

Prospective comparison of the HEP score and 4Ts score for the diagnosis of heparin-induced thrombocytopenia

Allyson M. Pishko,¹ Sara Fardin,² Daniel S. Lefler,¹ Koosha Paydary,³ Rolando Vega,¹ Gowthami M. Arepally,⁴ Mark Crowther,⁵ Lawrence Rice,⁶ Douglas B. Cines,^{1,7} and Adam Cuker^{1,7}

¹Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; ²Department of Medicine, Jersey Shore University Medical Center, Neptune, NJ; ³Department of Internal Medicine, John H. Stronger Jr Hospital of Cook County, Chicago IL; ⁴Division of Hematology, Department of Medicine, Duke University, Durham, NC; ⁵Department of Medicine, McMaster University, Hamilton, ON, Canada; ⁶Department of Medicine, Weill Cornell Medical College, Houston, TX; and ⁷Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

Key Points

- In a prospective study, the HEP score performed similarly to the 4Ts score for diagnosis of HIT.
- The HEP score was superior to the 4Ts score among less experienced clinicians and in patients in the ICU.

The HIT Expert Probability (HEP) score compared favorably with the 4Ts score in a retrospective study. We assessed the diagnostic accuracy of the HEP score compared with the 4Ts score in a prospective cohort of 310 patients with suspected heparin-induced thrombocytopenia (HIT). A member of the clinical team calculated the HEP score and 4Ts score. An independent panel adjudicated HIT status based on a clinical summary as well as the results of HIT laboratory testing. The prevalence of HIT in the study population was 14.7%. At a cutoff of ≥ 3 , the HEP score was 95.3% sensitive (95% confidence interval [CI], 84.2-99.4) and 35.7% specific (95% CI, 29.8-42.0) for HIT. A 4Ts score of ≥ 4 had a sensitivity of 97.7% (95% CI, 86.2-99.8) and specificity of 32.9% (95% CI, 27.2-39.1). The areas under the receiver operating characteristic (ROC) curves (AUCs) for the HEP score and 4Ts score were similar (0.81 [95% CI, 0.74-0.87] vs 0.76 [95% CI, 0.69-0.83]; $P = .12$). The HEP score exhibited a significantly higher AUC than the 4Ts score in patients in the intensive care unit (ICU) (0.86 vs 0.79; $P = .03$). Among trainee scorers, the HEP score performed significantly better than the 4Ts score (AUC, 0.80 vs 0.73; $P = .03$). Our data suggest that either the 4Ts score or the HEP score may be used in clinical practice. The HEP score may be preferable in ICU patients and among less experienced clinicians.

Introduction

Heparin-induced thrombocytopenia (HIT) is a prothrombotic disorder mediated by platelet-activating antibodies of the immunoglobulin G class that target multimolecular complexes of platelet factor 4 (PF4) and heparin. Accurate and timely diagnosis is paramount. Delays in diagnosis and initiation of appropriate therapy are associated with an initial 6.1% daily risk of thromboembolism, limb loss, and death and a cumulative thrombotic risk of 38% to 53% at 30 days.¹⁻³ Conversely, misdiagnosis needlessly exposes patients without HIT to discontinuation of "standard" heparin anticoagulants and initiation of costly alternative anticoagulants with their attendant bleeding risk.⁴

Several clinical scoring systems have been proposed to improve and standardize the diagnosis of HIT.⁵⁻⁷ In the 4Ts score, the most extensively studied of these scoring systems, patients are evaluated across 4 domains: magnitude of thrombocytopenia, timing of onset of thrombocytopenia, thrombosis or other clinical sequelae, and the likelihood of other causes of thrombocytopenia (supplemental Table 1).⁶ The 4Ts score has a negative predictive value (NPV) approaching 100%,⁸ but is limited by modest positive

predictive value (PPV) and significant interrater variability.^{7,9} Moreover, the performance of the 4Ts score may be poorer in medical-surgical intensive care unit (ICU) patients.^{10,11}

The HIT Expert Probability (HEP) score, a clinical scoring system based on the opinions of 26 HIT experts from North America, was designed to overcome the limitations of the 4Ts score. In addition to the 4 domains in the 4Ts score, it includes bleeding and explicit itemization of common other causes of thrombocytopenia (supplemental Table 1).⁷ The HEP score has been compared with the 4Ts score in 3 published studies. In a retrospective study of 50 patients with suspected HIT, the HEP score demonstrated improved PPV and interobserver agreement compared with the 4Ts score.⁷ In 2 other studies, a study of 47 patients from Thailand¹² and a study of 51 patients from a US center,¹³ the HEP score and 4Ts score exhibited similar operating characteristics. Important limitations of these studies included their small size and performance of scoring by study personnel rather than by clinicians directly involved in care of the patient.

In the present study, we aimed to assess the diagnostic accuracy of the HEP score and compare its performance to the 4Ts score in a real-world clinical setting.

Methods

Patients

Consecutive adult inpatients (≥ 18 years of age) were recruited from the Hospital of the University of Pennsylvania and an affiliated community hospital, Penn Presbyterian Medical Center, from June 2012 through January 2015. Patients with suspected acute HIT in whom diagnostic HIT laboratory testing was requested were eligible for enrollment. Demographic and clinical information was collected at time of enrollment, during the duration of the patient's hospital admission, and by telephone follow-up 30 days after enrollment using standardized case report forms. The protocol was approved by the University of Pennsylvania Institutional Review Board and conducted in compliance with the Declaration of Helsinki. All subjects or their surrogates provided written informed consent.

Assessment by HEP score and 4Ts score

At the time of study enrollment, a member of the patient's care team (usually a hematology fellow or attending hematologist) who was directly involved in the patient's care was asked to tabulate a HEP score and 4Ts score for the patient. Scorers were also asked to record how long it took to complete the HEP score. We did not collect data on time for completion of the 4Ts score. To ensure that scorers were not biased by the results of HIT laboratory testing, scoring was required to be completed prior to availability of HIT laboratory test results. Subjects were excluded if HIT laboratory test results became available before the scorer completed the 4Ts score and HEP score.

HIT laboratory testing

All patients underwent HIT laboratory testing with a polyspecific HIT antibody enzyme-linked immunosorbent assay (ELISA; Immucor GTI Diagnostics Inc, Waukesha, WI) and an in-house serotonin release assay (SRA) as previously described.¹⁴ Our SRA uses patient plasma and donor platelet-rich plasma rather than patient heat-inactivated serum and donor washed platelets, as used in other forms of the SRA.¹⁵ An ELISA result ≥ 0.4 optical density

(OD) units was considered positive in accordance with the manufacturer's instructions. The SRA was considered positive if there was $<5\%$ serotonin release in the absence of heparin and $>20\%$ release after addition of 0.1 or 0.5 U/mL heparin. If there was $\geq 5\%$ serotonin release in the absence of heparin, the result was classified as indeterminate.

Definition of HIT

In light of the high negative predictive value of the ELISA and SRA, all patients with a negative ELISA (<0.4 OD units) and a negative SRA were classified as not having HIT.¹⁶ Clinical case summaries of the remaining subjects were prepared using a standardized form. These case summaries included detailed clinical information about the hospital course, the platelet count trend, exposure to heparin and other medications, thrombotic events, and the results of the HIT antibody ELISA and SRA. Also included was 30-day follow-up information on vital status and diagnosis of thromboembolic events after discharge.

An independent adjudication panel of 3 HIT experts (G.M.A., M.C., and L.R.) reviewed the clinical case summaries and selected 1 of the following options for each case: (a) the patient had HIT; (b) I am uncertain whether the patient had HIT, but my suspicion is sufficiently high that I would manage the patient as though she/he had HIT; (c) I am uncertain whether the patient had HIT, but my suspicion is sufficiently low that I would manage the patient as though she/he did not have HIT; (d) the patient did not have HIT. Disagreement among adjudicators was resolved by discussion and consensus. In cases in which consensus could not be reached, the majority opinion was used. Subjects who were given a consensus or majority rating of (a) or (b) by the adjudicators were classified as HIT⁺. All others were classified as HIT⁻.

Statistical analysis

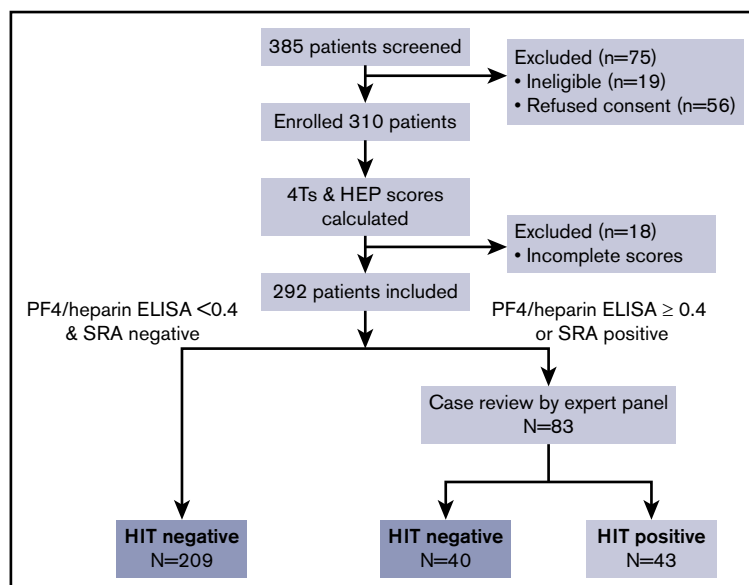
In a retrospective study, the HEP score demonstrated a sensitivity of 100% and a specificity of 60%.⁷ Assuming a prevalence of true HIT of 10% in our study population, we estimated that we would need to enroll 300 subjects to confirm a specificity of 60% with a 95% confidence interval (CI) of 54% to 66% and a sensitivity of 100% with a 95% CI of 86% to 100%.

Descriptive statistics were tabulated for subjects with HIT and without HIT. For nonnormally distributed continuous variables, the Wilcoxon rank sum test was used to compare characteristics between HIT⁺ and HIT⁻ patients. The Pearson χ^2 or Fisher's exact test was used to compare categorical variables between groups.

Diagnostic accuracy was assessed using nonparametric receiver operating characteristic (ROC) analysis. The area under the ROC curve (AUC) was calculated using the trapezoidal method. The standard error and 95% CI for the AUC were calculated using the method of DeLong et al.¹⁷ The method of DeLong et al for correlated samples was used to compare AUCs for the HEP score and 4Ts score.¹⁷ Operating characteristics of the HEP score at each cutpoint were calculated, including sensitivity, specificity, PPV, and NPV, and their exact binomial 95% CIs. A cutoff for the HEP score was selected from the ROC curve, which had high sensitivity ($>95\%$) and maximized specificity. The sensitivity and specificity of the HEP score at the selected cutoff and the 4Ts score at a cutoff of 4 were compared using the McNemar test or the exact McNemar test as appropriate.¹⁸

We conducted several prespecified subgroup analyses. We evaluated the diagnostic accuracy of the HEP score and 4Ts score

Figure 1. Patient enrollment and adjudication flow diagram.



separately in medical patients, surgical patients, cardiothoracic surgical patients, and ICU patients. We also assessed the operating characteristics of the 2 scoring systems based on whether the scorer was an attending physician or a trainee. Finally, to test the robustness of our reference standard, we conducted a sensitivity analysis in which the SRA alone was used as the reference standard.

For all analyses, $P < .05$ was considered statistically significant. Statistical analyses were performed using STATA, version 14 (College Station, TX).

Results

Patient enrollment

A total of 385 patients were screened for eligibility and 310 patients were enrolled in the study (Figure 1). Eighteen were excluded because the HEP score and/or 4Ts score was not completed prior to the availability of HIT laboratory test results, leaving 292 subjects available for analysis. Eighty-three (28.4%) had a HIT antibody ELISA ≥ 0.4 OD units and/or a positive SRA and were referred for adjudication by the expert panel.

Adjudicated HIT status

Consensus on the dichotomous classification of HIT⁺ or HIT⁻ status among the 3 adjudicators was achieved for 79 of the 83 subjects (95.2%) who underwent adjudication. For the remaining 4 subjects, the majority opinion was used to define HIT status.

Overall, 43 subjects (14.7%) met the prespecified definition of HIT and were classified as HIT⁺. Forty subjects did not have HIT in the estimation of the adjudication panel. These 40 subjects, along with the 209 subjects with negative HIT laboratory test results, were classified as HIT⁻ (Figure 1).

Patient characteristics

The demographic, clinical, and laboratory characteristics of the HIT⁺ and HIT⁻ subjects are summarized in Table 1. In general, demographic and clinical characteristics were similar between the

2 groups with the exception of thrombosis at enrollment, which was twice as common in the HIT⁺ cohort compared with the HIT⁻ cohort 65.1% vs 32.5% ($P < .0001$). The HIT ELISA was ≥ 1.0 OD units in 93.0% of HIT⁺ patients compared with 7.6% of HIT⁻ patients ($P < .0001$). The SRA was positive in 67.4% of HIT⁺ patients (29 of 43) and in 0.4% of HIT⁻ patients (1 of 249) ($P < .0001$).

Fourteen patients were judged to be HIT⁺ by the expert adjudicators even though the SRA result was not positive. Nine of the 14 patients had thrombosis. One HIT⁺ patient had an indeterminate SRA and an ELISA > 3.0 OD units. Ten of 13 subjects with a negative SRA had an ELISA ≥ 1.0 OD units and 3 had a weakly positive ELISA (0.4-0.99 OD units). Figure 2 shows the distribution of ELISA results according to HIT status and SRA result.

Scoring

A physician directly involved in the care of the patient calculated a 4Ts score and a HEP score for each subject. In total, 46 different scorers (73.9% trainees, 26.1% attending physicians) participated in the study. 103 (35.3%) subjects were scored by an attending physician and 189 (64.7%) by a trainee. Almost all (99.0%) of the subjects scored by an attending physician were scored by an attending hematologist. Almost all (95.2%) of the subjects scored by a trainee were scored by a hematology fellow.

The self-reported time for completion of the HEP score by all scorers ranged from 1 to 20 minutes, with a median of 4 minutes (IQR, 2-5). Among trainee scorers, the median time was 3 minutes (IQR, 2-5) compared with 5 minutes (IQR, 4-7) among attending scorers ($P < .0001$).

Diagnostic accuracy of 4Ts score and HEP score

The 4Ts score ranged from 0 to 8 points. The median 4Ts score was greater in HIT⁺ patients than in HIT⁻ patients (5 vs 4, $P < .0001$) (Figure 3A). A 4Ts score ≥ 4 was associated with a sensitivity of 97.7% (95% CI, 86.2%-99.8%) and specificity of 32.9% (95% CI, 27.2%-39.1%) and a positive and negative

Table 1. Demographic, clinical, and laboratory characteristics of patients with and without HIT

	HIT ⁻ , N = 249	HIT ⁺ , N = 43	P*
Demographic characteristics			
Age, median (IQR), y	67 (58-75)	64 (53-74)	.163
Male sex, n (%)	128 (51.4)	27 (62.8)	.167
Race, n (%)			.786
White	178 (71.5)	34 (79.1)	
Hispanic	4 (1.6)	0 (0.0)	
African American	60 (24.1)	9 (20.9)	
Asian	7 (2.8)	0 (0.0)	
Clinical characteristics at enrollment			
Type of patient, n (%)			.319†
Medical	113 (45.4)	16 (37.2)	
Surgical	136 (54.6)	27 (62.8)	
Cardiothoracic	109 (43.8)	16 (37.2)	
Orthopedic	1 (0.4)	0 (0.0)	
Trauma	5 (2.0)	1 (2.3)	
Vascular	4 (1.6)	1 (2.3)	
Other	17 (6.8)	9 (20.9)	
ICU, n (%)	148 (59.4)	22 (51.2)	.310‡
Medical	49 (19.6)	9 (20.9)	
Cardiothoracic	84 (33.7)	8 (18.6)	
General surgical	15 (6.0)	5 (11.6)	
Hospital type, n (%)			.091
Academic	158 (63.5)	33 (76.7)	
Community	91 (36.6)	10 (23.3)	
Type of heparin, n (%)			.567
UFH	160 (75.5)	31 (79.5)	
LMWH	25 (11.8)	3 (7.7)	
Both	27 (12.6)	5 (12.8)	
Platelet count at enrollment, median (IQR), ×10 ⁹ /L	60.5 (38-86.5)	51 (32-94)	.569
Thrombosis at enrollment,§ n (%)			<.001
Absent	168 (67.4)	15 (34.88)	
Present	81 (32.5)	28 (65.1)	
Venous	45 (18.1)	21 (48.8)	
Arterial	34 (13.6)	6 (13.9)	
Both	2 (0.8)	1 (2.3)	
HIT laboratory test results			
HIT ELISA (OD), n (%)			<.001
0-0.39	209 (83.9)	0 (0)	
0.4-0.99	21 (8.4)	3 (7.0)	
≥1.0	19 (7.6)	40 (93.0)	
SRA, n (%)			<.001
Positive	1 (0.4)	29 (67.4)	
Negative	244 (97.9)	13 (30.2)	
Indeterminate	2 (0.8)	1 (2.3)	
Missing	2 (0.8)	0 (0)	

Table 1. (continued)

	HIT ⁻ , N = 249	HIT ⁺ , N = 43	P*
In-hospital and 30-d outcomes			
Length of stay, median (IQR), d	16 (10-29)	22 (14-30)	.067
In-hospital mortality, n (%)			.969
Alive	186 (74.7)	32 (74.4)	
Deceased	63 (25.3)	11 (25.6)	
30-d mortality, n (%)			.893
Alive	161 (64.7)	27 (62.8)	
Deceased	75 (30.1)	13 (30.2)	
Lost to follow-up	13 (5.2)	3 (7.0)	

IQR, interquartile range; LMWH, low-molecular-weight heparin; UFH, unfractionated heparin.

*Wilcoxon rank sum test for continuous variables and the χ^2 or Fisher's exact test for categorical variables.

†P value represents statistical comparison of proportion of patients on medical or surgical service.

‡P value represents statistical comparison of proportion of patients on an ICU service.

§Only thromboses confirmed by imaging are included.

||P value represents statistical comparison of proportion of patients with thrombosis present.

predictive value of 20.1% (95% CI, 15.0%-26.3%) and 98.8% (95% CI, 92.5%-99.9%), respectively (Table 2).

The HEP score ranged from -7 to 15 points. The median HEP score was 8 in HIT⁺ patients and 5 in HIT⁻ patients ($P < .0001$) (Figure 3B). At a cutoff of ≥ 3 , the HEP score was 95.3% (95% CI, 84.2%-99.4%) sensitive and 35.7% (95% CI, 29.8%-42.0%) specific with a PPV of 20.4% (95% CI, 15.1%-26.6%) and NPV of 97.9% (95% CI, 92.3%-99.7%) (Table 2).

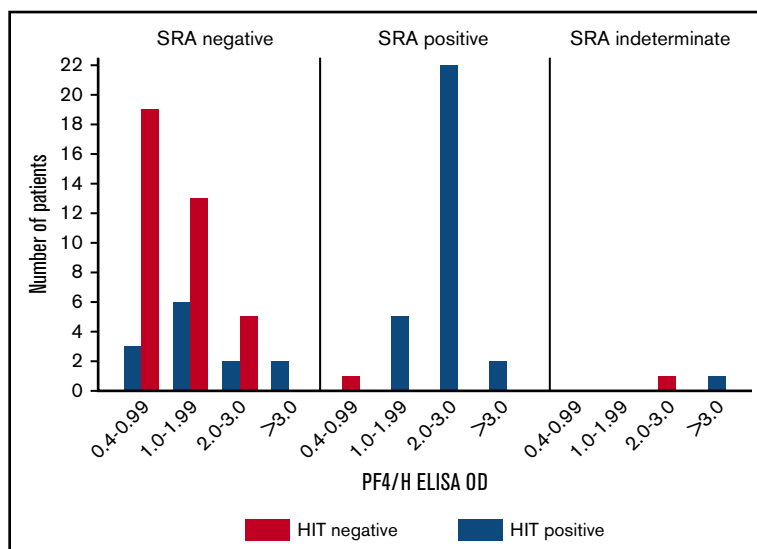
There was no difference between the HEP score and 4Ts scores at their respective cutoffs with respect to sensitivity (95.3% vs 97.7%; $P = 1.0$) or specificity (35.7% vs 32.9%; $P = .34$). The areas under the ROC curves for the HEP score and 4Ts score were also similar (0.81 [95% CI, 0.74-0.87] vs 0.76 [95% CI, 0.69-0.83]; $P = .12$) (Figure 4).

To test the robustness of the adjudication process, we conducted a sensitivity analysis in which SRA result, rather than adjudicated HIT status, was used as the reference standard. In this analysis, the areas under the ROC curves were similar to the results of the primary analysis for the HEP score (0.81; 95% CI, 0.73-0.90) and 4Ts score (0.77; 95% CI, 0.69-0.84). Using SRA positivity as the reference standard, the sensitivity of the HEP score at a cutoff ≥ 3 was similar to the 4Ts score at a cutoff ≥ 4 (93.3% [95% CI, 77.9%-99.2%] vs 100% [88.4%-100%]; $P = .5$). The specificities of the HEP score and 4Ts score were also similar at these thresholds using SRA positivity as the reference standard (33.8% [95% CI, 28.1%-39.9%] vs 31.5% [95% CI, 25.9%-37.6%], $P = .41$).

Subgroup analyses by patient population

We assessed the operating characteristics of the HEP score and 4Ts score in different patient populations (Table 2). The two scoring systems performed similarly in medical, surgical, and cardiothoracic surgical patients. The HEP score exhibited a significantly higher AUC than the 4Ts score in ICU patients (0.87 vs 0.79, $P = .03$).

Figure 2. HIT ELISA results according to SRA result and adjudicated HIT status. PF4/H, platelet factor 4/heparin.



At a cutoff of ≥ 3 for the HEP score and ≥ 4 for the 4Ts score, sensitivity was similar in this population (100% [95% CI, 84.6-100] vs 100% [95% CI, 81.5-100], $P = 1.0$). However, the HEP score exhibited significantly greater specificity (40.5% [95% CI, 32.6-42.8]

vs 31.1% [95% CI, 23.7-39.2], $P = .02$). At these cutoff values, the HEP score correctly classified 82/170 (48.2%) ICU patients whereas the 4Ts score correctly classified 68 of 170 of subjects (40.0%).

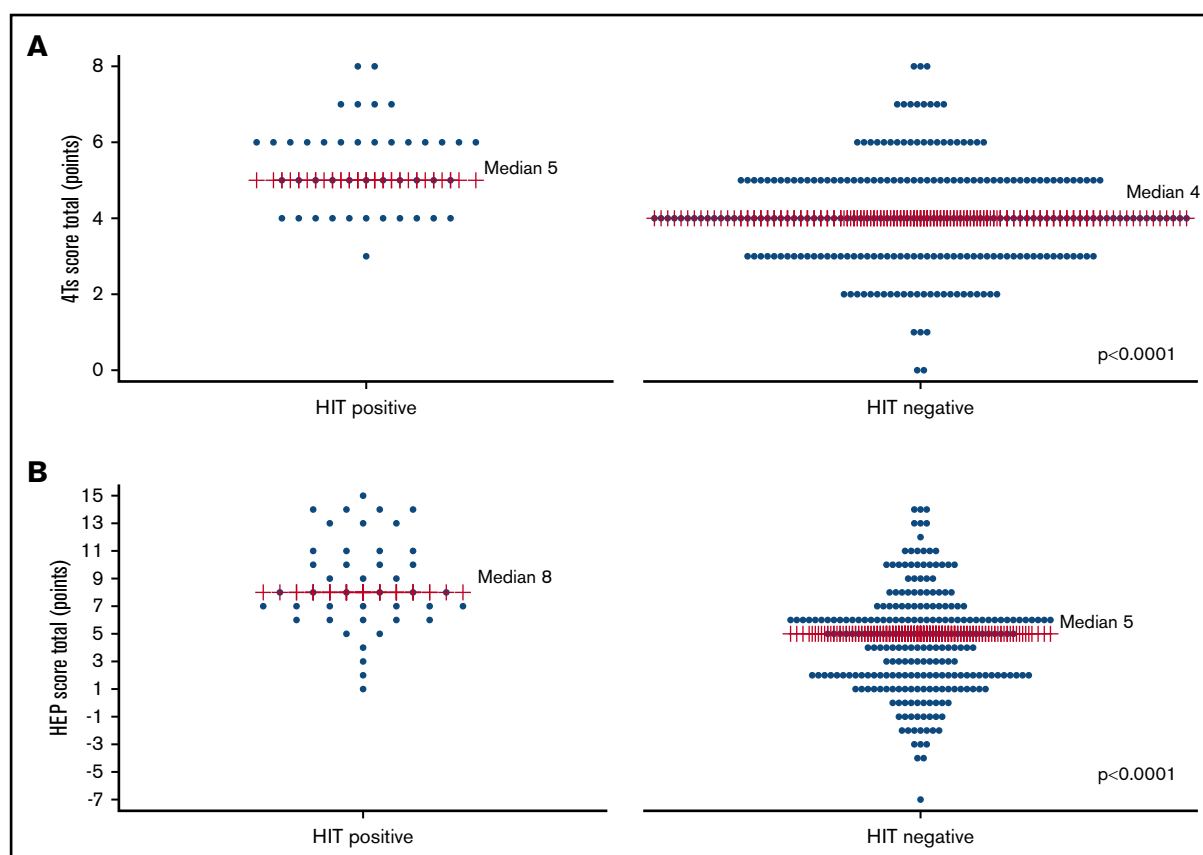


Figure 3. Plot of HEP score and 4Ts score by adjudicated HIT status. The distribution of 4Ts scores (A) and HEP scores (B) according to adjudicated HIT status are shown. Median scores are labeled and denoted with red lines. The median score was greater in HIT^+ patients than in HIT^- patients for both the HEP score and 4Ts score ($P < .0001$).

Table 2. Operating characteristics of the HEP score and 4Ts score by patient population

Model	AUC (95% CI)	P*	Cutoff	Sensitivity, % (95% CI)†	Specificity, % (95% CI)†	PPV, % (95% CI)†	NPV, % (95% CI)†
Total population, n = 292: HIT prevalence, 14.7%							
HEP score	0.81 (0.74-0.87)	.12	3	95.3 (84.2-99.4)	35.7 (29.8-42.0)	20.4 (15.1-26.6)	97.8 (92.3-99.7)
4Ts score	0.76 (0.67-0.84)		4	97.7 (86.2-99.8)	32.9 (27.2-39.1)	20.1 (15.0-26.3)	98.8 (92.5-99.9)
Medical patients, n = 129: HIT prevalence, 12%							
HEP score	0.81 (0.73-0.90)	.42	3	100 (79.4-100)	28.3 (20.2-37.6)	16.5 (9.73-25.4)	100 (89.1-100)
4Ts score	0.78 (0.67-0.88)		4	100 (79.4-100)	27.4 (19.5-36.6)	16.3 (9.6-25.2)	100 (88.0-100)
Surgical patients, n = 163: HIT prevalence, 17%							
HEP score	0.82 (0.72-0.91)	.13	3	92.6 (75.7-99.1)	41.9 (33.5-50.7)	24 (16.2-33.4)	96.6 (88.3-99.6)
4Ts score	0.76 (0.67-0.84)		4	96.3 (81.0-99.9)	37.5 (29.4-46.2)	23.4 (15.9-32.4)	98.1 (89.7-100)
Cardiothoracic surgery patients, n = 125: HIT prevalence, 13%							
HEP score	0.74 (0.61-0.87)	.43	3	87.3 (61.7-98.4)	44 (34.5-53.9)	18.7 (10.6-29.3)	96.0 (86.3-99.5)
4Ts score	0.69 (0.57-0.80)		4	93.8 (69.8-99.8)	39.4 (30.2-49.3)	18.5 (10.8-28.7)	97.3 (88.0-99.0)
ICU patients, n = 170: HIT prevalence, 13%							
HEP score	0.86 (0.80-0.93)	.03	3	100 (84.6-100)	40.5 (32.6-48.9)	20 (13.0-28.7)	100 (94.0-100)
4Ts score	0.79 (0.71-0.88)		4	100 (81.5-100)	31.1 (23.8-39.2)	17.7 (11.7-25.8)	100 (92.3-100)

*P value represents comparison of the AUC using the method of DeLong et al.¹⁷

†Exact binomial CIs displayed for sensitivity, specificity, PPV, and NPV.

Subgroup analyses by scorer experience

The operating characteristics of the HEP score and 4Ts score by scorer experience are displayed in Table 3. When attending hematologists served as scorers, the sensitivity and specificity of the two scoring systems was similar. There was no significant difference in the area under the ROC curve for the HEP score and 4Ts score (0.85 vs 0.84; $P = .86$). When trainees served as scorers, the area under the ROC curve was significantly greater for the HEP score than the 4Ts score (0.80 vs 0.73; $P = .03$). However, the sensitivity and specificity of the 4Ts score and HEP score were not significantly different ([96.9% vs 93.9%; $P = 1.0$] and [26.4% vs 32.7%; $P = .17$], respectively). In subjects scored by trainees, the HEP score correctly classified 82/189 (43.4%) while the 4Ts score correctly classified 74 of 189 patients (39.2%).

Discussion

In this prospective study, the HEP score exhibited similar diagnostic accuracy to the 4Ts score in an overall cohort of patients with suspected HIT. The HEP score demonstrated superior operating characteristics compared with the 4Ts score in ICU patients and among trainee scorers.

Prior evaluations of the HEP score were limited by their retrospective study design, small number of HIT⁺ patients, performance of scoring by study personnel rather than clinicians directly involved in the patient's care, and/or nonindependent adjudication of HIT status.^{7,12,13} Moreover, scorers in some studies had access to HIT laboratory test results at the time of scoring, which could have biased their scores.

We aimed to overcome these limitations by designing our study to mirror use of HIT scoring systems in real-world clinical practice.

Scoring was conducted by clinicians directly involved in the patient's care rather than by study personnel. Completion of scoring was required prior to availability of HIT laboratory test results. In addition, HIT status was determined by an independent adjudication panel of 3 HIT experts, who had no role in scoring or care of the patient.

Because the HEP score and 4Ts score demonstrated similar diagnostic accuracy, our findings suggest that use of either scoring system may be justified in clinical practice. Substantial overlap in scores between HIT⁺ and HIT⁻ subjects was observed with both scoring systems (Figure 3), highlighting the limitations of each and the challenge of clinical diagnosis in general. An advantage of the

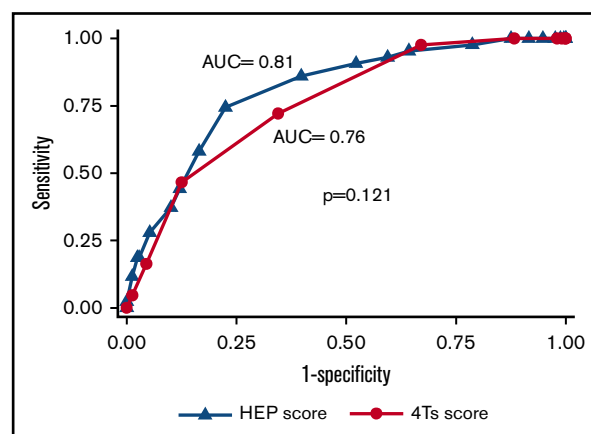
**Figure 4. ROC curves for the HEP score and 4Ts score.**

Table 3. Operating characteristics of the HEP score and 4Ts score by scorer experience

Model	AUC (95% CI)	P*	Cutoff	Sensitivity, % (95% CI)†	Specificity, % (95% CI)†	PPV, % (95% CI)†	NPV, % (95% CI)†
Attending physicians, 103 cases: HIT prevalence, 9.8%							
HEP score	0.85 (0.74-0.96)	.85	3	100 (69.2-100)	40.9 (30.8-51.5)	15.4 (7.6-26.5)	100 (90.7-100)
4Ts	0.84 (0.74-0.94)		4	100 (69.2-100)	43.5 (32.8-53.7)	15.9 (7.9-27.3)	100 (91.2-100)
Trainees, 189 cases: HIT prevalence, 18%							
HEP score	0.80 (0.72-0.88)	.02	3	93.9 (79.8-99.3)	32.7 (25.4-40.7)	22.8 (16.0-30.8)	96.2 (87-99.5)
4Ts	0.72 (0.63-0.81)		4	97.0 (84.2-99.9)	26.9 (20.1-34.6)	21.9 (15.5-29.5)	97.7 (87.7-99.9)

*P value represents comparison of AUC using the method of DeLong et al.¹⁷

†Exact binomial CIs displayed for sensitivity, specificity, PPV, and NPV.

4Ts score compared with the HEP score is that it is simpler and easier to calculate, though the HEP score did not prove to be unduly burdensome in our study with a median time to completion of 4 minutes.

The greater complexity of the HEP score may be advantageous in certain situations. For example, the HEP score showed enhanced operating characteristics compared with the 4Ts score among trainee scorers. Scorers with less experience may benefit from the explicit itemization of alternate causes of thrombocytopenia in the HEP score rather than the more subjective determination of likelihood of other causes of thrombocytopenia required by the 4Ts score (supplemental Table 1). Critically ill patients are complex and frequently have alternative explanations for thrombocytopenia such as infection, medications other than heparin, intravascular devices, and disseminated intravascular coagulation.^{10,11} The HEP score specifically prompts the rater to consider these possibilities, offering a potential explanation for the improved performance of the HEP score compared with the 4Ts score among medical and surgical ICU patients.

An important ancillary observation that emerged from our study is how difficult it can be to diagnose HIT, even among experts under idealized conditions. Despite having access to detailed clinical information including 30-day follow-up as well as the results of state-of-the-art HIT laboratory testing, the 3 expert adjudicators agreed on “definite HIT” or “definitely not HIT” in only 36 of the 83 patients who underwent adjudication. In the other 47 patients (56.6%), at least 1 of the adjudicators acknowledged uncertainty about the diagnosis, though a consensus diagnosis was achieved after discussion among the three adjudicators in almost all cases.

Our study has several important limitations. First, patients were enrolled from a single health system, which may limit the generalizability of our findings. However, we recruited patients from both an academic medical center and an affiliated community hospital and both the 4Ts score and HEP score performed similarly in these 2 settings (data not shown). Second, scoring was carried out almost exclusively by hematology fellows and attending hematologists. Although this reflects practice in our health system, our findings may not be applicable to institutions where initial clinical assessment for HIT is conducted primarily by nonhematologists (eg, hospitalists, intensivists, surgeons). Future studies are needed to evaluate and compare the performance of the HEP score and 4Ts score among nonhematologist clinicians. Third, HIT status could

have been misclassified in some subjects. Because there is no universally accepted gold standard for HIT, we used a rigorous reference standard involving an independent panel of 3 expert adjudicators, who based their diagnosis on detailed clinical information, HIT laboratory test results, and 30-day follow-up. A sensitivity analysis in which the SRA was used as the reference standard yielded similar results to those of the primary analysis, confirming the robustness of our reference standard. Fourth, our study was observational. An interventional study will be required to evaluate the impact of the HEP score as a clinical prediction rule on clinical outcomes.

In summary, our findings support the use of either the 4Ts score or HEP score in patients with suspected HIT. Our data suggest that the HEP score may be preferred among inexperienced clinicians and in patients with critical illness, although these subgroup findings should be considered preliminary and require independent confirmation in future studies. Implementation studies are needed to assess the feasibility and acceptability of the HEP score as well as its impact on clinical practice.

Acknowledgments

This work was supported by the National Institutes of Health, National Heart, Lung, and Blood Institute (T32-HL007971-16A1 [A.M.P.] and K23-HL112903 [A.C.]). A.M.P. was also supported by a 2018 Hemostasis and Thrombosis Research Society (HTRS)/Novo Nordisk Clinical Fellowship Award in Hemophilia and Rare Bleeding Disorders from the HTRS, which was supported by an educational grant from Novo Nordisk Inc.

Authorship

Contribution: A.C. and D.B.C. designed the research; A.M.P., A.C., D.S.L., K.P., S.F., and R.V. collected the data; G.M.A., M.C., and L.R. adjudicated clinical case summaries; A.M.P. analyzed results; A.M.P. and A.C. wrote the manuscript; and D.S.L., K.P., S.F., R.V., G.M.A., M.C., L.R., and D.B.C. critically reviewed and edited the manuscript.

Conflict-of-interest disclosure: A.M.P. has received research support from Novo Nordisk. L.R. has served on an advisory board or speaker's bureau for Pfizer and Novartis; has served on a data safety monitoring committee for Apellis; and has received research support from Baxalta and Alexion. A.C. has served as a consultant or advisory board member for Bioverativ, Genzyme, Kedrion, Stago, and Synergy and his institution has received research support on

his behalf from Alexion, Bayer, Bioverativ, Novo Nordisk, Pfizer, Shire, and Spark. D.B.C. has served as a consultant or advisory board member for Amgen, Rigol, Novartis, Ionis, Sanofi, and Bayer and has received research support from T2Biosystems, Syntimmune, and Momenta. M.C. has served on a data safety monitoring board for Bayer and Daiichi Sankyo; has served as a consultant or advisory board member for Shionogi, Octapharma, Bristol-Myers Squibb Canada, Pfizer, Alexion, and CSL Behring; and has received research funding from Bayer, Leo Pharma, and the Heart

and Stroke Foundation of Canada. The remaining authors declare no competing financial interests.

ORCID profiles: A.M.P., 0000-0001-9997-454X; G.M.A., 0000-0003-0496-5064; M.C., 0000-0003-4986-4873; D.B.C., 0000-0001-5986-504X; A.C., 0000-0002-3595-5697.

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

References

- 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.
- Warkentin TE, Kelton JG. A 14-year study of heparin-induced thrombocytopenia. *Am J Med*. 1996;101(5):502-507.
- Wallis DE, Workman DL, Lewis BE, Steen L, Pifarre R, Moran JF. Failure of early heparin cessation as treatment for heparin-induced thrombocytopenia. *Am J Med*. 1999;106(6):629-635.
- McMahon CM, Tanhehco YC, Cuker A. Inappropriate documentation of heparin allergy in the medical record because of misdiagnosis of heparin-induced thrombocytopenia: frequency and consequences. *J Thromb Haemost*. 2017;15(2):370-374.
- 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.
- 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.
- 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.
- Cuker A, Gimotty PA, Crowther MA, Warkentin TE. Predictive value of the 4Ts scoring system for heparin-induced thrombocytopenia: a systematic review and meta-analysis. *Blood*. 2012;120(20):4160-4167.
- Nagler M, Fabbro T, Willemin 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.
- Fiorenza MA, Frazee EN, Personett HA, Dierkhising RA, Schramm GE. Assessment of a modified 4T scoring system for heparin-induced thrombocytopenia in critically ill patients. *J Crit Care*. 2014;29(3):426-431.
- Crowther M, Cook D, Guyatt G, et al; PROTECT collaborators; Canadian Critical Care Trials Group; Australian and New Zealand Intensive Care Society Clinical Trials Group. Heparin-induced thrombocytopenia in the critically ill: interpreting the 4Ts test in a randomized trial. *J Crit Care*. 2014;29(3):470.e7-470.e15.
- Uaprasert N, Chanswangphuwana C, Akkawat B, Rojnuckarin P. Comparison of diagnostic performance of the heparin-induced thrombocytopenia expert probability and the 4Ts score in screening for heparin-induced thrombocytopenia. *Blood Coagul Fibrinolysis*. 2013;24(3):261-268.
- Joseph L, Gomes MP, Al Solaiman F, et al. External validation of the HIT Expert Probability (HEP) score. *Thromb Haemost*. 2015;113(3):633-640.
- Cines DB, Kaywin P, Bina M, Tomaski A, Schreiber AD. Heparin-associated thrombocytopenia. *N Engl J Med*. 1980;303(14):788-795.
- Sheridan D, Carter C, Kelton JG. A diagnostic test for heparin-induced thrombocytopenia. *Blood*. 1986;67(1):27-30.
- Husseinadeh HD, Gimotty PA, Pishko AM, Buckley M, Warkentin TE, Cuker A. Diagnostic accuracy of IgG-specific versus polyspecific enzyme-linked immunoassays in heparin-induced thrombocytopenia: a systematic review and meta-analysis. *J Thromb Haemost*. 2017;15(6):1203-1212.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837-845.
- Lachenbruch PA, Lynch CJ. Assessing screening tests: extensions of McNemar's test. *Stat Med*. 1998;17(19):2207-2217.