Ability of Lower-Extremity Injury Severity Scores to Predict Functional Outcome After Limb Salvage

By Thuan V. Ly, MD, Thomas G. Trivison, PhD, Renan C. Castillo, MS, Michael J. Bosse, MD,
Ellen J. MacKenzie, PhD, and the LEAP Study Group

Investigation performed at the Carolinas Medical Center, Charlotte, North Carolina

Background: Lower-extremity injury severity scoring systems were developed to assist surgeons in decision-making regarding whether to amputate or perform limb salvage after high-energy trauma to the lower extremity. These scoring systems have been shown to not be good predictors of limb amputation or salvage. This study was performed to evaluate the clinical utility of the five commonly used lower-extremity injury severity scoring systems as predictors of final functional outcome.

Methods: We analyzed data from a cohort of patients who participated in a multicenter prospective study of clinical and functional outcomes after high-energy lower-extremity trauma. Injury severity was assessed with use of the Mangled Extremity Severity Score; the Limb Salvage Index; the Predictive Salvage Index; the Nerve Injury, Ischemia, Soft-Tissue Injury, Skeletal Injury, Shock, and Age of Patient Score; and the Hannover Fracture Scale-98. Functional outcomes were measured with use of the physical and psychosocial domains of the Sickness Impact Profile at both six months and two years following hospital discharge. Four hundred and seven subjects for whom the reconstruction regimen was considered successful at six months were included in the analysis. We used partial correlation statistics and multiple linear regression models to quantify the association between injury severity scores and Sickness Impact Profile outcomes with the subjects' ages held constant.

Results: The mean age of the patients was thirty-six years (interquartile range, twenty-six to forty-four years); 75.2% were male and 24.8% were female. The median Sickness Impact Profile scores were 15.2 and 6.0 points at six and twenty-four months, respectively. The analysis showed that none of the scoring systems were predictive of the Sickness Impact Profile outcomes at six or twenty-four months to any reasonable degree. Likewise, none were predictive of patient recovery between six and twenty-four months postoperatively as measured by a change in the scores in either the physical or the psychosocial domain of the Sickness Impact Profile.

Conclusions: Currently available injury severity scores are not predictive of the functional recovery of patients who undergo successful limb reconstruction.

Level of Evidence: Prognostic Level I. See Instructions to Authors for a complete description of levels of evidence.

The decision to amputate or perform limb salvage of a severely injured lower extremity remains a difficult dilemma for orthopaedic surgeons, and multiple scoring systems¹⁻⁷ have been developed to assist them in making this decision. These published guidelines contain different criteria that are objective or subjective, or both. The five commonly used scoring systems are the Mangled Extremity Severity Score (MESS)^{1,3}; the Limb Salvage Index (LSI)⁵; the Predictive Sal-

vage Index (PSI)²; the Nerve Injury, Ischemia, Soft-Tissue Injury, Skeletal Injury, Shock, and Age of Patient Score (NISSSA)⁴; and the Hannover Fracture Scale-98 (HFS-98)^{6,7}.

Advances in orthopaedic, plastic, and vascular reconstruction have made it possible to treat many severely injured limbs with salvage procedures. We previously showed that the five commonly used lower-extremity injury severity scoring systems were not good predictors of limb amputation or sal-

Disclosure: In support of their research for or preparation of this work, one or more of the authors received, in any one year, outside funding or grants in excess of \$10,000 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases, the National Institutes of Health. Neither they nor a member of their immediate families received payments or other benefits or a commitment or agreement to provide such benefits from a commercial entity. No commercial entity paid or directed, or agreed to pay or direct, any benefits to any research fund, foundation, division, center, clinical practice, or other charitable or nonprofit organization with which the authors, or a member of their immediate families, are affiliated or associated.

vage⁸. It remained unclear, however, if these scoring systems might be predictors of the function of the salvaged limb. The systems were not designed to predict functional outcome, which likely explains why their association with long-term functional outcomes of treatment of severely injured limbs has been evaluated in only a few studies⁹⁻¹². These studies were small retrospective series in which validated patient-reported outcomes instruments were not employed¹³⁻¹⁵.

The primary aim of this study was to evaluate whether the five commonly used lower-extremity scoring systems could predict functional outcomes after limb salvage. A secondary aim was to determine if these lower-extremity scoring systems were able to predict whether patients for whom an amputation was recommended (on the basis of the threshold for amputation in the individual injury severity scoring systems) would ultimately do worse than patients for whom limb salvage was recommended. We hypothesized that increasing lower-extremity injury severity scores would be associated with worsening functional outcome scores. We also hypothesized that patients for whom limb amputation was recommended would have a worse functional outcome than patients for whom limb salvage was recommended.

Materials and Methods

We analyzed data from a cohort of patients who participated in the Lower Extremity Assessment Project (LEAP), which is a multicenter prospective study^{16,17}. The LEAP study consisted of 601 patients who were admitted to eight level-I trauma centers after high-energy lower-extremity trauma. Participants were between the ages of sixteen and sixty-nine years old and were enrolled between March 1994 and June 1997. Included injuries were (1) traumatic amputations below the distal aspect of the femur; (2) selected Gustilo¹⁸ type-IIIA tibial fractures; (3) Gustilo type-IIIB tibial fractures; (4) Gustilo type-IIIC tibial fractures; (5) dysvascular injuries below the distal aspect of the femur (not the foot), including dislocations of the knee, closed tibial fractures, and penetrating wounds with vascular injury noted on an arteriogram or ultrasound or at the time of surgery; (6) selected major soft-tissue injuries below the distal aspect of the femur (not the foot); and (7) selected severe injuries of the distal aspect of the tibia or the foot. Exclusion criteria were (1) an age younger than sixteen or older than sixty-nine years of age, (2) transfer to the level-I trauma center more than twenty-four hours after the time of injury or primary orthopaedic treatment elsewhere, (3) a Glasgow Coma Scale¹⁹ score of <15 points at twenty-one days after the injury or at the time of discharge if that was sooner, (4) a spinal cord deficit, (5) a prior leg or foot amputation or a patient who was unable to walk prior to the injury, (6) a third-degree burn measuring more than one handbreadth on the injured lower limb, (7) a documented psychiatric disorder or mental retardation, (8) an inability to speak English or Spanish, (9) active military duty, or (10) residence outside the catchment area of the trauma center. Complete details have previously been published^{16,17}.

For the current study, analyses were restricted to data on 407 limbs for which reconstruction regimens had been suc-

TABLE I Characteristics of the Subjects in the LEAP Study Treated Successfully with Limb Salvage (N = 407)

Variable	Mean	Standard Deviation
Age (yr)	35.6	12.3
SIP score (points)		
Physical domain		
6 mo	17.4	12.4
24 mo	9.5	9.8
Psychosocial domain		
6 mo	12.1	14.1
24 mo	9.9	14.6
Injury severity scores (points)		
MESS	4.4	1.4
LSI	2.8	1.4
PSI	5.6	1.8
NISSSA	7.2	1.6
HFS-98	5.3	1.9

cessful at six months after hospital discharge. Twenty-one of the included subjects had a bilateral injury (that is, had sustained trauma to both lower limbs), and fourteen of the twenty-one had undergone bilateral reconstruction regimens that were both successful at six months following hospital discharge. Injury severity was assessed with use of commonly applied predictive indices, including the MESS, LSI, PSI, NISSSA, and HFS-98.

The MESS was described by Johansen et al.³ in 1990. The components of this system are the degree of skeletal and soft-tissue injury, the severity of limb ischemia, systemic hypotension, and patient age. A MESS of ≥ 7 points predicted amputation with 100% accuracy. The LSI was proposed by Russell et al.⁵ in 1991. The seven components of the LSI include injury to an artery, deep vein, nerve, bone, skin, and muscle as well as warm ischemia time. An LSI of ≥ 6 points indicates that the limb should be amputated. The PSI was developed by Howe et al.² in 1987. The components that make up the PSI are the level of arterial injury, the degree of bone injury, the degree of muscle injury, and the time to surgery. The threshold for limb amputation is a score of ≥ 8 points. The NISSSA was described by McNamara et al. in 1994⁴. The elements that make up the NISSSA are nerve injury, ischemia, soft-tissue injury, skeletal injury, shock, and patient age. The threshold for limb amputation is a score of ≥ 11 points. The HFS-98 was developed in 1982⁶ and modified in 2001⁷. The components of the HFS-98 are bone loss, skin injury, muscle injury, wound contamination, periosteal stripping, local circulation, systematic circulation, and nerve function. A score of ≥ 11 points is the threshold for limb amputation. MESS, LSI, and PSI scores were available for all 407 limbs, whereas NISSSA and HFS-98 scores were available for 284 and 313 limbs, respectively.

Functional outcomes were measured with use of the physical and psychosocial domains of the Sickness Impact

Fig. 1

Physical SIP subscores versus limb salvage scores, divided into categories. Means and 95% confidence intervals are displayed. There is little evidence of a consistent association between these scores at six months (upper panel) and little resemblance between patterns observed at six months and twenty-four months (lower panel).

Profile (SIP) at both six months and two years following hospital discharge. The SIP is a validated and widely used self-reported measure of the health status of the patient¹³⁻¹⁵. It contains 136 statements detailing the patient's functional status in five independent domains, four psychosocial domains, and three physical domains. The five independent domains are sleep and rest, eating, work, home management, and recreation and pastimes. The psychosocial domains consist of social interaction, alertness behavior, emotional behavior, and communication. The physical domains include walking, mobility, and body care and movement. The overall SIP scores can range from 0 to 100 points. Scores of between 0 and 3 points indicate no disability, those between 4 and 9 points indicate mild disability, those between 10 and 19 points indicate moderate disability, and those of ≥ 20 points indicate severe disability¹⁴.

Exploratory and graphical analyses were employed to carry out empirical assessments of unadjusted associations. The physical and psychosocial domains of the SIP were as-

sessed as independent outcomes, each at six and twenty-four months. In addition, we calculated the change in the SIP subscores between six and twenty-four months as a measure of improvement in functional status over that time. Each of the five lower-extremity injury severity scores was independently assessed with respect to its association with SIP outcomes at six and twenty-four months as well as with the change over that time. We used partial correlation statistics and multiple regression analyses to quantify the age-adjusted associations between severity scores, treated as continuous variables, and SIP outcomes. Wald-type tests were employed to assess the significance of regression coefficients. Results were considered significant if the appropriate null hypothesis could be rejected at the 0.05 level.

Results

Summary statistics are given in Table I. The mean age of the subjects was thirty-six years, with an interquartile range (the values for the middle 50% of the subjects) of twenty-six to

forty-four years and a standard deviation of 12.3 years; 75.2% of the patients were male, and 24.8% were female. Overall, the subjects exhibited substantial improvement in the SIP score from six to twenty-four months after hospitalization. They displayed a mean physical SIP score (and standard deviation) of 17.4 ± 12.4 points at six months and 9.5 ± 9.8 points at twenty-four months. Subjects displayed similar declines (improvements) in the scores in the psychosocial domain, from a mean of 12.1 ± 14.1 points at six months to a mean of 9.9 ± 14.6 points at twenty-four months. Both SIP domains displayed mild right skew at both time points.

The majority of the scores on all of the injury severity scales were concentrated within a few values of the median score for that scale. For example, the interquartile range of the MESS scores was 3 to 5 points; in fact, 73% of the subjects had a MESS score of 3, 4, or 5 points.

Age-adjusted associations between injury severity scores (when treated as continuous covariates) and SIP outcomes were uniformly weak. For instance, we observed that the partial correlation (age-adjusted) between the MESS and the physical domain of the SIP was -0.08 ($p = 0.14$) six months after hospital discharge; associations between all other scores and the SIP physical subscore were at least as weak. Additionally, whatever association there was between several of the injury severity scores and the SIP outcomes at six months appeared to dissipate over time. For instance, the partial correlation between the MESS and SIP physical score was 0.01 ($p = 0.91$) at twenty-four months after hospital discharge. Finally, of all of the injury severity scores, only the MESS displayed a significant age-adjusted association with the change in the physical SIP subscore from six to twenty-four months following hospitalization; the MESS displayed a partial correlation of -0.13 ($p = 0.02$) with a decreasing physical SIP subscore over time. None of the scores displayed a significant association with changes in the psychosocial domain of the SIP. To assess the possibility of the influence of skew in the data on the results, we replicated unadjusted correlation analyses using a nonparametric alternative (Spearman rank correlation statistic), which is less sensitive to outliers. We observed no evidence that a change in the method influenced the results.

The results with regard to the physical subscore of the SIP are shown in Figure 1, and the results with regard to the psychosocial subscore were similar. There is only limited evidence of associations between the injury severity scores and SIP outcomes, and the trends are not necessarily coherent. For instance, at six months after hospital discharge, subjects with a MESS of ≤ 2 points had a mean physical SIP subscore of 21.0 points, subjects with a MESS of between 3 and 5 points had a mean physical SIP subscore of 17.4 points, and subjects with a MESS of between 6 and 11 points had a mean physical SIP subscore of 17.7 points. These results were in contrast to what we had expected to see, which was an increasing MESS associated with worsening physical SIP score. In other cases, the pattern was even less consistent. For instance, at six months after hospital discharge, the seventy-three subjects with an LSI of < 2 points had a mean psychosocial SIP subscore of 10.4

points, whereas the 294 subjects with an LSI of 2 to 4 points had a mean psychosocial SIP subscore of 12.7 points and the forty subjects with an LSI of ≥ 5 points had a mean psychosocial SIP subscore of 10.5 points. Thus, an increasing LSI was not consistently associated with either an increasing or a decreasing psychosocial SIP subscore. Moreover, when age effects were controlled, evidence contradicting the hypothesis that the true mean psychosocial SIP score was equal in all three LSI categories was not significant ($p = 0.23$). The results were similar for the physical SIP domain and for all injury severity scores. In addition, even when an association between an SIP score and an injury severity score (for instance, between the psychosocial SIP score and the MESS) approached significance at six months, evidence in favor of that association appeared to dissipate by twenty-four months.

Using multiple linear regression to control for age, we assessed the potential for differences in SIP outcomes between subjects for whom amputation was recommended on the basis of the limb severity scores (i.e., a MESS of ≥ 7 points, a PSI of ≥ 8 points, and/or an LSI of ≥ 6 points) and the subjects for whom amputation was not recommended. None of the three scores were significant in predicting either psychosocial or physical SIP subscores at either six or twenty-four months after hospitalization. For example, an amputation would be recommended for a patient with a MESS of 9 points, but patients with such a score did not necessarily do worse overall (as indicated by the SIP outcome) than patients who had a MESS of 2 points and thus would be advised to undergo limb salvage.

Discussion

The five lower-extremity injury severity scoring systems developed in the last several decades may provide orthopaedic and vascular surgeons with objective criteria to assist in the determination of the fate of a severely injured limb. Recent reports, however, have shown that these scoring systems are not reliable⁸⁻¹⁰. Technical advances in skeletal fixation, vascular repair, and soft-tissue coverage over the last two decades have allowed more severely injured limbs to be salvaged routinely. Following injury, the primary goal of treatment is to return patients to their preinjury functional status. We previously showed that, when we controlled for all injury, patient, and environmental factors, the functional outcomes, as measured with the SIP at two and seven years, did not differ between patients who had undergone limb salvage and those who had undergone amputation¹⁶. Both groups showed substantial disability and similar, low return-to-work rates (approximately 50%).

We are aware of few studies⁹⁻¹² in which investigators have attempted to associate these scoring systems with the functional outcome status of the limb. In our current study, we provided new information by evaluating the clinical utility of the five systems and their association with functional outcomes among subjects who had undergone a limb salvage. We observed no such association in this population.

These results are consistent with those of smaller retrospective studies. Durham et al.¹⁰ retrospectively assigned scores to twenty-three upper-extremity and fifty-one lower-extremity

injuries that had been treated over a ten-year period. Of thirty limbs that were treated with a salvage procedure, twenty-four had a successful outcome. The functional outcome, as defined by impairment, return to work, and disability, was determined at clinic visits and/or with a telephone interview. The authors concluded that none of the scoring systems predicted functional outcome. This analysis was affected by the small number of subjects and the limited measures of functional status.

In contrast, the results presented here were derived with a multidimensional self-reported health status outcomes tool. The SIP has been well tested for reliability and validity¹³⁻¹⁵. With a larger cohort (407 compared with twenty-four limbs) and a prospective evaluation, this study strengthens the conclusion of Durham et al.¹⁰ that the available lower-extremity injury scoring systems are not good predictors of functional outcome.

Our study suggests that lower-extremity injury severity scores are not predictive of the short-term (six-month) or longer-term (twenty-four-month) functional outcomes of limb salvage in patients who received aggressive treatment. The findings that the strongest and most significant association between a lower-extremity injury score (the MESS at six months) and physical function was an inverse association while the other scores showed little evidence of any association with functional outcome underline this conclusion. Despite the subjects showing improvements in their SIP scores over time, our data showed little to no association between the injury severity scores and this improvement with time.

We assessed differences in SIP outcomes between patients for whom amputation had been recommended on the basis of their scores on the MESS, PSI, and/or LSI and patients for whom amputation had not been recommended. The amputation threshold of these scoring systems displayed no utility in predicting SIP outcomes for patients undergoing reconstruction. In other words, the subjects who underwent reconstruction even though an amputation had been "recommended" did not do worse overall than subjects for whom amputation had not been recommended, suggesting that subjects who disregarded that recommendation did not have a resulting loss of function.

It should be noted, however, that subjects enrolled in the LEAP study received high-level care at level-I trauma centers. It

is possible that such care would overcome differences in outcomes over two years that could have been predicted with use of injury severity scores in the face of less aggressive treatment. On the other hand, the uniformity of care provided to the LEAP subjects arguably reduced the intersubject variation that might mask such associations.

The results of the present study indicate that the lower-extremity injury severity scoring systems are not predictive of functional recovery among patients who have undergone successful limb reconstruction. Practitioners should exercise caution when interpreting scores in the context of potential recovery from high-energy trauma. ■

NOTE: The authors acknowledge the efforts of the study coordinators and physical therapists at each of the eight LEAP study sites. Their dedication to the study's objectives and their commitment to quality data collection were essential to the success of the study. They include Julie Agel, ATC; Jennifer Avery, PT; Denise Bailey, PT; Wendell Bryan; Debbie Bullard; Carla Carpenter, PT; Elizabeth Chaparro, RN; Kate Corbin; Denise Darnell, RN, BSN; Stephanie Dickason, PT; Thomas DiPasquale, DO; Betty Harkin, PT; Michael Harrington, PT; Dolfi Herscovici, DO; Amy Holdren, RNC, ANP, MSN; Linda Howard, PT; Sarah Hutchings, BS; Marie Johnson, LPN; Melissa Jurewicz, PT; Donna Lampke, PT; Karen Lee, RN; Marianne Mars, PT; Maxine Mendoza-Welch, PA; J. Wayne Meredith, MD; Nan Morris, PT; Karen Murdock, PT; Andrew Pollak, MD; Pat Radey, RN; Sandy Shelton, PT; Sherry Simpson, PT; Steven Sims, MD; Douglas Smith, MD; Adam Starr, MD; Celia Wiegman, RN; John Wilber, MD; Stephanie Williams, PA; Philip Wolinsky, MD; Mary Woodman, BA; and Michelle Zimmerman, RN.

Thuan V. Ly, MD
Department of Orthopaedic Surgery, University of Minnesota,
Regions Hospital, 640 Jackson Street, MS11503L,
St. Paul, MN 55101-2925

Thomas G. Trivison, PhD
New England Research Institutes, 9 Galen Street,
Watertown, MA 02472

Renan C. Castillo, MS
Department of Health Policy and Management, Johns Hopkins
Bloomberg School of Public Health, 624 North Broadway Street,
Hampton House 544, Baltimore, MD 21205

Michael J. Bosse, MD
Department of Orthopaedic Surgery, Carolinas Medical Center,
P.O. Box 32861, Charlotte, NC 28232

Ellen J. MacKenzie, PhD
Johns Hopkins Bloomberg School of Public Health,
624 North Broadway Street, Room 462, Baltimore, MD 21205

References

1. Helfet DL, Howey T, Sanders R, Johansen K. Limb salvage versus amputation. Preliminary results of the Mangled Extremity Severity Score. *Clin Orthop Relat Res*. 1990;256:80-6.
2. Howe HR Jr, Poole GV Jr, Hansen KJ, Clark T, Plonk GW, Koman LA, Pennell TC. Salvage of lower extremities following combined orthopedic and vascular trauma. A predictive salvage index. *Am Surg*. 1987;53:205-8.
3. Johansen K, Daines M, Howey T, Helfet D, Hansen ST Jr. Objective criteria accurately predict amputation following lower extremity trauma. *J Trauma*. 1990;30:568-73.
4. McNamara MG, Heckman JD, Corley FG. Severe open fractures of the lower extremity: a retrospective evaluation of the Mangled Extremity Severity Score (MESS). *J Orthop Trauma*. 1994;8:81-7.
5. Russell WL, Sailors DM, Whittle TB, Fisher DF Jr, Burns RP. Limb salvage versus traumatic amputation. A decision based on a seven-part predictive index. *Ann Surg*. 1991;213:473-81.
6. Tscherne H, Oestern HJ. [A new classification of soft-tissue damage in open and closed fractures (author's transl)]. *Unfallheilkunde*. 1982;85:111-5. German.
7. Krettek C, Seekamp A, Köntopp H, Tscherne H. Hannover Fracture Scale '98—re-evaluation and new perspectives of an established extremity salvage score. *Injury*. 2001;32:317-28. Erratum in: *Injury*. 2001;32:611.
8. Bosse MJ, MacKenzie EJ, Kellam JF, Burgess AR, Webb LX, Swiontkowski MF, Sanders RW, Jones AL, McAndrew MP, Patterson BM, McCarthy ML, Cyril JK. A prospective evaluation of the clinical utility of the lower-extremity injury-severity scores. *J Bone Joint Surg Am*. 2001;83:3-14.
9. Bonanni F, Rhodes M, Lucke JF. The futility of predictive scoring of mangled lower extremities. *J Trauma*. 1993;34:99-104.
10. Durham RM, Mistry BM, Mazuski JE, Shapiro M, Jacobs D. Outcome and utility of scoring systems in the management of the mangled extremity. *Am J Surg*. 1996;172:569-74.
11. O'Sullivan ST, O'Sullivan M, Pasha N, O'Shaughnessy M, O'Connor TP. Is it possible to predict limb viability in complex Gustilo IIIB and IIIC tibial fractures? A comparison of two predictive indices. *Injury*. 1997;28:639-42.

12. Lin CH, Wei FC, Levin LS, Su JJ, Yeh WL. The functional outcome of lower-extremity fractures with vascular injury. *J Trauma*. 1997;43:480-5.

13. Bergner M, Bobbitt RA, Kressel S, Pollard WE, Gilson BS, Morris JR. The Sickness Impact Profile: conceptual formulation and methodology for the development of a health status measure. *Int J Health Serv*. 1976;6:393-415.

14. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. *Med Care*. 1981;19:787-805.

15. Jurkovich G, Mock C, MacKenzie E, Burgess A, Cushing B, deLateur B, McAndrew M, Morris J, Swiontkowski M. The Sickness Impact Profile as a tool to evaluate functional outcome in trauma patients. *J Trauma*. 1995;39:625-31.

16. Bosse MJ, MacKenzie EJ, Kellam JF, Burgess AR, Webb LX, Swiontkowski MF, Sanders RW, Jones AL, McAndrew MP, Patterson BM, McCarthy ML, Trivison TG,

Castillo RC. An analysis of outcomes of reconstruction or amputation after leg-threatening injuries. *N Engl J Med*. 2002;347:1924-31.

17. MacKenzie EJ, Bosse MJ, Kellam JF, Burgess AR, Webb LX, Swiontkowski MF, Sanders RW, Jones AL, McAndrew MP, Patterson TM, McCarthy ML. Characterization of patients with high-energy lower extremity trauma. *J Orthop Trauma*. 2000;14:455-66.

18. Gustilo RB, Mendoza RM, Williams DN. Problems in the management of type III (severe) open fractures: a new classification of type III open fractures. *J Trauma*. 1984;24:742-6.

19. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet*. 1974;2:81-4.