Original

Computer-assisted Diagnosis of Metastases Based on Bone Scintigraphy: Lesion Detection Compared Using BONENAVI 1 and 2

Hiroyasu Tomonaga[1], Takahito Nakajima[1], Yukiko Arisaka[1], Azusa Tokue[1], Tetsuya Higuchi[1] and Yoshito Tsushima[1]

1 Department of Diagnostic Radiology and Nuclear Medicine, Gunma University Graduate School of Medicine, 3-39-22 Showa-machi, Maebashi, Gunma 371-8511, Japan

Abstract

Background & Aims: This study aimed to evaluate the diagnostic performance and features of a computer-assisted diagnostic system. BONENAVI versions 1 (BN1) and 2 (BN2) were used to detect lesions on bone scintigraphy.
Methods: Bone scintigraphy of 33 prostate cancer and 27 breast cancer patients with bone metastases was evaluated. Spots detected and analyzed by BN1 and BN2 were compared with those of manual analyses by nuclear medicine physicians.
Results: The sensitivity of BN1 and BN2 was 99.2% and 97.0% for prostate cancer and 96.8% and 95.0% for breast cancer, respectively. The specificity was 64.7% and 68.0% for prostate cancer and 65.3% and 75.8% for breast cancer, respectively. Positive predictive values tended to be higher for BN2, and negative predictive values tended to be higher for BN1. BN2 showed fewer false positive spots. BN2 was superior to BN1 for detection of physiological uptake in the head and pelvis in breast cancer and in the cervical/lumbar spine in both breast and prostate cancer.
Conclusions: BN2 showed better diagnostic performance than BN1. Understanding the characteristics of lesion detection using each version of BONENAVI may be useful when evaluating new lesions on bone scintigraphy.

Article Information

Key words:
BONENAVI
bone metastasis,
computer-assisted diagnosis,
scintigraphy

Publication history:
Received: August 8, 2017
Revised: September 25, 2017
Accepted: September 28, 2017

Corresponding author:
Takahito Nakajima
Department of Diagnostic Radiology and Nuclear Medicine, Gunma University Graduate School of Medicine, 3-39-22 Showa-machi, Maebashi, Gunma 371-8511, Japan
Tel: +81-27-220-8401
E-mail: sojin@gunma-u.ac.jp

Introduction

The number of annual deaths in Japan due to cancer has increased because of population aging.[1] Metastatic bone lesions are common in breast and prostate cancer.[2] Bone scintigraphy is a highly-sensitive technique for detecting bone metastases,[3] but evaluating lesions can be difficult for inexperienced clinicians. BONENAVI (Fuji RI Pharmacy, Tokyo, Japan) is a computer-assisted diagnostic (CAD) software program for analysis of lesions detected by a bone scan. BONENAVI can assess 1) a patient's likelihood of having a metastatic lesion and 2) the treatment response of known metastatic lesions. The bone scan index (BSI) is employed as an imaging biomarker for evaluating treatment efficacy in bone metastases.[4] However, the detection and evaluation of individual bone lesions by physicians are still important for patient management.

BONENAVI reveals lesions highly suspected of being bone metastases as red spots, and uptake due to other pathologies as blue spots. BONENAVI was recently upgraded from version 1 (BN1) to version 2 (BN2). BN1 included data of 904 bone scans from a single Japanese hospital.[5] BN2 used a multi-center training database of 1,532 patients from 9 Japanese hospitals,[6] with the expectation that the larger number of hospitals and patients would help improve diagnostic accuracy.

In our facility, we initially used BN1 to analyze

Table 1 Patient distribution.

	patient	age (means±SD (range))
Prostate cancer	33 males	71.6±8.5 (56–89) years old
Breast cancer	27 females	57.9±9.1 (38–74) years old

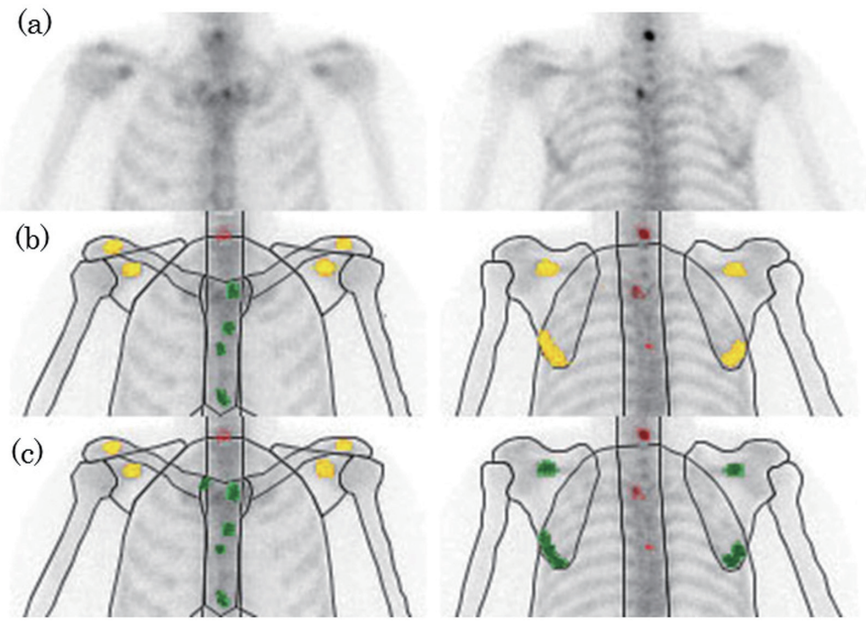

Fig. 1 Differences of area segmentation in the chest and upper extremity: (a) native bone scan, (b) BN1, and (c) BN2.
Physiological accumulations were recognized at the acromion, coracoid process, and lower scapular angle. Yellow spots belong to the upper extremity area and green spots to the chest area.

bone metastases and switched to BN2 when it became available. In this study, we analyzed data of a patient population using both versions to assess their reliability in lesion detection, and these results were compared to evaluations performed by nuclear medicine physicians. The present study aimed to clarify the features of each version for each body region.

Materials and methods

Patients

Patients (27 females and 33 males; Table 1) with one or more bone metastatic lesions of prostate or breast cancer who underwent bone scintigraphy from 2008 to 2010 (n=21) and from 2013 to 2015 (n=39) were recruited from other studies by independently applying BN1 or BN2 to acquired images. Data of the former and latter periods were analyzed using BN1 and BN2, respectively, in our hospital. The period from 2010 to 2013 was excluded from the analysis because the data for this period were not retained. Images were reanalyzed using both BN1 and BN2 in the present study.

Bone scintigraphy

Bone scintigraphy was performed approximately 3 h after an intravenous injection of 740 MBq technetium-99 m methylene diphosphonate (^{99m}Tc-MDP; Fujifilm RI Pharma Co., Ltd., Tokyo, Japan). Whole-body images were obtained using a gamma camera (ECAM Signature, Toshiba, Tokyo, Japan) equipped with low-energy, high-resolution parallel-hole collimators. The matrix size was 256 × 1024. The energy peak was centered at 140 keV with a 7.5% window. The whole body scan speed was 11 cm/min.

Data analysis

Bone scintigraphy images of all patients were analyzed using both BN1 and BN2. After BONENAVI automatically recognized hot spots that accumulated ^{99m}Tc-MDP, each hot spot was classified as being either a) a lesion with a high possibility of metastasis (red spot) or b) a lesion with a low possibility of metastasis (blue spot). All hot spots analyzed by BN1 or BN2 were retrospectively reviewed by two nuclear medicine physicians (YA, 13 years of experience; HT, 3 years of experience). When the two physicians did not agree, they reviewed the bone scan images together, referring to other modalities, such as computed tomography, magnetic resonance, or positron emission tomography, to reach a consensus. Discrepancies in the interpretations of hot spots were divided into two groups: spots recognized as high risk by BONENAVI despite being identified by the nuclear medicine physicians as low risk (low-risk spot as high: LStoH) and spots recognized as low risk by BONENAVI despite being

Table 2 Discrepancy ratios in lesion numbers detected manually and with BONENAVI.

(a)

			spine			extremity				
primary diseasc	version	head	cervical	thoracic	lumbar	upper	lower	chest	pelvis	total
prostate cancer	BN1	40/245 (16.3%)	7/32 (21.9%)	75/282 (26.6%)	50/165 (30.3%)	73/208 (35.1%)	1/22 (4.5%)	34/429 (7.9%)	50/282 (17.7%)	330/1,665 (19.8%)
	BN2	40/254 (15.7%)	6/33 (18.2%)	65/279 (23.3%)	31/175 (17.7%)	39/141 (27.7%)	5/22 (22.7%)	50/549 (9.1%)	48/254 (18.9%)	284/1,707 (16.6%)
breast cancer	BN1	61/359 (17.0%)	9/35 (25.7%)	35/251 (13.9%)	34/133 (25.6%)	65/183 (35.5%)	6/35 (17.1%)	24/333 (7.2%)	78/290 (26.9%)	312/1,619 (19.3%)
	BN2	37/356 (10.4%)	7/33 (21.2%)	44/242 (18.2%)	21/141 (14.9%)	43/125 (34.4%)	6/33 (18.2%)	35/423 (8.3%)	52/285 (18.2%)	245/1,638 (15.0%)

(b)

			spine			extremity				
primary diseasc	version	head	cervical	thoracic	lumbar	upper	lower	chest	pelvis	total
prostate cancer	BN1	0/245 (0.0%)	0/32 (0.0%)	0/282 (0.0%)	0/165 (0.0%)	1/208 (0.5%)	0/22 (0.0%)	4/429 (0.9%)	1/282 (0.4%)	6/1,665 (0.4%)
	BN2	0/254 (0.0%)	3/33 (9.1%)	0/279 (0.0%)	6/175 (3.4%)	0/141 (0.0%)	0/22 (0.0%)	10/549 (1.8%)	6/254 (2.4%)	25/1,707 (1.5%)
breast cancer	BN1	7/359 (1.9%)	0/35 (0.0%)	1/251 (0.4%)	1/133 (0.8%)	0/183 (0.0%)	1/35 (2.9%)	11/333 (3.3%)	2/290 (0.7%)	23/1,619 (1.4%)
	BN2	10/356 (2.8%)	3/33 (9.1%)	3/242 (1.2%)	3/141 (2.1%)	0/125 (0.0%)	1/33 (3.0%)	9/423 (2.1%)	2/285 (0.7%)	31/1,638 (1.9%)

The numbers and ratios of low-risk spots as high (LStoH) are shown in Table 2a, and those of high-risk spors as low (HStoL) are shown in Table 2b. There were generally fewer LStoH cases in BN2, whereas those in the lower extremity in patients with prostate cancer were increased. There were fewer HStoL cases than LStoH cases.

identified by the physicians as high risk (high-risk spot as low: HStoL).

Locations of the lesions were divided into the following 8 regions: head, spine (cervical, thoracic, or lumbar), extremity (upper or lower), chest, and pelvis. As it was differently defined depending on the BONENAVI version, the scapular region was included in the upper extremity region by BN1 and in the chest region by BN2 (Fig. 1). Based on corrected risk classification after analysis using BONENAVI, the physicians tallied the number of high-risk spots by region before and after the correction. These numbers were analyzed with regard to the types of carcinoma and versions of BONENAVI. Red spot ratios were calculated according to the following formula:

Ratio of red spots=Number of red spots/
Total number of hot spots

Statistical analysis

Comparisons among patient groups were performed using Student's t-test. The chi-square test was used to compare the hot spot numbers, matched or mismatched with interpretations by the physicians and BONENAVI. The sensitivity, specificity, and positive and negative predictive values were also calculated. A p-value of <0.05 was considered statistically significant. Statistical analysis was performed with Prism (version 6.0; GraphPad Software, San Diego, CA, USA).

Results

Number of hot spots

The number of hot spots ranged from 16 to 110 in 33 prostate cancer patients and from 24 to 101 in 27 breast cancer patients, showing no significant difference in the distribution ($p = 0.94$ and 0.43, respectively). The total number of hot spots and proportion of red spots are shown in Fig. 2. In carcinoma cases with

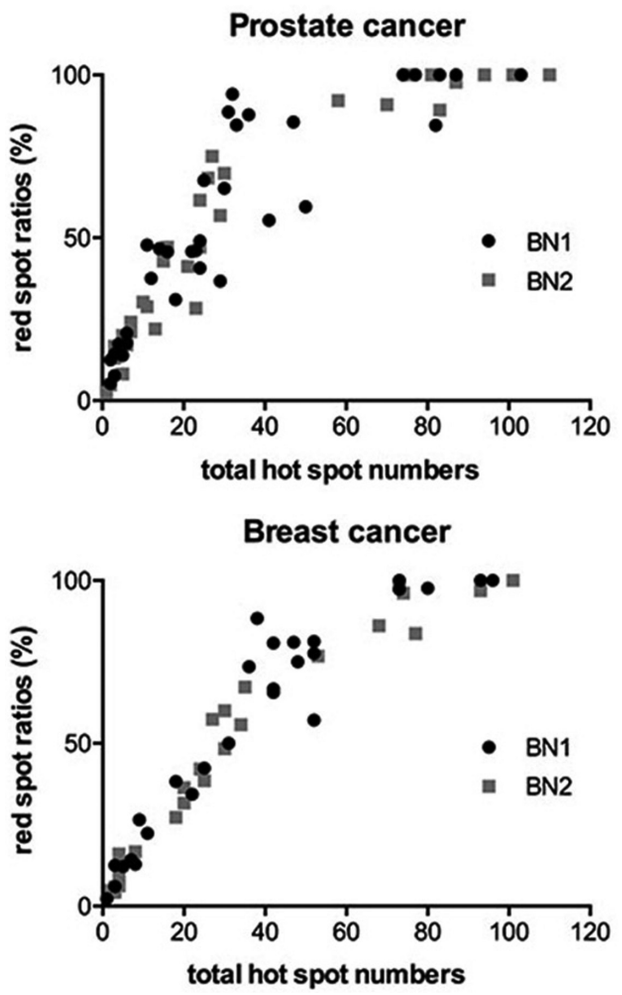

Fig. 2 Correlations between the ratio of red spots and total number of hot spots in prostate cancer (a) and breast cancer (b) patients. The ratio of red spots was calculated according to the following formula:
Ratio of red spots=Number of red spots/Total number of hot spots

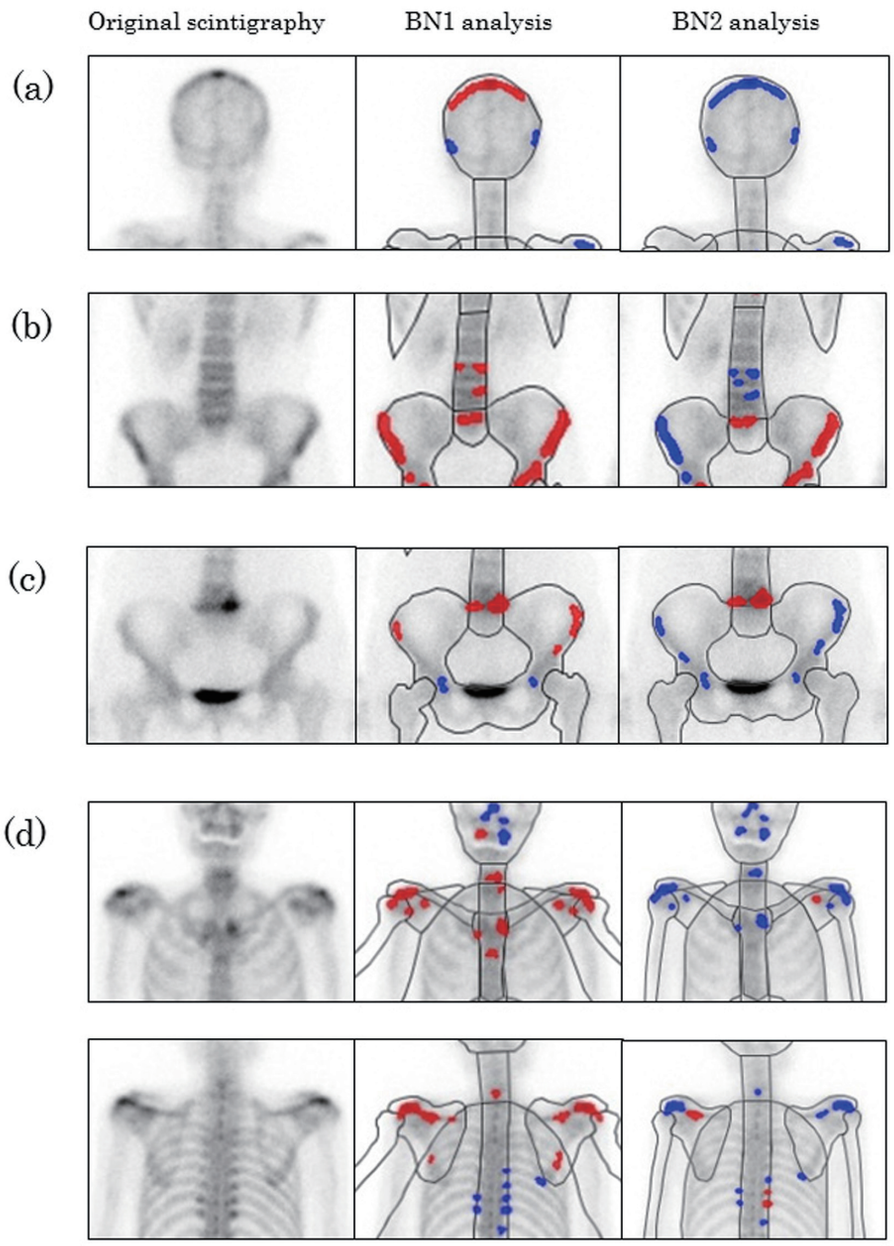

Fig. 3 Case presentations comparing BN1 and BN2 in the following areas: (a) skull, (b) lumbar spine, (c) pelvis, and (d) upper extremity. Red spots indicate high-risk lesions, and blue spots indicate low-risk lesions.

a significantly large number of bone metastases, almost all spots were determined to be of high risk, including spots that were judged to be of low risk by the physicians. The tendency was more notable in BN1.

False-positive spots: LStoH

The total ratios of LStoH cases are shown in Table 2a. False positives were increased in the head, spine, pelvis, and upper extremity regions. In the head region, false positives were increased on the scalp. The false positive ratio was lower in BN2 than in BN1 in breast cancer patients (all breast cancer patients were female, Fig. 3a). With an increase in false positives in the spine, the false positive ratio was lower in BN2 than in BN1 for any carcinoma (Fig. 3b). Similarly, false positives were increased at the anterior superior iliac spine and the false positive ratio was lower in B2 than in BN1 in breast cancer patients (Fig. 3c); however, the ratio was constant in prostate cancer patients. The false positive ratio was highest in the upper extremity for any type of carcinoma, with increased false positives observed at the shoulder joint (Fig. 3d). The ratios ranged from 27.7% (BN2 analysis of prostate cancer) to 35.5% (BN1 analysis of breast cancer), and the false positive ratio was lower in BN2 than in BN1 in prostate cancer patients.

False-negative spots: HStoL

The total ratios of HStoL cases are shown in Table 2b. In a comparison of breast cancer patients, the false negative rate was 1.4% for BN1 and 1.9% for BN2, with no significant difference ($p = 0.36$). In

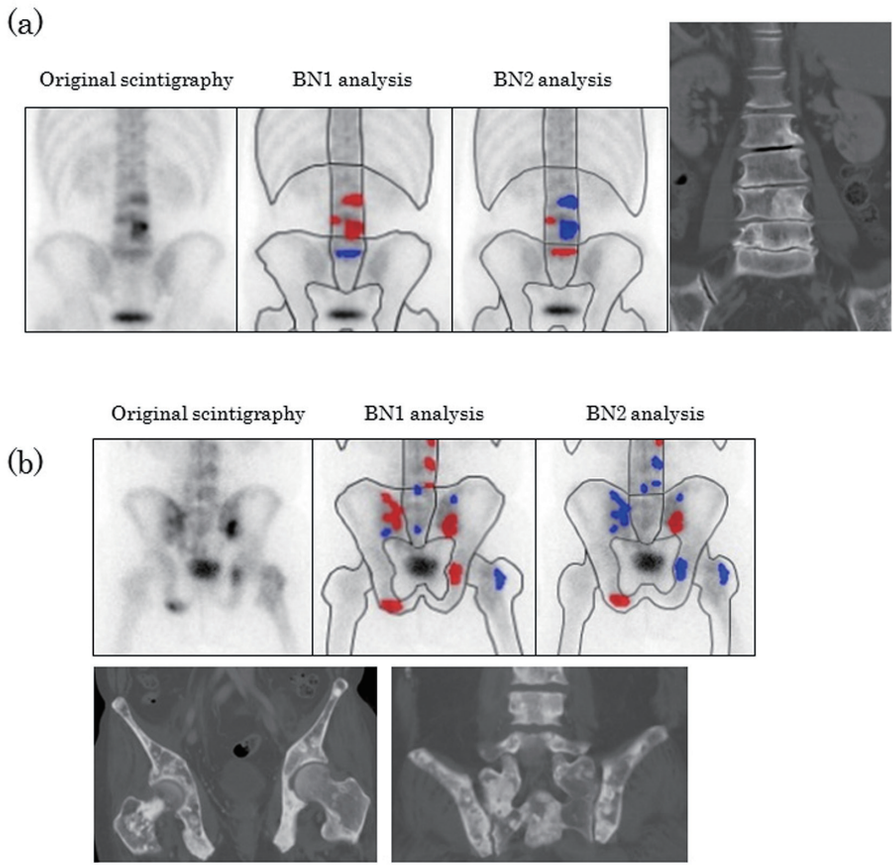

Fig. 4 (a) Accumulation in L2 and L4 vertebral bodies, determined as low risk (blue spot) in BN1, and high risk (red spot) in BN2. Irregular sclerotic lesions were observed at the same locations with computed tomography (CT), and bone metastases were suspected. The determination was correct in BN1, and false negative in BN2.
(b) Accumulation in the left acetabular cartridge and right sacroiliac joint region was determined as high risk in BN1 and low risk in BN2. CT showed irregular sclerotic lesions at the same locations, and these were considered bone metastases. In this case, only BN2 showed a false negative result.

Table 3 Lesion detection in patients with prostate and breast cancer using BONENAVI versions 1 and 2

		TP	FN	FP	TN	sensitivity	specificity	PPV	NPV
Prostate cancer	BN1	725	6	330	604	99.2%	64.7%	68.7%	99.0%
(n=33)	BN2	795	25	284	603	97.0%	68.0%	73.7%	96.0%
Breast cancer	BN1	697	23	312	587	96.8%	65.3%	69.1%	96.2%
(n=27)	BN2	593	31	245	769	95.0%	75.8%	70.8%	96.1%

This tabls shoes 1) nembers of true positive (TP), false negative (FN), and true negative (TN) lesions, and 2) ratios of sensitivity, positive predictive value (PPV) and negative predictive value (NPV).

prostate cancer patients, the false negative rate was 0.4% for BN1 and 1.5% for BN2; the increase was mildly significant ($p<0.01$). Although high risk was correctly determined with BN1 in a small number of cases, false negatives were also found in some BN2 cases. Two false negative cases identified by BN2 alone are shown in Fig. 4.

Lesion detection ability

Overall, BONENAVI showed high sensitivity, while specificity was relatively low. Compared to BN1, BN2 had slightly better specificity without a decrease in sensitivity in breast and prostate cancer patients (Table 3).

Discussion

Cases with multiple bone metastases yielded over 60 total hot spots using both BONENAVI versions. Analyses using BONENAVI indicated a high-risk spot ratio of nearly 100% (Fig. 2). Even areas of uptake considered low risk, such as degeneration, were included among the hot spots; however, when their number increased, the hot spots collectively tended to be considered high risk. When analyzing cases with a large number of bone metastases, it is necessary to consider

the presence of non-bone metastases, which can easily be determined as high risk lesions. The total number of LStoH (false-positive) spots decreased with BN2 in both prostate and breast cancer patients. Significant decreases in LStoH spots were particularly evident in the cervical and lumbar spine, female skull and pelvis, and male upper extremities. False negative spots were relatively rare, ranging from 0 % to 9.1%.

In anatomical region-based analysis, the number of hot spots was almost identical with use of BN1 and BN2, except for the number in the upper extremity and chest areas. Since the scapula was considered the upper extremity/chest by BN1/BN2, respectively, the number of hot spots identified by BN2 decreased in the upper extremity for each cancer, while the number increased in the chest.

The physicians and BONENAVI tended to disagree with regard to lesions in the skull, spine, upper extremities, and pelvis. Uptake of bone tracers by the middle-aged female skull is common and normal. This phenomenon is well known to physicians as physiological accumulation.[7] BN1 tended to recognize areas of normal uptake in the skull as high-risk lesions (Figs. 3a), while BN2 correctly recognized these areas as having a low malignant potential. For this reason, the false positive rate of the head region in breast cancer patients was lower in BN2 than in BN1. Areas of abnormal spinal uptake are common in clinical settings. Most are degenerative lesions with osteophytes. BONENAVI tended to recognize uptake in vertebral plates or osteophytes as malignant lesions. It was thought that spinous process hot spots at the centers of vertebrae were mostly recognized as low-risk lesions, but when the body position was rotated, it was difficult for BONENAVI to determine whether hot spots in the spinous processes were malignant. The number of misrecognitions decreased in BN2 (Fig. 3b). As shown in Fig. 3c, physiologic accumulation in the anterior superior iliac spine in patients with breast cancer was recognized as high risk in BN1; however, in many cases, this was correctly recognized as low risk in BN2; thus, the false positive rate declined from 26.9% (BN1) to 18.2% (BN2) ($p<0.05$). The false positive rate in the pelvic region of prostate cancer patients was 17.7% in BN1 and 18.9% in BN2, but the difference was not significant ($p = 0.77$). In the upper extremities, hot spots in the acromion and coracoid process tended to be recognized as high-risk lesions in BN1 in both prostate and breast cancer patients. BN2 correctly recognized these lesions as degenerative most of the time, but other spots around the shoulder were accurately diagnosed more often by the nuclear medicine physicians. Thus, as shown Fig. 3d, there were cases in which red spots remained as false positives around the shoulder joint.

As shown in Fig. 4, a few high risk cases that were correctly determined in BN1 were considered false negatives in BN2. Since BN1 recognized hot spots as high risk overall, the number of false positives increased, but bone metastases were rarely recognized as low risk. BN2 tended to correctly recognize low-risk lesions such as physiological accumulation, but in some cases bone metastasis was misidentified as physiological accumulation.

Kikushima et al. reported that the sensitivity and specificity in BN2 were 94% and 88% in male patients and 86% and 85% in female patients, respectively.[8] In terms of sensitivity, our findings were similar, but the specificity in our study was lower in both males and females than that reported by Kikushima. Our study had lower specificity because the evaluation method was different. Kikushima only considered whether a patient had one or more metastatic lesions, while we evaluated every spot in our study. Our method was more complicated; however, our results reflected more accurate evaluations.

High sensitivity implies that BONENAVI is suitable for detecting new lesions in patients with bone metastases. Sadik et al. reported that a CAD system using a Swedish database improved physician sensitivity for detection of metastases from 78% without the system to 88% with the system ($p<0.001$).[9] However, because the specificity for bone hot spots was relatively low, LStoH cases should be carefully interpreted by physicians. Our findings according to anatomical site can help physicians accurately interpret LStoH cases. When using BONENAVI to measure treatment effect, clinicians have concluded that accumulation significantly decreases when a spot changes from high risk to low risk.[5] However, when a new high-risk spot appears, clinicians must decide whether this represents physiological accumulation or a new metastatic lesion. Our study determined that BN2 still yielded misdiagnoses in newly detected lesions, particularly in the cervical and lumbar spines, female skull and pelvis, and male upper extremities.

This study has some limitations. We only counted the number of hot spots instead of evaluating BSI. While BSI can be used for evaluating bone metastatic lesions in the entire body, we included many small spots, and the extent of disease activity was not entirely reflected by the number of lesions. However, to detect new lesions, it is important to evaluate even small lesions. Another limitation was that bone biopsy was not performed. We decided whether each hot spot was an LStoH or HStoL by visual assessment and by evaluating additional computed tomography or magnetic resonance images. This might account for lower specificity than previously reported.

To conclude, BN2 had improved diagnostic accuracy compared to BN1, and both versions had high sensitivity for detecting metastatic bone lesions. However, BN2 still misinterpreted physiological accumulation of bone tracer as high-risk for bone metastases. The knowledge gained from this study will be useful for evaluating new lesions with BONENAVI in specific areas.

Acknowledgement

We would like to thank Ayako Taketomi-Takahashi for editing the language of this manuscript.

References

1. Yang L, Fujimoto J, Qiu D, et al. Trends in cancer mortality in the elderly in Japan, 1970-2007. Ann Oncol 2010; 21(2): 389-396.
2. Coleman RE. Clinical features of metastatic bone disease and risk of skeletal morbidity. Clin Cancer Res 2006; 12(20 Pt 2): 6243s-6249s.
3. Koizumi M, Yoshimoto M, Kasumi F, et al. Comparison between solitary and multiple skeletal metastatic lesions of breast cancer patients. Ann Oncol 2003; 14(8): 1234-1240.
4. Imbriaco M, Larson SM, Yeung HW, et al. A new parameter for measuring metastatic bone involvement by prostate cancer: the Bone Scan Index. Clin Cancer Res 1998; 4(7): 1765-1772.
5. Horikoshi H, Kikuchi A, Onoguchi M, et al. Computer-aided diagnosis system for bone scintigrams from Japanese patients: importance of training database. Ann Nucl Med 2012; 26(8): 622-626.
6. Nakajima K, Nakajima Y, Horikoshi H, et al. Enhanced diagnostic accuracy for quantitative bone scan using an artificial neural network system: a Japanese multi-center database project. EJNMMI Res 2013; 3(1): 83.
7. Sadeghi R, Zakavi SR, Ghasemov M, et al. Age-related changes in skull uptake on bone scintigraphy: a quantitative study. Nucl Med Rev Cent East Eur 2008; 11(2): 67-69.
8. Kikushima S, Hanawa N, Kotake F. Diagnostic performance of bone scintigraphy analyzed by three artificial neural network systems. Ann Nucl Med 2015; 29(2): 125-131.
9. Sadik M, Suurkula M, Höglund P, et al. Improved classifications of planar whole-body bone scans using a computer-assisted diagnosis system: a multicenter, multiple-reader, multiple-case study. J Nucl Med 2009; 50(3): 368-375.