Novel Computed Tomography Scan Scoring System Predicts the Need for Intervention after Splenic Injury

Burke T. Thompson, MD, Felipe Munera, MD, Stephen M. Cohn, MD, FACS, Alexandra A. MacLean, MD, John Cameron, MD, Luis Rivas, MD, and David Bajayo, MD

Background: The purpose of this study was to develop a computed tomography (CT) scan screening test to predict the need for intervention in patients with splenic injury.

Methods: CT scans of 20 patients with blunt injury to the spleen were reviewed to identify findings that correlated with the need for intervention (surgery or embolization). A screening test was created and then validated in CT scans from 56 consecutive patients.

Results: Three findings correlated with the need for intervention: 1) devascularization or laceration involving 50% or more of the splenic parenchyma, 2) contrast blush greater than one centimeter in diameter (from active extravasation of intravenous contrast material or pseudoaneurysm formation), and 3) a large hemoperitoneum. The sensitivity of the screening test was 100%, specificity was 88%, and overall accuracy was 93%.

Conclusions: These CT scan grading criteria appear to reliably predict the need for invasive management in patients with blunt injury to the spleen.

Key Words: Splenic injury, Injury grading scales

J Trauma. 2006;60:1083–1086.

Over the past two decades, the pendulum has swung from routine splenectomy for splenic injury to a policy of splenic preservation and nonoperative management.1–4 A recent multi-institutional study of blunt splenic injury in 1,488 adults found that 38.5% of adults with splenic injury went directly to surgery, whereas 61.5% were initially managed nonoperatively.5 The failure rate for the nonoperative approach was 11%. Unfortunately, radiologists utilizing the current American Association for the Surgery of Trauma (AAST) grading scale to evaluate abdominal computed tomography (CT) scans cannot reliably predict those patients needing intervention. Our group found a high degree of intra- and interrater variability in CT grading of splenic injury, even among highly experienced trauma radiologists.6 The development of a clinically relevant screening test, relating patterns of injury present on CT scans with the need for intervention, should improve management and ultimately the outcome of patients with injury to the spleen. We present a clinical screening test based on CT imaging results and demonstrate its utility in a series of adults with splenic injuries.

PATIENTS AND METHODS

Patients included in this study were admitted to the Ryder Trauma Center, Jackson Memorial Hospital, University of Miami, Miami, Florida.

In Phase I, 20 hemodynamically stable patients with blunt injury to the spleen, as visualized on a contrast CT scan of the abdomen and pelvis (PQ 6000, Marconi Systems, Highland Heights, OH), were selected consecutively from our database. B.T. and F.M. reviewed the CT scans to identify which of a number of findings correlated with the need for intervention including: magnitude of hemoperitoneum, amount of active extravasation of intravenous contrast material, size of pseudoaneurysm of the spleen, and percentage of splenic tissue involved in the lacerations. The indications for interventions were clinical instability, requirement of ongoing transfusions, and the development of worsening abdominal symptoms.

For Phase II, findings that consistently correlated with the performance of an intervention (surgery or embolization) formed the basis of the screening test. A CT scan having one or more of the findings was defined as positive. This screening test was validated in 56 consecutive patients retrospectively identified from our trauma database who were victims of blunt trauma, hemodynamically stable on admission, and had splenic injuries identified on CT scan. Trauma radiologists read each scan and scored them as positive or negative. The radiologists were blinded to the clinical course of the patients. The need for intervention for each patient (clinical instability, requirement of ongoing transfusions, and the development of worsening abdominal symptoms) was then compared with the radiology interpretation results. The sen-
Sensitivity, specificity, positive and negative predictive values, and accuracy of the grading system were determined.

Scans from Phase I were obtained on a single-detector helical CT (PQ 6000); scans from Phase II were acquired on a four-detector row scanner (MX8000; Philips Medical Systems, Andover, MA). The CT scanning parameters that were common to both phases included: 120 to 140 Kvp; 200 to 250 mAs; and field of view, 240 to 350 mm. The remaining parameters varied depending on the phase of the study. In Phase I, the single-detector scanner was used: slice thickness 8 mm, interval of reconstruction 7 mm, and pitch factor of 1:1 to 1.5:1.0. In Phase II, the multidetector scanner was used with the following parameters: effective collimation, 3.2 mm; reconstruction interval, 3 mm; and pitch, 6.

All patients received 100 mL of intravenous contrast material containing 320 mg of iodine per milliliter (iohexol, Optiray; Mallinckrodt Imaging, Hazelwood, MO) administered with a power injector (Liebel-Flarsheim CT 9000 Adv; Mallinckrodt) at a rate of 3 mL/sec. We administered 800 mL of iodinated contrast material (3% Gastroview; Mallinckrodt Medical, St Louis, MO) perorally or through the nasogastric tube. The scans were not delayed to wait for the contrast migration into the bowel.

The data were analyzed using STATA version 7.0 (Stata Corporation, College Station, TX) and Microsoft Excel from Office 2000 (Microsoft, Seattle, WA). The NCSS 2004 software package (Number Cruncher Statistical Systems, Kaysville, UT) was used to construct exact 95% confidence intervals, based on binomial probabilities, about the point estimates for sensitivity and specificity. A continuity correction, the subtraction of 0.5 from the numerator, was used in the calculation of the confidence intervals.

This project was approved by the University of Miami Human Subjects Institutional Review Board.

RESULTS

The characteristics of the Phase I group of 20 patients do not appear to differ significantly from the Phase II group of 56 patients. The mean ages of the two groups did not differ significantly: Phase I was 35 years old, Phase II was 38 years old ($p = 0.54$). The populations also had a similar sex distribution with 25% and 32% of Phase I and II patients, respectively, being female.

In Phase I, three specific findings appeared to be associated with the performance of a surgical or angiographic intervention: 1) devascularization or laceration involving 50% or more of the splenic parenchyma (Fig. 1); 2) a blush of contrast greater than 1 cm diameter (Fig. 2); and 3) a large hemoperitoneum, defined here as fluid seen in three or more of the following areas: Morrison’s pouch, right subphrenic space, left subphrenic space, perisplenic space, right gutter, left gutter, or pelvis (Fig. 3). The CT scans were also assessed...
for presence of a splenic pseudoaneurysm but this finding did not correlate with the need for intervention.

In Phase II, this new CT scan scoring system was tested and the need for intervention was ascertained in 56 individuals. The sensitivity of the grading method was 100% (95% confidence interval [CI], 85% to 100%); 41% of the patients had a positive CT scan and all of these patients had an intervention. The specificity of the test was 88% (95% CI, 72% to 95%) because 4 of the 56 patients had a positive CT scan yet did not require an intervention; three of these patients had a significant hemoperitoneum and one had a splenic laceration greater than 50%. Overall, the test had a negative predictive value of 100% (95% CI, 85% to 100%), positive predictive value of 85% (95% CI, 65% to 95%), and accuracy of 93%. Combinations of the findings were also analyzed. Seven patients had all three findings and all required surgical intervention. When combinations of two findings were examined, seven patients with both hemoperitoneum and a blush had a positive test and all seven patients required surgical intervention. Analysis of combinations of either a laceration and hemoperitoneum or a laceration and blush failed to produce any patients with a positive test. When the individual findings were examined, only patients with hemoperitoneum had a positive test resulting in eight interventions (seven surgical and one embolization; Table 1). Analysis of combined findings (laceration + blush + hemoperitoneum; hemoperitoneum + blush) resulted in specificity of 100%, but the sensitivities were reduced to 30% and 47%, respectively. In Table 2, the results of the analysis with this new grading system are compared with previously published results using the AAST grading system.6,7

**DISCUSSION**

We developed a novel CT scan grading system to predict need for intervention following splenic injury. Three findings correlated with the need for intervention: 1) devascularization or laceration involving 50% or more of the splenic parenchyma; 2) contrast blush greater than one centimeter in diameter (from active extravasation of intravenous contrast material or pseudoaneurysm formation); and 3) a large hemoperitoneum. When the new system was tested on 56 patients, it had a sensitivity of 100% and specificity of 88% when predicting the need for surgical or angiographic intervention. Of note, the combined finding of a large hemoperitoneum and active arterial extravasation, analogous to an expanding hematoma, appears to be equivalent to a grade V injury in its ability to predict intervention.

The screening test proposed in this study is fairly simple and we believe it can be taught to both radiologists and surgeons in a reasonable amount of time with a learning curve of approximately 10 cases. Our screening test was 100% sensitive and therefore it appeared to be an excellent method to exclude lesions requiring intervention. As with all grading systems that try to avoid failing to diagnose life-threatening lesions, our specificity was only 88%; there were four false-positive test results. Unfortunately, unnecessary interventions are not innocuous. For example, a laparotomy carries a 12% to 15% risk of incisional hernia development and risk of bowel obstruction.8 The three who tested positive for significant hemoperitoneum may have had confounding sources of intra-abdominal fluid. One false positive for percentage of splenic involvement could have represented a judgment error. This may be avoided with the new or multislice helical CT technology available to make more precise calculations of volume.9,10

This study has a number of potential limitations. We have not studied the interrater reliability of this grading method in a large group of clinicians. A recent study at this institution demonstrated the current AAST grading system has poor interrater reliability.6 Other limitations of this investigation are the relatively small size of the validation group and the use a retrospective review of prospectively collected data. Because of the small validation sample, there are fairly wide confidence intervals about the point estimates for both sensitivity (85% to 100%) and specificity (72% to 95%). A prospective validation on a larger population sample appears indicated before reaching conclusions regarding test performance or recommendations for its adoption as standard of care.

**REFERENCES**


---

**Table 1 Analysis of Individual and Combined Findings**

<table>
<thead>
<tr>
<th>Number of Findings</th>
<th>Laceration + Hemoperitoneum</th>
<th>Laceration + Blush</th>
<th>Hemoperitoneum + Blush</th>
<th>Laceration</th>
<th>Blush</th>
<th>Hemoperitoneum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with positive test</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Intervention</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Intervention (surgical(s)/embolization(e))</td>
<td>7 (s)</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 2 Comparison of Grading Systems**

<table>
<thead>
<tr>
<th>New Grading System (%)</th>
<th>AAST Grading System (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>100</td>
</tr>
<tr>
<td>Specificity</td>
<td>88</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>85</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>100</td>
</tr>
<tr>
<td>Accuracy</td>
<td>93</td>
</tr>
</tbody>
</table>

* Data from this study.
† Data from study in reference 10.


