Toward Better Use of Bone Scans Among Men With Early-stage Prostate Cancer

Selin Merdan, Paul R. Womble, David C. Miller, Christine Barnett, Zaojun Ye, Susan M. Linsell, James E. Montie, and Brian T. Denton

OBJECTIVE	To evaluate the performance of published guidelines compared with that of current practice for radiographic staging of men with newly diagnosed prostate cancer.
MATERIALS AND METHODS	Using data from the Michigan Urological Surgery Improvement Collaborative clinical registry, we identified 1509 men diagnosed with prostate cancer from March 2012 through June 2013. Clinical data included age, prostate-specific antigen (PSA) level, Gleason score (GS), clinical trial stage, number of biopsy cores, and bone scan (BS) results. We then fit a multivariate logistic regression model to examine the association between clinical variables and the occurrence of bone metastases. Because some patients did not undergo BS, we used established methods to correct for verification bias and estimate the diagnostic accuracy of published guidelines.
RESULTS	Among 416 men who received a BS, 48 (11.5%) had evidence of bone metastases. Patients with bone metastases were older, with higher PSA levels and GS (all $P < .05$). In multivariate analyses, PSA ($P < .001$) and GS ($P = .004$) were the only independent predictors of positive BS. Guidelines from the American Urological Association and the National Comprehensive Cancer Network demonstrated similar performance in detecting bone metastases in our population, with fewer negative study results than those of the European Association of Urology guideline. Applying the American Urological Association recommendations (ie, image when PSA level >20 ng/mL or GS ≥8) to current clinical practice, we estimate that <1% of positive study results would be missed, whereas the number of negative study results would be reduced by 38%.
CONCLUSION	Based on current practice patterns, more uniform application of existing guidelines would ensure that BS is performed for almost all men with bone metastases, while avoiding many negative imaging studies. UROLOGY 84: 793−798, 2014. © 2014 Elsevier Inc.

T he optimal treatment of men with newly diagnosed prostate cancer depends on the stage of disease at diagnosis. An important aspect of clinical staging is the detection of metastases, including spread to the bone. Accordingly, the performance of a radionuclide bone scan (BS) is pivotal to the diagnostic evaluation and treatment planning for some men with prostate cancer. At the same time, however, these studies are expensive and time consuming, and the overall yield (ie, likelihood of detecting metastases) is quite low for men with low- or intermediate-risk cancers. For these and other reasons, many express concern about the well-established and persistent variation in the use of staging BSs, including potentially unnecessary testing in many men at low risk for metastatic disease and the absence of testing for some men with high-risk cancers. Underscoring the significance of this issue, the American Urological Association (AUA) recently identified the avoidance of BSs in men with low-risk prostate cancer as its number 1 priority for the national Choosing Wisely program.[1]

Nonetheless, although existing clinical guidelines are clear about omitting BSs in men with low-risk cancers,[2-5] there is no consensus regarding the optimal use of imaging for men with higher risk but still clinically localized tumors. The net effect is that imaging practice patterns continue to vary widely, implying immediate opportunities

to improve value in this area of prostate cancer care. Many believe that an important next step in this process is to move away from recommendations based on the risk of recurrence after treatment (eg, D'Amico risk groups) and toward the identification and implementation of imaging criteria that most accurately forecast a positive study finding that would actually change clinical decision making.

In this context, we sought to identify predictors of a positive BS result in a population-based sample of men with newly diagnosed prostate cancer from the diverse academic and community practices in the Michigan Urological Surgery Improvement Collaborative (MUSIC).[6] We examined the association between routine clinical variables (eg, prostate-specific antigen [PSA], clinical trial stage) and the occurrence of bone metastases. Because not all men with newly diagnosed prostate cancer underwent a staging BS, we used an established method to correct for verification bias to evaluate the accuracy of published imaging guidelines for detection of bone metastases in this real-world patient population. We also estimated the percentage of patients with positive study results that would be missed, total percentage of negative study results, and change in the total number of BSs that can be expected from successful implementation of each clinical guideline compared with current practice.

MATERIALS AND METHODS

Study Population and Clinical Variables
Established in 2011 with funding from Blue Cross Blue Shield of Michigan,[7] MUSIC is a consortium of 32 practices from throughout Michigan (including >75% of urologists in the state) that aims to improve the quality and cost efficiency of care provided to men with prostate cancer. Each practice involved in MUSIC obtained an exemption or approval for participation from a local institutional review board.

All participating practices employ trained clinical abstractors to review the medical record and enter standardized data elements into a Web-based clinical registry. Included in the registry are all men seen in participating practices for prostate biopsy or newly diagnosed prostate cancer. The registry contains detailed clinical and demographic information, including patient age, serum PSA levels at diagnosis, clinical trial stage, biopsy Gleason score (GS), total number of biopsy cores, number of positive cores, and the receipt and results of staging BS ordered by the treating urologist. This analysis included 1519 patients with newly diagnosed prostate cancer seen at 19 practices in Michigan from March 2012 through June 2013.

Primary Outcome
The primary outcome variable for this analysis was the occurrence of a positive BS result. The final classification of a study result as positive or negative was determined by the local data abstractor, treating urologist, and clinical champion in each practice, according to the established criteria for MUSIC. For a sample of patients, BS results were also validated by members of the MUSIC Coordinating Center during regular on-site data audits performed at each participating practice.

Table 1. Patient characteristics

Variables	All Patients Without BS (N = 1103)	All Patients With BS (N = 416)	P Value
Age, y			.02
Mean (median)	64.2 (64.4)	68.2 (67.7)	
Range	40.4-95.8	41.8-90.5	
Clinical stage, n (%)			<.0001
T1	881 (79.9)	216 (51.9)	
T2	214 (19.4)	173 (41.6)	
T3/4	8 (0.7)	27 (6.5)	
PSA level, ng/mL			.003
Mean (median)	8.0 (5.2)	61.8 (7.7)	
Range	0.2-620.8	0.4-6873.4	
Ln(PSA + 1)			<.0001
Mean (median)	1.9 (1.8)	2.5 (2.2)	
Range	0.2-6.4	0.3-8.8	
PSA level, ng/mL, n (%)			—
≤10	1018 (92.3)	247 (59.4)	
10.1-20	58 (5.3)	81 (19.5)	
20.1-50	10 (0.9)	45 (10.8)	
50.1-100	12 (1.1)	20 (4.8)	
>100	5 (0.5)	23 (5.5)	
Biopsy Gleason score sum, n (%)			<.0001
≤6	488 (44.2)	33 (7.9)	
3 + 4	439 (39.8)	105 (25.2)	
4 + 3	137 (12.4)	58 (13.9)	
8-10	39 (3.6)	220 (52.9)	
Biopsy cores taken, n			.50
Mean (median)	12.5 (12.0)	12.9 (12.0)	
Range	4-82	1-78	
Positive cores, n			.0004
Mean (median)	3.2 (3.0)	6.3 (6.0)	
Range	0-20	1-16	
Positive cores, %			<.0001
Mean (median)	26.4 (21.1)	51.2 (50.0)	
Range	0-100	3.1-100	

BS, bone scan; PSA, prostate-specific antigen.

Statistical Analyses
As a first step, we compared the clinical and pathologic characteristics of patients with or without BSs. Differences between these 2 groups of patients in medians for quantitative variables and differences in distributions for categorical variables were compared using the Mann-Whitney U and chi-square tests, respectively. We next performed univariate and multivariate analyses to examine the association between a positive BS result and several routinely available clinical variables in the sample of patients who received a staging BS. The variables included in the models were the age at diagnosis, a natural logarithm of PSA + 1 [ln(PSA + 1)], biopsy GS (≤3 + 4 vs 4 + 3 vs 8-10), clinical stage (T1 vs T2 vs T3/4), and the percentage of positive biopsy cores (defined as the number of cores containing cancer over total number of cores sampled). The selection of these variables was based on both previously published studies and clinical experience. All statistical testing was 2-sided with a significance level of .05 and was performed using computerized software (SAS, version 9.3, SAS Institute Inc, Cary, NC).

Guideline Assessment and Correction for Verification Bias
Next, we evaluated the sensitivity and specificity of the European Association of Urology (EAU), AUA, and the National

Table 2. Univariate and multivariate logistic regression models predicting the presence of bone metastases at diagnosis

| Factors | Univariate Logistic Regression Model | | Multivariate Logistic Regression Model | | Overall |
	OR (95% CI)	P Value	OR (95% CI)	P Value	P Value
Age	1.04 (1.01-1.08)	.01	1.03 (0.99-1.06)	.14	.14
Ln(PSA + 1)	2.25 (1.76-2.88)	<.0001	2.00 (1.51-2.64)	<.0001	<.0001
Biopsy Gleason score sum					.004
≤3 + 4	Reference		Reference		
4 + 3	5.04 (0.90-28.31)	.07	3.30 (0.55-19.89)	.19	
8-10	16.05 (3.82-67.45)	.0002	9.53 (2.14-42.38)	.003	
Clinical trial stage					.43
T1	Reference		Reference		
T2	2.64 (1.31-5.33)	.007	1.61 (0.72-3.57)	.25	
T3/4	9.19 (3.51-24.03)	<.0001	1.91 (0.57-6.43)	.30	
Positive cores	13.32 (4.26-41.72)	<.0001	1.70 (0.42-6.90)	.46	.46

CI, confidence interval; OR, odds ratio; other abbreviation as in Table 1.

Comprehensive Cancer Network (NCCN) guidelines,[2,3,5] each of which recommends staging BSs only in certain patient subgroups. The EAU guideline recommends staging BSs in patients with GS ≥8, locally advanced disease, or PSA level >10 ng/mL.[5] According to the AUA guidelines, the BS is recommended for patients with poorly differentiated tumors or PSA level >20 ng/mL.[2] According to the NCCN guidelines, staging BSs should be performed in all patients with GS ≥8, cT3/4, cT1 and PSA level >20 ng/mL, or cT2 and PSA level >10 ng/mL.[3] Recently, Briganti et al[4] developed a risk stratification tool using the classification and regression tree (CART) technique to identify patients requiring a staging BS at diagnosis. Based on this analysis, BS should be performed in patients with biopsy GS ≥8 or PSA level >10 ng/mL and cT2/3. A table summarizing the guidelines is provided as Supplementary Materials.

A key consideration in this step is that patients who did not undergo a staging BS at diagnosis have unverified disease status because the presence or absence of bone metastases is not known with certainty. To address this and obtain more accurate estimates of sensitivity and specificity, we used the method of Begg and Greenes[8] to correct for verification bias. To apply this method, we estimated the probability of a positive BS result for all patients as a function of clinical variables using a multivariate logistic regression model. For each guideline, the predicted probabilities were summed separately for patients who were recommended and not recommended for staging BSs, yielding the estimated number of patients with a positive or negative BS result. The sensitivity and specificity for the entire sample were estimated using the equations defined in the study by Begg and Greenes.[8] These estimates are unbiased if the presence of metastatic disease is conditionally independent of whether a patient underwent BS.

RESULTS

Table 1 presents clinical characteristics of the 1509 patients with newly diagnosed prostate cancer. Among this group, 416 (27.6%) underwent a staging BS. Patients who received a staging BS had higher mean PSA values and higher percentages of positive cores compared with patients who did not receive a BS (all P ≤.001). Moreover, patients with BS were significantly older and showed a higher biopsy GS as well as a higher rate of locally advanced prostate cancer compared with patients without BS (all P ≤.001). Among the patients who received a BS,

Table 3. Performance characteristics of the EAU, AUA, and NCCN guidelines after correction for verification bias

Clinical Guidelines	Sensitivity, %	Specificity, %
EAU[5]	84.5	75.9
AUA[2]	81.3	82.0
NCCN[3]	82.3	80.9
CART by Briganti et al[4]	79.4	83.3

AUA, American Urological Association; CART, classification and regression tree; EAU, European Association of Urology; NCCN, National Comprehensive Cancer Network.

48 (11.5%) had a positive study result with evidence for bone metastases.

Table 2 summarizes results from univariate and multivariate analyses evaluating the relationship between clinical parameters and BS findings. There was a wide range of serum PSA values (0.4-6873.4 ng/mL; coefficient of variation, 651.2), and because of the dispersion in PSA levels, we used the natural logarithm transformation. In univariate logistic regression analyses, all variables were significant predictors of bone metastases (all P ≤.01). In multivariate analyses, only serum PSA and biopsy GS were significant predictors of a positive BS result (both P values ≤.004; Table 2). Illustrating this point, the adjusted odds of a positive BS result for patients with a biopsy GS of 4 + 3 = 7 are 3.30 times (95% confidence interval, 0.55-19.89) as great as that for patients with a GS of 3 + 4 = 7 or GS = 6, whereas for patients with a biopsy GS of 8-10, the odds of a positive BS result are 9.53 times (95% confidence interval, 2.14-42.38) the odds for patients in the reference group.

Guideline Assessment and Correction for Verification Bias

The verification bias–adjusted sensitivity and specificity of several existing guidelines are presented in Table 3. The EAU guideline had the highest sensitivity and the lowest specificity. CART by Briganti et al had the lowest sensitivity but the highest specificity. The performance of the AUA and NCCN guidelines was relatively similar to the performance of the CART by Briganti et al in terms

of sensitivity and specificity, with a maximum difference of approximately 3%.

We also used the multivariate logistic regression model to evaluate the performance of the guidelines with respect to the estimated number of positive BS results missed and the number of negative BS results. The results from this analysis are summarized in Table 4. The guidelines had <1% of positive study results missed; however, the EAU guidelines resulted in a much higher number of negative study results and the highest total number of studies performed. The AUA and NCCN guidelines and CART by Briganti et al had significantly fewer studies performed compared with the EAU guideline. Under the AUA guidelines, the average number of negative study results and the total number of studies would be reduced by 38% and 6%, respectively, compared with the current practice.

COMMENT

We used contemporary data from a large group of community and academic urology practices to investigate the association between routinely available clinical variables and the likelihood of a positive BS result among men with newly diagnosed prostate cancer. We found that serum PSA level and biopsy GS were the principal predictors of a positive study result among patients who received a BS. Furthermore, after accounting for the fact that not all patients underwent a staging BS, we demonstrated that the AUA and NCCN guidelines and the CART model by Briganti et al performed reasonably well in terms of sensitivity and specificity. The EAU guideline resulted in not only higher sensitivity but also substantially lower specificity.

Our work is consistent with previous investigations demonstrating that higher PSA levels and biopsy GS are associated with an increased risk for a positive BS result.[9-21] In contrast to our findings, some studies have also identified clinical stage as an important predictor of bone metastases at diagnosis.[4,9,10,22] In one recent study of 851 consecutive patients with imaging, Briganti et al[4] found that PSA level, biopsy GS, and clinical stage were significant predictors of a positive BS result and should therefore be used as criteria for selecting patients for such studies. Based on the findings of these studies, several organizations (EAU, AUA, and NCCN) updated their recommendations indicating the need for a staging BS only for newly diagnosed prostate cancer patients with certain unfavorable characteristics. However, despite the availability of these guidelines, there is still controversy over the referral criteria, and no consensus exists about the most accurate and cost-effective strategy. Accordingly, widespread implementation of these guidelines could improve care delivery by eliminating a large number of BSs that do not reliably contribute useful clinical information, while adding significant costs to the health-care system.

Because not all patients in the registry received a BS, our findings are susceptible to verification bias. We partially

Table 4. Performance characteristics of existing clinical guidelines

Guidelines	Patients With BS (n = 416)			Patients Without BS (N = 1103)			Entire Population (N = 1519)		
	No. of Patients to be Scanned, n (%)	No. of Positive BSs Missed, n (%)	No. of Negative BSs, n (%)	No. of Patients to be Scanned, n (%)	Expected No. of Positive BSs Missed, n (%)	Expected No. of Negative BSs, n (%)	No. of Patients to be Scanned, n (%)	Expected No. of Positive BSs Missed, n (%)	Expected No. of Negative BS, n (%)
EAU[5]	288 (69.2)	1 (0.1)	127 (30.5)	117 (10.6)	8 (0.7)	108 (9.8)	405 (26.7)	10 (0.7)	350 (23.0)
AUA[2]	255 (61.3)	1 (0.1)	160 (38.5)	59 (5.3)	9 (0.8)	51 (4.6)	314 (20.7)	12 (0.8)	261 (17.2)
NCCN[3]	265 (63.7)	1 (0.1)	150 (36.1)	67 (6.1)	9 (0.8)	59 (5.3)	332 (21.9)	12 (0.8)	278 (18.3)
CART by Briganti et al[4]	244 (58.7)	5 (1.2)	167 (40.1)	50 (4.5)	10 (0.9)	44 (4.0)	292 (19.2)	13 (0.9)	243 (16.0)

Abbreviations as in Tables 1 and 3.

mitigated the impact of this by using a bias correction procedure for estimating sensitivity and specificity developed by Begg and Greenes.[8] Because the underlying assumptions behind these formulas cannot be proven, the values of adjusted sensitivity and specificity must be viewed as estimates. Second, the results of this study depend on the characteristics of the patient population in practices in Michigan, and these may differ from those observed in other geographic regions. Another limitation is the possibility for correlation among clinical practices in MUSIC regarding imaging patterns for staging BSs. We addressed this issue by performing sensitivity analyses that implemented generalized estimating equations to account for potentially correlated data, and we noted no substantive changes in our principal findings. Therefore, we assumed that the possible correlation among clinical practices in MUSIC is not strong, and we fit our final logistic regression model.

Notwithstanding these limitations, our study has several strengths, as well as important clinical and policy implications. This is the first analysis to evaluate the performance of existing clinical guidelines for staging BSs in a population-based sample of patients seen in diverse academic and community practices. In addition, we provide specific estimates around the impact of specific guideline implementation (relative to existing practice patterns) with respect to the number of positive scan results missed, number of negative scan results, and total number of scans performed in a population. Such estimates may prove quite useful for clinicians, specialty societies, and other stakeholders seeking a satisfactory trade-off between the benefits and harms of using BSs for staging of patients newly diagnosed with prostate cancer.

Illustrating this point, our data indicate that adherence with recommendations to image with a BS only when the PSA level is >20 ng/mL or a GS is ≥8 would lead to an estimated decrease in the overall utilization of staging BSs by 6.6% compared with current imaging practices in Michigan. If these criteria were implemented across all MUSIC practices, we estimate that <1% of patients with bone metastases would not be imaged and that a large proportion of studies with negative results that are now being ordered could be safely omitted. Given the consistency of our empirical findings with recommendations from the AUA, many urologists in MUSIC have coalesced around a PSA level of >20 ng/mL or a GS of ≥8 as criteria for ordering staging BSs in patients with newly diagnosed prostate cancer. Moreover, we are taking purposeful steps to now implement these criteria statewide through the use of data feedback, reminder cards, and other established quality improvement strategies.

CONCLUSION

In this analysis of patients seen in diverse community and academic practices in Michigan, we identified serum PSA level and biopsy GS as significant predictors for the presence of bone metastases in newly diagnosed untreated prostate cancer patients. Our results also suggest that implementing recommendations where a staging BS is performed only in patients with a PSA level of >20 ng/mL or a GS of ≥8 would simultaneously result in fewer positive study results missed, fewer negative study results, and fewer BSs overall.

References

1. American Urological Association: Choosing Wisely. http://www.choosingwisely.org/. Accessed February 15, 2014.
2. Thompson I, Thrasher JB, Aus G, et al. Guideline for the management of clinically localized prostate cancer: 2007 update. *J Urol.* 2007;177:2106-2131.
3. National Comprehensive Cancer Network Clinical Recommendations. http://www.nccn.org/professionals/physician_gls/f_guidelines.asp. Accessed February 20, 2014.
4. Briganti A, Passoni N, Ferrari M, et al. When to perform bone scan in patients with newly diagnosed prostate cancer: external validation of the currently available guidelines and proposal of a novel risk stratification tool. *Eur Urol.* 2010;57:551-558.
5. European Association of Urology: EAU Clinical Guidelines. http://www.uroweb.org/guidelines/online-guidelines/. Accessed February 15, 2014.
6. Michigan Urological Surgical Improvement Collaborative (MUSIC). https://musicurology.com/Registry/index, 2014.
7. Blue Cross Blue Shield Michigan (BCBSM). http://www.bcbsm.com/. Accessed January 15, 2014.
8. Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. *Biometrics.* 1983;39:207-215.
9. Gleave ME, Coupland D, Drachenberg D, et al. Ability of serum prostate-specific antigen levels to predict normal bone scans in patients with newly diagnosed prostate cancer. *Urology.* 1996;47:708-712.
10. Lee N, Fawaaz R, Olsson CA, et al. Which patients with newly diagnosed prostate cancer need a radionuclide bone scan? An analysis based on 631 patients. *Int J Radiat Oncol Biol Phys.* 2000;48:1443-1446.
11. Kosuda S, Yoshimura I, Aizawa T, et al. Can initial prostate specific antigen determinations eliminate the need for bone scans in patients with newly diagnosed prostate carcinoma? A multicenter retrospective study in Japan. *Cancer.* 2002;94:964-972.
12. Ayyathurai R, Mahapatra R, Rajasundaram R, et al. A study on staging bone scans in newly diagnosed prostate cancer. *Urol Int.* 2006;76:209-212.
13. McArthur C, McLaughlin G, Meddings RN. Changing the referral criteria for bone scan in newly diagnosed prostate cancer patients. *Br J Radiol.* 2012;85:390-394.
14. O'Sullivan JM, Norman AR, Cook GJ, et al. Broadening the criteria for avoiding staging bone scans in prostate cancer: a retrospective study of patients at the Royal Marsden Hospital. *BJU Int.* 2003;92:685-689.
15. Ritenour CW, Abbott JT, Goodman M, et al. The utilization of Gleason grade as the primary criterion for ordering nuclear bone scan in newly diagnosed prostate cancer patients. *ScientificWorldJournal.* 2009;9:1040-1045.
16. Hirobe M, Takahashi A, Hisasue S, et al. Bone scanning—who needs it among patients with newly diagnosed prostate cancer? *Jpn J Clin Oncol.* 2007;37:788-792.

17. Tanaka N, Fujimoto K, Shinkai T, et al. Bone scan can be spared in asymptomatic prostate cancer patients with PSA of <=20 ng/ml and Gleason score of <=6 at the initial stage of diagnosis. *Jpn J Clin Oncol.* 2011;41:1209-1213.

18. Al-Ghazo MA, Ghalayini IF, Al-Azab RS, et al. Do all patients with newly diagnosed prostate cancer need staging radionuclide bone scan? A retrospective study. *Int Braz J Urol.* 2010;36:685-691; discussion 691-682.

19. Zaman MU, Fatima N, Sajjad Z. Metastasis on bone scan with low prostate specific antigen ($\leq$20 ng/ml) and Gleason's score (<8) in newly diagnosed Pakistani males with prostate cancer: should we follow Western guidelines? *Asian Pac J Cancer Prev.* 2011;12:1529-1532.

20. Lee SH, Chung MS, Park KK, et al. Is it suitable to eliminate bone scan for prostate cancer patients with PSA $\leq$20 ng/mL? *World J Urol.* 2012;30:265-269.

21. Wang Y, Guo J, Xu L, et al. Should bone scan be performed in Chinese prostate cancer patients at the time of diagnosis? *Urol Int.* 2013;91:160-164.

22. Bruwer G, Heyns CF, Allen FJ. Influence of local tumour stage and grade on reliability of serum prostate-specific antigen in predicting skeletal metastases in patients with adenocarcinoma of the prostate. *Eur Urol.* 1999;35:223-227.

APPENDIX

SUPPLEMENTARY DATA