Diagnosing Testicular Torsion before Urological Consultation and Imaging: Validation of the TWIST Score

Kunj R. Sheth, Melise Keays, Gwen M. Grimsby, Candace F. Granberg, Vani S. Menon, Daniel G. DaJusta, Lauren Ostrov, Martinez Hill, Emma Sanchez, David Kuppermann, Clanton B. Harrison, Micah A. Jacobs, Rong Huang, Berk Burgu, Halim Hennes, Bruce J. Schlomer and Linda A. Baker*,†

From the University of Texas Southwestern Medical Center (KRS) and Children's Health (LO, MH, ES, RH), Dallas, Texas (VSM, CBH, MAJ, HH, BJS, LAB), Children's Hospital of East Ontario, Ottawa, Ontario, Canada (MK), Phoenix Children's Hospital, Phoenix, Arizona (GMG), Mayo Clinic, Rochester, Minnesota (CFG), Nationwide Children's Hospital, Columbus, Ohio (BB), Harvard Medical School, Boston, Massachusetts (DGD), and Ankara Üniversitesi Tıp Fakültesi, Ankara, Turkey (DK)

Abbreviations and Acronyms

EMT = emergency medical technician

ER = emergency room

NIH = National Institutes of Health

NIRS = near-infrared spectroscopy

NPV = negative predictive value

PPV = positive predictive value

TWIST = Testicular Workup for Ischemia and Suspected Torsion

Accepted for publication January 15, 2016.

* Correspondence: Center for Pediatric Urology, Children's Health, 2350 Stemmons Fwy., 4th Floor, F4.04, Dallas, Texas 75207 (telephone: 214-456-2480; FAX: 214-456-8803); e-mail: linda.baker@childrens.com).

† Supported by NIH Grant R21DK092654.

See Editorial on page 1650.

Purpose: The TWIST (Testicular Workup for Ischemia and Suspected Torsion) score uses urological history and physical examination to assess risk of testis torsion. Parameters include testis swelling (2 points), hard testis (2), absent cremasteric reflex (1), nausea/vomiting (1) and high riding testis (1). While TWIST has been validated when scored by urologists, its diagnostic accuracy among nonurological providers is unknown. We assessed the usefulness of the TWIST score when determined by nonurological nonphysician providers, mirroring emergency room evaluation of acute scrotal pain.

Materials and Methods: Children with unilateral acute scrotum were prospectively enrolled in a National Institutes of Health clinical trial. After undergoing basic history and physical examination training, emergency medical technicians calculated TWIST score and determined Tanner stage per pictorial diagram. Clinical torsion was confirmed by surgical exploration. All data were captured into REDCap™ and ROC curves were used to evaluate the diagnostic usefulness of TWIST.

Results: Of 128 patients (mean age 11.3 years) 44 (13.0 years) had torsion. TWIST score cutoff values of 0 and 6 derived from ROC analysis identified 31 high, 57 intermediate and 40 low risk cases (positive predictive value 93.5%, negative predictive value 100%).

Conclusions: TWIST score assessed by nonurologists, such as emergency medical technicians, is accurate. Low risk patients do not require ultrasound to rule out torsion. High risk patients can proceed directly to surgery, with more than 50% avoiding ultrasound. In the future emergency medical technicians and/or emergency room triage personnel may be able to calculate TWIST score to guide radiological evaluation and immediate surgical intervention at initial assessment long before urological consultation.

Key Words: diagnosis, scrotum, spermatic cord torsion, ultrasonography

TESTICULAR torsion is among the few pediatric urological emergencies, accounting for 10% to 15% of acute scrotum presentations,[1,2] with an annual incidence of 3.8 per 100,000 pediatric patients.[3] Intervention within 4 to 8 hours is critical to prevent permanent testicular loss or atrophy from compromised testicular arterial flow.[4,5] Thereafter, the testicle often is

0022-5347/16/1956-1870/0
THE JOURNAL OF UROLOGY®

http://dx.doi.org/10.1016/j.juro.2016.01.101
Vol. 195, 1870-1876, June 2016
Printed in U.S.A.

unsalvageable, resulting in orchiectomy rates of 32% to 41%.[3,6] While delayed emergent presentation cannot be controlled, prompt and accurate diagnosis on arrival is essential to identify patients requiring surgical detorsion.

The classic presentation for testicular torsion in pubertal males is acute onset unilateral testicular pain with nausea and vomiting. An absent cremasteric reflex has been considered specific for testicular torsion,[7] although there are reports of torsion with cremasteric reflex present.[8] Testicular swelling, tenderness and high lie are nonspecific, often making diagnosis based solely on physical examination difficult.[9] Therefore, testicular ultrasound with Doppler is heavily relied on for confirmation.[10,11] Since awaiting imaging constitutes a time delay, risk scoring systems based on signs and symptoms only, such as TWIST, have been proposed.

Barbosa et al devised and validated the TWIST scoring system, which assigns a summed score (range 0 to 7 points) based on the absence (0) or presence of 5 variables, including testicular swelling (2), hard testicle (2), absent cremasteric reflex (1), nausea/vomiting (1) and high riding testis (1).[12] Their ROC analysis yielded all binary variables. Thus, for the categories of testicular swelling and hard testis patients could receive a score of only 0 or 2 points (no option for 1 point). Per their initial validation patients at high risk for torsion (TWIST score 5 or higher) could proceed straight to surgery without imaging as the PPV was 100%. Patients with intermediate risk (TWIST score 3 to 4) required ultrasound to evaluate for torsion, and patients at low risk for torsion (2 or less) did not require scrotal ultrasound as the NPV was 100%. However, TWIST score does not account well for physiological differences in children, control for interobserver variability or substitute for medicolegal need for ultrasound documentation. While data were collected by urologists in the study by Barbosa et al,[12] in practice these data would be obtained by an ER provider before urological consultation. We evaluated the usefulness of the TWIST score when measured by trained nonphysician, nonurological personnel, specifically EMTs, who are often the first medical providers to encounter patients in an emergency setting.

MATERIALS AND METHODS

Study Population

The study population was drawn from a NIH funded prospective series evaluating use of NIRS in the diagnosis of torsion.[13] We prospectively evaluated males 1 month to 21 years old presenting at a tertiary care pediatric ER with acute scrotum, defined as painful scrotum or testis, abdominal pain and/or waddling gait ("cowboy shuffle") from painful scrotum, between March 2013 and March 2015. Patients with synchronous bilateral testicular torsion or previously known testicular or scrotal pathology were excluded according to study protocol due to inability to use the contralateral testis as an internal control for NIRS measurements. We also excluded patients with chronic respiratory, hematological or vascular problems that could affect total body tissue oxygenation levels and thus NIRS measurements.

Study Design

Per NIH study protocol ER staff were instructed to page the on call EMT research personnel on arrival of any patient with acute scrotum. Potentially eligible study subjects were approached by the on call EMT and screened for inclusion and exclusion criteria. Informed consent and required study data were collected while patients were receiving care and undergoing ultrasound and/or surgery, avoiding any delay in care. Scrotal ultrasound was used as the gold standard for diagnosis of testicular torsion and was intended to be performed for all patients included in the study, unless the clinical suspicion of the physician was high enough to forgo ultrasound and proceed to surgery for emergent detorsion. Patients with no evidence of torsion on ultrasound were given a urology followup appointment in 2 weeks and ER warning signs.

For all enrolled study patients the research EMTs performed an evaluation, assigning binary components of the TWIST score and Tanner stage using a descriptive and pictorial table (fig. 1). The EMT received no training specific to scrotal examination or TWIST scoring, but rather relied on basic history and physical examination training learned during EMT certification. Patients without complete TWIST score components were excluded from the study.

Data Analysis

All collected data were entered into a REDCap database (Clinical and Translational Science Awards NIH Grant UL1TR001105)[14] and extracted as needed for analysis. Fisher exact, Wilcoxon rank sum and t-tests were used for comparisons. ROC curves were constructed to analyze and compare performance of the TWIST score as a diagnostic test for torsion.[15] Optimal cutoff values for low, intermediate and high risk groups were chosen to maximize performance of the test, taking into account clinically meaningful results to optimize NPV and PPV while limiting false-negative and false-positive results. All statistical analysis was performed using Stata®, version 12.

RESULTS

Of 316 patients assessed for eligibility 115 did not meet study criteria and 47 declined to participate. One patient was taken straight to surgery for high suspicion of torsion without ultrasound. Two patients without torsion did not undergo ultrasound, with imaging being refused by 1 and deferred in 1 due to low clinical suspicion. Both patients underwent urological followup and demonstrated no

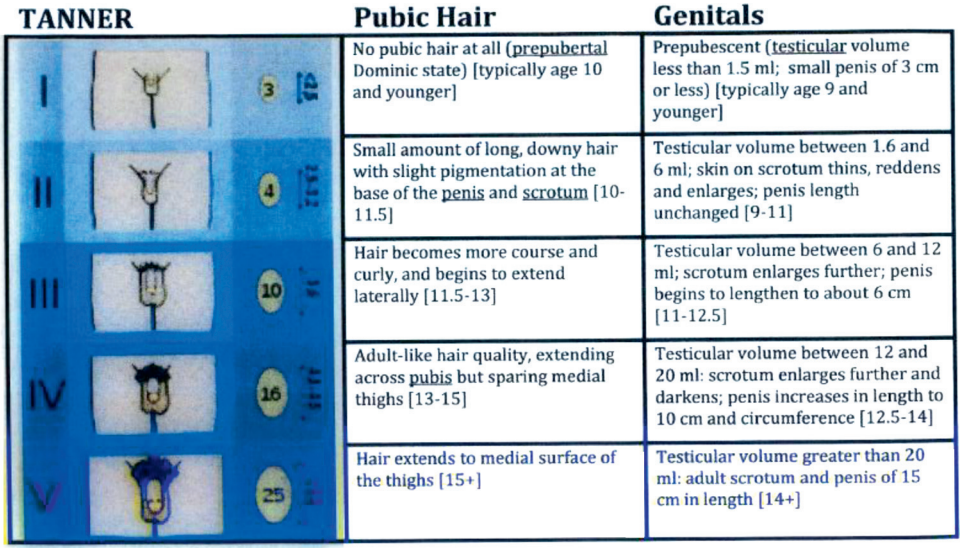

Figure 1. Chart used by EMTs to identify patient Tanner stage at evaluation

evidence of torsion. Of 154 enrolled patients 128 had all TWIST data available (fig. 2), of whom 44 (34.4%) were diagnosed with torsion that was surgically confirmed. Of patients not diagnosed with torsion in the ER none presented with missed torsion, although only 45% returned for clinical followup. Demographic characteristics of patients with and without torsion are outlined in table 1. Patients with torsion were older, were more likely to be white or black, and had higher Tanner stage and shorter duration of pain before arrival.

TWIST score components and total score distribution are outlined in table 2. Median TWIST score was 6 in patients with torsion and 1 in those without torsion ($p < 0.001$). ROC curve using all points in the TWIST score had an AUC of

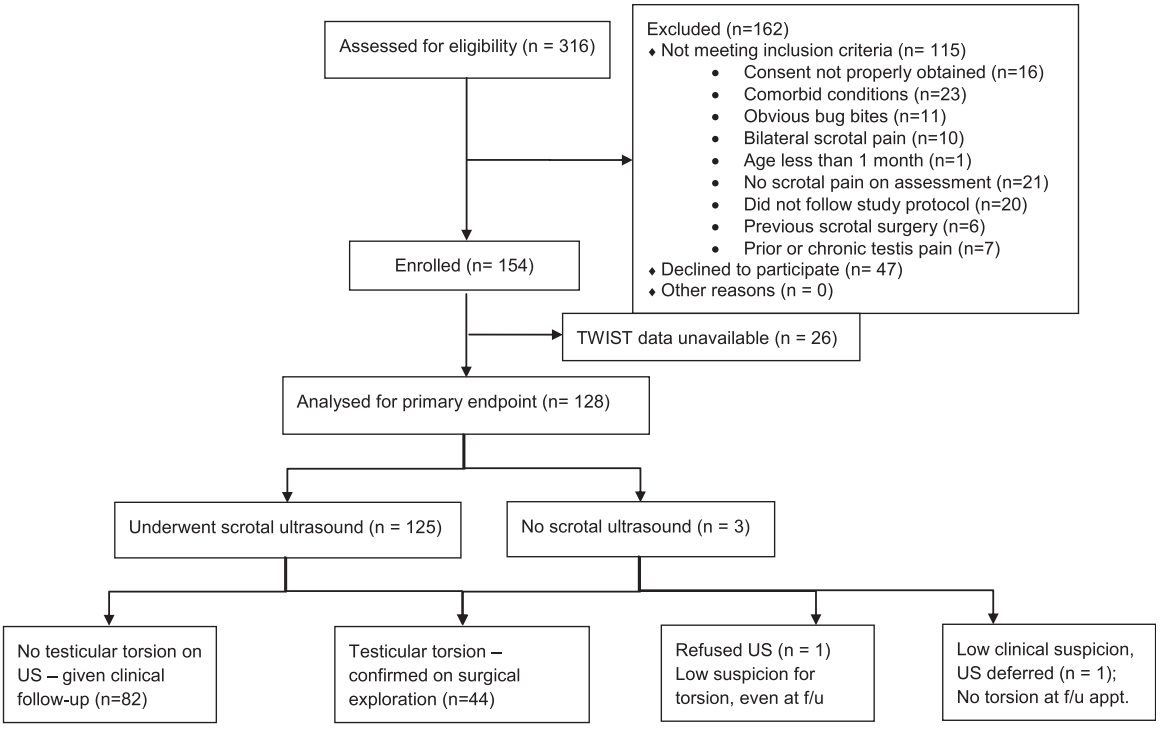

Figure 2. Study cohort diagram. *US,* ultrasound. *f/u,* followup.

Table 1. *Patient characteristics*

	Torsion	No Torsion	p Value
Mean ± SD age (yrs)	13.0 ± 4.0	10.4 ± 4.4	0.001 (2-tailed t-test)
No. race/total No. (%):			
White	11/44 (25.0)	16/84 (19.1)	0.003 (Fisher exact test)
Hispanic	18/44 (40.9)	58/84 (69.1)	
Black	14/44 (31.8)	9/84 (10.7)	
Asian	1/44 (2.3)	0/84 (0)	
Other	0/44 (0)	1/84 (1.2)	
No. Tanner stage/total No. (%):*			
1 (mean age 6.3 yrs)	6/42 (14.3)	28/84 (33.3)	<0.001 (Wilcoxon rank sum test)
2 (mean age 10.8 yrs)	4/42 (9.5)	31/84 (36.9)	
3 (mean age 13.7 yrs)	12/42 (28.6)	11/84 (13.1)	
4 (mean age 15.3 yrs)	17/42 (40.5)	10/84 (11.9)	
5 (mean age 15.0 yrs)	3/42 (7.1)	4/84 (4.8)	
Median hrs pain before arrival (range)	17.3 (0.6–129.3)	29.2 (0.9–346.1)	0.02 (Wilcoxon rank sum test)

*Tanner stage was missing in 2 patients with torsion.

0.95 (95% CI 0.91–0.98, fig. 3, *A*). Clinically meaningful TWIST score cutoff values of 0 and 6 were used to categorize patients into low risk (0 points), intermediate risk (1 to 5) and high risk (6 to 7) with an optimized AUC of 0.90 (95% CI 0.85–0.94, fig. 3, *B*). There were no patients with a TWIST score of 0 who had torsion, giving a NPV of 100% and specificity of 47.6%. Of 31 patients with a TWIST score of 6 or greater 29 had torsion, for a PPV of 93.5% and a sensitivity of 65.9%. The 2 patients with torsion with a TWIST score of 1 for nausea/vomiting underwent manual detorsion elsewhere, and operative room findings remained consistent with diagnosis.

For Tanner stage 3 to 5 patients a high risk TWIST score had a PPV of 100% and sensitivity of 65.6% (fig. 4, *A*). In contrast, for Tanner stage 1 to 2 patients a high risk TWIST score had a PPV of 77.8% and sensitivity of 70.0% (fig. 4, *B*). The 2 patients with a high risk TWIST score without testicular torsion were Tanner stages 1 and 2 (1 each) with torsion of the appendix testis.

Table 2. *TWIST score results*

	Torsion	No torsion
No. TWIST score/total No. (%):		
0	0/44 (0)	40/84 (47.6)
1	3/44 (6.8)	16/84 (19.1)
2	0/44 (0)	9/84 (10.7)
3	2/44 (4.6)	9/84 (10.7)
4	2/44 (4.6)	4/84 (4.8)
5	8/44 (18.2)	4/84 (4.8)
6	14/44 (31.8)	2/84 (2.4)
7	15/44 (34.1)	0/84 (0)
No. TWIST score risk category/total No. (%):		
Low (0 points)	0/44 (0)	40/84 (47.6)
Intermediate (1–5 points)	15/44 (34.1)	42/84 (50.0)
High (6–7 points)	29/44 (65.9)	2/84 (2.4)

p <0.001 (Wilcoxon rank sum test) for patients with vs without torsion. There was no significant difference in breakdown of TWIST score components for patients with torsion when substratified by Tanner stage.

DISCUSSION

We assessed TWIST score obtained by trained nonphysician, nonurological providers. In our population the TWIST score performed well as a diagnostic test for torsion, although not as well as previously reported and with different optimal cutoff values.[12] Based on our results, we devised an algorithm to evaluate patients presenting emergently with acute scrotum (fig. 5).

In our population no torsions were missed on followup, and all patients undergoing surgery had torsion or testicular ischemia in the absence of torsion, indicating recent detorsion (spontaneous, manual at ER or with anesthesia). In current practice ultrasound is increasingly used to guide diagnosis of testicular torsion, with reported 100% sensitivity, 97.9% specificity and 98.1% diagnostic accuracy.[2,10,11] Thus, ultrasonography served as the gold standard in our study population although constituting a 30 to 60-minute delay in diagnosis.

There is a growing effort to return to traditional history and physical examination findings to diagnose torsion, decreasing reliance on imaging, minimizing cost and facilitating rapid surgical intervention.[16–19] The TWIST score is easy to calculate with a simple patient evaluation. In our study 29 of 44 testicular torsions (65.9%) were detectable by high risk TWIST score (6 to 7 points) and only 2 of 31 in the high risk group without testicular torsion (6.5%) would undergo a negative/unnecessary surgical exploration. Per our ROC analysis the high risk score cutoff was 6 rather than 5, which has been reported previously.[12] In our population 4 of 12 patients with a TWIST score of 5 (33.3%) did not have testicular torsion, which would lead to an unacceptably high negative exploration rate. Alternatively a high risk cutoff of 7 points would yield a 100% PPV but decrease sensitivity to 34.1%, leading to an optimal cutoff of 6. In addition, our low risk group cutoff score of 0 was different than previously published. Of 128 children 40 (31.3%) comprised the low risk group with no cases of testicular torsion (100% NPV).

If the goal of a low risk TWIST score is to avoid use of ultrasound, then a NPV of 100% is necessary as a missed torsion is unacceptable. A scrotal ultrasound could be avoided in all low (31.3%) and high risk patients (24.2%), comprising more than 50% of patients. Barbosa et al found ultrasound unnecessary in approximately 80% of patients.[12]

Torsion has a bimodal age distribution with the first peak in the neonatal period and the second peak around puberty.[20] While postpubertal children

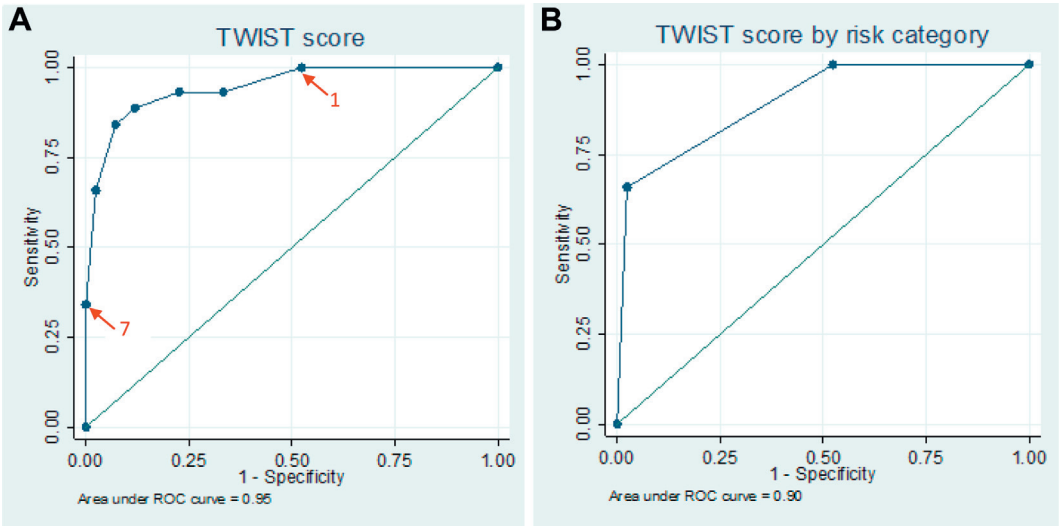

Figure 3. *A*, ROC curve for TWIST score shows AUC of 0.95 (95% CI 0.91–0.98). *B*, ROC curve for TWIST score risk categories reveals AUC of 0.90 (95% CI 0.85–0.94).

usually present with severe testicular pain, identifying typical torsion symptoms and performing sonography to appropriately diagnose torsion are more challenging in prepubescent children.[6,21] In our study the 2 patients without torsion who were in the high risk group were Tanner stages 1 and 2 (1 each) with a diagnosis of torsion of the appendix testis that led to a high TWIST score. In Tanner stage 3 to 5 patients the high risk TWIST score had a PPV of 100%, signifying that the TWIST score performs better for peripubertal or postpubertal children. TWIST also performs well in Tanner 1 to 2 patients, although some of these children with torsion of the appendix testis will be categorized as being at high risk for torsion. Therefore, ultrasound should be considered for these patients to help confirm diagnosis.

One key difference and an asset in this validation of the TWIST score that may explain the differences in group stratification is that all of the involved components were collected by nonphysician, nonurological personnel. At presentation a urologist is typically not immediately available in the ER, and the initial evaluation and decision regarding whether to obtain ultrasound are often rendered by nonurologists. Thus, we suggest that the cutoff values seen in our study are more valid for use in the ER. Implementation and evaluation

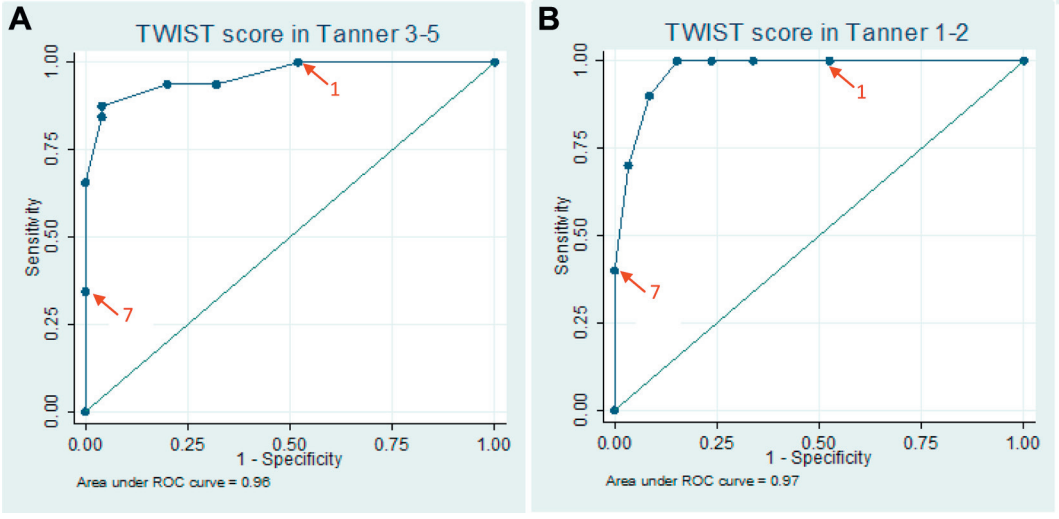

Figure 4. *A*, ROC curve for TWIST score in Tanner 2 to 5 patients demonstrates AUC of 0.95 (95% CI 0.91–0.99). *B*, ROC curve for TWIST score in Tanner 1 patients shows AUC of 0.96 (95% CI 0.91–1.00). Risk category cutoff points (0 and 6) are circled in both ROC curves.

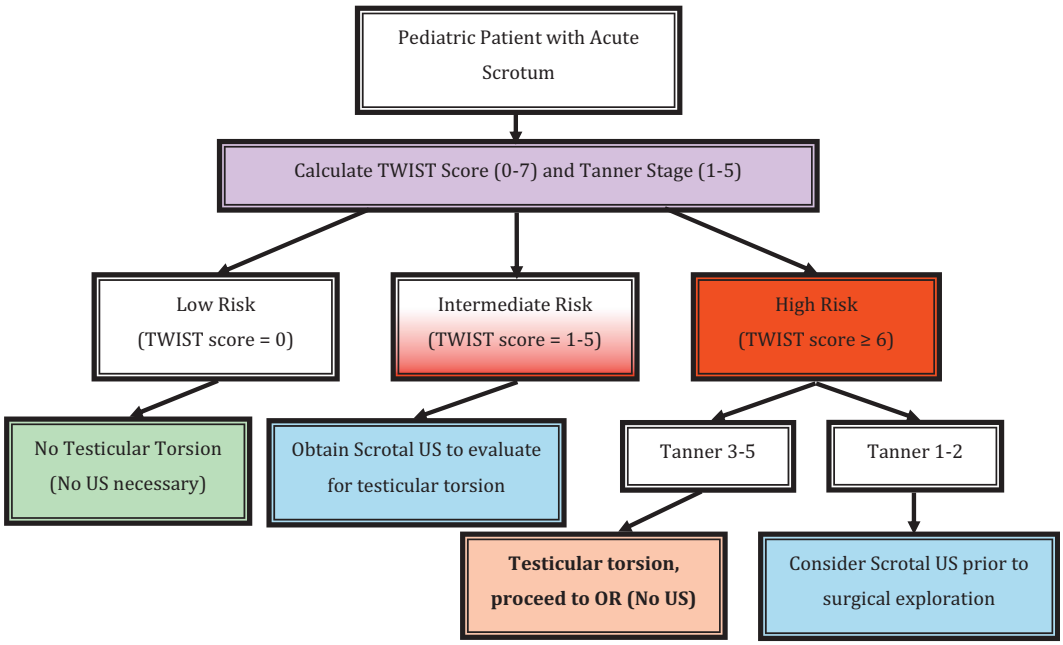

Figure 5. Proposed diagnostic algorithm for testicular torsion in patients with acute scrotum. *US*, ultrasound. *OR*, operating room.

of the TWIST score by ER providers and triage nurses following our diagnostic algorithm will be the next step to potentially expedite urological consultation and minimize time to surgery. Taking this concept a step further, validating EMT generated TWIST scoring opens the door to early risk stratification in the field or during transit to the hospital, analogous to Glasgow Coma Scale use. Such rapid triage could allow for expedited care on arrival for patients with TWIST scores of 6 or 7, bypassing the ER with direct transport for surgery. Of course, this time-saving approach would require further investigation and evaluation before routine implementation.

A limitation of this NIH sponsored study is that it was powered for NIRS performance to diagnose testicular torsion. Thus, this secondary outcome subanalysis evaluating TWIST score performance is limited by the small number of patients. Similarly individuals without torsion in the ER had poor followup, which could lead to misclassification of non-torsion cases. Furthermore, the TWIST score does not account for time from initial symptom onset. Generally as torsion progresses, more signs and symptoms associated with the TWIST score will be present.

Our tertiary care center is often a referral center for pediatric patients with testicular pain, and our torsion rate in this study of 34.4% is much higher than that reported in the literature (10% to 15%).[1,2] Many patients are transferred from elsewhere, prolonging the duration of symptoms. Thus, our study cohort may represent a biased group with prolonged torsion, enabling the TWIST score to be more diagnostic.

The TWIST score does not incorporate severity of pain in risk stratifying cases. While pain severity may be hard to quantify for patients, clinical suspicion tends to be greater when a patient presents with sudden onset severe pain. Finally, the results are applicable only to patients who were not excluded from our study due to prior testicular pathology or other medical comorbidities. However, strengths of the study include the prospective, internally controlled study design and the patient evaluation with TWIST scoring by nonphysician, nonurological staff.

CONCLUSIONS

TWIST score was highly predictive in our population when evaluated by EMTs, especially in Tanner 3 to 5 patients, where positive and negative predictive values were 100%. Therefore, our proposed algorithm can potentially guide emergency room physicians and staff to triage patients presenting with acute scrotum. Due to difficulty in definitive torsion diagnosis in Tanner stage 1 and 2 cases, we recommend obtaining an ultrasound even in high risk TWIST cases. Since low risk patients do not require ultrasound to rule out torsion and high risk patients who are Tanner stage 3 to 5 can proceed directly to surgery, ultrasound is safely avoided in more than 50% of patients.

REFERENCES

1. McAndrew HF, Pemberton R, Kikiros CS et al: The incidence and investigation of acute scrotal problems in children. Pediatr Surg Int 2002; **18:** 435.
2. Liang T, Metcalfe P, Sevcik W et al: Retrospective review of diagnosis and treatment in children presenting to the pediatric department with acute scrotum. AJR Am J Roentgenol 2013; **200:** W444.
3. Zhao LC, Lautz TB, Meeks JJ et al: Pediatric testicular torsion epidemiology using a national database: incidence, risk of orchiectomy and possible measures toward improving the quality of care. J Urol 2011; **186:** 2009.
4. Visser AJ and Heyns CF: Testicular function after torsion of the spermatic cord. BJU Int 2003; **92:** 200.
5. Kapoor S: Testicular torsion: a race against time. Int J Clin Pract 2008; **62:** 821.
6. Cost NG, Bush NC, Barber TD et al: Pediatric testicular torsion: demographics of national orchiopexy versus orchiectomy rates. J Urol, suppl., 2011; **185:** 2459.
7. Rabinowitz R: The importance of the cremasteric reflex in acute scrotal swelling in children. J Urol 1984; **132:** 89.
8. Nelson CP, Williams JF and Bloom DA: The cremasteric reflex: a useful but imperfect sign in testicular torsion. J Pediatr Surg 2003; **38:** 1248.
9. Kadish HA and Bolte RG: A retrospective review of pediatric patients with epididymitis, testicular torsion, and torsion of testicular appendages. Pediatrics 1998; **102:** 73.
10. Liguori G, Bucci S, Zordani A et al: Role of US in acute scrotal pain. World J Urol 2011; **29:** 639.
11. Gunther P, Schenk JP, Wunsch R et al: Acute testicular torsion in children: the role of sonography in the diagnostic workup. Eur Radiol 2006; **16:** 2527.
12. Barbosa JA, Tiseo BC, Barayan GA et al: Development and initial validation of a scoring system to diagnose testicular torsion in children. J Urol 2013; **189:** 1859.
13. Schlomer BJ, Keays M, Grimsby G et al: Transscrotal near infrared spectroscopy in the emergency department to diagnose testicular torsion in pediatric patients presenting with acute scrotum. J Urol, suppl., 2015; **193:** e463.
14. Harris PA, Taylor R, Thielke R et al: Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009; **42:** 377.
15. DeLong ER, DeLong DM and Clarke-Pearson DL: Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988; **44:** 837.
16. Boettcher M, Bergholz R, Krebs TF et al: Clinical predictors of testicular torsion in children. Urology 2012; **79:** 670.
17. Boettcher M, Krebs T, Bergholz R et al: Clinical and sonographic features predict testicular torsion in children: a prospective study. BJU Int 2013; **112:** 1201.
18. Srinivasan A, Cinman N, Feber KM et al: History and physical examination findings predictive of testicular torsion: an attempt to promote clinical diagnosis by house staff. J Pediatr Urol 2011; **7:** 470.
19. Beni-Israel T, Goldman M, Bar Chaim S et al: Clinical predictors for testicular torsion as seen in the pediatric ED. Am J Emerg Med 2010; **28:** 786.
20. Sharp VJ, Kieran K and Arlen AM: Testicular torsion: diagnosis, evaluation, and management. Am Fam Physician 2013; **88:** 835.
21. Patriquin HB, Yazbeck S, Trinh B et al: Testicular torsion in infants and children: diagnosis with Doppler sonography. Radiology 1993; **188:** 781.