3-1-2016

CT-based Structural Rigidity Analysis Is More Accurate Than Mirels Scoring for Fracture Prediction in Metastatic Femoral Lesions.

TA Damron
Department of Orthopedic Surgery, Upstate Bone and Joint Center, State University of New York Upstate Medical University

A Nazarian
Department of Orthopaedic Surgery, Center for Advanced Orthopaedic Studies, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

V Entezari
Department of Orthopaedic Surgery, Center for Advanced Orthopaedic Studies, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

C Brown
Department of Orthopedic Surgery, Upstate Bone and Joint Center, State University of New York Upstate Medical University, East Syracuse, NY

W Grant
Biostatistics, Department of Emergency Medicine, Upstate Medical University, Syracuse, NY

Clin Orthop Relat Res (2016) 474:643–651
DOI 10.1007/s11999-015-4453-0

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

SYMPOSIUM: 2014 MUSCULOSKELETAL TUMOR SOCIETY

CT-based Structural Rigidity Analysis Is More Accurate Than Mirels Scoring for Fracture Prediction in Metastatic Femoral Lesions

Timothy A. Damron MD, Ara Nazarian PhD, Vahid Entezari MD, Carlos Brown MD, William Grant EdD, Nathan Calderon MSc, David Zurakowski PhD, Richard M. Terek MD, Megan E. Anderson MD, Edward Y. Cheng MD, Albert J. Aboulafia MD, Mark C. Gebhardt MD, Brian D. Snyder MD, PhD

Published online: 14 July 2015

Abstract

Background Controversy continues regarding the appropriate assessment of fracture risk in long bone lesions affected by disseminated malignancy.

Questions/purposes The purpose of this ongoing Musculoskeletal Tumor Society-sponsored, multi-institutional prospective cross-sectional clinical study is to compare CT-based structural rigidity analysis (CTRA) with physician-derived Mirels scoring for predicting pathologic fracture in femoral bone lesions. We hypothesized CTRA would be superior to Mirels in predicting fracture risk within the first year based on (1) sensitivity, specificity, positive predictive value, and negative predictive value; (2) receiver operator characteristic (ROC) analysis; and (3) fracture prediction after controlling for potential confounding variables such as age and lesion size.

Methods Consented patients with femoral metastatic lesions were assigned Mirels scores by the individual enrolling orthopaedic oncologist based on plain radiographs and then underwent CT scans of both femurs with a phantom of known density. The CTRA was then performed. Between 2004 and 2008, six study centers performed CTRA on 125 patients. The general indications for this test were femoral metastatic lesions potentially at risk of fracture. The enrolling physician was allowed the choice of prophylactic stabilization or nonsurgical treatment, and the local treating oncology team along with the patient made this decision. Of those 125 patients, 78 (62%) did not undergo prophylactic stabilization and had followup sufficient for inclusion, which was fracture through the lesion within 12 months of CTRA, death within 12 months of CTRA, or 12-month survival after CTRA without fracture, whereas 15 (12%) were lost to followup and could not be studied here. The mean patient age was 61 years (SD, 14 years). There were 46 women. Sixty-four of the lesions were located in the proximal femur, 13

This study has been partially funded by a grant from the Musculoskeletal Tumor Society (co-Principal Investigators TAD and BDS) and grants from the Carol Baldwin Breast Cancer Research Foundation (TAD), Boston Children's Hospital Orthopaedic Surgery Foundation (BDS) and the National Institutes of Health (NIH T32 COMET Program (AN: AR055885) and NIH-LRP (AN: L30 AR056606).
All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.
Clinical Orthopaedics and Related Research® neither advocates nor endorses the use of any treatment, drug, or device. Readers are encouraged to always seek additional information, including FDA-approval status, of any drug or device prior to clinical use.
Each author certifies that his or her institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.
This work was performed at all authors' institutions.

T. A. Damron (✉), C. Brown
Department of Orthopedic Surgery, Upstate Bone and Joint Center, State University of New York Upstate Medical University, 6620 Fly Road, Suite 200, East Syracuse, NY 13057, USA
e-mail: damront@upstate.edu

A. Nazarian, V. Entezari, B. D. Snyder
Department of Orthopaedic Surgery, Center for Advanced Orthopaedic Studies, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

N. Calderon
Baylor College of Medicine, Houston, TX, USA

D. Zurakowski
Departments of Anesthesiology and Surgery, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA

R. M. Terek
Department of Orthopaedic Surgery, Rhode Island Hospital, Providence, RI, USA

were in the diaphysis, and four were distal. Osteolytic lesions prevailed (48 lesions) over mixed (31 lesions) and osteoblastic (15 lesions). The most common primary cancers were breast (25 lesions), lung (14 lesions), and myeloma (11 lesions). CTRA was compared with Mirels based on sensitivity/specificity analysis, ROC, and fracture prediction by multivariate analysis. For the CTRA, reduction greater than 35% in axial, bending, or torsional rigidities at the lesion was considered at risk for fracture, whereas a Mirels score of 9 or above, as suggested in the original manuscript, was used as the definition of impending fracture.

Results CTRA provided higher sensitivity (100% versus 66.7%), specificity (60.6% versus 47.9%), positive predictive value (17.6% versus 9.8%), and negative predictive value (100% versus 94.4%) compared with the classic Mirels definition of impending fracture (≥ 9), although there was considerable overlap in the confidence intervals. ROC curve analysis found CTRA to be better than the Mirels score regardless of what Mirels score cutoff was used. After controlling for potential confounding variables including age, lesion size, and Mirels scores, multivariable logistic regression indicated that CTRA was a better predictor of fracture (likelihood ratio test $= 10.49$, $p < 0.001$).

Conclusions CT-based structural rigidity analysis is better than Mirels score in predicting femoral impending pathologic fracture. CTRA appears to provide a substantial advance in the accuracy of predicting pathological femur fracture over currently used clinical and radiographic criteria.

Level of Evidence Level III, diagnostic study.

Introduction

One of the roles of the orthopaedic surgeon in the setting of disseminated bone involvement by metastatic disease, myeloma, and lymphoma is the prediction of fracture risk. Classically, this determination has been based on the clinician's assessment of the plain radiographs, but these criteria have proven lacking in both sensitivity and specificity [1–4, 6, 8, 9, 12–15, 17, 25, 28–30]. Although Mirels criteria (also based on plain radiographs) have improved sensitivity of fracture prediction in the appendicular skeleton, specificity has remained poor [5, 6, 10, 11, 21, 23, 24]. Hence, highly sensitive and more specific criteria are needed for the evaluation of long bone lesions at potential risk for fracture in the setting of disseminated malignancy.

The use of CT-based structural rigidity analysis (CTRA) has proven successful in fracture prediction for both benign bone lesions and in the spine for patients with disseminated malignancy [19, 26, 27]. However, to date, it has not been reported for use in prediction of long bone fracture risk in the latter setting.

The purpose of this ongoing Musculoskeletal Tumor Society-sponsored, multi-institutional prospective clinical study is to compare CTRA with physician-derived Mirels scoring for fracture risk prediction in femoral bone lesions. We hypothesized CTRA would be superior to Mirels in terms of the 1-year risk of fracture based on (1) sensitivity, specificity, positive predictive value, and negative predictive value; (2) receiver operator characteristic (ROC) analysis; and (3) fracture prediction after controlling for potential confounding variables such as age and lesion size.

Patients and Methods

The STROBE checklist for cross-sectional studies was used in manuscript preparation. The study design included prospective enrollment and data collection and cross-sectional analysis. Patients were enrolled, evaluated, and followed until 12 months or death, whichever came first. Initial enrollment began during July 2008 and continues to date. For the current analysis, patients enrolled from July 2008 through December 2012 were included.

The setting for enrollment involved institutions located within the continental United States and Canada. Data analysis was conducted at the Center for Advanced Orthopaedic Studies, Department of Orthopaedic Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA. Data contributors to the current report included the Departments of Orthopedic Surgery at each of the following institutions: Upstate Medical University, Syracuse, NY, USA; Rhode Island Hospital, Providence, RI, USA; Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA; the University of Minnesota Medical Center, Minneapolis, MN, USA; Sinai Hospital, Baltimore, MD, USA; and the Edwards Comprehensive Cancer Center, Huntington, WV,

M. E. Anderson, M. C. Gebhardt
Department of Orthopaedic Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

E. Y. Cheng
Department of Orthopaedic Surgery, University of Minnesota Medical Center, Minneapolis, MN, USA

A. J. Aboulafia
Department of Orthopaedic Surgery, Sinai Hospital, Baltimore, MD, USA

B. D. Snyder
Department of Orthopaedic Surgery, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA

W. Grant
Biostatistics, Department of Emergency Medicine, Upstate Medical University, Syracuse, NY, USA

USA. The McGill University Health Centre, Montreal, Quebec, Canada, and the Section of Orthopaedic Oncology at the University of Texas MD Anderson Cancer Center, Houston, TX, USA, also participated. Although each of these institutions has enrolled subjects in the study, only six institutions' enrolled subjects were able to be analyzed according to the inclusion criteria for this study: Upstate Medical University, Beth Israel Deaconess Medical Center, Rhode Island Hospital, University of Minnesota Medical Center, Edwards Comprehensive Cancer Center, and Sinai Hospital of Baltimore.

Each institution involved in the study obtained approval from their local institutional review board for the protection of human subjects. The local orthopaedic oncologist served as the local primary investigator (PI) and the point of access for potential study subjects. Initial eligibility criteria in the overall study included patients evaluated for femoral bone involvement by disseminated bone disease, primarily including metastatic carcinoma, myeloma, and lymphoma. Patients were included regardless of their prognosis and prior and expected treatment.

Between 2008 and 2012, six study centers were actively enrolling patients who met inclusion criteria. During this period, 125 patients consented and underwent CTRA. Of those, 32 (26%) underwent prophylactic fixation, leaving 93 patients for potential analysis here, and of those, 78 (62%) had followup sufficient for inclusion in this study, which was fracture through the lesion within 12 months of CTRA, death within 12 months of CTRA, or 12-month survival after CTRA without fracture, whereas 15 (12%) were lost to followup and could not be studied here. Consenting subjects underwent CT scans of both femurs from hip to knee using a three-rod phantom (0, 500, and 1000 hydroxyapatite mg/cm^3) provided to each local PI. The phantom allowed for the uniform conversion of pixel density values from DICOM units to volumetric bone mineral density across all enrolling sites to eliminate differences based on variations in scanners and techniques. The local PI provided the anonymized clinical data, including patient demographics (age, gender, primary cancer, bone site involved within the femur) and assigned Mirels scores as part of their initial evaluation for impending fracture. Each of the four elements of the Mirels scoring was individually recorded (1–3 points each for size, site, nature, and pain level) as was a total score of up to 12 points [23, 24]. The physician was then asked to assign a clinical plan based on the initial clinical and radiographic assessment before the CTRA evaluation. After the subjects' CTRA findings were reported back to the local PI, the local PI was asked to report on the plan based on the CTRA results and the actual plan that had been or was to be executed. The enrolling physician was allowed the choice of prophylactic stabilization or nonsurgical treatment, and the local treating oncology team along with the patient made this decision. Data regarding the effects of CTRA on decision-making are not part of the current analysis. However, the excluded patient characteristics are described subsequently.

Participants enrolled thus far include 125 patients (156 lesions) from six investigator sites (see Table 1 for patient demographics and lesion characteristics). For the purposes of this study, exclusion criteria were subjects who underwent prophylactic fixation within the first year and subjects alive with less than 1-year followup. Each subject included in the current analysis had one of the three following clinical endpoints: (1) completion of 12 months of clinical followup without fracture of the analyzed femur; (2) death within the 12 months of the index CTRA evaluation; or (3) fracture of the analyzed femur. Twelve months was chosen as the surveillance cutoff because the typically progressive natural history of metastatic disease beyond 1 year would potentially confound the analysis.

Of these 125 patients, 32 patients (40 of the total 156 lesions) were deemed by the local orthopaedic oncologist clinically to be at such high risk for pathological fracture that prophylactic stabilization was performed. Twenty-two lesions in 16 patients were classified as ineligible for inclusion, including two lesions in two patients who discontinued the study, 16 lesions in 10 patients (one of whom had a contralateral femoral lesion that fractured early and was included in the final analysis) who were lost to followup before 12 months, and four lesions in four patients for which no CTRA was able to be performed. Thus, 78 patients (94 lesions) met inclusion criteria: 41 patients (48 lesions) completed 12 months of followup without fracture, six patients (6% of lesions) sustained a pathologic fracture through the lesion, and 31 patients (40 lesions) died without sustaining a femur fracture. The mean age was 60 years (range, 22–82 years). Forty-six were females. Sixty-four lesions were located in the proximal femur, 13 in the diaphysis, and four in the distal femur. Osteolytic lesions prevailed (48 lesions) over mixed (31 lesions) and osteoblastic (15 lesions). The most common primary cancers were breast (25 lesions), lung (14 lesions), and myeloma (11 lesions). The Mirels score on average for these 94 lesions was 8.7 (score of 5 in three lesions, score of 6 in eight lesions, score of 7 in eight lesions, score of 8 in 24 lesions, score of 9 in 23 lesions, score of 10 in 14 lesions, score of 11 in eight lesions, and score of 12 in six lesions). Forty of the 94 (43%) lesions were predicted to be at risk for fracture by the CTRA.

Because the exclusion of the 40 lesions that were prophylactically stabilized has the potential for biasing the study, the nature of these excluded lesions is described here for comparison to the patients included for analysis. Mean age of these excluded patients was 65 years (range,

Table 1. Patient demographics and lesion characteristics

Patient demographics			
Eligible patients, number		Ineligible patients, number	
78		47	
Age (years)		Age (years)	
Mean (SD)	60 (13.6)	Mean (SD)	63.1 (12.6)
Median (range)	61.6 (23–82)	Median (range)	63 (37–88)
Sex, number (%)		Sex, number (%)	
Female	46 (59)	Female	21 (44.7)
Male	31 (39.7)	Male	25 (53.2)
Unknown	1 (1.3)	Unknown	1 (2.1)
CTRA prediction of fracture, number (%)		CTRA prediction of fracture, number (%)	
No risk	47 (60.3)	No risk	21 (44.7)
Risk	31 (39.7)	Risk	24 (51.1)
		No SRA	2 (4.2)
Disposition, number (%)		Disposition, number (%)	
Completed 1 year	41 (52.6)	Stabilized	32 (68.1)
Death without fracture	31 (39.7)	Lost to followup	11 (23.4)
Fracture	6 (7.7)	Discontinued	4 (8.5)
Primary cancers, λ, patients, number (%)		Primary cancers, λ, patients, number (%)	
λ	14	λ	13
Anal/colorectal	1 (1.3)	Anal/colorectal	1 (2.1)
Bladder	1 (1.3)	Bladder	1 (2.1)
Breast	22 (28.2)	Breast	10 (21.3)
Histiocytosis	1 (1.3)	Gastrointestinal	1 (2.1)
Liposarcoma	1 (1.3)	Lung	12 (25.5)
Liver	1 (1.3)	Lymphoma	3 (6.5)
Lung	10 (12.8)	Myeloma	5 (10.6)
Lymphoma	1 (1.3)	Pheochromocytoma	1 (2.1)
Melanoma	1 (1.3)	Prostate	3 (6.5)
Myeloma	10 (12.8)	Renal	4 (8.5)
Prostate	6 (7.7)	Salivary gland	1 (2.1)
Renal	10 (12.8)	Tonsils	1 (2.1)
Thyroid	1 (1.3)	Unknown	4 (8.5)
Unknown	12 (15.4)		

CTRA = CT-based structural rigidity analysis.

45–88 years). Twenty were women. The proximal femur was the location for 29, the shaft for four, and the distal femur for four. Thirty of these 40 prophylactically stabilized lesions were lytic, nine were mixed, and one was blastic. The most common primary cancers were lung (10 lesions), breast (nine lesions), myeloma, renal, and lymphoma (five lesions each). The Mirels score on average for these 40 lesions was 9.84 (score of 7 in two lesions, 8 in seven lesions, 9 in seven lesions, 10 in eight lesions, 11 in seven lesions, and 12 in seven lesions). Twenty-three of the 40 (58%) prophylactically stabilized lesions were predicted to be at risk for fracture by CTRA.

All of the six fractures occurred within 4 months of the CTRA evaluation and were correctly predicted by CTRA reductions of $\geq$ 35% and all had Mirels scores $\geq$ 8 points (8, 8, 10, 12, 12, 12). Five of the six fractures occurred through lytic proximal femoral lesions and one through a diaphyseal mixed lesion. Although the study protocol dictated clinical and radiographic followup every 4 months during the first year after the baseline CTRA to monitor for progression, at least one of the fracture cases was noted to have progressed to a larger size on the fracture films. In the remainder of the fracture cases, it was difficult to appreciate clear radiographic progression of the lesion or there was simply no documentation of it having occurred. Over the intervening year of followup, no additional fractures occurred.

With respect to the primary variables tested, both CTRA and Mirels definitions of impending fractures were

evaluated. For the CTRA, the CT scan data for each subject were analyzed centrally at the Center for Advanced Orthopaedic Studies at Beth Israel Deaconess Medical Center and Harvard Medical School using the previously developed, validated, and reported CTRA technique [19, 26, 27]. This involved incorporating the density of the bone (derived from the comparison with the known density of the phantom) and the cross-sectional areal characteristics at the point of maximal weakness to derive bending, torsion, and axial rigidity. These parameters were then compared with a gender- and size-matched femur from a normative CT database because patients in this oncology population were considered to potentially have occult bilateral involvement. For the CTRA, reduction greater than 35% in axial, bending, or torsional rigidities at the lesion was considered at risk for fracture. This cutoff was based on the findings of the original study in which fracture prediction was validated using CTRA for benign bone lesions [19, 26, 27]. Hence, the definition of an impending pathologic fracture based on CTRA for the purposes of this study was a $> 35\%$ reduction in any of the three loading parameters.

For the Mirels scoring provided by the local orthopaedic oncologist attending PI, three different definitions of impending pathologic fracture were analyzed based on the findings in Mirels' original manuscript [24]. In that study, patients with a total of 8 points were felt to be at "borderline" risk and those with 9 points or above were felt to be at increased risk and therefore formed their basis for the original classic Mirels definition of an impending pathologic fracture [23, 24]. In this study, we examined the following three definitions of impending pathologic fracture: (1) Mirels total score ≥ 8; (2) Mirels total score ≥ 9; and (3) Mirels total score ≥ 10.

The outcome variable analyzed was occurrence of a pathologic fracture and this was considered to have occurred if the local PI documented it within the first year of followup. A pathologic fracture was considered not to have occurred if the local PI reported followup of 12 months without fracture or death of the patient without preceding fracture.

Statistical analysis was accomplished under the guidance of two biostatisticians (WG, DZ). In this cross-sectional study, descriptive statistics were used in the initial analysis of the data to describe the patient population as well as the group of patients who underwent prophylactic stabilization [16]. For the first hypothesis, sensitivity, specificity, positive (PPV) and negative predictive values (NPV) were determined for each of the designated Mirels score categories as well as CTRA using 2×2 tables (Table 2). The calculated sensitivity, specificity, PPV, and NPV were then compared by calculation of 95% confidence intervals (CIs). For the second hypothesis, ROC curve analysis was used to assess the comparative prognostic accuracy of the CTRA and Mirels using the areas under the curve (AUC). Simply stated, accuracy of any test depends on its ability to separate each patient/lesion into those with and without the condition in question, and accuracy is measured by the area under the ROC curve. A perfect test is represented by an area of 1, whereas a meaningless test is represented by an area of 0.5. The traditional academic point system for classifying a diagnostic test considers the test "excellent" if the AUC is 0.90 to 1.0, "good" if the AUC is 0.80 to 0.90, "fair" if the AUC is 0.70 to 0.80, "poor" if the AUC is 0.60 to 0.70, and "failure" if the AUC is 0.5 to 0.6. Statistical comparison of the AUC for CTRA versus Mirels score was done using the two-sided Hanley, McHeil approach [7]. For the third hypothesis, multivariable logistic regression using backward selection was applied to confirm whether continuous CTRA is a significant predictor ($p < 0.05$) of fracture independent of patient age, lesion size, and continuous Mirels score with the likelihood ratio test used to assess significance. Data were arrayed for verification and final specification using Microsoft Excel (Microsoft, Inc, Redmond, WA, USA), and all statistical analyses were performed using MedCalc (Version 12.7; MedCalc Software, Ostend, Belgium) and IBM SPSS Statistics (Version 22.0; IBM, Armonk, NY, USA).

Table 2. Two-by-two tables for CTRA and each of the Mirels cutoffs

A		
CTRA	Fracture	
	Yes	No
Yes	6	35
No	0	53

B		
Mirels ≥ 9	Fracture	
	Yes	No
Yes	4	47
No	2	41

C		
Mirels ≥ 8	Fracture	
	Yes	No
Yes	6	68
No	0	20

D		
Mirels ≥ 10	Fracture	
	Yes	No
Yes	4	23
No	2	65

CTRA = CT-based structural rigidity analysis.

Table 3. Sensitivity, specificity, and positive and negative predictive values according to prediction tools

Statistical parameter	CTRA	Mirels (≥ 8)	Mirels (≥ 9)	Mirels (≥ 10)
Sensitivity	100%	100%	67%	67%
(95% CI)	(54%–100%)	(54%–100%)	(22%–96%)	(22%–96%)
Specificity	61%	24%	48%	78%
(95% CI)	(48%–72%)	(15%–36%)	(36%–60%)	(66%–87%)
PPV	18%	10%	10%	20%
(95% CI)	(7%–35%)	(4%–21%)	(2%–23%)	(6%–44%)
NPV	100%	100%	94%	97%
(95% CI)	(92%–100%)	(81%–100%)	(81%–99%)	(88%–100%)

CTRA = CT-based structural rigidity analysis; CI = confidence interval; PPV = positive predictive value; NPV = negative predictive value.

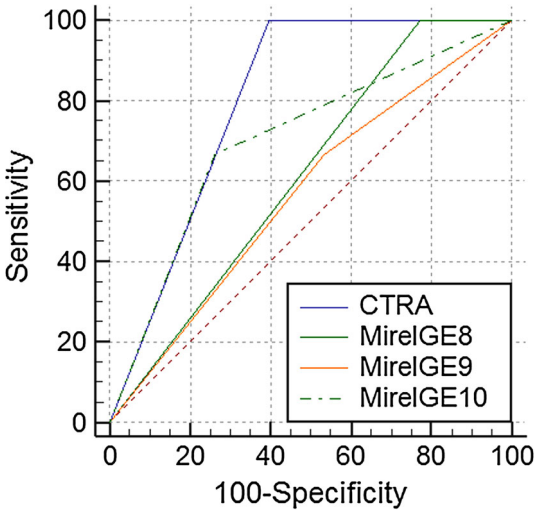

Fig. 1 ROC curves comparing the four fracture prediction tools show the greatest AUC for CTRA and the least for Mirels > 9. MirelsGE8 = Mirels total score of ≥ 8; MirelsGE9 = Mirels total score of ≥ 9; MirelsGE10 = Mirels total score of ≥ 10.

Results

Comparison of the sensitivity, specificity, PPV, and NPV showed that all parameters were greater for CTRA when compared with Mirels score ≥ 9 (Table 3). Both CTRA and a Mirels score ≥ 8 were 100% sensitive for predicting a pathological fracture, but a Mirels score ≥ 9 or ≥ 10 predicted only four of the six (67%) fractures that occurred. Specificity was greater for CTRA (61%) than Mirels of ≥ 9 (48%). Positive predictive value was greater for CTRA (18%) compared with Mirels score ≥ 9 (9%). Negative predictive value was also greater for CTRA (100%) compared with Mirels score ≥ 9 (94%). The small number of fractures resulted in wide and overlapping CIs.

Receiver operating characteristic analysis found CTRA to be superior to Mirels scoring regardless of what Mirels score cutoff was used (Fig. 1). Receiver operating characteristic curves (Table 4) indicated that CTRA provided

Table 4. Area under the receiver operating characteristic curve using CTRA and Mirels scoring to predict fracture

Impending fracture criteria	AUC	SE[7]	95% CI*
CTRA	0.801	0.0262	0.706–0.876
Mirels ≥ 8	0.614	0.0225	0.508–0.712
MIrels ≥ 9	0.566	0.109	0.460–0.668
MIrels ≥ 10	0.703	0.108	0.600–0.793

* Binomial exact; CTRA = CT-based structural rigidity analysis; AUC = area under the curve; SE = standard error; CI = confidence interval.

"good" accuracy (AUC = 0.801; 95% CI, 0.706–0.876; p = 0.014), whereas a Mirels score ≥ 9 was no better than chance for predicting the occurrence of a pathological fracture (AUC = 0.56; 95% CI, 0.460–0.668). Mirels score ≥ 10 provided "fair" accuracy (AUC = 0.740; 95% CI, 0.505–0.978; p = 0.049), and a Mirels score ≥ 8 provided "poor" accuracy (AUC = 0.614). Statistically comparing the AUC for each, CTRA was more accurate for predicting a pathological fracture than Mirels score ≥ 9 (p = 0.03) and also for Mirels score ≥ 8 (p < 0.0001).

After controlling for potential confounding variables including age, lesion size, and Mirels scores, multivariable logistic regression indicated that continuous CTRA was a better predictor of fracture (likelihood ratio test = 10.49, p < 0.001). Age (p = 0.83), lesion size (p = 0.50), and continuous Mirels score (p = 0.16) were not predictors of fracture in the multivariate analysis.

Discussion

Metastatic disease to bone affects between 280,000 and 600,000 patients annually in the United States [20]. The primary cancers that account for 70% to 80% of this disease are breast and lung in women and prostate and lung in men. The femur and humerus are the most common

appendicular sites involved, and the primary skeletal-related event is pathologic fracture. In the appendicular skeleton affected by metastatic disease, one of the major orthopaedic challenges is the prediction of fracture risk [6, 8, 29], because prophylactic stabilization may be used to prevent imminent fracture [30]. In the current study, the hypotheses that CTRA would be a better predictor of fracture risk than Mirels score based on (1) sensitivity, specificity, PPV, and NPV; (2) ROC analysis; and (3) fracture prediction after controlling for potential confounding variables such as age and lesion size were shown to hold. Furthermore, ROC analysis suggests that CTRA is good in terms of accuracy and is significantly better than the Mirels classic cutoff of 9, which failed to be better than chance at fracture prediction in this study. To our knowledge, this is the first prospective, multi-institutional evaluation of Mirels testing and also the first demonstration of an improvement on the now over 25-year-old Mirels testing scheme for prediction of impending pathologic fracture in long bones.

One of the main limitations of this study was the fact that the local site investigators were allowed to treat independent of the CTRA results. Consequently, 32 patients (40 lesions) were excluded because, regardless of the CTRA results, the investigators chose to proceed with prophylactic stabilization procedures, effectively eliminating these patients from consideration for observation to determine their risk for fracture. Those patients were treated surgically based on the surgeon's preference and personal experience, Mirels scoring, or some combination. The patients who were treated prophylactically and therefore excluded from analysis had higher mean Mirels scores (9.8 versus 8.7) and were more likely to be at risk for fracture by CTRA (58% versus 43%), both a reflection of the means by which the physicians made their determination of the need for surgery. However, all other demographic features were similar, suggesting that this excluded subset was simply in general a higher risk subset of the overall group. This elimination of the patients at highest risk for fracture no doubt contributed to the very low fracture rate among the patients in the study population (six of 94 lesions; 78 patients). The effect of this bias on current results cannot be predicted with certainty but is likely to have lowered the PPV of CTRA. Because the accuracy of CTRA in femoral metastatic disease remained unproven before the current study, the investigators considered it unethical to require the participating surgeons to treat based on CTRA results alone. Subsequent prospective analysis using a study design requiring the surgeons to more closely adhere to the results of the CTRA would be desirable. Although no record was kept of the specific numbers of patients with whom study participation was offered but declined, the clinician with the largest enrollment (TAD), accounting for approximately two-thirds of the patients and lesions enrolled, had less than 10 additional patients lost for this reason. Hence, based on this high percentage of enrolled patients in the institution with the largest enrollment, the applicability to the broader population of patients with metastatic disease would not seem to have been compromised. The loss of another 13% of patients (12 of 93) to either discontinuation of the study or loss to followup introduces further potential bias. The current results are also restricted to patients with disseminated malignancy affecting the femur and cannot necessarily be extrapolated to other long bones, although subsequent evaluation of the humerus would appear warranted. A small number of patients (four) could not be included in the analysis because of the presence of contralateral lesions and the unavailability of a suitable control femur to compare. This limitation of the technique should be less relevant as the normative proximal femoral CT database continues to accrue. One must also remember that CTRA does not reflect biology or potential for changes over time, either positive or negative. Hence, it cannot distinguish between necrotic and viable bone and tumor and nontumor tissue. It also cannot predict whether bone will heal with treatment or whether tumor will progress. Hence, patients who are predicted not to be at risk of fracture should be followed clinically and radiographically to monitor for changes. Finally, the study design did not allow for any means by which to establish intraobserver or interobserver consistency of the Mirels scoring. However, our own work has shown interobserver error to be acceptable [5, 11]. In addition, the study design we chose highlighted the comparison between a practicing orthopaedic oncologist's assigned Mirels scoring, as would be done in his or her clinical practice, as compared with the CTRA.

The rationale for predicting impending fracture is that there is treatment that will reliably prevent the skeletal related event of pathologic fracture. In the femur, that procedure is typically an intramedullary stabilization procedure. Preventing fracture reliably spares the patient not only the morbidity of pathologic fracture, including both the immediate pain and need for emergent hospitalization, but also subsequent potential for delayed or nonunion, failure of internal fixation, and potential need for subsequent reoperation. Furthermore, those patients who undergo prophylactic stabilization have on average lower blood loss, shorter hospitalization, a higher rate of home discharge (versus extended care facility), and a threefold higher rate of aid-free ambulation compared with patients who undergo pathologic fracture fixation [30]. In this population of unfortunate patients with a limited lifespan, making an accurate determination of fracture risk is crucial not only to provide adequate and timely surgical

management to those patients with truly imminent fractures, but also to avoid unnecessary surgery in those who can be treated with simpler means, including systemic treatment, bisphosphonates, and/or radiotherapy.

Historic means of fracture prediction, based on clinical and plain radiographic criteria, are inaccurate and poorly specific [1–4, 6, 8, 9, 12–15, 17, 25, 28–30]. The Mirels scale, which has been widely quoted and variably used for prediction of impending pathologic fracture, is based on a four-factor (pain, size, site, nature), 12-point scale relying on subjective assessment of pain and plain radiographic bone involvement [23, 24]. Although this scale is reported to be highly sensitive, groups of reported patients with the suggested definition of pathologic fracture (9 of 12 points) actually fracture in only approximately one-third of cases [5, 6, 11, 21]. Hence, up to two-thirds of patients predicted as being at increased risk by Mirels criteria do not fracture and potentially receive unnecessary surgery. Other studies suggest Mirels is less predictive than other criteria but agree that a more specific predictor accounting for comorbidities and prognosis is needed [10, 29]. Although finite element modeling has been used successfully for accurate fracture risk prediction in the laboratory, it has yet to be applied in human subjects and suffers from the potential to be a cumbersome, slow means of evaluation for individual patients in real time [18, 22].

Although at this time CTRA as an analysis software tool is not widely available across the country, any CT scanner may be used to supply data for analysis. To date, all scans used for the reported study were done with a phantom containing known bone densities, but work is underway to develop techniques to use CT scans without the phantoms. Still, the data currently have to be transferred to the Beth Israel Deaconess Laboratory for analysis on the software, from where a report is generated and sent to the institution generating the CT data. The time for processing the data from generation of the CT data to receipt of the final report is variable, depending on the availability of laboratory staff at Beth Israel Deaconess to run the analysis, the number of pending studies to be analyzed, and the potential need for obtaining normative data from the database for direct comparison. The authors hope to make the technique more widely available using the results of this study to establish its validity.

Based on the findings of the current study, in the setting of femoral metastatic bone disease, CTRA using a $\geq$ 35% reduction in rigidity compared with the contralateral (or appropriately matched) normal femur as the threshold improves sensitivity, specificity, and PPV, and NPV over Mirels using the classic threshold of $\geq$ 9 points as the definition of an impending pathologic fracture. Furthermore, CTRA maintains 100% sensitivity, which was not true when the Mirels cutoff was raised to $\geq$ 10 as the definition of impending fracture. Hence, in this analysis, CTRA with a threshold of 35% reduction in rigidity appears to be a sensitive, specific, accurate, and durable tool to predict fracture risk in patients with femoral metastatic disease. Although the use of a Mirels score of $\geq$ 10 points yielded marginally better PPV and specificity with nearly the same NPV compared with CTRA, the loss of sensitivity is notable (67% versus 100%), lowering the value of Mirels $\geq$ 10 as a screening tool. Overall, based on the ROC analysis, the accuracy of CTRA was better than that of Mirels.

With this evidence in hand, further enrollment in the current study is being encouraged to improve the robustness of the statistics. In addition, studies are underway to increase the more widespread use of CTRA by making it available to institutions without the need for a phantom to standardize the CT data with each patient evaluation.

Acknowledgments We thank John A. Hipp PhD, who was integral in the development of the CTRA technique, Felix Cheung MD, Patrick Lin MD, and Robert Turcotte MD, who were active in getting the study through the institutional review board at their respective institutions, and the members of the Musculoskeletal Tumor Society and its leadership for the initial and ongoing support of this multi-institutional study in part through contributions to the Orthopaedic Education and Research Foundation.

References

1. Beals R, Lawton G, Snell W. Prophylactic internal fixation of the femur in metastatic breast cancer. *Cancer.* 1971;28:1350–1354.
2. Bunting R. Pathologic fracture risk in rehabilitation of patients with bony metastases. *Clin Orthop Relat Res.* 1985;192:222–227.
3. Cheng D, Seitz C, Eyre H. Nonoperative management of femoral, humeral, and acetabular metastases in patients with breast carcinoma. *Cancer.* 1980;45:1533–1537.
4. Cumming D, Cumming J, Vince A, Benson R. Metastatic bone disease: the requirement for improvement in a multidisciplinary approach. *Int Orthop.* 2009;33:493–496.
5. Damron TA, Morgan H, Prakash D, Grant W, Aronowitz J, Heiner J. Critical evaluation of Mirels' rating system for impending pathologic fractures. *Clin Orthop Relat Res.* 2003;415(Suppl): S201–207.
6. Damron TA, Ward WG. Risk of pathologic fracture: assessment. *Clin Orthop Relat Res.* 2003;415(Suppl):S208–211.
7. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44:837–845.
8. Eastley N, Newey M, Ashford RU. Skeletal metastases–the role of the orthopaedic and spinal surgeon. *Surg Oncol.* 2012;21:216–222.
9. Eilkins R, Sim F, Springfield D. Metastatic disease of the femur. *Orthopedics.* 1992;15:621–630.
10. El-Husseiny M, Coleman N. Inter- and intra-observer variation in classification systems for impending fractures of bone metastases. *Skeletal Radiol.* 2010;39:155–160.
11. Evans AR, Bottros J, Grant W, Chen BY, Damron TA. Mirels' rating for humerus lesions is both reproducible and valid. *Clin Orthop Relat Res.* 2008;466:1279–1284.

12. Fidler M. Incidence of fracture through metastases in long bones. *Acta Orthop Scand.* 1981;52:623–627.
13. Haidukewych GJ. Metastatic disease around the hip: maintaining quality of life. *J Bone Joint Surg Br.* 2012;94(Suppl A):22–25.
14. Harrington KD. New trends in management of lower extremity metastases. *Clin Orthop Relat Res.* 1982;169:53–61.
15. Hayes WC, Piazza SJ, Zysset PK. Biomechanics of fracture risk prediction of the hip and spine by quantitative computed tomography. *Radiol Clin North* Am. 1991;29:1–18.
16. Hulley SB, Cummings SR, Browner WS, Grady W, Hearst N, Newman TB. *Designing Clinical Research: An Epidemiologic Approach*. 2nd ed. Philadelphia, PA, USA: Lippincott & Williams; 2001.
17. Keene JS, Sellinger DS, McBeath AA, Engber WD. Metastatic breast cancer in the femur: a search for the lesion at risk of fracture. *Clin Orthop Relat Res.* 1986;203:282–288.
18. Keyak JH, Kaneko TS, Rossi SA, Pejcic MR, Tehranzadeh J, Skinner HB. Predicting the strength of femoral shafts with and without metastatic lesions. *Clin Orthop Relat Res.* 2005;439:161–170.
19. Leong NL, Anderson ME, Gebhardt MC, Snyder BD. Computed tomography-based structural analysis for predicting fracture risk in children with benign skeletal neoplasms: comparison of specificity with that of plain radiographs. *J Bone Joint Surg Am.* 2010;92:1827–1833.
20. Li S, Peng Y, Weinhandl ED, Blaes AH, Cetin K, Chia VM, Stryker S, Pinzone JJ, Acquavella JF, Arneson TJ. Estimated number of prevalent cases of metastatic bone disease in the US adult population. *Clin Epidemiol.* 2012:4:87–93.
21. Mac Niocaill RF, Quinlan JF, Stapleton RD, Hurson B, Dudeney S, O'Toole GC. Inter- and intra-observer variability associated with the use of the Mirels' scoring system for metastatic bone lesions. *Int Orthop.* 2011;35:83–86.
22. Mann KA, Lee J, Arrington SA, Damron TA, Allen MJ. Predicting distal femur bone strength in a murine model of tumor osteolysis. *Clin Orthop Relat Res.* 2008;466:1271–1278.
23. Mirels H. Metastatic disease in long bones. *Clin Orthop Relat Res.* 1989;249:256–264.
24. Mirels H. Metastatic disease in long bones: a proposed scoring system for diagnosing impending pathologic fractures. 1989. *Clin Orthop Relat Res.* 2003;415(Suppl):S4–13.
25. Parrish F, Murray J. Surgical treatment for secondary neoplastic fractures. A retrospective study of 96 patients. *J Bone Joint Surg Am.* 1970;52:665–686.
26. Snyder BD, Cordio MA, Nazarian A, Kwak SD, Chang DJ, Entezari V, Zurakowski D, Parker LM. Noninvasive prediction of fracture risk in patients with metastatic cancer to the spine. *Clin Cancer Res.* 2009;15:7676–7683.
27. Snyder BD, Hauser-Kara DA, Hipp JA, Zurakowski D, Hecht AC, Gebhardt MC. Predicting fracture through benign skeletal lesions with quantitative computed tomography. *J Bone Joint Surg Am.* 2006;88:55–70.
28. Thompson R. Impending fracture associated with bone destruction. *Orthopedics.* 1992;15:547–550.
29. Van der Linden YM, Dijkstra PD, Kroon HM, Lok JJ, Noordijk EM, Leer JW, Marijnen CA. Comparative analysis of risk factors for pathological fracture with femoral metastases. *J Bone Joint Surg Br.* 2004;86:566–573.
30. Ward WG, Holsenbeck S, Dorey FJ, Spang J, Howe D. Metastatic disease of the femur: surgical treatment. *Clin Orthop Relat Res.* 2003;415(Suppl):S230–244.