Transactions of the Royal Society of Tropical Medicine and Hygiene (2006) **100**, 775—784

available at www.sciencedirect.com

SCIENCE DIRECT®

journal homepage: www.elsevierhealth.com/journals/trst

A systematic review and meta-analysis of the diagnostic accuracy of rapid immunochromatographic assays for the detection of dengue virus IgM antibodies during acute infection

Stuart D. Blacksell[a,b,c], Jenny A. Doust[d], Paul N. Newton[a,c], Sharon J. Peacock[a,b], Nicholas P.J. Day[a,b,c,*], Arjen M. Dondorp[a,b]

[a] *Centre for Clinical Vaccinology and Tropical Medicine, Nuffield Department of Clinical Medicine, Churchill Hospital, University of Oxford, Oxford, UK*
[b] *Wellcome Trust—Mahosot University—Oxford Tropical Medicine Programme, Faculty of Tropical Medicine, Mahidol University, 420/6 Rajvithi Road, Bangkok, 10400, Thailand*
[c] *Wellcome Trust—Mahosot Hospital—Oxford Tropical Medicine Research Collaboration, Microbiology Laboratory, Mahosot Hospital, Vientiane, Lao PDR*
[d] *Centre for General Practice, School of Medicine, University of Queensland, Herston, Queensland, Australia*

Received 26 July 2005; received in revised form 7 October 2005; accepted 25 October 2005
Available online 23 March 2006

KEYWORDS
Dengue;
Diagnosis;
Immunochromatographic test;
IgM;
Antibody

Summary A meta-analysis of rapid (≤60 min) dengue diagnostic assays was conducted to determine accuracy and identify causes of between-study heterogeneity. A systematic review identified 302 potentially suitable studies, of which 11 were selected for meta-analysis. All selected studies evaluated the immunochromatographic test (ICT) manufactured by Panbio Pty Ltd. Individual study results for sensitivity ranged from 0.45 to 1.0, specificity 0.57—1.0, diagnostic odds ratio 4.5—1287, and positive:negative likelihood ratios 2.3—59 and 0.01—0.56, respectively. Results indicated that the ICT evaluated in the selected studies can both rule in and rule out disease but is more accurate when samples are collected later in the acute phase of infection. Limitations of this meta-analysis were significant between-study heterogeneity caused by inconsistencies in evaluation methodologies, and the evaluation of only the Panbio ICT. It is recommended that additional, standardized evaluations are required for other dengue ICTs.

* Corresponding author. Present address: Wellcome Trust—Mahidol University—Oxford Tropical Medicine Programme, Faculty of Tropical Medicine, Mahidol University, 420/6 Rajvithi Road, Bangkok, 10400, Thailand. Tel.: +66 2 3549172; fax: +66 2 3549169.
E-mail address: nickd@tropmedres.ac (N.P.J. Day).

0035-9203/$ — see front matter © 2005 Royal Society of Tropical Medicine and Hygiene. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.trstmh.2005.10.018

1. Introduction

Dengue fever (DF), dengue haemorrhagic fever (DHF) and dengue shock syndrome (DSS) represent a spectrum of disease resulting from the transmission of dengue virus by the mosquito vector *Aedes aegypti*. The spectrum of clinical illness is broad and ranges in severity from mild symptoms to death. Dengue infection is clinically similar to many other acute febrile illnesses, such as malaria, rickettsioses, chikungunya and leptospirosis (Gubler, 1998), and serology plays an important role in patient diagnosis and in guiding management (Guzman and Kouri, 2004). However, antibodies from related flaviviruses, such as Japanese encephalitis virus (JEV) may cross-react and cause a false-positive result (Guzman and Kouri, 2004; Shu and Huang, 2004). DHF and DSS, the more severe clinical manifestations of dengue infection, are thought to occur more commonly in those with a second or subsequent infection (Halstead, 2003). Laboratory diagnosis may provide an indication of whether or not a given infection represents the first or subsequent episode, alerting the clinician to the possibility of more severe disease in those with reinfection (Halstead, 2003).

Diagnosis relies on serological tests based on the detection of dengue-specific IgM antibodies during the acute phase of infection, a four-fold rise in antibody titre in paired serum collections, or a single serum with a haemagglutination inhibition (HAI) assay titre of $\geq$1:2560 (Guzman and Kouri, 2004). The development of rapid immunochromatographic tests (ICTs) represents a potentially important advance, as it provides a mechanism for simple and rapid bedside serological testing (Price, 2001; Vaughn et al., 1998). Many manufacturers of dengue ICTs also claim that their products can differentiate between primary and secondary dengue infection. Several studies have compared dengue ICTs with reference assays, but their diagnostic accuracy has not been reliably established, largely because of the multiplicity of evaluation methodologies used (Shu and Huang, 2004). Here, we present a meta-analysis to: (1) determine the accuracy of rapid ($\leq$60 min), point-of-care dengue assays for the diagnosis of acute dengue infection when compared with reference assays; and (2) evaluate and explain heterogeneity between studies included in the meta-analysis.

2. Materials and methods

2.1. Search strategy

Searches were conducted on the following databases: MEDLINE (1966—March 2004), EMBASE (1994—March 2004), Cochrane Library, Web of Science and SCIRUS. The search terms used were 'dengue', 'IgM', combined with a previously validated filter for diagnostic studies (van der Weijden et al., 1997) using the terms 'diagnosis', 'sensitivity', 'specificity', 'monitoring', 'ROC', 'reference value'. Reference lists of each of the selected articles were hand-searched for additional studies. No language restrictions were imposed.

2.2. Study selection using standardized quality assessment criteria

Abstracts of identified studies were printed, and if potentially relevant were obtained as full-text articles. Articles were assessed using the standardized quality tool QUADAS (Whiting et al., 2003) by two researchers (S.D.B. and A.M.D.) and given a score out of 14. Studies were excluded if they had any of the following characteristics: (1) use of inappropriate reference assays to assign true positive/true negative status to study samples, including 'in-house' assays for which the diagnostic accuracy had not been previously established; (2) inappropriate study population (such as convalescent samples only); (3) the study was limited to the detection of IgG rather than IgM and IgG; (4) the number of study samples was insufficient; (5) incomplete description of samples, such that it was impossible to determine the timing of sample collection; (6) errors or inconsistencies in the published study data; (7) the exclusion of indeterminate results; (8) partial verification of the study samples or the use of multiple reference assays; (9) the assay took more than 60 min to perform, such as immunoblot (IBT)-style assays.

2.3. Data extraction

Data were extracted independently from the accepted studies by two researchers (S.D.B. and A.M.D.) and recorded onto a standard form. Discrepancies were resolved by mediation. If multiple result sets were included in a single study, only IgM results for admission sera that were compared with a valid reference comparator were extracted. In the case of primary and secondary infections, the infection status diagnosis was assigned using the criteria described by the individual study.

2.4. Statistical analysis

The 'gold standard' (or reference) assay was compared with the index test to define true-positive (Tp), false-positive (Fp), false-negative (Fn) and true-negative (Tn) values. The measures of diagnostic accuracy, sensitivity (Sn), specificity (Sp), positive likelihood ratio (LR+), negative likelihood ratio (LR−) and diagnostic odds ratio (DOR), were then calculated (Habbema et al., 2002). Individual study results were pooled to generate an overall estimate of diagnostic accuracy. Chi-squared and I^2 (Higgins et al., 2003) statistics were calculated before pooling to detect any significant heterogeneity overall and between subgroups. A χ^2 result of $P < 0.1$ was considered significant, given the low power of the test. I^2 values have a continuous scale of 0—100%, with 0% defining no heterogeneity and 25, 50 and 75% having been tentatively assigned as limits of low, medium and high heterogeneity (Higgins et al., 2003). If heterogeneity was not significant, a Mantel—Haenszel fixed effects model (Mantel and Haenszel, 1959) was used to calculate results and, when significant, a random effects model was used (DerSimonian and Laird, 1986). Summary receiver operator characteristic (SROC) (Littenberg and Moses, 1993) were also calculated to give a final area under the curve (AUC) value for pooled and subgroup analyses. Analyses were performed using Stata™