnostic accuracy of different immunoassay classes as characterized by positive and negative likelihood ratios. Likelihood ratios are powerful measures describing how many times more (or less) likely a test result is in patients *with* the disease in contrast to patients *without* the disease\(^\text{80}\). In the context of HIT, a test with a +LR above 10 (corresponding to a specificity of 90%) and a –LR below 0.05 (corresponding to a sensitivity of 95%) is considered as favorable.

**Figure 5:** Probability of HIT after conducting an immunoassay at varying pre-test probabilities. Physicians are interested to know, how likely a disease is with a positive test result and how unlikely it is with a negative test results. The *post-test probability* depends largely on the *pre-test probability* (corresponding to the prevalence or the result of a clinical scoring system), but also on the likelihood ratios of the test.
Figure 1

Database searching: n=2714
- MEDLINE: n=648
- EMBASE: n=2066
- Cochrane Database: n=0

Manual review: n=2

Records identified: n=2716
Duplicates removed: n=455

Records screened: n=2261
Records excluded: n=2003

Full-text articles assessed for eligibility: n=258

Studies excluded: n=209
- Double publication: n=37
- No immunoassay: n=78
- No „suspected-HIT“ setting: n=24
- Reference standard not clearly defined: n=31
- Insufficient data for analysis: n=39

Studies included: n=49
(128 test evaluations)
Figure 2

Risk of Bias and Applicability Concerns:

- **Patient Selection**
  - High: Red
  - Unclear: Yellow
  - Low: Green

- **Index Test**
  - High: Red
  - Unclear: Yellow
  - Low: Green

- **Reference Standard**
  - High: Red
  - Unclear: Yellow
  - Low: Green

- **Flow and Timing**
  - High: Red
  - Unclear: Yellow
  - Low: Green

- **Applicability Concerns**
  - High: Red
  - Unclear: Yellow
  - Low: Green
Figure 4

- Polyspecific ELISA, low threshold: Positive likelihood ratio = 8.11 (5.87, 10.34), Negative likelihood ratio = 0.03 (0.01, 0.06)
- Polyspecific ELISA, intermediate threshold: Positive likelihood ratio = 18.85 (11.42, 26.27), Negative likelihood ratio = 0.01 (-0.00, 0.03)
- IgG-specific ELISA, low threshold: Positive likelihood ratio = 6.75 (3.98, 9.51), Negative likelihood ratio = 0.02 (-0.00, 0.04)
- IgG-specific ELISA, intermediate threshold: Positive likelihood ratio = 14.11 (6.34, 21.88), Negative likelihood ratio = 0.09 (0.05, 0.14)
- Particle gel immunoassay: Positive likelihood ratio = 15.29 (-0.30, 30.88), Negative likelihood ratio = 0.04 (-0.00, 0.08)
- Lateral flow immunoassay: Positive likelihood ratio = 10.13 (5.18, 15.08), Negative likelihood ratio = 0.02 (-0.02, 0.06)
- Polyspecific CLIA, low threshold: Positive likelihood ratio = 6.88 (4.27, 9.48), Negative likelihood ratio = 0.01 (-0.01, 0.04)
- Polyspecific CLIA, high threshold: Positive likelihood ratio = 39.54 (7.40, 71.69), Negative likelihood ratio = 0.02 (-0.04, 0.07)
- IgG-specific CLIA, low threshold: Positive likelihood ratio = 18.28 (8.33, 28.24), Negative likelihood ratio = 0.01 (-0.03, 0.06)
Figure 5
Diagnostic value of immunoassays for heparin-induced thrombocytopenia: a systematic review and meta-analysis

Michael Nagler, Lucas M. Bachmann, Hugo ten Cate and Arina ten Cate-Hoek

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include digital object identifier (DOIs) and date of initial publication.