Stephen Lloyd, MD, CCFP(EM)

Acute Ankle Injuries: Clinical/Radiologic Assessment in Diagnosis

SUMMARY

Acute ankle injuries represent a large burden of illness in the primary care/emergency setting. Concern has been expressed in the medical literature about the unnecessary use of X-rays in the assessment of these patients. A literature review was conducted to address the question: "Can clinical examination alone accurately diagnose fractures and sprains in patients with acute ankle injuries?" Published reports offer inconsistent conclusions on the value of clinical criteria in the diagnosis of fractures or sprains. These inconsistent conclusions appear to result from numerous biases, different patient populations studied, varied interpretations of normal and abnormal results, and different groups of examiners. The only criteria associated with the diagnosis of fracture in more than one study are age greater than 40 years, inability to bear weight on examination, "point" or "bone" tenderness, and swelling. Until consistent results are available, no change in the rate of X-ray referral should be expected. (*Can Fam Physician* 1988; 34:2261–2265.)

RÉSUMÉ

Les traumatismes aigus de la cheville constituent une bonne part des pathologies que rencontre le médecin de première ligne en situation d'urgence. La littérature médicale s'est préoccupée de l'utilisation abusive des radiographies pour évaluer ces patients. Afin de pousser l'étude de cette question, on a procédé à une révision de la littérature: «Est-il possible que le seul examen physique soit suffisamment précis pour poser un diagnostic de fracture et d'entorse chez les patients qui se présentent avec des traumatismes aigus de la cheville?» Les rapports publiés sont inconsistants quant à la valeur des critères cliniques pour poser un diagnostic de fracture ou d'entorse. L'inconsistance de ces conclusions semble causée par de nombreux biais, les différences entre les populations étudiées, la variété des interprétations entre ce qui est normal et anormal et les différences entre les groupes de chercheurs. Les seuls critères que l'on retrouve dans plus d'une étude et qui sont reliés au diagnostic de fracture sont l'âge du patient (plus de 40 ans), l'impossibilité de mise en charge lors de l'examen et un gonflement et une sensibilité osseuse à un point précis. Aussi longtemps que les résultats ne seront pas plus consistants, on ne peut espérer modifier les taux de demande de radiographies.

Key words: acute ankle injuries, clinical diagnosis, epidemiologic methods

Dr. Lloyd is an Assistant Professor in the Department of Family Medicine, Division of Emergency Medicine at McMaster University, Hamilton. Requests for reprints to: Stephen Lloyd, M.D., C.C.F.P. (EM), Room 2R17, Department of Family Medicine, Division of Emergency Medicine, McMaster University, 1200 Main Street West, Hamilton, Ont. L8N 3Z5

ACUTE ANKLE INJURIES represent a large burden of illness in the primary care setting and require both efficient and effective diagnosis. Sprains and strains are the sixth-to ninth-most-frequent cause for visits to family physicians and the number-one injury in many occupations.[1] Ankle injuries account for a large proportion (3%–12%) of visits to emergency departments.[2,3] Ninety per cent of patients with acute ankle injuries are X-rayed, and the overwhelming majority of such injuries are sprains.[4]

Numerous authors have expressed concern about the unnecessary use of X-rays for patients presenting with acute ankle injuries.[2–9] A few investigators offer protocols using specific clinical criteria as guides to reducing the rate of unnecessary referral to

radiology.[2-4,9] These concerns and proposals have not altered physician practice, and the vast majority of patients presenting with acute ankle injuries are still referred for radiographic assessment.

A literature review was carried out to address the question: "Can clinical criteria alone accurately diagnose fractures and sprains in patients with acute ankle injuries?"

Methods

Six separate data bases were searched to answer the question posed above; an office file on acute injuries, a recently published emergency-medicine textbook,[10] MEDLINE, Science Citation Index, reference lists of identified citations, and recently published issues of popular journals (*CMAJ*, *British Medical Journal*, *JAMA*, *Lancet* and *Injury*). Although the office file and MEDLINE data bases proved most valuable and efficient, no single source identified all citations.

Relevant citations were critically appraised according to epidemiologic methods and previously published guides designed to help clinicians decide whether or not to use a diagnostic test (Chart 1).

Chart 1
Guides for Reading Articles To Learn Whether To Use a (new) Diagnostic Test

1. Was there an independent, "blind" comparison with a "gold standard" of diagnosis?
2. Did the patient sample include an appropriate spectrum of mild and severe, treated and untreated disease, plus individuals with different but commonly confused disorders?
3. Was the setting for the study, as well as the filter through which study patients passed, adequately described?
4. Was the reproducibility of the test result (precision) and its interpretation (observer variation) determined?
5. Was the term "normal" defined sensibly?
6. If the test is advocated as part of a cluster or sequence of tests, was its contribution to the overall validity of the cluster or sequence determined?
7. Were the tactics for carrying out the test described in sufficient detail to permit their exact replication?
8. Was the "utility" of the test determined?

Source: See reference 12.

Results

The search strategy used identified well over 100 citations relating to acute ankle injuries. Most of these citations relate to management and treatment of specific injuries. Ten citations address the use of clinical criteria in the diagnosis of these injuries and compare these criteria with the "gold standard" of radiographic imaging.[2-9,11,13]

Inclusion criteria, demographic data, and prognostic factors were not well handled in most studies. When assembled, inception cohorts were loosely defined, on initial evaluation, by patient age,[4,5] time from injury to evaluation,[3,4] or as "simple sprain".[11] Demographic data include only age in four studies,[4-7] and age and sex in

Table 1
Relative Risk of Fracture If "Test" Is Positive

		Study						
		2	3	4[a]	5	6	7	10
Age	< 40							
	> 40	4.2		2.27			2.75	
Inability to bear weight at exam		4.8		3.8			2.75	
Degree of swelling	Severe				3.2			
	Moderate			2.8		48.0	1.75	
	Slight							
	None							
Ecchymosis present				0.4	4.6		1.3	2.8
Type of pain	Bone[b]	6.8		7.6			4.21	
	Soft tissue						0.1	
	Not specified						2.42	
Limitation of range of motion				3.0	1.5		0.9	

a. Above 48 years of age in this study.
b. Point tenderness described as distal fibular or malleolar.

Table 2
Likelihood Ratio For Positive Test Result

		Study					
		2	3	5	6	7	10
Age	< 40					0.82	
	> 40	2.41				3.13	
Inability to bear weight at exam		3.92				2.17	
Degree of swelling	Severe			14.2	13.6	1.9	
	Moderate		2.0		3.2	0.89	
	Slight					0.42	
	None					1.4	
Ecchymosis present				3.6		1.3	1.2
Type of pain	Bone[a]	1.9				2.3	
	Soft tissue					0.32	
	Not specified					2.8	
Limitation of range of motion						0.98	

a. Point tenderness described as distal fibular or malleolar.

another;[2] no data are given in five.[2,3,8,9,11,13] Only two studies adjust for extraneous prognostic factors; of these, one considers medical history suggesting increased risk of fracture and impaired sensation of the extremity,[4] and the other considers previous history of injury.[2]

The examiner's level of training and experience varies from one study to another. Emergency physicians conduct the examination in four reports,[2,9,11,13] radiologists in two,[3,6] physician assistants in one,[4] and a variety of physicians and assistants in another;[5] one study fails to identify the person who carried out the examinations.[7]

The prevalence of disease (fracture) is consistent from study to study, according to setting. The emergency department is the setting in seven studies, and in these studies fracture prevalence is 12%–17.5%.[2,4,5,8,9,11,13] The radiology department is the setting for two studies, which evaluate "radiographic evidence of swelling". In the radiology studies fracture prevalence is 15%–35%.[3,6] One study, which fails to identify its setting, reports a 12.7% prevalence of fracture.[7]

Although confounding (Appendix) does not appear to be a major problem in the studies reviewed, bias and inter-observer variability may account for systematic error. Age and underlying illness can act as confounders. The descriptive cross-section survey design of the studies provides no protection from numerous biases. Although one study controlled for bias through "qualitative analysis of all injuries", the analysis is not described.[9] Patient selection bias could have influenced all studies.[2–9,11,13] Observation bias was a concern in four studies.[3,4,7,8] Only one study tests for inter-observer variation,[4] and only one "blinded" the radiologist to the result of the clinical exam.[2]

Statistical significance is not considered in every study but in only five of the 10,[2,4,7,9,11] and all studies do not evaluate the same clinical criteria. Although investigators associate various clinical signs or symptoms with sprains or fractures, they do not always provide data to support their conclusions.

Only seven of 10 studies provide data for analysis by readers, in support of their conclusions.[2–7,10] Relative risk of fracture with a positive test can be calculated from seven reports (Table 1) and likelihood ratios from six (Table 2). However, the different studies do not evaluate the same clinical criteria, and most fail to define their respective criteria to allow for comparison. Brand and colleagues present data that allow rough estimates of relative risk but insufficient data to calculate likelihood ratios.[4]

More than one author provide data for analysis of six clinical criteria. These data allow calculation of relative risk and likelihood ratios (Appendix). The six criteria are: age; inability to bear weight on examination; the presence of ecchymosis; limitation of movement in the injured ankle; type of pain/"point" tenderness; and degree of swelling. Relative risk of fracture is consistently high for patients who are more than 40 years of age, are unable to bear weight on examination, and have "point" or "bone" tenderness (Table 1). Relative risk of fracture is elevated in patients with slight to severe swelling, but the degree of risk varies greatly from study to study (Table 1). There is no consistency in relative risk of fracture for patients with ecchymosis or limited range of movement on examination (Table 2).

Only three clinical criteria were found to generate high likelihood ratios for a positive test result for fracture in more than one study; these criteria were age of more than 40 years; inability to bear weight on examination; and swelling (Table 2). Two of the three studies generating high likelihood ratios for swelling looked at radiographic films rather than patients.[3,6]

Discussion

Although data from a few studies show consistent association between specific clinical criteria and fractures, this is not the rule. Authors' conclusions about the diagnostic value of various clinical criteria are generally inconsistent. Some investigators suggest that age is important,[2,3,5,7,11] while others report inability to bear weight as significant.[2,4,7,9] A few authors find swelling important;[2–6,8,9,11] others report point tenderness[2,4,5,7–9,11] and/or bruising[2,4,5,11] as significant. Limited range of motion is reported as key to diagnosis by some researchers.[4,5,7,8] Data

Appendix

Confounding: A confounder must both be associated with the exposure and, independent of that exposure, be a risk factor for the disease.[15]

Relative risk (RR): Incidence of disease in exposed subjects relative to the incidence of disease in those non-exposed.[16]

RR = fracture in those exposed/fracture in those non-exposed

$$= \frac{a/a+b}{d/c+d}$$

Likelihood ratio (LR): The odds that a given level of test result would be expected in a patient with (in contrast to one without) the disease.[17]

LR for a positive "test" result = sensitivity/1 – specificity
LR for a negative "test" result = 1 – sensitivity/specificity

		fracture		
		+	–	
"test" i.e. age > 40 yrs	+	a	b	PPV = a/a+b
	–	c	d	NPV = d/c+d
		a/a+c	d/b+d	Prevalence = a+c/a+b+c+d

sensitivity specificity

Note: PPV = positive predictive value.
NPV = negative predictive value.

are not supplied to support most authors' conclusions. Only age greater than 40 years, inability to bear weight on examination, "point" or "bone" tenderness, and swelling are supported by published data and reported by more than one investigator (Tables 1 and 2).

Why are conclusions about the importance of different clinical criteria inconsistent? Although cross-sectional surveys using gold standards for comparison are appropriate designs for evaluating diagnostic tests, they cannot protect against numerous biases such as patient selection and referral-filter. While bias may be a problem in the studies reviewed, these studies are seriously flawed for other reasons. Most of them fail to describe adequately their patient population. Most fail to describe their examination technique or their definition of normal and abnormal results. With the exception of one study,[4] the examination was not tested for intra-or inter-observer variation. There is also tremendous variation in the examiner's training and experience from study to study.

To generalize the results of any study, a clear picture is required of the patients' characteristics, including age, sex, occupation, underlying illness, and previous injuries. Readers need to know how the examination is carried out, and how "normal" or "abnormal" is defined. The findings on clinical examination should be tested to determine intra-and inter-observer variation. Characteristics of the examiner can influence the examinee, as well as the examination, and the examiner's level of experience and training are important considerations. The setting in which a study is carried out can greatly influence the patients being examined, the examination, and the examiner.[14]

Most of the studies reviewed provided limited[2,4–7] or no[3,8,9,11,13] information identifying the patients studied. Failure to assemble an inception cohort can lead to inconsistent results and may explain the inconsistent conclusions of the studies reviewed. Varying degrees of examiner experience and training may also contribute to the inconsistent results.

A major problem in the studies reviewed is their failure to describe the technique of their examinations and to define the terms "normal" and "abnormal". Without information on the method of examination and specification of which findings are considered normal or abnormal, the results are of limited value. Without clear information on the examination, investigators cannot expect precision or reliability of "test" results. Failure to describe the examination could also explain the inconsistent conclusions reported.

The setting is important because it influences both the patients and the physicians. The setting is quite consistent in the studies reviewed, and this consistency is reflected in the stable fracture prevalence in emergency-department patients. (12%–17.5%). A change in setting may explain the aberrantly high fracture prevalence (possibly resulting from referral-filter bias) reported in one of the two radiology-department studies that examined X-ray films rather than patients.[3]

If we consider the previously published guides to determine whether to use a diagnostic test (Table 1), the studies reviewed fare poorly. Only one study used a blind comparison with a gold standard.[2] We do not know whether the samples include an appropriate spectrum of mild and severe injuries, but a few studies suggest that mild and severe cases are excluded through referral-filtre.[4,8,11] The setting alternates between the emergency department and the radiology department. The filters through which study patients pass are not described. Only one study tests for observer variation,[4] and the precision (reproducibility) of the examination is not evaluated. The definitions of normal and abnormal test results are not clearly defined. The use of specific criteria in the "cluster of clinical criteria" is considered in only one

study.[4] Tactics for carrying out the examination are not described in any report. Although the "utility" of specific criteria are evaluated in a few studies[2–4,9] and found to reduce significantly the rate of unnecessary referral to radiology, the criteria used to decide on referral are different in each study.

Conclusions

While numerous authors have expressed concern about the unnecessary use of X-rays for patients presenting to the emergency department with acute ankle injuries, published reports on the value of clinical criteria in diagnosis of fracture or sprain offer inconsistent conclusions and provide little alternative to radiographic assessment. The only criteria associated with the diagnosis of fracture in more than one study are age of more than 40 years, inability to bear weight on examination, "point" or "bone" tenderness, and swelling (Tables 1 & 2). The inconsistent conclusions from one study to the next appear to result from numerous biases, different patient populations studied, use of dissimilar examination techniques, varied interpretations of normal or abnormal results, and different groups of examiners.

Finding the same four criteria associated with fractures in different reports is encouraging and suggests that they may be used in diagnosis. This hypothesis should be tested in studies that control for the aforementioned flaws. Until consistent results are available from better design studies, no change in the rate of X-ray referral for patients presenting with acute ankle injuries should be expected. ■

References

1. Kennedy AF, Chambers LW. Fact book on health status (burden of illness) of citizens in the Regional Municipality of Hamilton-Wentworth. Hamilton: The Regional Municipality 1987 (August).

2. Dunlop MG, Beattie TF, White GK, et al. Guidelines for selective radiologic assessment of inversion ankle injuries. *Br Med J* 1986; 293:603–5.

3. Cockshott WP, Jenkin JK, Pui M. Limiting the use of routine radiography for acute ankle injuries. *Can Med Assoc J* 1983; 129:129–31.

4. Brand DA, Frazier WH, Kohlhepp WC, et al. Protocol For selecting patients With injured extremities who need X-rays. *New Eng J Med* 1982; 306:333–9.

5. Brooks SC, Potter BT, Rainey JB. Inversion injuries of the ankle: clinical assessment and radiographic review. *Br Med J* 1981; 282:607–8.

6. DeLacey G, Bradbrooke S. Rationalizing requests for X-ray examination of acute ankle injuries. *Br Med J* 1979; 1:1597–8.

7. Vargish T, Clarke WR, Young RA, et al. Ankle injury: indications for the selective use of X-rays. *Injury* 1983; 6:507–9.

8. Lettin AW. Diagnosis and treatment of sprained ankle. *Br Med J* 1963; 1:1056–60.

9. Gleadhill DN, Thomson JY, Simms P. Can more efficient use be made of X-ray examinations in the accident and emergency department? *Br Med J* 1987; 294:943–7.

10. Meyeda DV. Trauma: musculoskeletal injuries. Ankle and foot. In: Rosen P (ed). *Emergency medicine: concepts and clinical practice*. St. Louis, Mo., Washington, DC., Toronto: CV Mosby Company, 1988: vol 2:897–930.

11. Montague AP, McQuillan RF. Clinical assessment of apparently sprained ankle and detection of fracture. *Injury* 1985; 16:545–6.

12. Sackett DL, Haynes RB, Tugwell P. How to read a clinical journal. In: Sackett DL, Haynes RB, Tugwell P, eds. *Clinical epidemioloy: a basic science for clinical medicine*. Boston, MA., Toronto: Little, Brown and Company, 1985:293.

13. Svenson J. Need for radiographs in the acutely injured ankle. (Letter). *Lancet* 1988; 1:244–5.

14. Goldsmith CH. Sources of variation in measurement. Course document, McMaster University, Department of Clinical Epidemiology and Biostatistics, MS721, 1988. Hamilton: The Department, 1988.

15. Hennekens CH, Buring JE. Statistical association and cause-effect relationships. In: Mayrent SL, ed. *Epidemiology in medicine*. Boston MA., Toronto: Little, Brown and Company, 1987:35–7.

16. Hennekens CH, Buring JE. Measures of disease frequency and association. In: Mayrent SL, ed. *Epidemiology in medicine*. Boston, MA., Toronto: Little, Brown and Company, 1987:77–82.

17. Sackett DL, Haynes RB, Tugwell P. Interpretation of diagnostic data. In: Sackett DL, Haynes RB, Tugwell P, eds. *Clinical epidemiology: a basic science for clinical medicine*. Boston MA., Toronto: Little, Brown and Company, 1985:108–26.