

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

Adam Cuker,<sup>1,2</sup> Phyllis A. Gimotty,<sup>3</sup> Mark A. Crowther,<sup>4,5</sup> and Theodore E. Warkentin<sup>4,5</sup>

Departments of <sup>1</sup>Medicine, <sup>2</sup>Pathology and Laboratory Medicine, and <sup>3</sup>Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; and Departments of <sup>4</sup>Medicine and <sup>5</sup>Pathology 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.<sup>1,2</sup> Management involves cessation of heparin, avoidance or postponement of an oral vitamin K antagonist until platelet count recovery, and initiation of a nonheparin anticoagulant.<sup>3-5</sup> 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.<sup>6</sup> Misdiagnosis of HIT, conversely, may result in exposure of thrombocytopenic patients to alternative anticoagulants and their attendant ~1% daily risk of major hemorrhage<sup>7,8</sup> or in thrombosis from unnecessary suspension of heparin.<sup>9</sup>

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.<sup>10</sup> Laboratory tests for HIT are of 2 varieties. Widely available immunoassays are sensitive and simple to perform but yield frequent false-positive results.<sup>11</sup> Washed platelet functional assays such as the <sup>14</sup>C-serotonin release assay (SRA) have much greater specificity but require radioisotope and reactive donor platelets, reagents that are unfeasible for most clinical laboratories.<sup>12</sup> 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 thrombocy-

topenia; (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.<sup>13,14</sup> 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,<sup>13</sup> through December 2011 was performed using the keywords: (heparin-induced thrombocytopenia) AND (4Ts 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

Submitted July 11, 2012; accepted September 11, 2012. Prepublished online as *Blood* First Edition paper, September 18, 2012; DOI 10.1182/blood-2012-07-443051.

The publication costs of this article were defrayed in part by page charge

payment. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734.

© 2012 by The American Society of Hematology

**Table 1. The 4Ts scoring system**

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

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.

HIT, who were evaluated by both the 4Ts and a reference standard (defined subsequently) against which the 4Ts could be compared, were eligible for inclusion. Studies were excluded if the reference standard was an immunosay because of the limited specificity of such tests.<sup>11,12</sup>

### Data extraction and quality assessment

Key characteristics were extracted from eligible studies and recorded in an evidence table by one reviewer (A.C.). These included: author, year of publication, study design, characteristics of study participants (patient population, age, and sex), setting, reference standard, disease prevalence, and number of subjects found to be positive and negative for HIT relative to the reference standard in each 4Ts probability category. Information was requested from authors when it was not included in the published report. Two reviewers (A.C. and T.E.W.) independently evaluated study quality using QUADAS-2,<sup>15</sup> a standardized tool for the quality assessment of studies of diagnostic accuracy. The tool is composed of 4 domains: patient selection, index test, reference standard, and flow and timing. Each domain is assessed in terms of risk of bias; the first 3 domains are also assessed with respect to applicability to clinical practice.<sup>15</sup> Disagreements between reviewers were resolved by discussion and consensus.

### Data synthesis and analysis

The principal summary measures of our meta-analysis were the positive predictive value (PPV) of a high (score  $\geq$  6), intermediate (4-5), and combined high and intermediate ( $\geq$  4) probability 4Ts score and the negative predictive value (NPV) of a low probability score ( $\leq$  3). Predictive values, sensitivities, specificities, and corresponding 95% confidence intervals were tabulated for each 4Ts probability category in each study. The binomial-normal model for meta-analysis of proportions<sup>16</sup> was used to determine pooled estimates of these measures. Subgroup analyses assumed a different binomial proportion for the 2 strata. Proc NLMIXED from SAS/STAT Version 9.3 software (SAS) was used for all analyses.  $P < .05$  was considered statistically significant.

To explore heterogeneity among studies, we prespecified several subanalyses. First, because predictive values are known to be influenced by disease prevalence,<sup>16</sup> we compared the performance of the 4Ts in studies with a low ( $\leq$  0.10) and high ( $>$  0.10) prevalence of HIT. We used 0.10 as the cut-point for this analysis because it represented the median prevalence among eligible studies. Second, we hypothesized that the 4Ts would show better predictive value when performed by study personnel trained and practiced in use of the model than when conducted by referring clinicians. To test this hypothesis, we compared studies that used these 2 scoring methods. Third, patient population (eg, cardiovascular surgery, medical) is known to influence the accuracy of HIT diagnostic tests.<sup>4</sup> We planned a subgroup analysis to examine the effect of this variable on predictive value of the 4Ts. Lastly, several eligible studies used a version of the 4Ts that differed slightly from the standard model shown in Table 1. We performed a posteriori an analysis of only those studies that used the standard model to test whether predictive values associated with this model were in line with our overall results.

## 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 which<sup>14</sup> reported the results of 2 distinct studies, for a total of 13 eligible studies (Figure 1). Of these, 8 were prospective<sup>14,17-22</sup> and 5 were retrospective<sup>10,23-26</sup> cohort studies. No randomized controlled trials were identified.

### Study characteristics

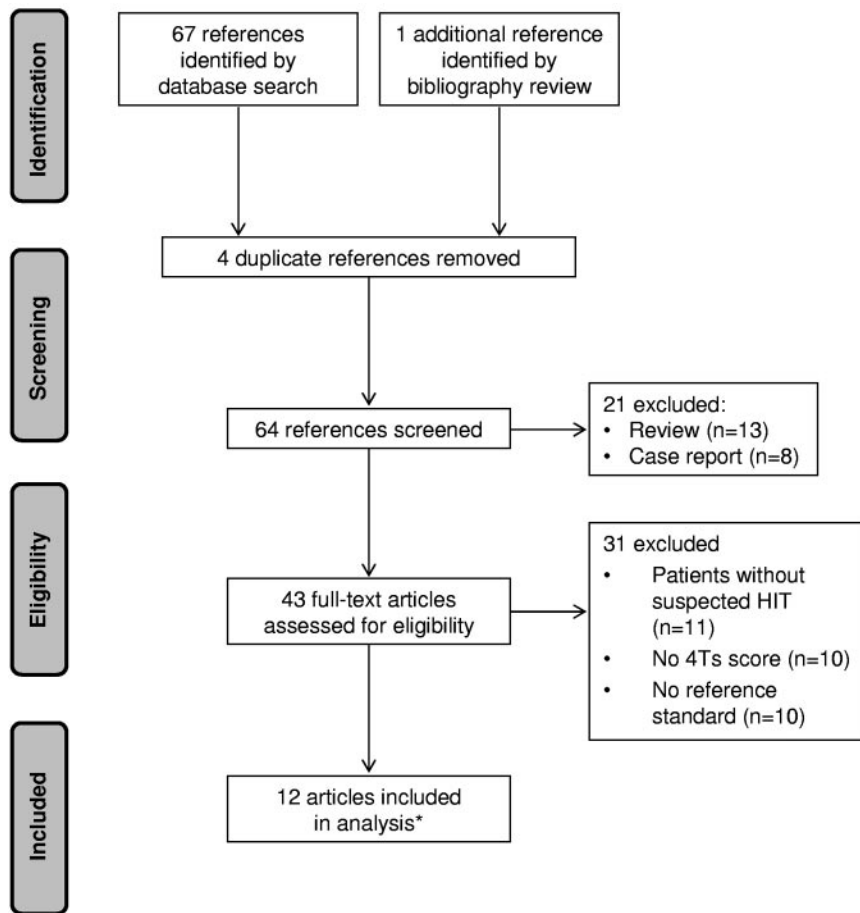
Study characteristics are summarized in Table 2. Studies were conducted in 8 countries and collectively enrolled 3068 patients. Sample size ranged from 43 to 1291 and prevalence of HIT from 0.04 to 0.42. Study populations were similar with respect to age and gender. Four studies restricted eligibility to cardiovascular (CV) surgical or critically ill subjects.<sup>17,24-26</sup> The remainder enrolled unselected medical and surgical patients with suspected HIT, although distribution by patient population varied. For example, European studies<sup>14,18,21</sup> included proportionately more medical and fewer CV surgical patients than North American series.<sup>10,14</sup>

The reference standard for defining HIT also differed among studies. One study used a clinical standard,<sup>10</sup> 10 studies used a laboratory standard,<sup>14,18-21,23-26</sup> and 2 studies used a standard incorporating both clinical and laboratory criteria.<sup>17,22</sup> The reference standard included a functional assay in all but one of the studies,<sup>10</sup> 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.<sup>14</sup> Patients in one study<sup>17</sup> were scored using an earlier rendering of the model.<sup>13</sup> Scores in 2 other studies were calculated using the Greifswald modification of the 4Ts.<sup>14</sup> 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 criteria<sup>15</sup> raised several potential methodologic limitations (Table 3). For instance,



**Figure 1. PRISMA flow diagram of literature search.**  
\*One of the included articles reported 2 eligible studies.

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 (0.970-1.000; Figure 2D). The pooled estimates of sensitivity and specificity of the 4Ts at a cut-off of  $\geq 4$  were 0.99 (0.86-1.00) 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 ( $\leq 0.10$ ,  $n = 8$ ) 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 (0.883-1.000)] 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,  $\sim 80\%$  of subjects were scored by the referring clinician<sup>21</sup>; 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

**Table 2. Characteristics of eligible studies**

Reference	Setting (n)	Patient characteristics			Prevalence	Reference standard
		Population (%)	Median age, y	No. (%) female		
Lillo-Le Louët et al <sup>17</sup>	France (84)	CV surgery (100)	67	37 (44)	0.42	1. Platelet count recovery within 5 days of heparin cessation AND 2. SRA > 20% OR no other explanation for thrombocytopenia
Lo et al <sup>14</sup>	Canada (100)	CV surgery (50) ICU (10) Medical (20) Non-CV surgery (20)	68	38 (38)	0.16	1. Positive polyspecific EIA AND 2. SRA > 50% (IgG-specific EIA > 1.0 OD units if SRA indeterminate)
Lo et al <sup>14</sup>	Germany (236)	CV surgery (18) ICU (20) Medical (47) Non-CV surgery (15)	55	106 (45)	0.08	1. Positive polyspecific EIA AND 2. HIPA
Pouplard et al <sup>18</sup>	France (213)	CV surgery (14) Medical (42) Non-CV surgery (19) NR (25)	69	73 (34)	0.10	SRA > 20%
Bryant et al <sup>19</sup>	Australia (246)	CV surgery (43) Medical (49) Non-CV surgery (8)	62	92 (37)	0.04	SRA > 20%
Denys et al <sup>23</sup>	Belgium (102)	CV surgery (51) Medical (42) Non-CV surgery (9)	70.5	45 (44)	0.10	Flow cytometric assay for CD62p
Bakchoul et al <sup>20</sup>	Germany (500)	NR	NR	NR	0.07	HIPA
Crowther et al <sup>24</sup>	Canada (49)	Medical ICU (90) Surgical ICU (10)	65	24 (49)	0.04	SRA > 50%
Cuker et al <sup>10</sup>	United States (50)	CV surgery (46) Medical (40) Non-CV surgery (14)	65	30 (60)	0.14	Consensus diagnosis of expert panel
Berry et al <sup>25</sup>	United States (104)	CV surgery (46) Non-CV surgery (54)	60	37 (36)	0.19	SRA > 20%
Nellen et al <sup>21</sup>	Switzerland (1291)	CV surgery (7) ICU (20) Medical (54) Non-CV surgery (16) Other (3)	68	566 (44)	0.07	PAT
Tawfik et al <sup>22</sup>	Egypt (50)	Medical (40) CV surgery (36) Non-CV surgery (24)	64	19 (38)	0.10	1. SRA > 2 SD above negative control AND 2. Clinical course after HIT laboratory testing consistent with HIT
Demma et al <sup>26</sup>	United States (43)	CV surgery (100)	65	20 (47)	0.26	SRA ≥ 20%

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 ( $P = .086, .32, .20, \text{ and } .87$ , 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 ( $> 0.20$ ,  $n = 8$ ) 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  $\leq 0.18$ ; in the remaining studies, the proportion was  $\geq 0.36$  (Table 2). One study did not collect information on patient population and was excluded from the analysis.<sup>20</sup> 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 [0.996 (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). The point estimate of PPV in the high CV surgery studies was greater than in the low CV surgery studies for 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.

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<sup>14,17,20</sup> were excluded. Random effects point estimates of the PPV of a high probability (0.63, 0.40-0.81), intermediate probability (0.14, 0.09-0.22), and combined high and intermediate probability (0.22, 0.14-0.31) 4Ts score were similar to our overall results. The NPV of a low probability 4Ts score remained high after exclusion of studies that did not use the standard 4Ts model (0.995, 0.954-0.999).

**Table 3. Study quality assessment by QUADAS-2 criteria**

Reference	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Lillo-Le Louët et al <sup>17</sup>	H	U	U	L	H	H	L
Lo et al <sup>14</sup> (Canada)	L	L	L	L	L	H	L
Lo et al <sup>14</sup> (Germany)	L	L	U	L	L	L	L
Pouplard et al <sup>18</sup>	L	L	L	L	L	U	L
Bryant et al <sup>19</sup>	L	L	L	L	L	H	L
Denys et al <sup>23</sup>	L	L	U	L	L	H	L
Bakchoul et al <sup>20</sup>	L	L	U	L	L	H	L
Crowther et al <sup>24</sup>	L	L	L	L	H	H	L
Cuker et al <sup>10</sup>	L	L	L	L	L	H	L
Berry et al <sup>25</sup>	H	L	L	H	L	U	L
Nellen et al <sup>21</sup>	L	L	U	L	L	L	L
Tawfik et al <sup>22</sup>	L	L	U	L	L	L	L
Demma et al <sup>26</sup>	H	U	L	H	L	H	L

H indicates high; U, unclear; and L, low.

## Discussion

In our systematic review and meta-analysis of the predictive value of the 4Ts, a low probability score ( $\leq 3$ ) 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 ( $\leq 0.10$ ) 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.<sup>27-29</sup> 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.<sup>10</sup> In another series, only 10% of pDTI recipients had a positive SRA.<sup>7</sup> 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

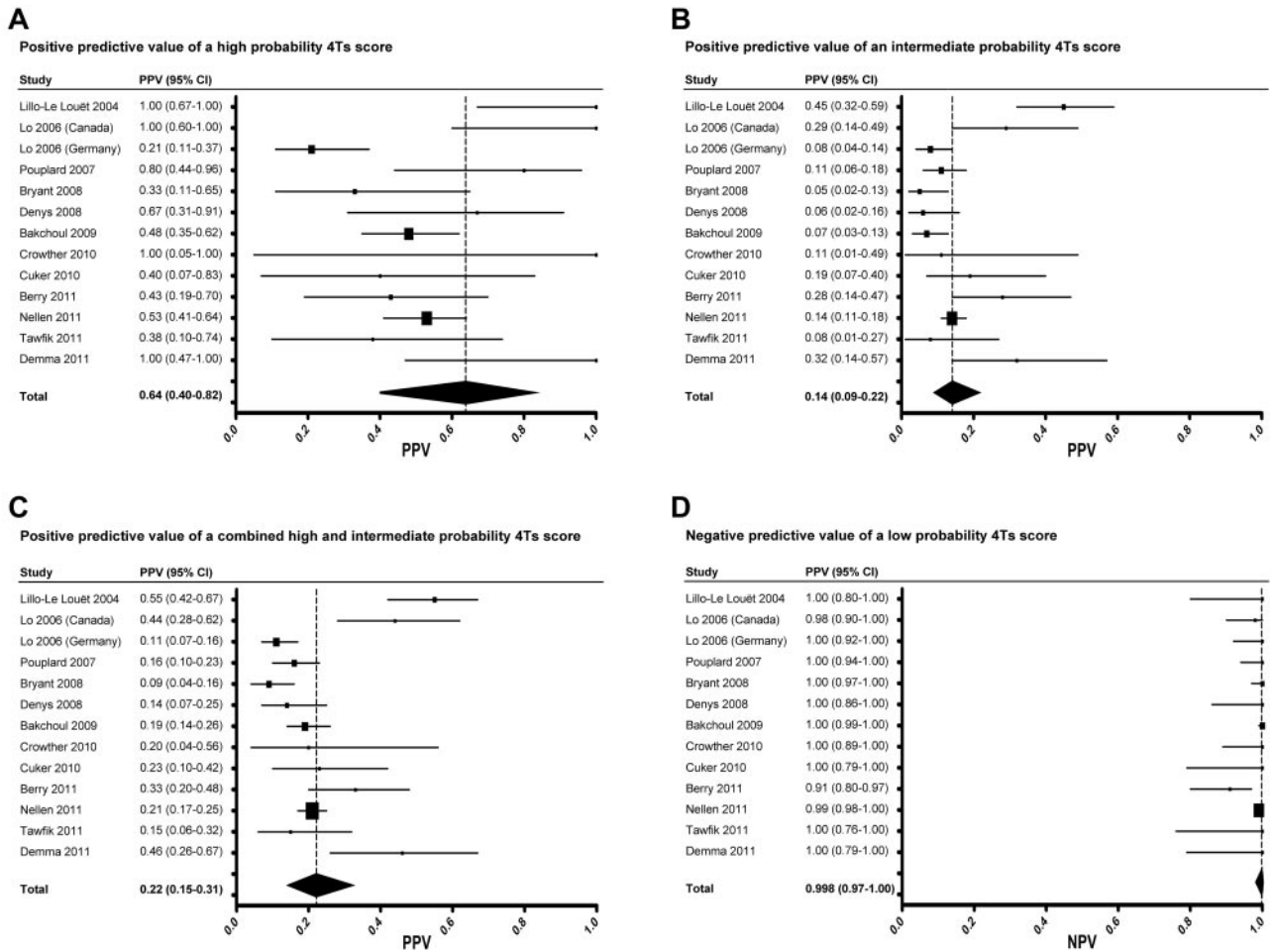
consequences. These drugs are associated with a 1% incidence rate of major bleeding per treatment day,<sup>8</sup> a risk compounded by the absence of reversal agents, and are ~ 100-fold more costly than an equivalent course of unfractionated heparin.<sup>30</sup>

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.<sup>4</sup>

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  $\geq 0.98$  in all but one of the included studies (Figure 2D). In this study,<sup>25</sup> the estimate of NPV (0.91, 0.80-0.97) may

**Table 4. Number of patients with and without HIT in each 4Ts probability category**

Reference	High probability		Intermediate probability		Low probability	
	HIT <sup>+</sup>	HIT <sup>-</sup>	HIT <sup>+</sup>	HIT <sup>-</sup>	HIT <sup>+</sup>	HIT <sup>-</sup>
Lillo-Le Louët et al <sup>17</sup>	11	0	24	29	0	20
Lo et al <sup>14</sup> (Canada)	8	0	8	20	1	63
Lo et al <sup>14</sup> (Germany)	9	33	11	128	0	55
Pouplard et al <sup>18</sup>	8	2	14	115	0	74
Bryant et al <sup>19</sup>	4	8	5	87	0	142
Denys et al <sup>23</sup>	6	3	4	58	0	31
Bakchoul et al <sup>20</sup>	26	28	9	121	0	316
Crowther et al <sup>24</sup>	1	0	1	8	0	39
Cuker et al <sup>10</sup>	2	3	5	21	0	19
Berry et al <sup>25</sup>	6	8	9	23	5	53
Nellen et al <sup>21</sup>	39	35	50	308	7	852
Tawfik et al <sup>22</sup>	3	5	2	24	0	16
Demma et al <sup>26</sup>	5	0	6	13	0	19
Total	128	125	148	955	13	1699

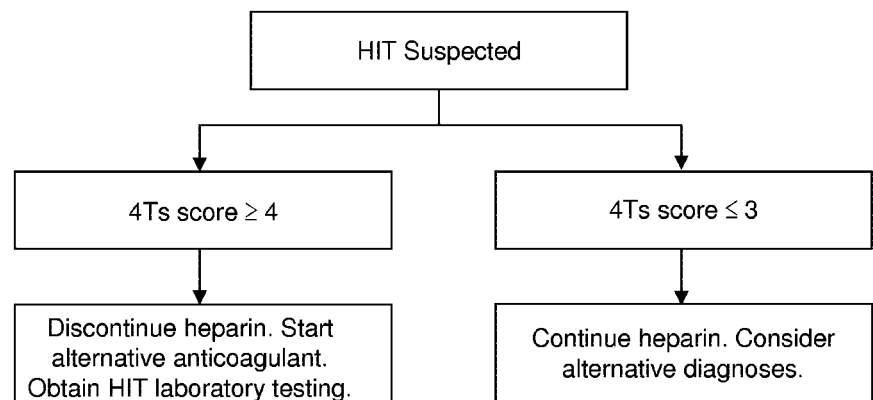


**Figure 2.** Forest plots of PPVs and NPVs of the 4Ts score. Forest plots of the PPV of a high probability (A), intermediate probability (B), and combined high and intermediate probability (C) 4Ts score. Forest plot of the NPV of a low probability 4Ts score (D).

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.<sup>31</sup> 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.<sup>32</sup>

Several limitations of our study deserve mention. Most meta-analyses of diagnostic tests use sensitivity and specificity as the principal summary measures.<sup>33</sup> To enhance the meaningfulness of our results to clinicians who evaluate and care for patients with



**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<sup>16</sup> and may be more intuitive to physicians.<sup>34,35</sup> A limitation of predictive values is their dependence on disease prevalence.<sup>16</sup> To address this limitation, we performed subanalyses of studies with a low ( $\leq 0.10$ ) 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<sup>3,4</sup> but lack standardization.<sup>36</sup> 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  $\kappa$  coefficients ranging between 0.5 and 0.7.<sup>10,37,38</sup> Inter-rater agreement was assessed in only one eligible study<sup>10</sup> 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.<sup>10,39</sup> 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 Health-care; 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

- Amiral J, Bridey F, Dreyfus M, et al. Platelet factor 4 complexed to heparin is the target for antibodies generated in heparin-induced thrombocytopenia. *Thromb Haemost*. 1992;68(1):95-96.
- Greinacher A, Potzsch B, Amiral J, Dummel V, Eichner A, Mueller-Eckhardt C. Heparin-associated thrombocytopenia: isolation of the antibody and characterization of a multimolecular PF4-heparin complex as the major antigen. *Thromb Haemost*. 1994;71(2):247-251.
- Linkins LA, Dans AL, Moores LK, et al. Treatment and prevention of heparin-induced thrombocytopenia: antithrombotic therapy and prevention of thrombosis, 9th Ed: American College of Chest Physicians Evidence-based Clinical Practice Guidelines. *Chest*. 2012;141(2 Suppl):e495S-e530S.
- Cuker A, Cines DB. How I treat heparin-induced thrombocytopenia. *Blood*. 2012;119(10):2209-2218.
- Warkentin TE, Elavathil LJ, Hayward CP, Johnston MA, Russett JI, Kelton JG. The pathogenesis of venous limb gangrene associated with heparin-induced thrombocytopenia. *Ann Intern Med*. 1997;127(9):804-812.
- Greinacher A, Eichler P, Lubenow N, Kwasny H, Luz M. Heparin-induced thrombocytopenia with thromboembolic complications: meta-analysis of 2 prospective trials to assess the value of parenteral treatment with lepirudin and its therapeutic aPTT range. *Blood*. 2000;96(3):846-851.
- Cuker A. Heparin-induced thrombocytopenia (HIT) in 2011: an epidemic of overdiagnosis. *Thromb Haemost*. 2011;106(6):993-994.
- Warkentin TE. HIT: treatment easier, prevention harder. *Blood*. 2012;119(5):1099-1100.
- Smythe MA, Warkentin TE, Woodhouse AL, Zakalik D. Venous limb gangrene and fatal hemorrhage: adverse consequences of HIT "overdiagnosis" in a patient with antiphospholipid syndrome. *Am J Hematol*. 2011;86(2):188-191.
- Cuker A, Arepally G, Crowther MA, et al. The HIT Expert Probability (HEP) Score: a novel pre-test probability model for heparin-induced thrombocytopenia based on broad expert opinion. *J Thromb Haemost*. 2010;8(12):2642-2650.
- Lo GK, Sigouin CS, Warkentin TE. What is the potential for overdiagnosis of heparin-induced thrombocytopenia? *Am J Hematol*. 2007;82(12):1037-1043.
- Pouplard C, Amiral J, Borg JY, Laporte-Simitsidis S,

- Delahousse B, Gruel Y. Decision analysis for use of platelet aggregation test, carbon 14-serotonin release assay, and heparin-platelet factor 4 enzyme-linked immunosorbent assay for diagnosis of heparin-induced thrombocytopenia. *Am J Clin Pathol*. 1999;111(5):700-706.
13. Warkentin TE, Heddle NM. Laboratory diagnosis of immune heparin-induced thrombocytopenia. *Curr Hematol Rep*. 2003;2(2):148-157.
  14. Lo GK, Juhl D, Warkentin TE, Sigouin CS, Eichler P, Greinacher A. Evaluation of pretest clinical score (4 T's) for the diagnosis of heparin-induced thrombocytopenia in two clinical settings. *J Thromb Haemost*. 2006;4(4):759-765.
  15. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536.
  16. Chu H, Nie L, Cole SR, Poole C. Meta-analysis of diagnostic accuracy studies accounting for disease prevalence: alternative parameterizations and model selection. *Statist Med*. 2009;28(18):2384-2399.
  17. Lillo-Le Louët A, Boutouyrie P, Alhenc-Gelas M, et al. Diagnostic score for heparin-induced thrombocytopenia after cardiopulmonary bypass. *J Thromb Haemost*. 2004;2(11):1882-1888.
  18. Pouplard C, Gueret P, Fouassier M, et al. Prospective evaluation of the '4Ts' score and particle gel immunoassay specific to heparin/PF4 for the diagnosis of heparin-induced thrombocytopenia. *J Thromb Haemost*. 2007;5(7):1373-1379.
  19. Bryant A, Low J, Austin S, Joseph JE. Timely diagnosis and management of heparin-induced thrombocytopenia in a frequent request, low incidence single centre using clinical 4T's score and particle gel immunoassay. *Br J Haematol*. 2008;143(5):721-726.
  20. Bakchoul T, Giptner A, Najaoui A, Bein G, Santoso S, Sachs UJH. Prospective evaluation of PF4/heparin immunoassays for the diagnosis of heparin-induced thrombocytopenia. *J Thromb Haemost*. 2009;7(8):1260-1265.
  21. Nellen V, Sulzer I, Barizzi G, Lammle B, Alberio L. Rapid exclusion or confirmation of heparin-induced thrombocytopenia: a single-center experience with 1291 patients. *Haematologica*. 2012;97(1):89-97.
  22. Tawfik NM, Hegazy MA, Hassan EA, Ramadan YK, Nasr AS. Egyptian experience of reliability of 4T's score in diagnosis of heparin induced thrombocytopenia syndrome. *Blood Coagul Fibrinolysis*. 2011;22(8):701-705.
  23. Denys B, Stove V, Philippé J, Devreese K. A clinical-laboratory approach contributing to a rapid and reliable diagnosis of heparin-induced thrombocytopenia. *Thromb Res*. 2008;123(1):137-145.
  24. Crowther MA, Cook DJ, Albert M, et al. The 4Ts scoring system for heparin-induced thrombocytopenia in medical-surgical intensive care unit patients. *J Crit Care*. 2010;25(2):287-293.
  25. Berry C, Tcherniantchouk O, Ley EJ, et al. Overdiagnosis of heparin-induced thrombocytopenia in surgical ICU patients. *J Am Coll Surg*. 2011;213(1):10-17.
  26. Demma LJ, Winkler AM, Levy JH. A diagnosis of heparin-induced thrombocytopenia with combined clinical and laboratory methods in cardiothoracic surgical intensive care unit patients. *Anesth Analg*. 2011;113(4):697-702.
  27. Schallmoser K, Drexler C, Rohde E, et al. The particle gel immunoassay as a rapid test to rule out heparin-induced thrombocytopenia? *J Thorac Cardiovasc Surg*. 2009;137(3):781-783.
  28. Juhl D, Eichler P, Lubenow N, Strobel U, Wessel A, Greinacher A. Incidence and clinical significance of anti-PF4/heparin antibodies of the IgG, IgM, and IgA class in 755 consecutive patient samples referred for diagnostic testing for heparin-induced thrombocytopenia. *Eur J Haematol*. 2006;76(5):420-426.
  29. Sachs UJ, von Hesberg J, Santoso S, Bein G, Bakchoul T. Evaluation of a new nanoparticle-based lateral-flow immunoassay for the exclusion of heparin-induced thrombocytopenia. *Thromb Haemost*. 2011;106(6):1197-1202.
  30. Patrick AR, Winkelmayer WC, Avorn J, Fischer MA. Strategies for the management of suspected heparin-induced thrombocytopenia: a cost-effectiveness analysis. *Pharmacoeconomics*. 2007;25(11):949-961.
  31. Warkentin TE. How I diagnose and manage HIT. *Hematology Am Soc Hematol Educ Program*. 2011;2011:143-149.
  32. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII. How to use articles about clinical decision rules. Evidence-based Medicine Working Group. *JAMA*. 2000;284(1):79-84.
  33. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM; Cochrane Diagnostic Test Accuracy Working Group: systematic reviews of diagnostic test accuracy. *Ann Intern Med*. 2008;149(12):889-897.
  34. Steurer J, Fischer JE, Bachmann LM, Koller M, ter Riet G. Communicating accuracy of tests to general practitioners: a controlled study. *BMJ*. 2002;324(7341):824-826.
  35. Puhan MA, Steurer J, Bachmann LM, ter Riet G. A randomized trial of ways to describe test accuracy: the effect on physicians' post-test probability estimates. *Ann Intern Med*. 2005;143(3):184-189.
  36. Price EA, Hayward CP, Moffat KA, Moore JC, Warkentin TE, Zehnder JL. Laboratory testing for heparin-induced thrombocytopenia is inconsistent in North America: a survey of North American specialized coagulation laboratories. *Thromb Haemost*. 2007;98(6):1357-1361.
  37. Strutt JK, Mackery JE, Johnson SM, Sylvia LM. Assessment of the 4Ts pretest clinical scoring system as a predictor of heparin-induced thrombocytopenia. *Pharmacotherapy*. 2011;31(2):138-145.
  38. Nagler M, Fabbro T, Wuillemin WA. Prospective evaluation of the interobserver reliability of the 4Ts score in patients with suspected heparin-induced thrombocytopenia. *J Thromb Haemost*. 2012;10(1):151-152.
  39. Warkentin TE, Linkins LA. Non-necrotizing heparin-induced skin lesions and the 4T's score. *J Thromb Haemost*. 2010;8(7):1483-1485.