Predictive value of the 4Ts scoring system for heparin-induced thrombocytopenia: a systematic review and meta-analysis

Adam Cuker,1,2 Phyllis A. Gimotty,3 Mark A. Crowther,4,5 and Theodore E. Warkentin4,5

Departments of 1Medicine, 2Pathology and Laboratory Medicine, and 3Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; and Departments of 4Medicine and 5Pathology and Molecular Medicine, McMaster University, Hamilton, ON

The 4Ts is a pretest clinical scoring system for heparin-induced thrombocytopenia (HIT). Although widely used in clinical practice, its predictive value for HIT in diverse settings and patient populations is unknown. We performed a systematic review and meta-analysis to estimate the predictive value of the 4Ts in patients with suspected HIT. We searched PubMed, Cochrane Database, and ISI Web of Science for studies that included patients with suspected HIT, who were evaluated by both the 4Ts and a reference standard against which the 4Ts could be compared. Quality of eligible studies was assessed by QUADAS-2 criteria. Thirteen studies, collectively involving 3068 patients, fulfilled eligibility criteria. A total of 1712 (55.8%) patients were classified by 4Ts score as having a low probability of HIT. The negative predictive value of a low probability 4Ts score was 0.998 (95% CI, 0.970-1.000) and remained high irrespective of the party responsible for scoring, the prevalence of HIT, or the composition of the study population. The positive predictive value of an intermediate and high probability 4Ts score was 0.14 (0.09-0.22) and 0.64 (0.40-0.82), respectively. A low probability 4Ts score appears to be a robust means of excluding HIT. Patients with intermediate and high probability scores require further evaluation. (Blood. 2012;120(20):4160-4167)

Introduction

Heparin-induced thrombocytopenia (HIT) is a prothrombotic and potentially fatal adverse drug reaction mediated by platelet-activating antibodies against multimolecular complexes of platelet factor 4 (PF4) and heparin.1,2 Management involves cessation of heparin, avoidance or postponement of an oral vitamin K antagonist until platelet count recovery, and initiation of a nonheparin anticoagulant.3,5 Accurate diagnosis and prompt commencement of therapy are paramount. Delays in treatment are associated with an initial 5%-10% daily risk of thrombosis, amputation, or death.6 Misdiagnosis of HIT, conversely, may result in exposure of thrombocypenic patients to alternative anticoagulants and their attendant ~ 1% daily risk of major hemorrhage7,8 or in thrombosis from unnecessary suspension of heparin.9

The diagnosis of HIT, which rests on both clinical assessment and laboratory testing, remains challenging despite these high stakes. Clinical evaluation is complex and imprecise, even among experienced diagnosticians.10 Laboratory tests for HIT are of 2 varieties. Widely available immunoassays are sensitive and simple to perform but yield frequent false-positive results.11 Washed platelet functional assays such as the 14C-serotonin release assay (SRA) have much greater specificity but require radioisotope and reactive donor platelets, reagents that are unfeasible for most clinical laboratories.12 Consequently, such assays are available to the majority of clinicians only as “send-out” tests and do not yield results in a timeframe necessary to inform initial clinical decision-making.

The 4Ts (see Table 1) is a pretest scoring system for HIT that was developed to improve and standardize clinical diagnosis. It incorporates 4 typical features of HIT: (1) magnitude of thrombocytopenia; (2) timing of thrombocytopenia with respect to heparin exposure; (3) thrombosis or other sequelae of HIT; and (4) likelihood of other causes of thrombocytopenia. The system yields an integer score between 0 and 8 with scores of 0-3, 4-5, and 6-8 classified as low, intermediate, and high pretest probability for HIT, respectively.13,14 The 4Ts is widely used in clinical practice, and several single-center experiences with the model have been reported. However, the generalizability of these studies to other settings and patient populations is uncertain. The objective of this systematic review and meta-analysis was to estimate the predictive value of the 4Ts in a heterogeneous group of patients with suspected HIT.

Methods

Data sources and searches

We performed a systematic review of the literature to examine best evidence for the predictive value of the 4Ts. A search of PubMed, Cochrane Database, and ISI Web of Science from March 2003, the month of initial publication of the 4Ts model,13 through December 2011 was performed using the keywords: (heparin-induced thrombocytopenia) AND (4Ts OR 4 T's OR 4T's OR 4 Ts OR 4T's OR 4 T's OR clinical score). There were no language restrictions.

Study selection

Articles were examined by 1 reviewer (A.C.), first by title and abstract, then by review of the complete paper as indicated. Additional articles were sought through review of bibliographies. Studies of patients with suspected...
Table 1. The 4Ts scoring system

<table>
<thead>
<tr>
<th>4Ts category</th>
<th>2 points</th>
<th>1 point</th>
<th>0 points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thrombocytopenia</td>
<td>Platelet count fall &gt; 50% and platelet nadir ≥ 20</td>
<td>Platelet count 30-50% or platelet nadir 10-19</td>
<td>Platelet count fall &lt; 30% or platelet nadir &lt; 10</td>
</tr>
<tr>
<td>Timing of platelet count fall</td>
<td>Clear onset days 5-10 or platelet fall ≤ 1 day (prior heparin exposure within 30 days)</td>
<td>Consistent with days 5-10 fall, but not clear (eg, missing platelet counts); onset after day 10; or fall ≤ 1 day (prior heparin exposure 30-100 days ago)</td>
<td>Platelet count ≤ 4 days without recent exposure</td>
</tr>
<tr>
<td>Thrombosis or other sequelae</td>
<td>New thrombosis (confirmed); skin necrosis; acute systemic reaction postintravenous unfractionated heparin bolus</td>
<td>Progressive or recurrent thrombosis; non-necrotizing (erythematous) skin lesions; suspected thrombosis (not proven)</td>
<td>None</td>
</tr>
<tr>
<td>Other causes of thrombocytopenia</td>
<td>None apparent</td>
<td>Possible</td>
<td>Definite</td>
</tr>
</tbody>
</table>

The 4Ts score is the sum of the values for each of the 4 categories. Scores of 1-3, 4-5, and 6-8 are considered to correspond to a low, intermediate, and high probability of HIT, respectively.

Results

Study selection

The initial literature search yielded 67 articles. One additional reference was identified from the bibliographies. Fifty-six articles were excluded: 4 were duplicates, 11 enrolled patients without suspected HIT, 10 did not report 4Ts scores, 10 did not define a reference standard or did not apply the standard to all subjects, 13 were reviews, and 8 were case reports. Twelve articles met selection criteria, one of which14 reported the results of 2 distinct studies, for a total of 13 eligible studies (Figure 1). Of these, 8 were prospective14,17,22 and 5 were retrospective10,23-26 cohort studies. No randomized controlled trials were identified.

Study characteristics

Study characteristics are summarized in Table 2. Studies were conducted in 8 countries and collectively enrolled 3068 patients. Sample size ranged from 43 to 1291 and prevalence of HIT from 0.04 to 0.42. Study populations were similar with respect to age and gender. Four studies restricted eligibility to cardiovascular (CV) surgical or critically ill subjects.17,24-26 The remainder enrolled unscreened medical and surgical patients with suspected HIT, although distribution by patient population varied. For example, European studies14,18,22 included proportionately more medical and fewer CV surgical patients than North American series.10,14

The reference standard for defining HIT also differed among studies. One study used a clinical standard,10 10 studies used a laboratory standard,14,18,21,23-26 and 2 studies used a standard incorporating both clinical and laboratory criteria.17,22 The reference standard included a functional assay in all but one of the studies,10 but choice of assay, methodology, cut-off, and interpretation of assay results varied (Table 2).

Several iterations of the 4Ts have been published. Ten of the 13 studies applied the standard 4Ts model shown in Table 1.14 Patients in one study17 were scored using an earlier rendering of the model.13 Scores in 2 other studies were calculated using the Greifswald modification of the 4Ts.14 The 3 versions of the 4Ts are essentially identical with respect to the “thrombosis or other sequelae” and “other causes of thrombocytopenia” categories but have different criteria related to platelet nadir and timing of platelet count fall.

Study quality

Assessment of study quality by QUADAS-2 criteria15 raised several potential methodologic limitations (Table 3). For instance,
4Ts scoring in 10 studies was performed by study personnel, often through retrospective chart review. Because this approach differs from anticipated use of the 4Ts in patient care (scoring by referring physicians through bedside evaluation), concern about the applicability of such studies to clinical practice was judged to be high. Two studies did not state whether 4Ts scorers were blinded to the reference standard. Because awareness of the result of the reference standard could bias scoring, these studies were deemed to carry an unclear risk of bias in this domain.

**Predictive value of the 4Ts**

Table 4 shows the number of patients with and without HIT (relative to the reference standard) in each 4Ts probability category. Of 3068 total patients, 1712 (55.8%) were classified by 4Ts score as having a low probability, 1103 (36.0%) as having an intermediate probability, and 253 (8.2%) as having a high probability of HIT. The overall random effects estimate of the prevalence of HIT among included studies was 0.11 (95% CI, 0.07-0.17). The PPV of high, intermediate, and combined high and intermediate probability 4Ts scores was 0.64 (0.40-0.82), 0.14 (0.09-0.22), and 0.22 (0.15-0.31), respectively (Figure 2A-C). The NPV of a low probability score was 0.998 (0.86-1.00; Figure 2D). The pooled estimates of sensitivity and specificity of the 4Ts at a cut-off of 0.11 were 0.99 (0.97-1.00) and 0.54 (0.43-0.66), respectively.

**Analysis of heterogeneity**

Visual inspection of Forest plots (Figure 2) revealed largely overlapping 95% CIs, suggesting reasonable homogeneity among studies. To further evaluate heterogeneity, we performed several subanalyses. First, using a cut-off of 0.10, the median prevalence among eligible studies, we compared studies with a high (> 0.10, n = 5) and low (≤ 0.10, n = 8) prevalence of HIT. The pooled random effects estimates of prevalence were 0.23 (0.13-0.37) and 0.07 (0.06-0.09) in the high and low prevalence studies, respectively. As expected, PPV increased with increasing prevalence. PPV was significantly higher in the high prevalence studies than in the low prevalence studies for intermediate probability [0.32 (0.20-0.46) vs 0.09 (0.06-0.13), P < .001] and combined high and intermediate probability 4Ts scores [0.40 (0.27-0.56) vs 0.15 (0.12-0.20), P < .001], but not for high probability scores [0.83 (0.05-1.00) vs 0.49 (0.31-0.64), P = .074] because of the large variability in PPV among the high prevalence studies. In contrast, the NPV of a low probability 4Ts score was similar and high among both the low prevalence [0.998 (0.883-1.00)] and high prevalence [0.983 (0.771-0.999)] studies (P = .054). These relationships held when different prevalence cut-points were used in the analysis (data not shown).

Studies were also compared on the basis of whether scoring was performed by study personnel (n = 10) or referring clinician (n = 3). In one study, ~80% of subjects were scored by the referring clinician; this study was included in the referring clinician subgroup. The random effects pooled estimate of PPV was greater in the study personnel subgroup than in the referring clinician subgroup for high probability [0.75 (0.43-0.92) vs 0.37 (0.19-0.60)], intermediate probability [0.16 (0.09-0.27) vs 0.11 (0.07-0.18)], and combined intermediate and high probability [0.25 (0.16-0.38) vs 0.16 (0.10-0.24)] 4Ts scores. The NPV of a low probability 4Ts score was similar in the 2 groups.
None of these differences was statistically significant ($P = 0.09$, $0.32$, $0.20$, and $0.87$, respectively).

We planned a third analysis to evaluate the predictive value of the 4Ts in different patient populations. Because studies did not routinely report patient level data on patient population, we used study level data (Table 2) for the analysis. We compared studies with a relatively large proportion of CV surgery patients ($\geq 0.20$, $n = 8$) with those in which CV surgery patients composed a smaller fraction of the overall study population ($\leq 0.20$, $n = 4$). We chose 0.20 as a cut-point because it served as a natural divide among eligible studies. In 4 studies, the proportion of CV surgery patients was $\approx 0.18$; in the remaining studies, the proportion was $\approx 0.36$ (Table 2). One study did not collect information on patient population and was excluded from the analysis.\textsuperscript{20} The overall proportion of CV surgery patients in the high and low CV surgery groups was 0.54 and 0.09, respectively. The NPV of a low probability 4Ts score was similar in the 2 groups [0.998 (0.877-1.000) vs 0.993 (0.978-0.998), $P = .49$]. This remained true when the cut-point used to define the 2 groups was varied (data not shown).

Lastly, we conducted a post hoc sensitivity analysis of the 10 studies that used the standard 4Ts scoring system shown in Table 1. Studies using a different version of the model\textsuperscript{14,17,20} were excluded. Random effects point estimates of the PPV of a high probability [0.74 (0.32-0.94) vs 0.52 (0.13-0.89)], intermediate probability [0.18 (0.09-0.33) vs 0.12 (0.07-0.18)], and combined high and intermediate probability [0.27 (0.15-0.43) vs 0.16 (0.10-0.25)] 4Ts scores, although these differences were not statistically significant.

### Table 2. Characteristics of eligible studies

<table>
<thead>
<tr>
<th>Reference</th>
<th>Setting (n)</th>
<th>Population (%)</th>
<th>Patient characteristics</th>
<th>Prevalence</th>
<th>Reference standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lillo-Le Louët et al\textsuperscript{17}</td>
<td>France (84)</td>
<td>CV surgery (100)</td>
<td>Median age, y: 67</td>
<td>No. (%) female: 37 (44)</td>
<td>0.42</td>
</tr>
<tr>
<td>Lo et al\textsuperscript{14}</td>
<td>Canada (100)</td>
<td>CV surgery (50)</td>
<td>Median age, y: 68</td>
<td>No. (%) female: 38 (38)</td>
<td>0.16</td>
</tr>
<tr>
<td>Lo et al\textsuperscript{14}</td>
<td>Germany (236)</td>
<td>CV surgery (18)</td>
<td>Median age, y: 55</td>
<td>No. (%) female: 106 (45)</td>
<td>0.08</td>
</tr>
<tr>
<td>Pouplard et al\textsuperscript{18}</td>
<td>France (213)</td>
<td>CV surgery (14)</td>
<td>Median age, y: 69</td>
<td>No. (%) female: 73 (34)</td>
<td>0.10</td>
</tr>
<tr>
<td>Bryant et al\textsuperscript{19}</td>
<td>Australia (246)</td>
<td>CV surgery (43)</td>
<td>Median age, y: 62</td>
<td>No. (%) female: 92 (37)</td>
<td>0.04</td>
</tr>
<tr>
<td>Denys et al\textsuperscript{23}</td>
<td>Belgium (102)</td>
<td>CV surgery (51)</td>
<td>Median age, y: 70.5</td>
<td>No. (%) female: 45 (44)</td>
<td>0.10</td>
</tr>
<tr>
<td>Bakchoul et al\textsuperscript{20}</td>
<td>Germany (500)</td>
<td>CV surgery (18)</td>
<td>Median age, y: 68</td>
<td>No. (%) female: 566 (44)</td>
<td>0.07</td>
</tr>
<tr>
<td>Crowther et al\textsuperscript{24}</td>
<td>Canada (49)</td>
<td>CV surgery (10)</td>
<td>Median age, y: 65</td>
<td>No. (%) female: 24 (49)</td>
<td>0.04</td>
</tr>
<tr>
<td>Cuker et al\textsuperscript{10}</td>
<td>United States (50)</td>
<td>CV surgery (46)</td>
<td>Median age, y: 65</td>
<td>No. (%) female: 30 (60)</td>
<td>0.14</td>
</tr>
<tr>
<td>Berry et al\textsuperscript{25}</td>
<td>United States (104)</td>
<td>CV surgery (46)</td>
<td>Median age, y: 60</td>
<td>No. (%) female: 37 (36)</td>
<td>0.19</td>
</tr>
<tr>
<td>Nellen et al\textsuperscript{21}</td>
<td>Switzerland (1291)</td>
<td>CV surgery (7)</td>
<td>Median age, y: 68</td>
<td>No. (%) female: 566 (44)</td>
<td>0.07</td>
</tr>
<tr>
<td>Tawfik et al\textsuperscript{22}</td>
<td>Egypt (50)</td>
<td>CV surgery (36)</td>
<td>Median age, y: 64</td>
<td>No. (%) female: 19 (38)</td>
<td>0.10</td>
</tr>
<tr>
<td>Demma et al\textsuperscript{26}</td>
<td>United States (43)</td>
<td>CV surgery (100)</td>
<td>Median age, y: 65</td>
<td>No. (%) female: 20 (47)</td>
<td>0.26</td>
</tr>
</tbody>
</table>

ICU indicates intensive care unit; NR, not reported; EIA, enzyme immunoassay; HIPA, heparin–induced platelet activation assay; and PAT, heparin-induced platelet aggregation test.

[0.999 (0.913-1.000) vs 0.992 (0.982-0.997)]. None of these differences was statistically significant ($P = .086$, $0.32$, $0.20$, and $0.87$, respectively).
Discussion

In our systematic review and meta-analysis of the predictive value of the 4Ts, a low probability score (≤ 3) was associated with a high NPV for HIT (0.998; 95% CI, 0.970-1.000; Figure 2D). The NPV of the 4Ts remained high irrespective of the party responsible for scoring, the prevalence of HIT, or the proportion of CV surgery patients in the study population. The PPV of the model was more limited. The PPV of high, intermediate, and combined high and intermediate probability scores was 0.64 (0.40-0.82), 0.14 (0.09-0.22), and 0.22 (0.15-0.31), respectively (Figure 2) and was reduced in studies with a relatively low prevalence (≤ 0.10) of disease.

Consistent with the pooled estimate of HIT prevalence in our meta-analysis (0.11, 0.07-0.17), recent cohort studies suggest that only 7%-12% of patients referred for laboratory testing in clinical practice have a positive functional assay.27-29 A substantial fraction of the remaining, functional assay-negative patients receive “empiric” treatment for HIT. In one single-institution study, 35% of patients ultimately determined to have a negative SRA were treated with a parenteral direct thrombin inhibitor (pDTI) before return of laboratory results.10 In another series, only 10% of pDTI recipients had a positive SRA.7 In centers where functional assays are not available and clinicians must rely on less specific immunoassays, misdiagnosis and overtreatment are probably even more frequent.

Overuse of pDTIs carries potential adverse medical and economic consequences. These drugs are associated with a 1% incidence rate of major bleeding per treatment day,8 a risk compounded by the absence of reversal agents, and are ~ 100-fold more costly than an equivalent course of unfractionated heparin.30

Our findings suggest that the 4Ts score may be used to direct the initial evaluation and management of patients with suspected HIT and to curtail overtesting and overtreatment. In light of the high NPV of the model, we propose that it may be possible to exclude HIT and continue heparin in patients with a low probability 4Ts score without need for HIT laboratory testing or treatment (Figure 3). Given that 55.8% of patients in the studies included in our meta-analysis had a low probability 4Ts score, it is probable that implementation of such a decision rule would effect a major reduction in testing and unnecessary treatment. In patients with an intermediate or high probability score, we propose withdrawal of heparin, initiation of an alternative anticoagulant, and acquisition of HIT laboratory testing (Figure 3). In view of the modest PPV of the 4Ts, a substantial number of patients without true HIT will still undergo testing and treatment by this approach, underscoring the need for better diagnostic tools.4

A concern raised by our proposed algorithm (Figure 3) is the possibility of missing patients with true HIT despite a low probability 4Ts score. Reassuringly, our findings suggest that such patients are rare. The point estimate of the NPV of a low probability 4Ts score was ≥ 0.98 in all but one of the included studies (Figure 2D). In this study,25 the estimate of NPV (0.91, 0.80-0.97) may

Table 3. Study quality assessment by QUADAS-2 criteria

<table>
<thead>
<tr>
<th>Reference</th>
<th>Patient selection</th>
<th>Index test</th>
<th>Reference standard</th>
<th>Flow and timing</th>
<th>Patient selection</th>
<th>Index test</th>
<th>Reference standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lillo-Le Louët et al17</td>
<td>H</td>
<td>U</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Lo et al14 (Canada)</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Lo et al14 (Germany)</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Pouplard et al15</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
</tr>
<tr>
<td>Bryant et al19</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Deyns et al23</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Bakchoul et al20</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Crowther et al24</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Berry et al25</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>U</td>
</tr>
<tr>
<td>Nellen et al21</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Tawfik et al22</td>
<td>L</td>
<td>L</td>
<td>U</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Demma et al26</td>
<td>H</td>
<td>U</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
</tbody>
</table>

H indicates high; U, unclear; and L, low.
have been biased by inclusion of only those patients with suspected HIT and a positive immunoassay (selection bias). Had the 506 persons with suspected HIT and a negative immunoassay not been excluded from this study, it is likely that the NPV of a low probability 4Ts score would have been more in line with other eligible studies (Figure 2D). It is also conceivable that some patients with a low probability 4Ts score were misclassified as having HIT by the reference standard (classification bias). Indeed, seroprevalence studies demonstrate that, depending on the patient population, up to half of patients with a positive functional assay do not evince clinical evidence of HIT. Nevertheless, even rare cases of missed HIT may have dire consequences. The decision to initiate a pDTI or other alternative anticoagulant must therefore remain individualized and be based on a careful weighing of the anticipated risks and benefits. The ultimate impact of our proposed approach (Figure 3) on physician behavior and clinical outcomes requires evaluation in a properly designed, randomized comparison with standard, intuition-based diagnosis and management.

Several limitations of our study deserve mention. Most meta-analyses of diagnostic tests use sensitivity and specificity as the principal summary measures. To enhance the meaningfulness of our results to clinicians who evaluate and care for patients with

---

**Figure 2.** Forest plots of PPVs and NPVs of the 4Ts score. Forest plots of the PPV of a high probability (A), intermediate probability (B), and combined high and intermediate probability (C) 4Ts score. Forest plot of the NPV of a low probability 4Ts score (D).

**Figure 3.** Proposed algorithm for the initial evaluation and management of patients with suspected HIT.
suspected HIT, we elected to base our analysis on predictive values, which more fully reflect the clinical utility of a test within a given patient population and may be more intuitive to physicians. A limitation of predictive values is their dependence on disease prevalence. To address this limitation, we performed subanalyses of studies with a low (≤ 0.10) and high (>0.10) prevalence of HIT. As expected, PPV was greater in the high prevalence studies. The NPV of a low probability 4Ts score, however, remained high in both groups, suggesting the utility of the 4Ts as a rule-out test for HIT irrespective of disease prevalence. This remained true when different prevalence cut-points were used in the analysis.

Another limitation of our meta-analysis is lack of a uniform reference standard. There is no universally accepted gold standard for the diagnosis of HIT. Functional laboratory assays, such as the SRA, are considered to have the greatest accuracy among HIT diagnostic tests but lack standardization. Eligible studies in our meta-analysis used a diversity of reference standards (Table 2) and test methods, an important potential source of heterogeneity. Reassuringly, the NPV of a low probability 4Ts score remained high, regardless of the reference standard (Figure 2D). We consider HIT to be a clinicopathologic diagnosis and recommend that future studies of the 4Ts use a reference standard that combines rigorous laboratory (eg, a washed platelet functional assay) and clinical (eg expert adjudication) criteria.

Finally, limitations in inter-rater agreement with the 4Ts have been observed with κ coefficients ranging between 0.5 and 0.7. Inter-rater agreement was assessed in only one eligible study and therefore could not be addressed in our meta-analysis, but is an important potential shortcoming of the 4Ts that merits further investigation. Newer scoring models, including a revised version of the 4Ts, include more detailed and explicit itemization of clinical features in an attempt to clarify their meaning and enhance reproducibility among raters. In clinical practice, scoring is likely to be performed by the referring clinician (rather than by study personnel). The high NPV of the 4Ts was maintained in studies that used this approach (0.992, 0.982-0.997), supporting its potential utility in patient care.

In conclusion, our findings suggest that a low probability 4Ts score is a robust means of excluding HIT. Integration of assessment by the 4Ts in the evaluation and initial management of patients with suspected HIT (Figure 3) may reduce overtesting, overdiagnosis, and overtreatment of this disorder. This approach requires investigation in a randomized comparison with intuition-based diagnosis. To overcome limitations of prior studies, such a trial should incorporate use of a single version of the 4Ts, scoring by clinical provider rather than study personnel, and use of a rigorous clinicopathologic reference standard.

Acknowledgments

The authors thank the authors of the studies included in our meta-analysis for generously providing additional information about their studies at our request. This work was supported by the National Institutes of Health (K23HL112903, A.C.). M.A.C. holds a career investigator award from the Heart and Stroke Foundation of Ontario and holds the Leo Pharma Chair in Thromboembolism Research at McMaster University.

Authorship

Contribution: A.C. conceived and designed the study, searched the literature, evaluated the quality of eligible studies, analyzed and interpreted the data, and wrote the manuscript; P.A.G. analyzed and interpreted the data and edited the manuscript; M.A.C. interpreted the data and edited the manuscript; and T.E.W. evaluated the quality of eligible studies, interpreted the data, and edited the manuscript.

Conflict-of-interest disclosure: A.C. has provided consulting services to Baxter, Bayer, Biogen-Idec, Canyon Pharmaceuticals, CSL Behring, and Genzyme; has received research funding from Baxter, Bayer, and Novo Nordisk; and has provided expert witness testimony relating to heparin-induced thrombocytopenia. T.E.W. has provided consulting services to Canyon Pharmaceuticals, GlaxoSmithKline, Paringenix, and W.L. Gore; has received lecture honoraria from Canyon Pharmaceuticals and Pfizer Canada; has received research funding from GlaxoSmithKline for study of the 4Ts scoring system; has received royalties from Informa Healthcare; and has provided expert witness testimony relating to heparin-induced thrombocytopenia. M.A.C. has sat on advisory boards for Leo Pharma, Pfizer, Bayer, Boehringer Ingelheim, Alexion, CSL Behring, and Artisan Pharma; has prepared educational materials for Pfizer, Octapharma, and CSL Behring; and has provided expert witness testimony for Bayer. His institution has received funding for research projects from Boehringer Ingelheim, Octapharma, Pfizer, and Leo Pharma. P.A.G. declares no competing financial interests.

Correspondence: Adam Cuker, Hospital of the University of Pennsylvania, 3 Dulles, 3400 Spruce St, Philadelphia, PA 19104; e-mail: adam.cuker@uphs.upenn.edu.

References

12. Poupard C, Amiral J, Borg JY, Laporte-Simitsidis S,


Predictive value of the 4Ts scoring system for heparin-induced thrombocytopenia: a systematic review and meta-analysis

Adam Cuker, Phyllis A. Gimotty, Mark A. Crowther and Theodore E. Warkentin