Markers of dengue severity: a systematic review of cytokines and chemokines

Yie Hou Lee,¹,² Wei-Yee Leong³ and Annelies Wilder-Smith³

¹KK Research Centre, KK Women's and Children's Hospital, 100 Bukit Timah Road, Singapore 229899, Singapore
²Infectious Diseases Interdisciplinary Group, Singapore-MIT Alliance for Research and Technology, 1 CREATE Way, #03-12/13/14 Enterprise Wing, Singapore 138602, Singapore
³Lee Kong Chian School of Medicine, Nanyang Technological University, 11 Mandalay Road, Singapore 308232, Singapore

The prognosis of dengue remains a challenge in the early, objective triage of patients with dengue fever of differing severity. Circulating immuno-modulating proteins have brought new possibilities as prognostic markers of severe dengue (SD). This systematic review is devoted to understanding the potential utility of blood-based cytokines and chemokines as prognostication markers of SD based on the current literature. PubMed and Embase were searched. Of 794 candidate articles, 685 abstracts were screened against our exclusion/inclusion criteria and 25 (3.6%) studies met the quality assessments. A total of 18 studies were retrospective observational and 2 were prospective cohort studies. Elevated IL-10, up to day 7 of fever onset, stood out as a candidate prognostic marker for SD using the 1997 and 2009 World Health Organization (WHO) case definitions. IFNγ was another potential prognostic marker of SD (1997 WHO case definition), but its levels varied between studies. Significant heterogeneity in methodologies and patient cohorts prevent ready application of IL-10 and IFNγ as prognostic markers to other dengue populations. Our results suggest that the current non-randomized studies are delivering inconsistent messages and higher-quality studies, with consistent methodologies and validation in independent patient cohorts, are needed to delineate confounding variables. Major gaps identified were full accounting and transparency of sampling days, dengue virus type, infection status and age group.

Introduction

Dengue is an emerging arboviral threat globally as its spread and incidence is moving on an upward trajectory, infecting some 390 million people yearly (Bhatt et al., 2013). Dengue can be asymptomatic, or present as a self-limiting dengue fever (DF), and in certain patient subsets the disease may precipitate into the severe, potentially life-threatening forms of dengue — dengue haemorrhagic fever (DHF) and dengue shock syndrome (DSS) or severe dengue (SD). In the 1997 World Health Organization (WHO) case definition, patients were diagnosed as having DHF if they had fever, thrombocytopenia, bleeding and evidence of plasma leakage (hypoproteinaemia, change in haematocrit of more than 20% or clinical fluid accumulation) (WHO, 1997). DSS has all the features of DHF plus circulatory failure in the form of rapid and weak pulse, and narrow pulse pressure (<20 mm Hg), or age-specific hypotension and cold, clammy skin and restlessness. In the 2009 WHO case definition, patients were diagnosed as SD if they had severe plasma leakage, severe bleeding and severe organ involvement (aspartate aminotransferase or alanine aminotransferase ≥1000 U/L, impaired consciousness, or failure of heart and other organs) (WHO, 2009). Dengue can be caused by any of the four dengue types of the genus Flavivirus. The propensity to develop SD is plausibly brought about by a number of factors, including heterologous secondary dengue infection, dengue type, age or a combination of these (Kyle & Harris, 2008; OhAinle et al., 2011). There are no effective antiviral therapeutics to treat dengue and the current most advanced vaccine has an imbalanced efficacy against the four dengue types (Capeding et al., 2014; Wilder-Smith & Massad, 2016). In view of this, early prognosis of SD, preferably less than 96 h from onset of fever, can guide patient triage, allow informed clinical decisions, and reduce disease morbidity and mortality.

The dynamic nature of dengue presents significant challenges in clinical management and on health services,
especially during an outbreak. Currently, routine haematological and biochemical measurements in dengue patients collectively correlate poorly with eventual clinical outcome, there are no tests to differentiate those who will have DF from those who will progress to and deteriorate to SD. In the absence of any effective treatment against the disease, proper fluid management is critical. Guidelines produced by the WHO list a number of warning signs to help inform clinical decision making, but they all have poor specificity resulting in large numbers being admitted unnecessarily (Thein et al., 2013). Over-hospitalization of dengue patients is attributed to in part by the inability to prognosticate early in the febrile phase, and resulted in excessive hospitalization rates and costs (Lee et al., 2013). The unpredictable nature of outbreaks often overwhelms already fragile health-care systems. Prognosis is also important for monitoring dengue patients who display probable warning signs of health deterioration or complications that warrant further investigation. A prognostic marker is defined as a clinical or biological characteristic that is objectively measurable and that provides information on the likely outcome of the disease in an untreated individual (Hayes et al., 1998).

Faced with the limited utility of current methods of SD prognosis, there has been interest in investigating the utility of systemic soluble factors as potential markers of SD prognosis. The two main prevailing hypotheses of DHF/DSS pathogenesis, antibody-dependent enhancement (ADE) and original antigenic sin (or T cell immunopathology), centre on an imbalanced immune system during secondary dengue infection (Halstead & O'Rourke, 1977; Mongkolsapaya et al., 2003) triggering an exaggerated and imbalanced inflammatory cascade (Srikiatkhamhorn & Green, 2010). In addition to secondary infections, the waning maternally-derived anti-dengue virus (DENV) IgG antibodies and/or an altered cytokine profile may explain SD in infants with primary infections (Libraty et al., 2009; Nguyen et al., 2004). Soluble factors emanating from immune cells (T cells, B cells, macrophages, mast cells, dendritic cells, etc.), platelets, stromal and endothelial cells in the form of cytokines and chemokines act as signalling molecules synergizing with one another to orchestrate cell growth and proliferation, differentiation/maturation, and immunity; hence, modulating host responses to infections (Fink et al., 2006). Cytokines and chemokines are small proteins typically ranging from 8 to 40 kDa, and in this review we broadly refer to them as immuno-modulating proteins. Cytokines are secreted proteins that play a role in cell signalling, in the induction, inhibition and effector phases of immune and inflammatory responses. Chemokines are a subset of small cytokines that recruit and exert chemotactic migration of other cells to a localized area to exert a variety of biological effects, including inflammation and homeostasis. Immuno-modulating protein profiles change with the clinical course of dengue, differ between DF and SD (Kumar et al., 2012; Rathakrishnan et al., 2012), and are believed to have direct impact on the manifestations of increased vascular permeability, plasma leakage and thrombocytopenia (Green & Rothman, 2006; Murphy & Whitehead, 2011). Immuno-modulating proteins have been proposed to cause a shift from the predominant TH1-type response in DF to the TH2-type in severe DHF (Fink et al., 2006); because molecular signalling seemingly precedes gross morphological or observable clinical symptoms, the potential use of immuno-modulating proteins as early prognostic markers is especially welcoming (Lee & Ooi, 2013). However, studies in humans have resulted in varied responses and remained conflicting, with no objective consolidation of data having taken place so far. Hence, this systematic review aims to be a collection and summary of primary research articles that focuses on the ostensible utility of soluble immuno-modulating markers for the prognosis of SD.

**Results**

**Study selection**

We ran searches on PubMed and Embase, and a total of 794 references were retrieved: PubMed (n=451) and Embase (n=343). After excluding 109 duplicates, 685 records were screened on the basis of title and abstract against our exclusion/inclusion criteria, thereby identifying 54 potentially eligible records (Table S1, available in the online Supplementary Material). Two studies were excluded because the main text could not be retrieved or had incomplete methods. Another 30 records were excluded (4 low quality, 25 medium quality and 1 with ambiguous sampling time). Some of the studies were further moderated based on empirical evidence leading to six studies being downgraded and one study being upgraded (Table S1). A total of 24 studies were identified as eventual eligible articles of high quality. The study selection flow chart is shown in Fig. 1. We restricted the search to ‘English’ language and ‘Human’ subjects. Three studies were excluded despite good study design due to a lack of comparison to DHF/DSS or SD (medium to high quality). Detection bias (blinding of outcome assessors) was relatively homogeneous across these non-randomized studies (NRS) and leaves all studies at approximately the same level of risk of bias and, thus, not considered.

**Study characteristics**

A total of 19 out of 24 studies were case–control studies (Arias et al., 2014; Bozza et al., 2008; Brasier et al., 2012; Butthep et al., 2012; Chen et al., 2005; Furuta et al., 2012; Green et al., 1999a, b; Guerrero et al., 2013; Houghton-Triviño et al., 2010; De La Cruz Hernández et al., 2014; Laur et al., 1998; Levy et al., 2010; Malavige et al., 2013a; Nguyen et al., 2004; Del Moral-Hernández et al., 2014; Pérez et al., 2004; Soundravally et al., 2014; Wang et al., 2007), and the remaining 5 studies were prospective cohort studies (de-Oliveira-Pinto et al., 2012; Kumar et al., 2012; Kurane et al., 1991, 1993; Suharti et al., 2003). Comparison groups were healthy controls, patients with other febrile illness (OFI) or non-severe disease patients (DF or non-SD); however, comparisons made between SD and non-SD cases were considered most ideal for the purpose of this review. Study characteristics of the
24 studies are reported in Table 1. One study did not specify which WHO case definition was applied, but this was verified with the authors and the manuscript was subsequently accepted for further evaluation (de-Oliveira-Pinto et al., 2012).

Ten studies noted affirmatively that their first blood sampling was performed at the febrile phase, which we defined here as <5 days or <96 h from fever onset (Brasier et al., 2012; Butthep et al., 2012; Green et al., 1999a, b; Kumar et al., 2012; De La Cruz Hernández et al., 2014; Malavige et al., 2013a; Pérez et al., 2004; Soundravally et al., 2014; Wang et al., 2007). The remaining studies had wider blood sampling windows, ranging from 1 to 10 days from symptom onset (Arias et al., 2014; Bozza et al., 2008; Chen et al., 2005; de-Oliveira-Pinto et al., 2012; Guerrero et al., 2013; Houghton-Triviño et al., 2010; Kurane et al., 1991, 1993; Laur et al., 1998; Levy et al., 2010; Del Moral-Hernández et al., 2014; Nguyen et al., 2004; Suharti et al., 2003). A majority of studies (21/24, 87.5%) employed ELISAs for quantification of single cytokine/chemokine, or multiplexed suspension bead immunoassays for broader profiling. Three studies (12.5%) used cytometric bead arrays (Guerrero et al., 2013; Houghton-Triviño et al., 2010; De La Cruz Hernández et al., 2014).

### Assessment of study design eligibility and risk of bias in individual studies

Table S2 summarizes the risk of study design eligibility and bias of selected studies according to the Cochrane NRS Methods group. First, all studies compared DHF, DSS or SD with DF, dengue without warning signs (DwoWS) or dengue with warning signs (DwWS) groups, as this was part of our quality assessment criterion (Table 2). Therefore, risk of bias in this item was negative. We considered the risk related to retrospective design as moderate, because the outcome assessments were retrospective and the generation of the hypothesis was prospective. Detection risk was considered low because patients were categorized according to canonical and accepted definitions of dengue severity clinical symptoms (WHO, 1997, 2009) and not by patients’ preference or marker (assay) outcomes. One study applied modified WHO guidelines (Bozza et al., 2008). However,
Table 1. Characteristics of included studies

<table>
<thead>
<tr>
<th>No.</th>
<th>Author and publication year</th>
<th>Study design</th>
<th>Sample size (N)</th>
<th>Age</th>
<th>Measured immuno-modulator</th>
<th>Sampling time and follow-up (where applicable)</th>
<th>Ascertainment of dengue diagnosis</th>
<th>Endpoint measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bozza et al. (2008)</td>
<td>CC</td>
<td>DF=20, DHF=39, total=59</td>
<td>DF, 28–48 years; DHF, 23–33 years</td>
<td>IL-1β, IL-2, IL-4, IL-6, IL-7, IL-13, GM-CSF, MCP-1, MIP-1β, TNFα</td>
<td>3 and 10 days after disease onset</td>
<td>WHO (1997)*</td>
<td>(1) Increased IL-1β, IFNγ, IL-4, IL-7, IL-13, GM-CSF in DHF compared to DF; (2) increased MIP-1β levels in DF compared to DHF; (3) MIP-1β and IFNγ were independent variables associated with disease outcome – MIP-1β increased during mild dengue with OR=0.181, while IFNγ was associated with severity with OR=1.138</td>
</tr>
<tr>
<td>2</td>
<td>Brasier et al. (2012)</td>
<td>CC</td>
<td>DF=38, DHF=13, total=51</td>
<td>DF, 15.8±7.8 years; DHF, 19 ±13.4 years</td>
<td>IL-2, IL-6, IL-10, IFNγ, IP-10, MIP-1α, TNFα, TRAIL, VEGF</td>
<td>Day 1 upon fever onset</td>
<td>WHO (1997)</td>
<td>(1) Increased log2-IL-10 and log2-IL-6 in DHF compared to DF; (2) increased IL-10 concentration associated with increased probability of DHF (using logistic regression model)</td>
</tr>
<tr>
<td>3</td>
<td>Butthep et al. (2012)</td>
<td>CC</td>
<td>OH=15, DF=51, DHF=98, total=164</td>
<td>Not stated</td>
<td>EGF, IL-1α, IL-1β, IL-2, IL-4, IL-6, IL-8, IFNγ, MCP-1, TNFα, VEGF</td>
<td>~5 days from defervescence</td>
<td>WHO (1997)</td>
<td>(1) Increased IL-4 in DHF II and DSS compared with DF, DHFII and OH only at day –1; (2) increased levels of IL-6 and IL-8 in DSS than in DF, DHFII, DHFII and OH on day +1 and day +2 compared with the other groups at day –2 to +2; (3) IFNγ and IL-10 – highest level was detected in DSS compared with DF, DHFII and DHF on day –1, whereas the lowest level was detected in OH; (4) increased TNFα in DF, DHFII and DSS than in OH from day –2 to 0 except for DSS on day 0; (5) increased IL-1β and IL-2 in DF than in DHF and OH from day –1 to +2, the lowest IL-2 was found in DSS except on day +2;</td>
</tr>
<tr>
<td>4</td>
<td>Chen et al. (2005)</td>
<td>CC</td>
<td>DF=66, DHF=33 (mild DHF=24/33, severe DHF=9/33); total=99</td>
<td>DF, 20–81 years; DHF, 30–76 years</td>
<td>IFNα, IFNγ, IL-10, IL-13</td>
<td>1–7 days upon symptom onset</td>
<td>WHO (1997)</td>
<td>(1) Increased IFNα in DF compared to DENV (P=0.01); (2) increased IL-10 in DSS compared to DF (P=0.03); (3) no difference in IFNα and IL-13 between DF and DHF; (4) significant correlation between disease severity and IL-10 (P&lt;0.001) but not IFNα, IFNγ or IL-13</td>
</tr>
<tr>
<td>5</td>
<td>De La Cruz Hernández et al. (2014)</td>
<td>CC</td>
<td>DF=116, DHF=88; total=204</td>
<td>Not stated</td>
<td>IFNα</td>
<td>First 5 days post-fever onset</td>
<td>WHO (1997)</td>
<td>Increased IFNα levels in DF than DHF for both DENV1 (P=0.0032) and DENV2 (P=0.0233)</td>
</tr>
<tr>
<td>6</td>
<td>Del Moral-Hernández et al. (2014)</td>
<td>CC</td>
<td>Controls=81, DF=70, DHF=80; total=231</td>
<td>0–95 years</td>
<td>sTM, VEGF</td>
<td>1–10 days post-fever onset</td>
<td>WHO (1997)</td>
<td>(1) Increased sTM level in DENV&gt; DF (P&lt;0.001), DF&gt; controls and DENV&gt; controls (P&lt;0.001); (2) increased VEGF in DF compared to DENV (P=0.005); DF compared to control and DENV compared to control (P=0.0001); (3) increased sTM levels in DENV2 compared to DENV1, no significant difference for VEGF with serotype</td>
</tr>
<tr>
<td>7</td>
<td>Furuta et al. (2012)</td>
<td>CC</td>
<td>DF=19, DHF=43, DSS=41; total=103</td>
<td>6 months–15 years</td>
<td>IL-9, IL-17m VEGF (VEGF-A), sVEGFR-1, sVEGFR-2</td>
<td>Blood samples collected at time of admission (day 0) and twice during the following 4 days (day 2 and 4)</td>
<td>WHO (1997)</td>
<td>(1) Increased VEGF in DF and DSS than DF and controls (P&lt;0.01) at day 0; (2) increased sVEGFR1 in DSS than DF and control (P&lt;0.01); (3) decreased sVEGFR2 in DF and DSS compared to DF and control (P&lt;0.01);</td>
</tr>
</tbody>
</table>
Table 1. cont.

<table>
<thead>
<tr>
<th>No.</th>
<th>Author and publication year</th>
<th>Study design</th>
<th>Sample size (N)</th>
<th>Age</th>
<th>Measured immunomodulator</th>
<th>Sampling time and follow-up (where applicable)</th>
<th>Ascertainment of dengue diagnosis</th>
<th>Endpoint measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Green et al. (1999a)</td>
<td>CC</td>
<td>OH=112, DF=32, DHF=28; total=172</td>
<td>6.8 years</td>
<td>IL-1β, TNF-α, IL-6, IL-4, IFN-γ, sIL2R, sCD8, sCD4, sTNFRI, sTNFRII</td>
<td>&lt;72h, equivalent to -2 days before defervescence</td>
<td>WHO (1997)</td>
<td>(4) increased IL-9 and IL-17 levels in DHF and DSS compared to DF and controls at day 0</td>
</tr>
<tr>
<td>9</td>
<td>Green et al. (1999b)</td>
<td>CC</td>
<td>OH=112, DF=32, DHF=28; total=172</td>
<td>6.8 years</td>
<td>IL-10, IL-12 p70, IL-12 (p40+p70)</td>
<td>&lt;72h, equivalent to -2 days before defervescence</td>
<td>WHO (1997)</td>
<td>(1) Increased mean plasma sTNFRI levels in DHF than DF from -2 days before defervescence (P&lt;0.01); (2) increased sIL2R and sCD8 levels in DHF compared with DF+1 after defervescence (P&lt;0.001 and &lt;0.05, respectively)</td>
</tr>
<tr>
<td>10</td>
<td>Houghton-Trivíno et al. (2010)</td>
<td>CC</td>
<td>DF=21, DHF I/II=4, DHF III/IV=13; total=38</td>
<td>0.3–55 years</td>
<td>IL-1β, IL-2, IL-4, IL-5, IL-8, IL-10, IL-12, IFNγ, sST2, TNFα</td>
<td>2–7 days of the disease</td>
<td>WHO (1997)</td>
<td>(1) Higher mean plasma IL-10 in DHF children compared to DF as early as 2 days before defervescence (P&lt;0.05); (2) higher mean plasma IL-12 (p40+p70) DHF compared to DF from fever day -2 (P&lt;0.05)</td>
</tr>
<tr>
<td>11</td>
<td>Kumar et al. (2012)</td>
<td>PC</td>
<td>DF=44, DHF=18; total=62</td>
<td>DF, 39 years; DHF, 40 years</td>
<td>IL-1βα, IL-1β, IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12, IFN-γ, sIL-2R, sCD8, sTNFRI, sVEGF</td>
<td>&lt;72h upon onset of fever</td>
<td>WHO (1997)</td>
<td>(1) Maximally increased in dengue patients during febrile phase – IFNγ, IP-10, IL-4, IL-10, IL-13, IL-1β, IL-6 and IL-8; (2) increased IP-10, IL-4, IL-9, IL-10, IL-1ra in dengue patients compared to controls during febrile phase; (3) decreased IFNγ, IL-4, IL-17 in DHF compared to DF during febrile stage</td>
</tr>
<tr>
<td>12</td>
<td>Kurane et al. (1991)</td>
<td>PC Afebrile controls=97, DF=10, DHF=59; total=166</td>
<td>6–13 years (97 healthy children)</td>
<td>IFNγ, sIL-2R, sCD4, sCD8.</td>
<td>Specimens collected within 24 h of admission to hospital and daily until discharge</td>
<td>Specimens collected within 24 h of admission to hospital and daily until discharge (1) Specimens collected within 24 h of admission to hospital and daily until discharge; (2) a convalescent specimen was collected from each child 7–10 days after hospital admission</td>
<td>WHO (1996); note – hospitalized cases of dengue that did not meet DHF criteria were classified as DF</td>
<td>(1) Increased sIL-2 R (P&lt;0.05), sCD4 (P&lt;0.001) and sCD8 (P&lt;0.001) in DHF than DF; (2) IFNγ and IL-2 levels were not different between DHF and DF</td>
</tr>
<tr>
<td>13</td>
<td>Kurane et al. (1993)</td>
<td>PC Controls=30, DF=10, DHF=55</td>
<td>Case, 5–14 years; control, 6–11 years; mean age (years) – DHF=9.1, DF=9.9, controls=7.9</td>
<td>IFNα</td>
<td>(1) Specimens collected within 24 h of admission to hospital and daily until discharge; (2) a convalescent specimen was collected from each child 7–10 days after hospital admission</td>
<td>(1) Specimens collected within 24 h of admission to hospital and daily until discharge; (2) a convalescent specimen was collected from each child 7–10 days after hospital admission</td>
<td>WHO (1986)</td>
<td>(1) Increased IFNα in DHF than in controls on days 2–4, days 6–7 and days 10–20; (2) increased IFNα in DHF than in controls on day 1 and day 3, but levels were not high on day 4–20 after onset of fever; (3) mean levels of IFNα in DHF patients were highest 2 days before defervescence, and decreased gradually until day of defervescence, IFNα levels did not change during day 0–19; (4) IFNα levels in DF patients were high 1 day before and on the day of defervescence, but these levels were not high after fever subsided; (5) IFNα level did not differ among different groups with different DHF grades</td>
</tr>
<tr>
<td>14</td>
<td>Laur et al. (1998)</td>
<td>CC</td>
<td>DF=106, DHF=17;</td>
<td>1 month to 15 years (mean 85.9)</td>
<td>TNFα, TGfβ-1</td>
<td>Days 1 to 8 post-fever onset</td>
<td>WHO (1997)</td>
<td>(1) Increased TGfβ-1 in DHF than in DF; (2) TNFα levels did not differ significantly between children with</td>
</tr>
</tbody>
</table>

Cytokines and chemokines as dengue prognosis biomarkers.
<table>
<thead>
<tr>
<th>No.</th>
<th>Author and publication year</th>
<th>Study design</th>
<th>Sample size (N)</th>
<th>Age</th>
<th>Measured immunomodulator</th>
<th>Sampling time and follow-up (where applicable)</th>
<th>Ascertainment of dengue diagnosis</th>
<th>Endpoint measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>Levy et al. (2010)</td>
<td>CC</td>
<td>DF=36, DHF=34; total=70</td>
<td>3–53 years (not differentiated among DF and DHF)</td>
<td>IL-1β, IL-6, TNFα</td>
<td>1 to 6 days post-fever onset</td>
<td>WHO (1997)</td>
<td>(1) Increased IL-6 in DHF than DF (P&lt;0.001); (2) increased TNFα than DF and DHF-I primary infection (P&lt;0.01)</td>
</tr>
<tr>
<td>16</td>
<td>Nyugen et al. (2004)</td>
<td>CC</td>
<td>DHF=85, DSS=22; total=107</td>
<td>1–11 months (mean age of DSS was significantly higher than DHF)</td>
<td>IL-2, IL-4, IL-6, IL-10, IFNγ, TNFα</td>
<td>3–7 days post-fever onset</td>
<td>WHO (1997)</td>
<td>(1) Increased IFNγ from infants with DHF/DSS than healthy controls; (2) increased IFNγ in DHF/DSS on days 4–6 after onset of fever and rapidly decreased on day 7 during convalescence; (3) increased TNFα in infants with DHF/DSS in acute phase compared to controls; (4) increased TNFα on days 4–7 after onset of fever and decreased on days 8–19; (5) increased IL-10 and IL-6 in DHF/DSS infants compared to controls during acute phase</td>
</tr>
<tr>
<td>17</td>
<td>Pérez et al. (2004)</td>
<td>CC</td>
<td>Controls=12, DF=28, DHF=6; total=46</td>
<td>16–59 years old (not differentiated among controls, DF and DHF)</td>
<td>IL-10, IL-12 (p70+p40), RANTES</td>
<td>Every 48 h – 1st (between first and third days post-fever onset), 2nd (fourth or fifth day) and 3rd (after sixth day of clinical evolution)</td>
<td>WHO (1997)</td>
<td>(1) Increased IL-10 in DHF than DF (P&lt;0.03) in 1st and 2nd determination; (2) lower RANTES in DHF than DF in 2nd and 3rd determination</td>
</tr>
<tr>
<td>18</td>
<td>Soundravally et al. (2014)</td>
<td>CC</td>
<td>DF=27, DHF=30, DSS=24; total=81</td>
<td>16–67 years; DHF=35 (20–59 years); DSS=24 (16–65 years)</td>
<td>TNFα, IFNγ</td>
<td>4 or 5 days post-fever onset</td>
<td>WHO (1997)</td>
<td>(1) Decreased TNFα in DF vs DSS (P&lt;0.001), increased IFNγ in DF vs DHF (P&lt;0.001), increased IFNγ in DF vs DSS (P=0.028); (2) decreased TNFα in DF vs DSS (P&lt;0.001); (3) decreased TNFα/IFNγ ratio in DF vs DSS (P=0.0099); (4) decreased TNFα/IFNγ ratio in DF vs DSS (P=0.0001)</td>
</tr>
<tr>
<td>19</td>
<td>Suharti et al. (2003)</td>
<td>PC</td>
<td>DHH III=43, DHH IV=7; total=50</td>
<td>6.6±2.8 (3–13) years</td>
<td>TNFα, IL-1β, IL-1Ra, IFNγ, IL-6</td>
<td>4.2±1 (2–7) days in the acute phase of the disease</td>
<td>WHO (1997)</td>
<td>(1) Increased IL-1Ra in non-survivor DSS (802.2±566.4 vs 1566.9±675.6, P&lt;0.0005); (2) IL-1Ra was significantly associated with mortality, P=0.007; (3) increased IL-6 in non-survivor DSS compared to survivors (219.30±38.262.0 vs 172.1±956.8, P&lt;0.0001)</td>
</tr>
<tr>
<td>20</td>
<td>Wang et al. (2007)</td>
<td>CC</td>
<td>DF=44, DHF 1/II=20, DHF III/IV=5; total=69</td>
<td>50.2±2.7 years; DHF, 48.54±3.6 years</td>
<td>TNFα, sTNFRI and sTNFRII</td>
<td>DF, 4.09±0.32 days; DHH, 4.91±0.44 days post-fever onset</td>
<td>WHO (1997)</td>
<td>(1) Increased plasma sTNFRI in DHF than DF (5826±1270 vs 2846±212 pg ml⁻¹; P&lt;0.04); (2) increased TNFα levels in DHF than DF (36.57±19.47 vs 349±30.6 pg ml⁻¹) (p&lt;NS)</td>
</tr>
<tr>
<td>21</td>
<td>Arias et al. (2014)</td>
<td>CC</td>
<td>DwoWS=12, DwWS=10, SD=8, non-dengue</td>
<td>Control, 18 (2–42 years); DwWS, 20 (5–33 years); DwoWS, 12 (1–7) months</td>
<td>IL-1β, IL-6, IL-12, IL-17, sTRAIL, sST2, sTNFRI, sTNFRII, TNFα</td>
<td>(1) Acute phase, 1–6 days after onset of symptoms; (2) convalescent, 7–26</td>
<td>WHO (2009)</td>
<td>(1) IL-6, IL-17, sTNFRI, sTNFRII, sST2 – control &gt; DwoWS &gt; DwWS &gt; SD; (2) increased IL-12 and sTRAIL in DwoWS compared to control, DwWS and SD (P&lt;0.05)</td>
</tr>
<tr>
<td>No.</td>
<td>Author and publication year</td>
<td>Study design</td>
<td>Sample size (N)</td>
<td>Age</td>
<td>Measured immuno-modulator</td>
<td>Sampling time and follow-up (where applicable)</td>
<td>Ascertainment of dengue diagnosis</td>
<td>Endpoint measured</td>
</tr>
<tr>
<td>-----</td>
<td>-----------------------------</td>
<td>--------------</td>
<td>----------------</td>
<td>-----</td>
<td>---------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>22</td>
<td>de Oliveira-Pinto et al. (2012)</td>
<td>PC</td>
<td>Controls=16, DF=33, WS/severe=40; total=89; convalescent samples (n=26) were not taken into account</td>
<td>DF mean age, 37.33 years; WS/severe mean age, 38.2 years; controls mean age, 31.9 years</td>
<td>CCL2/MCP-1, CCL4/MIP-1β, CCL5/RANTES</td>
<td>days after onset of symptoms; (3) recovery phase 2–9 days after first symptoms</td>
<td>Hybrid WHO (1997) and WHO (2009)</td>
<td>(1) Increased CCL2/MCP-1 in DF than WS/severe (P=0.002); (2) increased CCL4/MIP-1β in DF patients than WS/severe (P=0.04)</td>
</tr>
<tr>
<td>25</td>
<td>Guerrero et al. (2013)</td>
<td>CC</td>
<td>Controls=23, DwoWS=17, DwWS=21, SD=28; total=89</td>
<td>6–144 months (6 months – 12 years)</td>
<td>IL-1β, IL-6, IL-8, IL-10, IL-12p70, IL-33, sST2, TNFα</td>
<td>3–6 days post-fever onset</td>
<td>WHO (2009)</td>
<td>(1) Increased sST2 in SD than DwoWS and DwWS (P&lt;0.001); (2) increased IL-6 and IL-8 in children with DwWS or SD compared to children with DwoWS (P&lt;0.01); (3) increased IL-10 in DwoWS, DwWS, SD than controls (P&lt;0.01)</td>
</tr>
<tr>
<td>24</td>
<td>Malavige et al. (2013a)</td>
<td>CC</td>
<td>Controls=15, non-SD=219, SD=40; total=274</td>
<td>Mean age (26.8 years) – SD=4.8, non-SD=5.25 (not differentiated among controls, non-SD and SD)</td>
<td>IL-10, IL-21, MIF</td>
<td>(1) Initial blood sample was obtained during day 4 and 5 of illness (day 1 was considered as first day of fever); (2) second blood sample was obtained in 65/259 patients during critical phase (2 days after obtaining the first blood sample)</td>
<td>WHO (2009)</td>
<td>Increased serum IL-10 in SD than non-SD</td>
</tr>
</tbody>
</table>

CC, case–control study; OR, odds ratio; PC, prospective cohort; WS, warning sign.

we consider the confounding risk of prognostic factors moderate to high, because the type of immune-modulator under study was not uniform across the studies. A Kappa index of 0.91 was derived between two reviewers (W. Y. L. and Y. H. L.) suggesting substantial agreement.

Assessment of study design eligibility and risk of bias across studies

In this systematic review, which summarized mostly observational NRS, bias due to outcome differences may have arisen but we considered it as low to moderate as the usage and application of WHO case definition guidelines were enforced, and any potential differing opinions of disease severity minimized. However, comparability between studies utilizing the 1986/1997 or 2009 WHO revised case classification over time may be affected since the overlap between old and new case definitions are substantially different and as one or the other gains acceptance.

Prognostication marker candidates: WHO (1997) DHF I–IV

WHO (1986)/WHO (1997) was the most commonly applied case definition in the studies (20/24, 83 %) (Table 1). Studies differed in the types of immuno-modulating proteins investigated, and the most frequently profiled cytokines and chemokines were TNFα (12/20, 60 %), IFNγ (11/20, 55 %), IL-6 (10/20, 50 %), IL-10 (9/20, 45 %), IL-4 (6/20, 30 %) and IFNα (3/20, 15 %). Despite being the most investigated cytokine, TNFα was significantly elevated in DHF patients in only five studies (Houghton-Triviño et al., 2010; Levy et al., 2010; Nguyen et al., 2004; Soundravally et al., 2014; Wang et al., 2007), and one study found no statistical difference in TNFα levels between DF and DHF patients (Laur et al., 1998). The rest of studies reported no statistical difference or did not detect TNFα. By contrast, IL-10 stood out as the cytokine for which concentration was elevated in DHF relative to DF patients consistently in seven studies utilizing the WHO (1986)/WHO (1997) case definition (Table 3) (Brasier et al., 2012; Butthep et al., 2012; Chen et al., 2005; Green et al., 1999b; Houghton-Triviño et al., 2010; Nguyen et al., 2004; Pérez et al., 2004). The association of dengue severity with circulating levels of IFNγ was more heterogeneous (Table 4) (Bozza et al., 2008; Butthep et al., 2012; Chen et al., 2005; Kumar et al., 2012; Nguyen et al., 2004; Soundravally et al., 2014). Three studies noted increased IFNγ in DHF and/or DSS patients relative to DF patients or healthy controls (Bozza et al., 2008; Butthep et al., 2012; Nguyen et al., 2004), whereas three studies noted decreased IFNγ levels in DHF/DSS over DF patients (Chen et al., 2005; Kumar et al., 2012; Soundravally et al., 2014). Bozza et al. (2008) noted deficiency phase IFNγ levels in adults (age 15–73 years) were associated with disease severity (odds ratio=1.138, 95 % confidence interval 0.41 to 1.245, P=0.0046). However, Green et al. (1999a) reported no significant difference in febrile IFNγ between DF and DHF paediatric patients. In congruence, in infants (age <1 year), IFNγ (together with IL-2, IL-4, IL-6, IL-10, TNFα) could not differentiate DHF I/II from DHF III/IV (Nguyen et al., 2004).

Prognostication marker candidates: WHO (1997) DHF III/IV

One study looked specifically at immuno-modulatory markers of predicting DSS (DHF III/IV) mortality (Suharti et al., 2003). The study population was restricted to only DENV3-infected paediatric patients who were diagnosed as DSS in a prospective study. In the acute phase of dengue infection, plasma IL-1Rα levels were significantly higher in DSS non-survivors (n=7) compared to DSS survivors (n=43; 802.2±566.4 vs 1566.9±675.6 pg ml⁻¹, P=0.0005). Multiple logistic regression identified IL-1Rα as having significant association with mortality on the day of admission (P=0.007).

Prognostication marker candidates: WHO (2009) SD

Of the selected 24 studies, 4 studies (16.7 %) employed the WHO (2009) case definition (Arias et al., 2014; de-Oliveira-Pinto et al., 2012; Guerrero et al., 2013; Malavige et al., 2013a). Among them, one reported elevation in IL-10 (Malavige et al., 2013a), congruent with the seven WHO (1997) guideline-utilizing studies that reported elevations in IL-10 in DHF patients (Table 3), and one study reported elevated IL-10 levels in DwoWS, DwWS and SD patients compared to controls (Guerrero et al., 2013). Two studies reported elevations in sST2 (IL-1R1L1) (Arias et al., 2014; Guerrero et al., 2013) in SD patients in levels above those of DwoWS and DwWS (Arias et al., 2014; Guerrero et al., 2013). The same studies reported increased IL-6 levels in SD and DwWS relative to patients with DwoWS. Elevations in TNFRI and TNRFII in DwoWS, DwWS and SD (Arias et al., 2014) were similar to those observed in DHF patients (Wang et al., 2007).

Discussion

Our systematic review of studies published from 1998 to 2014 evaluated cytokines and chemokines as potential early markers to prognosticate the development of SD. We evaluated and documented judgements about evidence quality from NRS and incorporated this into reporting of differentially changed cytokines and/or chemokines for outcomes based on WHO (1986)/WHO (1997) or WHO (2009) dengue case definitions. Twenty-four studies were identified to be of suitably high quality, but with major gaps in descriptions of patient demographics, DENV types and infection status (primary, secondary or tertiary infections). Although a consensus has yet to emerge, IL-10 levels stood out as a potential marker of SD (DHF and SD) that was significantly increased regardless of age (range 1 month to 76 years) or infection status (primary or secondary).

Eight studies utilizing WHO (1986)/WHO (1997) or WHO (2009) case definitions demonstrated elevated IL-10 levels in SD patients. Interestingly, several studies, including a
Cytokines and chemokines as dengue prognosis biomarkers

recent publication, showed similar IL-10 levels between patients with and without SD (Bozza et al., 2008; Cui et al., 2016; Guerrero et al., 2013; Kumar et al., 2012). The reason for this discrepancy is not entirely clear. IL-10 is an important anti-inflammatory cytokine and general suppressor of immune reactions, inhibiting IL-1, IL-6, IL-10 itself, IL-12, IL-18, CSF and TNFα, as well as inhibiting the synthesis of IL-2, IL-3, GM-CSF, TNFα and IFN-γ (D’Andrea et al., 1993). IL-10 may contribute to disease severity through NS1-induced IL-10 production by monocytes, which in turn suppresses dengue-specific T cell responses (Adikari et al., 2016; Malavige et al., 2013b). Notably, data from animal studies suggest that NS1 activates immunological cascades in monocytes and macrophages leading to the pathologies observed in SD (Chen et al., 2013). IL-10 has been secreted by a variety of cells including CD4+ and CD8+ T cells, B cells, macrophages, monocytes, eosinophils and mast cells (O’Garra & Vieira, 2007). Measurements of circulating levels of IL-10 reflect the sum total of IL-10 produced from all cells in the body at any one point. Production of IL-10 may change according to the predominant cell type specific to the day of infection to regulate viral clearance immunopathology (reviewed by Tsai et al., 2013b) as different immune cells are invoked at different stages of dengue infections (Boonnak et al., 2008; Clyde et al., 2006). The timing of IL-10 production is dynamic and varies throughout the disease course, exemplified by IL-10 levels peaking around early defervescence in DHF patients, but less so in DF patients (Adikari et al., 2016; Butthep et al., 2012; Kumar et al., 2012; Libraty et al., 2002a), suggesting that the biggest difference between DF and DHF patients may be observed around the day of early defervescence. Mechanistically, the intrinsic ADE of DENV infections may modulate the severity of dengue via increased IL-10 production and subsequent enhancement of the Suppressor of Cytokine Signaling (SOCS) system (Chareonsrisuthigul et al., 2007; Suhrbier & La Linn, 2003; Ubol et al., 2010). Furthermore, in vitro studies suggest that in ADE-DENV infections, the role of IL-10 changes in the early and later stages of infection from anti-viral to immunoregulation (Halstead et al., 2010; Tsai et al., 2013b). Two studies noted significant elevations in IL-10 and IFNγ in DHF and DSS patients compared to DF patients (P<0.01) (Butthep et al., 2012; Nguyen et al., 2004), whereas one study showed elevated IL-10 levels (DF 11.7±3.5 vs DHF 117.0±52.8 pg ml⁻¹; P=0.03) but decreased IFNγ in DHF compared to DF (DF 130.2±16.9 vs DHF 84.2±12.6 pg ml⁻¹; P=0.01) (Chen et al., 2005). By contrast, three other studies reported increased IL-10 levels in DHF patients, but no difference in IFNγ between DF and DHF patients (Brasier et al., 2012; Green et al., 1999a; Houghton-Triviño et al., 2010). Note that these studies had different sampling timings. In addition, we found no consistent trend of any potential interactions between IL-10 with the following cytokines: IL-2, IL-6, IL-12 and TNFα. This is congruent with other studies as cytokine changes are often mixed in dengue infections in different populations (Chaturvedi et al., 2000), with the exception of IL-10 as it is shown in this systematic review. Collectively, the exact sampling timing of IL-10 may be a critical aspect in determining its maximal potential as a marker for SD prognosis.

The discrepancy of IFNγ in the five studies could be due to the rapid kinetics in its circulating levels, and depending on which infection day blood was sampled and analysed the IFNγ levels could be different (Green et al., 1999a; Libraty et al., 2002b; Nguyen et al., 2004). Reports from humans and from animal models suggest that IFNγ controls infections through viral clearance and limiting virus replication (Costa et al., 2012; Horras et al., 2011; Pal et al., 2014; Shresta et al., 2004). Notably, dengue viral titres correlate with disease severity (Endy et al., 2004; Vaughn et al., 2000), but are plausibly influenced by DENV type and infection status (Duyan et al., 2011). Levels of IFNγ can be attenuated by IL-10 through SOCS-3 blockage of STAT1-IFNγ receptor interaction in intrinsic ADE-DENV infection (Chareonsrisuthigul et al., 2007; Ubol et al., 2010). Indeed, the absence of IFNγ in a mouse model of dengue led to primary dengue-infection-induced lethality (Costa et al., 2012). Thus, depending on the infection status and DENV type, in combination with timing of IL-10 production, IFNγ levels may be affected, leading to altered viral clearance, prolonged infection and consequently determining disease severity. IFNγ levels peak before defervescence in DHF patients and peak after defervescence in DF patients, because of this the maximum difference in the febrile phase of dengue suggests an advantage of using IFNγ as an early marker of SD. However, the implication and utility of IFNγ needs further study with properly defined sampling windows and DENV types taken into consideration.

To our knowledge, this is the first systematic review of immuno-modulating proteins as prognostic markers of SD sequelae. Two earlier published reviews did not apply and justify quality assessments (Chaturvedi et al., 2000; Yacoub & Wills, 2014). This is worth highlighting as the restriction of studies according to a set of inclusion/exclusion criteria, and the introduction of quality assessments, retained studies with better-defined methodologies. Longitudinal studies, where daily measurements of studied immuno-modulating proteins are taken, may best capture the dynamic kinetics of these proteins in dengue. Although a few immuno-modulating proteins are potentially promising as prognostic markers, prospective observation cohort trials, such as the one by International Research Consortium on Dengue Risk Assessment, Management, and Surveillance (IDAMS; https://ClinicalTrials.gov; identifier: NCT01550016), may aid further endorsement of the utility of an acute phase prognostic marker for SD. In addition, immuno-modulating proteins, in reflecting the cardinal symptoms of SD, plasma leakage and thrombocytopenia, should be given research priority (Yacoub & Wills, 2014; Zapata et al., 2014). One promising example is TNFα, which was reported to induce endothelial cell apoptosis (Chen et al., 2007) and was also noted in this review as an elevated
cytokine in DHF. Other potential circulating markers of dengue severity, such as proteases (Koraka et al., 2010; St John et al., 2013; van de Weg et al., 2014), soluble adhesion molecules (Cardier et al., 2006) or metabolites (Cui et al., 2013, 2016), may show potential as alternative soluble prognostic markers of SD, but face similar challenges in the extensive demands of large study cohorts. Capturing soluble factors that correlate and potentially foretell pathognomonic symptoms may not only increase the specificity and sensitivity of prognosing SD, but also avoid the non-specific ‘cytokine storms’ observed in other acute infectious diseases, such as influenza and malaria (Clark, 2007).

There were several limitations in our review process. First, there were inadequate reports on the details of patient demographics, DENV type and patients’ infection status (primary, secondary or tertiary infection) in the studies. Inclusion of such information may be necessary to understand why results may be heterogeneous. Critically, sampling windows (days of illness from onset of fever) were at times not documented, or varied widely, sometimes even within the same study. Second, endpoint measurements of these studies varied substantially, differing in the types of immuno-modulating proteins under study. Not all studies investigated the same type of immuno-modulating proteins in relation to dengue severity. Furthermore, data presented in tables or figures in the published manuscripts rarely report concentration means and/or the range of the significantly different immuno-modulating proteins. This severely limited any statistical estimation of odds ratios, sensitivity and specificity of the prognosis marker candidates and prevents meta-analysis. Such wide variability in studies has been reported elsewhere (Potts & Rothman, 2008). Third, after strictly adhering to internal assessments of the quality of the primary studies, further moderation led to six studies being downgraded and one study being upgraded (Table S1). A study may have scored as high quality but may have been weighed down due to an unlisted criterion, for example, large variation in the studied cytokines (Chen et al., 2006). Although this could potentially introduce bias, we included this empirical exercise to ensure retention or exclusion of studies that were under- or over-ranked but still added value to the review.

Table 2. Quality assessment criteria matrix

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study design (control, DF, DHF)</td>
<td>Only control vs DF</td>
</tr>
<tr>
<td>Timing of sampling (days)</td>
<td>No or ambiguous description</td>
</tr>
<tr>
<td>Age</td>
<td>No or ambiguous description</td>
</tr>
<tr>
<td>Data collection</td>
<td>No or ambiguous description</td>
</tr>
<tr>
<td>Diagnostic criteria [WHO (1997) or WHO, (2009)]</td>
<td>No or ambiguous description</td>
</tr>
<tr>
<td>Inclusion/exclusion criteria</td>
<td>No or ambiguous description</td>
</tr>
<tr>
<td>Infection status</td>
<td>No or ambiguous description</td>
</tr>
<tr>
<td>DENV type</td>
<td>No or ambiguous description</td>
</tr>
<tr>
<td>Statistical analysis</td>
<td>No or ambiguous description</td>
</tr>
<tr>
<td>Data statistics (P value provision)</td>
<td>No or ambiguous description</td>
</tr>
</tbody>
</table>

Conclusions

This systematic review provides the basis for future high-quality studies urgently required to identify key prognostic immuno-modulating protein markers of SD. They should be performed according to REMARK criteria (McShane et al., 2005) to avoid the methodological deficiencies discovered in this systematic review, mostly with substantial heterogeneity in the study populations and prognostic endpoints. Studies in larger population sizes should reflect in detail the study population’s age, infection status and DENV types, with standardization on first sample collection timing (<96 h from fever onset), detailed patient infection information and prognostic endpoint measurements.

The use of soluble immuno-modulating proteins as prognostic markers of SD remains under investigation and yet a consensus marker or marker signature for prognosing SD is desirable. Is this achievable? Current evidence suggests that different patient subpopulations – age groups (adults vs children vs geriatric), DENV types 1 to 4, and infection status – respond differently to dengue infections (Hammond et al., 2005; OhAinle et al., 2011; Tricou et al., 2011). Accordingly, different markers may be needed to optimally prognose SD in different patient subpopulations (Lee & Ooi, 2013), and potentially aid the design of dengue vaccines.

Methods

Search strategy and data extraction

The primary literature search was conducted according to PRISMA (Preferred Reporting Items for Systematic Reviews
<table>
<thead>
<tr>
<th>Publication</th>
<th>Sample size</th>
<th>Age (years)</th>
<th>Infection status</th>
<th>Information on DENV type available?</th>
<th>First sampling time</th>
<th>DHF IL-10 change relative to DF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DF</td>
<td>DHF</td>
<td>DF</td>
<td>DHF</td>
<td>DF/ DHF</td>
<td>Primary (%)</td>
</tr>
<tr>
<td>Braiser et al. (2012)</td>
<td>38</td>
<td>13</td>
<td>15.76±7.8</td>
<td>19±13.4</td>
<td>DF</td>
<td>Not stated</td>
</tr>
<tr>
<td>Butthep et al. (2012)</td>
<td>51</td>
<td>98</td>
<td>Not stated</td>
<td>Not stated</td>
<td>DF</td>
<td>Not stated</td>
</tr>
<tr>
<td>Chen et al. (2005)</td>
<td>66</td>
<td>33</td>
<td>46.8 (20–81)</td>
<td>57.8 (30–76)</td>
<td>DF</td>
<td>75</td>
</tr>
<tr>
<td>Green et al. (1999b)</td>
<td>22</td>
<td>20</td>
<td>8.1 (3.2)*</td>
<td>8.1 (2.6)*</td>
<td>DF</td>
<td>43</td>
</tr>
<tr>
<td>Houghton-Triviño et al. (2010)</td>
<td>21</td>
<td>17</td>
<td>26 (3–55)</td>
<td>4 (0.3–12)</td>
<td>DF</td>
<td>23.8</td>
</tr>
<tr>
<td>Malavige et al. (2013a)†</td>
<td>40</td>
<td>219</td>
<td>Non-SD: 27.4 (11.4)</td>
<td>SD: 23.3 (8.8)</td>
<td>DF</td>
<td>11.8</td>
</tr>
<tr>
<td>Malavige et al. (2013a)‡</td>
<td>22</td>
<td></td>
<td></td>
<td>6.4 (1–11)</td>
<td>8.2 (4–11)</td>
<td>DHF I/ II</td>
</tr>
<tr>
<td>Malavige et al. (2013a)§</td>
<td>39</td>
<td>100</td>
<td>DHF II/ III/ IV</td>
<td>19.6</td>
<td>0.9</td>
<td>0</td>
</tr>
</tbody>
</table>

*The information provided here is a subset of patients described in Kalayanarooj et al. (1997).
†Comparison made between non-SD and SD, following the WHO (2009) case definition.
‡Comparison was DHF III/IV relative to DHF I/II.
§DHF I/II.
||DHF III/IV (DSS).
Table 4. Studies with significantly different IFNγ levels in dengue patients

<table>
<thead>
<tr>
<th>Publication</th>
<th>Sample size</th>
<th>Age (years)</th>
<th>Infection status</th>
<th>Information on DENV type available?</th>
<th>First sampling time</th>
<th>DHF IFNγ change relative to DF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DF DHF</td>
<td>DF DHF</td>
<td>DF/DHF</td>
<td>Primary (%) Secondary (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bozza et al. (2008)</td>
<td>20 39</td>
<td>28–48 23–53</td>
<td>DF DHF</td>
<td>74 58</td>
<td>3–10 days after disease onset</td>
<td>Increase</td>
</tr>
<tr>
<td>Butthep et al. (2012)</td>
<td>51 98</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Not stated Not stated</td>
<td>–3 days from defervescence</td>
<td>Increase</td>
</tr>
<tr>
<td>Chen et al. (2005)</td>
<td>66 33</td>
<td>46.8 (20–81) 57.8 (30–76)</td>
<td>DF DHF</td>
<td>75 43</td>
<td>1–7 days post-symptom onset</td>
<td>Decrease</td>
</tr>
<tr>
<td>Nguyen et al. (2004)*</td>
<td>85† 22‡</td>
<td>6.4 (1–11) 8.2 (4–11)</td>
<td>DHF I/II DHF III/IV</td>
<td>75.7 19.6</td>
<td>3–7 days post-fever onset</td>
<td>Increased in DHF I–IV compared to healthy controls</td>
</tr>
<tr>
<td>Kumar et al. (2012)</td>
<td>44 18</td>
<td>39 (13.02) 40 (14.08)</td>
<td>DF DHF</td>
<td>56.8 61.1</td>
<td>&lt;72 h post-fever onset</td>
<td>Decrease</td>
</tr>
<tr>
<td>Soundravally et al. (2014)</td>
<td>27 30</td>
<td>22 (16–67) 35 (20–59)</td>
<td>DF DHF</td>
<td>Not stated Not stated</td>
<td>4 or 5 days post-fever onset</td>
<td>Decrease</td>
</tr>
</tbody>
</table>

*Comparison was between healthy controls, DHF I/II and DHF III/IV.
†DHF I/II.
‡DHF III/IV (DSS).
and Meta-Analyses) guidelines (Moher et al., 2009), via PubMed and Embase databases, for original research articles with restriction to human studies and English language. The search terms used for both databases were as followed: 'dengue AND (Cytokine OR Chemokine) AND (predict* OR prognos* OR correlat* OR associat* OR indicat*)'. We designed search strategies as shown in Table 5.

Two independent reviewers (W. Y. L. and Y. H. L.) screened the results to identify relevant literature based on titles and abstracts (when available), followed by another evaluation according to the inclusion and exclusion criteria. Articles with reported epidemiology, clinical signs, laboratory parameters and prognosis markers of SD outcome were included.

Our exclusion criteria included animal models, cell lines (in vitro studies), ex vivo cell studies, vaccine or anti-viral trials, genetic markers studies, studies without controls or inappropriate controls and also review articles. Of note, this review is devoted to the utility of human circulating cytokines and chemokines as prognostication markers; hence, other proteins such as immunoglobulins, acute phase proteins and proteases were excluded. When both reviewers (W. Y. L. and Y. H. L.) agreed on the final selected title and abstracts, the full texts of the articles were obtained and independently reviewed for eligibility.

Data extraction was performed by a single reviewer (W. Y. L.) using a data extraction sheet and was checked by another reviewer (Y. H. L.). The data extracted included the name of first author, year of publication, age, gender, disease outcome – DF/DHF/DSS or DwoWS, DwWS and SD – sample size, study design, location of study, specific markers investigated and results of the analysis.

### Risk of bias and quality assessment

Two reviewers (W. Y. L. and Y. H. L.) independently evaluated the risk of bias in selected studies by assessing studies according to a checklist developed by the Cochrane NRS Methods group (Cochrane Group, 2012) as shown in Table S3. Assessment of outcome was not masked in all studies due to the need to diagnose dengue severity.

Subsequently, the quality of selected studies was assessed across a matrix of ten metrics, namely (i) study design, (ii) time of sampling, (iii) age, (iv) data collection, (v) diagnostic criteria, (vi) inclusion/exclusion criteria, (vii) infection status, (viii) DENV type, (ix) statistical analysis, (x) data statistics. The quality assessment score system was based on a modified version of the Newcastle–Ottawa scale (NOS) (Wells et al., 2013) and is described in Table 2. The maximum score is 12, and a score of 9–12 is considered as high quality, 5–8 as moderate quality and 1–4 as low quality. The measure of agreement between the two reviewers (W. Y. L. and Y. H. L.) was assessed using the Kappa index (MedCalc version 12.5). None of the studies were randomized clinical trials and...
because the studies were observational studies (case–control, cohort), we assessed the quality of evidence using a list of ten criteria modified from the NOS (Table 2), rather than Grading of Recommendations, Assessment, Development and Evaluations (GRADE) (Guyatt et al., 2008).

Ascertainment of outcomes relied on WHO (1986), WHO (1997) or WHO (2009) dengue case definition guidelines. The WHO (1986) and WHO (1997) case definition guidelines were grouped together during analysis, but were analysed separately from the WHO (2009) case definition guidelines due to reported differences (Leo et al., 2013; Narvaez et al., 2011; Tsai et al., 2013a). There was the possibility of bias or discrepancy in the classification of dengue severity; therefore, two reviewers (W. Y. L. and Y. H. L.) assessed the studies either for explicit description of the WHO guidelines used or for labelling of the groups’ categorization (DF, DHF and DSS correspond to the WHO (1986) and WHO (1997) case definitions, whereas non-SD, SD, DwoWS and DwWS correspond to the WHO (2009) case definition).

Odds ratios or statistical mean/median values could not be identified in most studies. No pooling was done given the extra sources of methodological diversity (types of prognostic factors being measured) and bias. Hence, NRS are expected to be more heterogeneous than randomized trials.

Acknowledgements

This work was supported by a research grant from the KK Women’s and Children’s Hospital given to Y. H. L.

References


