Evaluating the performance of common reference laboratory tests for acute dengue diagnosis: a systematic review and meta-analysis of RT-PCR, NS1 ELISA, and IgM ELISA

Kamla Pillay, Suzanne H Keddie*, Elizabeth Fitchett, Cassandra Akinde, Oliver Bärenbold, John Bradley, Jane Falconer, Ruth H Keogh, Zhia Ning Lim, Behrouz Nezafat Maldonado, Laura Maynard-Smith, Ellen Sugrue, Okuda Taylor, Heidi Hopkins†, Audrey Dubot-Pérès†*

Summary

Background **Dengue fever is listed among the top ten global health threats by WHO. Prompt identification of dengue virus can guide clinical management and outbreak response, yet laboratory diagnosis is complex, costly, and lacks consensus on performance evaluation. This systematic review aims to provide reliable diagnostic accuracy estimates in order to inform global guidance and evaluate novel rapid diagnostic tests.**

Methods **In this systematic review and meta-analysis, we searched nine literature databases on Feb 16, 2021, for reports on five common reference tests for dengue infection: NS1 ELISA, IgM ELISA, IgG ELISA, RT-PCR, and viral neutralisation test. Articles were included if they reported primary data from more than five participants to complete 2×2 tables comparing one of these tests (on human serum) with any comparator. Diagnostic accuracy was estimated using Bayesian random-effect meta-analysis, which does not require a gold-standard comparator. Risk of bias was assessed using QUADAS-2. This review is registered with PROSPERO (CRD42022341552).**

Findings **Data were extracted from 161 articles, allowing analysis of multiple timeframes for three tests of interest. Pooled sensitivities of RT-PCR (0–4 days after symptom onset), NS1 ELISA (0–4 days), and IgM ELISA (1–7 days) were 95% (95% credible interval 77–99), 90% (68–98), and 71% (57–84), respectively. The corresponding pooled estimates of specificity were 89% (60–98), 93% (71–99), and 91% (82–95). A subanalysis of only studies at low risk of bias demonstrated similar estimates.**

Interpretation **IgM ELISA shows poor diagnostic accuracy early in the symptom course. NS1 ELISA shows similar diagnostic accuracy to RT-PCR, which has important implications for global public health policy, given its relatively low cost and accessibility.**

Funding **None.**

Introduction

Dengue fever is identified by WHO as one of the top ten threats to global health.[1] Dengue incidence has increased in recent years, with 5·2 million reported cases in 2019, a ten-fold increase since 2000.[2] Actual cases are likely underestimated as many are self-limiting.[3] WHO categorises dengue into three clinical groups: with warning signs, without warning signs, and severe.[4] Dengue can also be categorised as primary, defined as an individual's first dengue virus (DENV) infection, or secondary, referring to DENV infection in an individual who has been previously infected by another DENV serotype.[5]

Diagnostic methods for suspected DENV infection are setting-dependent. WHO defines laboratory confirmation of a probable case of dengue as one of: RT-PCR or viral culture positive; IgM or IgG seroconversion in paired serum samples; or four-fold IgG titre increase in paired serum samples.[4] Laboratory results highly suggestive of acute dengue include IgM positivity, IgG positivity with a haemagglutination inhibition titre of 1280 or greater, or detection of NS1 antigen by ELISA or rapid test, all of which relate to a single serum sample.[4] Novel rapid diagnostic tests (RDTs) for DENV are increasingly used, reflecting the need for timely diagnostics that do not require laboratory capacity in high-incidence settings, and align with the REASSURED criteria (highlighting real-time connectivity, ease of use, and environmental impact, in addition to the traditional ASSURED criteria).[6] Multiplex RDTs enable simultaneous testing for co-circulating pathogens with similar symptoms. To develop, evaluate, and regulate RDTs, the diagnostic accuracy of reference tests must be characterised.

Diagnostic accuracy of tests for acute dengue depend on multiple factors; critical among these is the timing of the test relative to disease progression, often measured in days post-onset of symptoms (DPO). Dengue viraemia is estimated to peak at 0–4 DPO. While methods that identify the virus, such as RT-PCR, are thought to be most accurate during this timeframe, methods focused on antibody response, such as IgM ELISA, peak later in the disease and remain positive for a longer period, limiting utility in detecting acute infection.[7]

Lancet Microbe 2025; 6: 101088

Published Online April 7, 2025 https://doi.org/10.1016/j.lanmic.2025.101088

*Contributed equally and share first authorship
†Contributed equally and share senior authorship

Faculty of Infectious and Tropical Diseases (K Pillay MSc, E Fitchett MPH, E Sugrue MSc, H Hopkins MD), **Faculty of Epidemiology and Population Health** (S H Keddie PhD, O Bärenbold PhD, Prof J Bradley PhD, Prof R H Keogh PhD), **and Library, Archive & Open Research Services** (J Falconer MA), **London School of Hygiene & Tropical Medicine, London, UK; DIATROPIX, Institut Pasteur de Dakar, Dakar, Senegal** (E Fitchett); **The Neo Child Initiative for Africa, Lagos, Nigeria** (C Akinde MSc); **Zayed Centre for Research, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK** (Z N Lim MRes); **Centre for Paediatrics and Child Health, Imperial College London, London, UK** (B Nezafat Maldonado MSc); **Division of Infection and Immunity, University College London, London, UK** (L Maynard-Smith MBBS); **Family Medicine Unit, Boniface Park Medical Centre, Toronto, ON, Canada** (O Taylor MSc); **Unité des Virus Émergents (UVE: Aix-Marseille Univ, Università di Corsica, IRD 190, Inserm 1207, IRBA), Marseille, France** (A Dubot-Pérès PhD)

Correspondence to:
Dr Kamla Pillay, Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK

Research in context

Evidence before this study

A PubMed search was performed to identify diagnostic accuracy of five common tests for acute dengue infection. No large-scale recent meta-analyses of diagnostic accuracy were identified. Furthermore, only traditional diagnostic accuracy meta-analytical methods had been used to estimate diagnostic accuracy of reference laboratory tests for dengue virus (DENV). The lack of a perfect comparator test is a major limitation of these methods. Estimates of diagnostic accuracy for each individual test vary based on which comparator is used. In clinical practice, DENV RT-PCR is seen as a gold standard, but performing this assay is expensive and difficult in resource-limited settings, and the short viraemic period in DENV infection limits its utility.

Added value of this study

This systematic review and meta-analysis is novel in terms of methods, size, and scope. To our knowledge, this is the largest review of the accuracy of dengue laboratory diagnostics to date. We used Bayesian effect latent-class analysis to compare tests of interest with any comparator, enabling us to capture a broad range of real-world data. We screened 11 048 articles and analysed data from 161 full texts to estimate the diagnostic accuracy of five reference laboratory tests for dengue.

Implications of all the available evidence

Our results indicate that IgM ELISA should not be used as a single test in the first 4 days of symptom onset, due to poor diagnostic accuracy in this period. Furthermore, we found that the diagnostic accuracy of NS1 ELISA is similar to RT-PCR, with implications for public health policy given ELISA's relatively low cost and ease of access. These estimates can also be used in future diagnostic accuracy studies, for example in the evaluation of novel rapid diagnostic tests, which have the potential to further improve dengue diagnosis in high-need settings.

The IgM antibody response to dengue starts around 4 DPO, and remains high for 2–3 months post-infection, while the IgG antibody response starts around 5 DPO.[5,8] Antibody responses can be more rapid in secondary infection. While NS1 antigen presence is related to viraemia, NS1 is detectable up to 10 DPO.[9]

See Online for appendix 1

Under-reporting and misdiagnosis of dengue due to other febrile illnesses, including co-circulating flaviviruses, necessitate quality-assured diagnostics for outbreak surveillance and assessment of control interventions.[9,10] Although published evaluations of dengue diagnostics exist, the lack of consensus on a gold standard and subsequent variation in reference standards makes interpretation of these difficult. To overcome these limitations, we undertook a large-scale systematic review and meta-analysis of the performance of five reference laboratory tests for dengue diagnosis: NS1, IgM, and IgG ELISA, RT-PCR, and viral neutralisation test (VNT). These tests were chosen because of their presence in the WHO and Pan American Health Organization (PAHO) guidelines, their use in reference laboratory testing for dengue, and the relative paucity of recently published accuracy estimates.[2,4,11] The primary purpose of this systematic review and meta-analysis is to publish accuracy estimates to inform the evaluation of novel tests. We estimate sensitivity and specificity using a Bayesian framework that assumes that the comparator tests are imperfect.

Methods

Search strategy and selection criteria

A search strategy was compiled in the OvidSP MEDLINE database. This included strings of terms, synonyms, and controlled vocabulary terms to reflect two concepts: dengue and diagnostic test of interest (TOI; ie, PCR, ELISA, or VNT). The search strategy was refined until the results retrieved reflected the scope of the project. Nine databases were searched on Feb 16, 2021: OvidSP MEDLINE, OvidSP Embase, OvidSP Global Health, Wiley Central Database of Controlled Trials, Web of Science Science Citation Index, Web of Science Social Sciences Citation Index, Elsevier Scopus, EBSCOhost Africa-Wide Information, and WHO Global Index Medicus (appendix 1 pp 1–16).

Results were initially limited to articles published from 2000 to the search date. This range was adjusted to a 10-year period (Jan 1, 2011, to Feb 16, 2021) at the full-text screening stage.

Citations were imported into EndNote 21 and duplicates removed. Unique results were imported into a systematic review software, CADIMA,[12] for abstract screening.

Papers were selected if they reported primary evidence from over five participants, were in English, Spanish, French, or Portuguese, were published in a peer-reviewed journal, and presented results for at least one TOI (NS1, IgG, or IgM ELISA, RT-PCR, or VNT) against a comparator using the same human serum sample. For serology (ie, IgG or IgM ELISA), paired samples were used if the first sample was tested against the comparator. Papers also needed to report accuracy data (sensitivity and specificity with denominators or a 2×2 table).

Abstracts were screened against the selection criteria by at least one of four reviewers (KP, EF, ZNL, BNM), using CADIMA.[12] Before screening, all reviewers conducted a consistency check of 50 abstracts with good concordance (kappa >0·80). Disagreements were discussed to clarify and harmonise the approach. Additionally, 10% of abstracts were screened by two reviewers independently to monitor consistency, which remained high (kappa >0·80). Abstracts marked as "unclear" were included for full-text screening.

Full-text articles were screened by at least one of seven reviewers (KP, SHK, EF, LM-S, ES, OT, CA). Consistency checks for every reviewer were compared with outcome based on review by KP and SHK, with good concordance

(kappa >0·80). Any full texts marked "unclear" were reviewed by a second reviewer (KP, SHK). A total of 570 (33%) full texts were reviewed independently by two reviewers. Discordant results were flagged for discussion in group meetings with a third reviewer, during which a consensus decision was made.

Data were extracted in duplicate by two independent reviewers (KP, SHK). Discordance between reviewers was minimal and any disagreements resolved by discussion. The following data were extracted: country of data collection, DENV serotype, participant population (including study selection criteria and demographic and clinical details), TOI and comparator used including brand and methodology, duration of sample storage, DPO, and a 2×2 table between TOI and comparator. For articles with multiple comparator tests, 2×2 tables were extracted for each pairwise comparison.

Data analysis

The quality of each study included was assessed using the Quality Assessment of Studies of Diagnostic Accuracy Approach-2 (QUADAS-2).[13] We modified this tool to make it suitable for our review question (appendix 1 pp 53–54). Each study was assessed in duplicate (KP and ES).

Meta-analysis estimated the diagnostic accuracy of the TOIs for acute dengue infection, with acute defined as symptom onset within the previous 2 weeks.[14] Data on timing of diagnostic testing days post-onset of symptoms were categorised into three groups depending on the design of the original studies, hereafter referred to as DPO subgroups: 0–4 days, 1–7 days, and an overall category of all acute symptomatic cases. The first two groups are mutually exclusive; no study was included in both the 0–4-day and 1–7-day groups. An additional DPO subgroup of 5–14 days was considered for IgM ELISA only. These were chosen to reflect the virological and immunological events used as diagnostic test targets and the resultant differences in diagnostic accuracy of TOIs at different DPO.[10,15–17] For NS1 ELISA and RT-PCR, the 0–4 days subgroup represents the primary analysis model. For IgM and IgG ELISA, the primary analysis model is the 0–7 days subgroup. An additional analysis of the most frequently reported IgM ELISA brand (Panbio Capture IgM ELISA [Panbio, Brisbane, QLD, Australia]) was performed for the all acute symptomatic group.

Where data were stated to be from symptomatic individuals but no DPO was given, these data were included in the all acute symptomatic analysis. Data from asymptomatic individuals, data where neither DPO nor symptom status was reported and data from symptomatic individuals with a stated DPO range that exceeded 14 days were excluded from all analyses. If a DPO subgroup had fewer than four studies, no meta-analysis was carried out.

Where articles included data from multiple comparator tests from the same individuals and DPO subgroup, multiple 2×2 tables were extracted. For all analyses, each 2×2 table represented a comparison between a TOI and a reference test within a unique cohort of individuals (hereafter referred to as a "study"; one published article may therefore provide multiple "studies" if reporting numerous diagnostic test comparisons).

We estimated pooled and predicted sensitivity and specificity, including 95% credible intervals (CrI), for each TOI. Pooled estimates represent the median accuracy across studies included in this analysis while predicted estimates represent the expected accuracy in another hypothetical future study. We used an extension to the Hierarchical Summary Receiver Operating Characteristic Model,[18] which relies on Bayesian latent-class analysis and is recommended in the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.[19] This approach assumes that the sensitivity and specificity of the TOI from each study lies on an underlying receiver operating characteristic curve and takes into account within-study and between-study variation as well as accounting for imperfect comparator tests.

Heterogeneity was explored by comparing the difference in pooled and predicted estimates[20] as well as by comparing the primary analysis DPO subgroup with the other subgroups. Initial plans to explore heterogeneity by participant age and study country were not possible due to small subgroup numbers. Sensitivity analyses included only studies at low risk of bias and with high external validity (appendix 1 p 54). Analyses were carried out in R 4.2.0 using Stan.[21] A full model specification is in appendix 1 (p 16).

Role of the funding source

There was no funding source for this study.

Results

Search results identified 28 043 articles, reducing to 11 048 after removing duplicates, to 2943 after abstract screening and to 2136 after restricting the publication date range (figure 1). After full-text review, 193 articles were identified as eligible (98 for RT-PCR, 123 for IgM ELISA, 64 for IgG ELISA, 67 for NS1 ELISA, and 12 for VNT). The included articles contributed 214, 122, 112, 219, and 28 of the 2×2 comparisons (studies) for IgM, IgG, and NS1 ELISA, RT-PCR, and VNT, respectively.

A full list of studies included in the meta-analyses is available in appendix 1 (pp 45–53). The results presented here are for IgM ELISA, NS1 ELISA, and RT-PCR (n=245 studies). The results for IgG ELISA were limited by lack of reporting of convalescent samples, substantially reducing their utility. There were insufficient data available to perform a meta-analysis for VNT performance. IgG ELISA and VNT results can be found in appendix 1 (pp 55–56) and appendix 2 (p 6).

See Online for appendix 2

IgM ELISA had the largest number of studies (n=107), representing 67 828 individuals (table). Included studies reported data from five of six WHO regions. Overall, half (137 [49%] of 282) were from the South-East Asia region and 124 (91%) of these were from India.

37 (15%) of 245 studies reported data separately for children and adults (eight [13%] of 60 for NS1 ELISA,

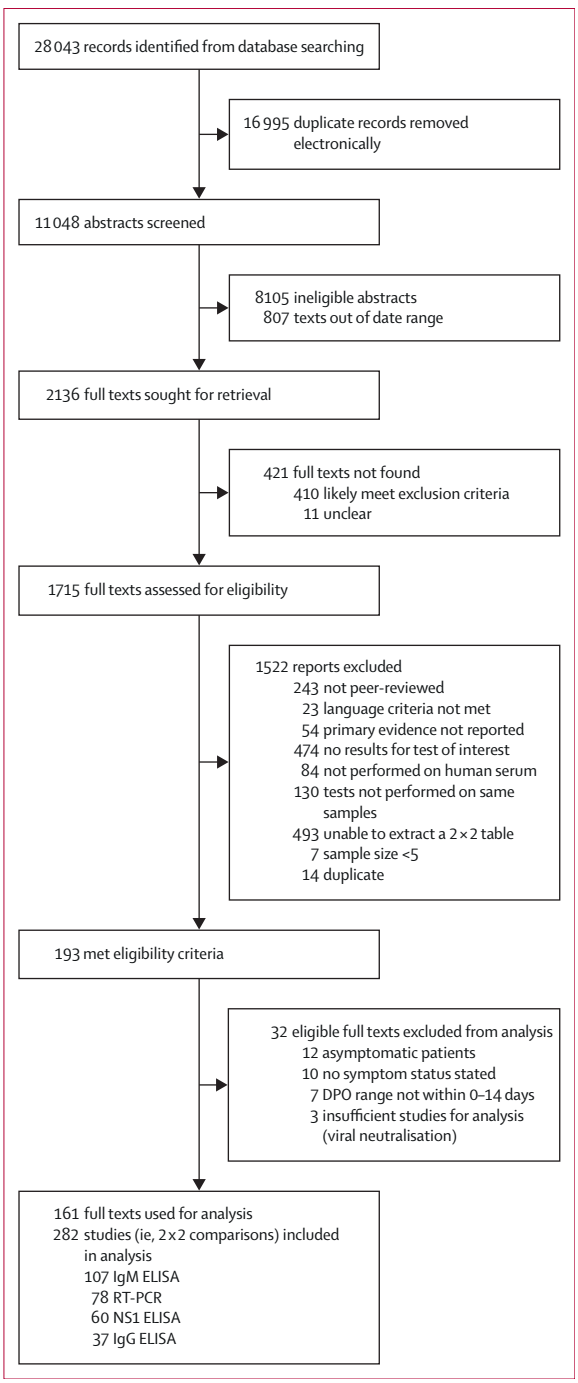

Figure 1: **Study selection**
DPO=days post-onset of symptoms.

11 [14%] of 78 for RT-PCR, and 18 [17%] of 107 for IgM ELISA). Similarly, 25 (14%) of 245 studies reported data separately for inpatients and outpatients (12 [20%] of 60 for NS1 ELISA, 12 [15%] of 78 for RT-PCR, and 21 [20%] of 107) for IgM ELISA; table).

Extracted data showed variability regarding the timing of diagnostic testing after onset of fever (figure 2). 101 (41%) of 245 studies stated that testing for acute dengue was performed on symptomatic individuals without listing DPO. This proportion was similar across TOIs (22 [37%] of 60 for NS1 ELISA, 37 [47%] of 78 for RT-PCR, and 42 [39%] of 107 for IgM ELISA).

Estimated sensitivity and specificity from our analysis for each TOI is shown in figure 3. Individual study estimates for both sensitivity and specificity for each TOI in each DPO subgroup are presented in appendix 1 (pp 45–53), which demonstrate the heterogeneity of results and variation in comparator tests.

Comparing results reported at 0–4 DPO, 1–7 DPO, and the all acute symptomatic group showed that sensitivity of RT-PCR and NS1 ELISA decreased with time after symptom onset. RT-PCR sensitivity decreased from 90% (95% CrI 68–98) to 86% (68–96) at 0–4 and 0–7 DPO, respectively. Similarly, NS1 ELISA decreased from 95% (77–99) to 85% (69–93). The overall point estimates for sensitivity and specificity of RT-PCR remained above 85% and 90%, respectively. Similarly, NS1 ELISA point estimates for sensitivity remained above 85% irrespective of DPO subgroup, with specificity above 90% (figure 3). Forest plots showing all studies for each TOI for each timeframe are in appendix 2 (pp 1–7).

Subgroup analyses of results reported at 1–7 DPO compared with 0–4 DPO demonstrated a marked difference in IgM ELISA sensitivity, at 71% (95% CrI 57–84) and 17% (3–51), respectively (figure 3). Comparing the IgM ELISA primary analysis model (1–7 DPO) with the model containing all acute symptomatic data, sensitivity and specificity estimates were lower at 62% (45–75) and 85% (76-91%), respectively. Sensitivity of IgM ELISA at 5–14 DPO versus 1–7 DPO was 82% (49–96) and 82% (45–96), respectively. The sensitivity and specificity of the PanBio Capture IgM ELISA (for all acute dengue) were 86% (68–94) and 87% (72–95), respectively (appendix 2 p 7).

Heterogeneity was seen across all meta-analyses as indicated by wider CrI for the predicted estimates when compared with the pooled estimates. For all TOIs, the model with the narrowest CrI for the predicted estimates was the primary analysis model and the widest CrI was seen for the all acute symptomatic model (figure 3).

Figure 4 summarises the risk of bias and applicability concerns for studies from each TOI. Overall, 107 (44%) of 245 studies were judged as having a low risk of bias and low concern regarding applicability. By TOI, this was 26 (43%) of 60 for NS1 ELISA, 29 (37%) of 78 for RT-PCR, and 52 (49%) of 107 for IgM ELISA. The lowest risk of bias was found in the index and comparator tests. 95 (39%) of 245 studies had high risk of bias from the flow and timing domain, mainly due to testing in the acute period with no specific DPO range.

148 (60%) of 245 studies had low applicability concerns regarding the timeline. A meta-analysis for the all acute symptomatic subgroup of studies with a low risk of both bias and applicability concerns is in appendix 1 (pp 54–55). This shows similar results to those presented in figure 3.

	Brand or technique, n (%)*	Study size, median (IQR)	Cohort, n (%)	Patient location, n (%)	Number of studies by WHO region, n (%)	DPO subgroup (number of studies; number of individuals)
NS1 ELISA (n=60)	Panbio 22 (37%); Platelia 13 (22%); J mitra 10 (17%)	193 (96–533)	Adults 4 (7%); children 4 (7%); both 27 (45%); not stated 25 (42%)	Outpatients 7 (15%); inpatients 5 (8%); both 6 (10%); not stated 42 (70%)	AFR 0 (0%) AMR 9 (15%); EUR 1 (2%); EMR 3 (5%); SEAR 38 (63%); WPR 7 (12%); multiple 2 (0%)	0–4 days (10; 1814); 1–7 days (16; 7949); all acute symptomatic (60; 31 084)
RT-PCR (n=78)	Real-time 22 (28%); conventional 2 (3%); not stated 54 (69%)	183 (98–307)	Adults 4 (5%); children 7 (9%); both 36 (46%); not stated 31 (40%)	Outpatients 5 (6%); inpatients 7 (9%); both 13 (17%); not stated 53 (68%)	AFR 4 (5%) AMR 23 (29%); EUR 1 (1%); EMR 2 (3%); SEAR 29 (37%); WPR 17 (22%); multiple 2 (0%)	0–4 days (4; 622); 1–7 days (27; 7519); all acute symptomatic (78; 21 402)
IgM ELISA (n=107)	Panbio 34 (32%); NIV 24 (22%); J mitra 8 (7%)	164 (67–461)	Adults 9 (8%); children 9 (8%); both 51 (48%); not stated 38 (36%)	Outpatients 9 (8%); inpatients 12 (11%); both 21 (20%); not stated 65 (61%)	AFR 5 (5%) AMR 21 (20%); EUR 2 (2%); EMR 4 (4%); SEAR 57 (53%); WPR 14 (13%); multiple 5 (5%)	0–4 days (7; 1209); 1–7 days (35; 20 057); all acute symptomatic (107; 69 488)
All (n=245)	NA	176 (80–376)	Adults 17 (7%); children 20 (8%); both 114 (47%); not stated 94 (38%)	Outpatients 21 (9%); inpatients 24 (10%); both 40 (16%); not stated 160 (65%)	AFR 9 (4%); AMR 53 (22%); EUR 4 (2%); EMR 9 (4%); SEAR 124 (48%); WPR 38 (17%); multiple 9 (4%)	0–4 days (21; 3645); 1–7 days (78; 35 525); all acute symptomatic (245; 1 219 74)

n refers to a study (defined for this review as a 2×2 comparison between a test of interest and a reference test within a unique cohort of individuals), not an article. WHO regions: African region (AFR), region of the Americas (AMR), European region (EUR), Eastern Mediterranean region (EMR), South-East Asia region (SEAR), Western Pacific region (WPR). DPO=days post-onset. NA=not applicable. NIV=National Institute of Virology. *The three most common brands or techniques for each test of interest are displayed. For companies or organisations listed in the table, product details are as follows: NS1 ELISA—NS1 antigen capture ELISA from Panbio (Brisbane, QLD, Australia), Platelia NS1 Ag from Bio-Rad (Hercules, CA, USA), Dengue NS1 Ag Microlisa from J Mitra (New Delhi, India); IgM ELISA—Dengue IgM Capture ELISA from Panbio (Brisbane, QLD, Australia), In-house IgM ELISA from NIV (Pune, India), Dengue IgM Microlisa from J Mitra (New Delhi, India).

Table: **Characteristics of studies included in analyses by test of interest**

Figure 2: **Reported DPO of test date by test of interest and study**

DPO subgroups are: 0–4 days, 1–7 days, and all acute symptomatic, which includes studies that did not report a DPO range but did state that testing was performed on symptomatic individuals for acute dengue (shown as grey bars spanning 0–14 days DPO). One additional analysis of 5–14 DPO was performed for IgM ELISA. Vertical dashed lines indicate DPO subgroup upper limits. The red arrow indicates the window of biological detection. DPO=days post-onset of symptoms.

Discussion

We report, to our knowledge, the most comprehensive review to date of the accuracy of common reference laboratory tests for acute dengue detection. Our analysis estimates the sensitivity and specificity of NS1 ELISA (0–4 DPO) at 90% (95% CrI 68–98) and 93% (71–99), respectively. Our results suggest NS1 ELISA to be more sensitive and less specific than previous published data

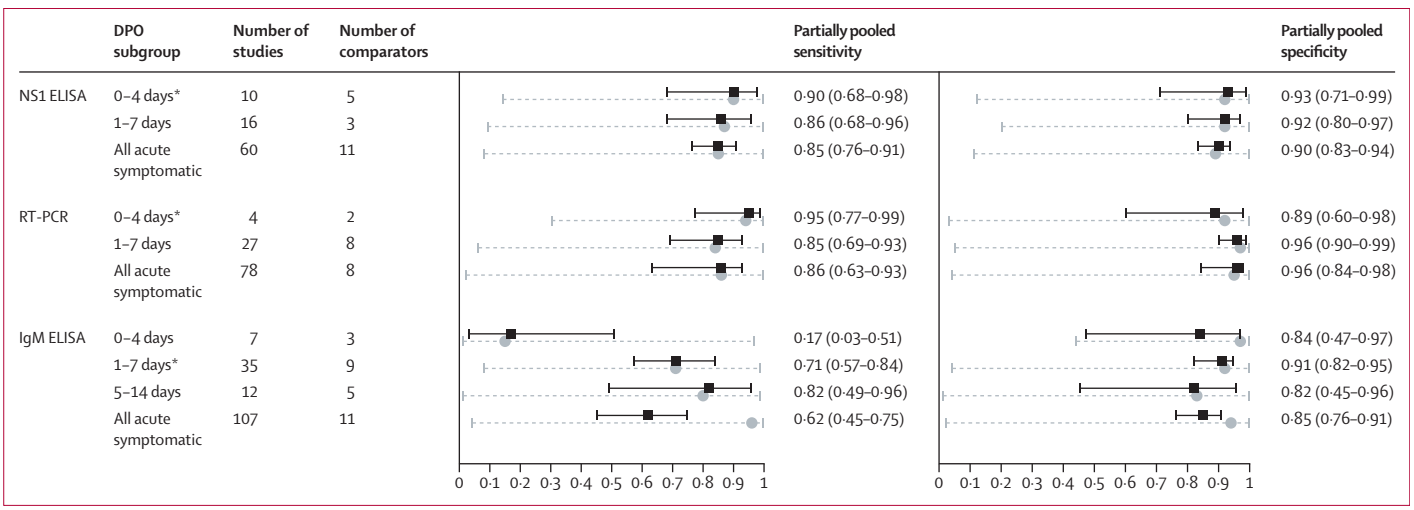

Figure 3: **Summary forest plot of pooled (black solid line) and predicted (grey dashed line) estimates of sensitivity and specificity from separate meta-analyses for each test of interest**
Stratified by DPO subgroup. Data in parentheses are 95% credible intervals. DPO=days post-onset of symptoms. *Indicates primary analysis subgroup for each test of interest.

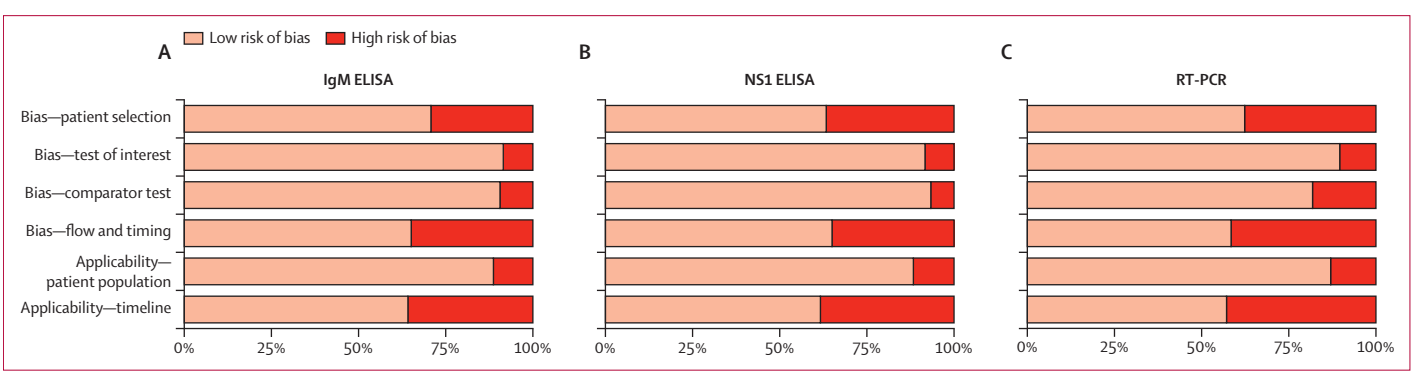

Figure 4: **Summary of the risk of bias and applicability concerns across six domains in the accuracy of dengue diagnostics for the detection of acute dengue**
Four domains for risk of bias and two domains for concerns for applicability were used, adjusted from QUADAS-2.[13]

(one study estimated sensitivity of 74% [95% CI 63–82] and specificity of 99% [97–100]).[22] Generally, previous analyses of NS1 ELISA accuracy have used RT-PCR as a gold-standard comparator test,[23,24] making the assumption that RT-PCR is 100% sensitive and 100% specific. If this assumption is incorrect, previous estimates of NS1 ELISA diagnostic accuracy will have been underestimated (assuming conditional independence). One study compared NS1 ELISA and RT-PCR using a third diagnostic method as a gold-standard comparator and found that the two tests had similar diagnostic accuracy, reflecting our findings.[25] In comparison with the 0–4 DPO subgroup, where RT-PCR has a better sensitivity than NS1 ELISA, in the 1–7 DPO subgroup, NS1 ELISA was found to have a similar sensitivity (86% [95% CrI 68–96]) to RT-PCR (85% [69–93]). These results highlight NS1 ELISA as a practical alternative to RT-PCR, given the persistence of the NS1 antigen in the bloodstream and the common presentation of patients after 4 DPO.

IgM ELISA was found to be 17% sensitive and 84% specific at 0–4 DPO, reflecting the lag in DENV IgM antibody titres in response to infection.[26,27] Our results suggest that IgM ELISA in a single acute sample must be interpreted with DPO in mind given the increase in sensitivity after 4 DPO, and, as per WHO guidance, should not be used as a confirmatory test.[4] At 1–7 DPO, pooled sensitivity and specificity of IgM ELISA were 71% (95% CrI 57–84) and 91% (82–95), respectively, reflecting increasing antibody titres. These results are likely to underestimate the actual diagnostic accuracy of IgM ELISA in the 5–7 DPO subgroup, given the inclusion of samples from less than 5 DPO. One previously reported estimate of IgM ELISA sensitivity and specificity was 38·1% and 100%, respectively.[23] Notably, although these estimates were not stratified by DPO, the majority of patients included in the analysis (44 of 86) were in the 1–3 DPO range, which might explain the poor sensitivity reported. Similarly, the higher sensitivity of 82% (95% CrI 49–96) and specificity of 82% (45–96) of IgM ELISA at 5–14 DPO in our study is physiologically

consistent, although the CrI is wide and overlaps with the 1–7 DPO results, reflecting the smaller number of studies in this subgroup. These results confirm the clinical relevance of IgM detection during this common timeframe for health-seeking behaviour and peak IgM response.

However, IgM titres rise less in response to secondary infection compared with primary DENV infection, limiting the utility of this test particularly in dengue-endemic regions.[26] Few studies in this review distinguished primary from secondary infections,[28–30] preventing sub-analysis. While IgM ELISA is often used as a single test in the acute phase, its accuracy might be increased by using paired samples, interpretation as a ratio with IgG ELISA, or using a single measurement in combination with other diagnostic methods as part of a diagnostic algorithm.[10,27] Meta-analyses to ascertain diagnostic accuracy of dengue algorithms should be a priority.

Differences in the diagnostic performance of commercial ELISA tests for NS1 and IgM ELISA for acute dengue infection have been demonstrated.[31] While there were insufficient data to perform additional analyses for each unique commercial test at each timeframe, a sub-analysis of the PanBio Capture IgM ELISA within the all acute symptomatic group suggested higher sensitivity (86% [95% CrI 68–94]) compared with all brands (62% [45–75]), although CrIs overlapped. Our findings highlight the need for comparative studies of commercial ELISAs using well-characterised serum panels to guide assay selection.

Strengths of this review include the comprehensive and robust search strategy, covering five reference laboratory tests for dengue diagnosis, across 11 bibliographic databases, which identified studies from 43 countries with diverse comparators. Furthermore the methodology used was adherent to QUADAS-2 and PRISMA guidelines. We used a Bayesian analysis that did not assume that any comparator was perfect. Although this approach has previously been used for rapid tests for dengue,[32] to our knowledge it has not previously been applied to a dataset this large for our TOIs.

The included studies had multiple limitations. Firstly, dengue diagnostics have been shown to have optimal performance at certain DPOs, depending on the diagnostic target.[10] However, studies reported overlapping DPO ranges for acute dengue from 1 DPO to 36 DPO for all TOIs (figure 2). We were therefore unable to stratify results into discrete DPO windows (ie, 0–4, 5–7, and 8–14 DPO), which would have strengthened interpretability. This inconsistent reporting highlights a lack of consensus on timeframes for acute dengue diagnosis, which hinders the ability to perform meta-analyses. Secondly, heterogeneity in reporting of important variables limited our ability to perform subgroup analysis. Such variables include DENV serotype, patient population (including clinical suspicion of dengue infection, which would affect the pre-test probability), primary or secondary infection, brand of TOI, and geographical location of study. The DPO subgroup analysis, although adding depth to our conclusions, resulted in data paucity for some timeframes. Notably, the 0–4 DPO RT-PCR was based on only four studies, and should be interpreted with caution. Lack of differentiation between primary and secondary infections in particular could affect our estimates of diagnostic accuracy, given that previous antibody production is likely to alter the diagnostic accuracy of IgM and NS1 ELISA in secondary infection when compared with primary infection. Given the previously described limitations in data reporting, a sub-analysis of these groups could not be performed. However, since primary and secondary infections are often indistinguishable at the time of presentation, our estimates of diagnostic accuracy remain useful in clinical practice. Thirdly, in order to maximise the data available for analysis, we included any clinical study that met the inclusion criteria, rather than restricting inclusion to diagnostic accuracy studies. Consequently, many studies poorly described aspects of test methodology and population group studied, highlighted by the results of QUADAS-2, with potential impacts on the overall estimates of both sensitivity and specificity.

Our meta-analytical approach also has limitations. While our analysis framework improves upon previous analyses by accounting for imperfect comparator tests, there are other model assumptions necessary that can lead to biased estimates of sensitivity and specificity, notably the choice of prior distributions for the comparator tests' diagnostic accuracy and the assumption of a two-state latent-class model.[33]

To summarise, our results show that IgM ELISA should not be used as a single test in the first days of symptom onset, due to poor diagnostic accuracy in this period. Additionally, we find that the diagnostic accuracy of NS1 ELISA is similar to RT-PCR. Our estimates provide a basis for future accuracy assessments of dengue tests designed for near-patient use, for example in the evaluation of novel RDTs and diagnostic algorithms, using both Bayesian and traditional analysis approaches. These findings have implications for dengue diagnostic test development, evaluation, and regulation, as well as global health and surveillance policy, especially in relation to the utility of NS1 ELISA.

Contributors

HH, EF, and KP conceived the study. KP, SHK, EF, LM-S, ES, OT, BNM, ZNL, and CA assessed the eligibility of the studies, KP, SHK, and ES extracted the data, and KP and ES assessed the methodological quality of the included studies. KP, ES, EF, and SHK verified the data reported in this study. JF developed the search strategy and conducted the literature search. SHK carried out the statistical analysis. JB, RHK, and OB advised on the statistical analysis. KP and SHK prepared the original draft of the manuscript, with considerable input from EF, HH, and AD-P. AD-P advised throughout on technical aspects of dengue diagnostic testing, and HH on methodological issues for the review. All authors reviewed, edited, and approved the final manuscript. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Declaration of interests

We declare no competing interests.

Data sharing

All model code can be found at: https://github.com/shk313/diagnostic-test-metaanalysis/tree/main/Dengue.

Acknowledgments

SHK is supported by the Medical Research Council London Intercollegiate Doctoral Training Partnership Studentship (MR/N013638/1).

References

1 WHO. Ten threats to global health in 2019. https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019 (accessed Jan 11, 2023).
2 WHO. Disease outbreak news: Dengue—global situation. Dec 21, 2023. https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON498 (accessed Jan 11, 2024).
3 Zeng Z, Zhan J, Chen L, Chen H, Cheng S. Global, regional, and national dengue burden from 1990 to 2017: a systematic analysis based on the global burden of disease study 2017. *EClinicalMedicine* 2021; **32:** 100712.
4 WHO. Dengue outbreak toolbox. 2019. https://www.who.int/docs/default-source/outbreak-toolkit/updates-documents_july-5/dengue-outbreak-data-collection-toolbox—inis-3-july-1.pdf?sfvrsn=ec3ffcf3_2#:~:text=WHO%20surveillance%20case%20definition,leukopenia%2C%20or%20any%20warning%20sign (accessed Jan 11, 2023).
5 Halstead SB. Dengue antibody-dependent enhancement: knowns and unknowns. *Microbiol Spectr* 2014; **2:** 2.6.30.
6 Land KJ, Boeras DI, Chen X-S, Ramsay AR, Peeling RW. REASSURED diagnostics to inform disease control strategies, strengthen health systems and improve patient outcomes. *Nat Microbiol* 2019; **4:** 46–54.
7 WHO. Dengue: guidelines for diagnosis, treatment, prevention and control, new edition. World Health Organization, 2009.
8 Guzman MG, Halstead SB, Artsob H, et al. Dengue: a continuing global threat. *Nat Rev Microbiol* 2010; **8** (suppl)**:** S7–16.
9 Kularatne SAM. Dengue fever. *BMJ* 2015; **351:** h4661.
10 Raafat N, Blacksell SD, Maude RJ. A review of dengue diagnostics and implications for surveillance and control. *Trans R Soc Trop Med Hyg* 2019; **113:** 653–60.
11 PAHO. Guidelines for the clinical diagnosis and treatment of dengue, chikungunya, and Zika. Pan American Health Organization, 2021.
12 Kohl C, McIntosh EJ, Unger S, et al. Online tools supporting the conduct and reporting of systematic reviews and systematic maps: a case study on CADIMA and review of existing tools. *Environ Evid* 2018; **7:** 8.
13 Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; **155:** 529–36.
14 WHO. Dengue and severe dengue. 2023. https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue (accessed Dec 5, 2023).
15 Duyen HTL, Ngoc TV, Ha T, et al. Kinetics of plasma viremia and soluble nonstructural protein 1 concentrations in dengue: differential effects according to serotype and immune status. *J Infect Dis* 2011; **203:** 1292–300.
16 Erra EO, Korhonen EM, Voutilainen L, Huhtamo E, Vapalahti O, Kantele A. Dengue in travelers: kinetics of viremia and NS1 antigenemia and their associations with clinical parameters. *PLoS One* 2013; **8:** e65900.
17 Tricou V, Minh NN, Farrar J, Tran HT, Simmons CP. Kinetics of viremia and NS1 antigenemia are shaped by immune status and virus serotype in adults with dengue. *PLoS Negl Trop Dis* 2011; **5:** e1309.
18 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:** 1285–93.
19 Takwoingi Y, Dendukuri N, Schiller I, Rücker G, Jones HE, Partlett C. Chapter 10: Undertaking meta-analysis. In: Deeks JJ, Bossuyt PM, Leeflang MM, Takwoingi Y, eds. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 2.0 (updated July 2023). https://training.cochrane.org/handbook-diagnostic-test-accuracy/current (accessed Nov 25, 2023).
20 Takwoingi Y, Dendukuri N, Schiller I, Rücker G, Jones HE, Partlett C. Chapter 11: Presenting findings. In: Deeks JJ, Bossuyt PM, Leeflang MM, Takwoingi Y, eds. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 2.0 (updated July 2023). https://training.cochrane.org/handbook-diagnostic-test-accuracy/current (accessed Nov 25, 2023).
21 Stan Development Team. Stan modeling language user's guide and reference manual. 2017. http://mc-stan.org/users/documentation/ (accessed Jan 10, 2024).
22 da Costa VG, Marques-Silva AC, Moreli ML. A meta-analysis of the diagnostic accuracy of two commercial NS1 antigen ELISA tests for early dengue virus detection. *PLoS One* 2014; **9:** e94655.
23 Luvira V, Thawornkuno C, Lawpoolsri S, et al. Diagnostic performance of dengue NS1 and antibodies by serum concentration technique. *Trop Med Infect Dis* 2023; **8:** 117.
24 Iqbal G, Javed H, Raza FA, Gohar UF, Fatima W, Khurshid M. Diagnosis of acute dengue virus infection using enzyme-linked immunosorbent assay and real-time PCR. *Can J Infect Dis Med Microbiol* 2023; **2023:** 3995366.
25 Ahmed NH, Broor S. Comparison of NS1 antigen detection ELISA, real time RT-PCR and virus isolation for rapid diagnosis of dengue infection in acute phase. *J Vector Borne Dis* 2014; **51:** 194–99.
26 Chanama S, Anantapreecha S, A-nuegoonpipat A, Sa-gnasang A, Kurane I, Sawanpanyalert P. Analysis of specific IgM responses in secondary dengue virus infections: levels and positive rates in comparison with primary infections. *J Clin Virol* 2004; **31:** 185–89.
27 Hunsperger EA, Muñoz-Jordán J, Beltran M, et al. Performance of dengue diagnostic tests in a single-specimen diagnostic algorithm. *J Infect Dis* 2016; **214:** 836–44.
28 Abhishek KS, Siddiqui O, Chakravarti A. Increasing number of secondary dengue cases: a concern. *Int J Res Med Sci* 2019; **7:** 1558–61.
29 Mathews AA, Boppe A, Thomas M. Role of NS1 antigen in the diagnosis of dengue viral infection. *J Commun Dis* 2017; **49:** 52–55.
30 Paudel D, Jarman R, Limkittikul K, et al. Comparison of real-time SYBR green dengue assay with real-time taqman RT-PCR dengue assay and the conventional nested PCR for diagnosis of primary and secondary dengue infection. *N Am J Med Sci* 2011; **3:** 478–85.
31 Lee H, Ryu JH, Park HS, et al. Comparison of six commercial diagnostic tests for the detection of dengue virus non-structural-1 antigen and IgM/IgG antibodies. *Ann Lab Med* 2019; **39:** 566–71.
32 Pan-ngum W, Blacksell SD, Lubell Y, et al. Estimating the true accuracy of diagnostic tests for dengue infection using Bayesian latent class models. *PLoS One* 2013; **8:** e50765.
33 Schofield MR, Maze MJ, Crump JA, Rubach MP, Galloway R, Sharples KJ. On the robustness of latent class models for diagnostic testing with no gold standard. *Stat Med* 2021; **40:** 4751–63.