Interpretative Variability and Its Impact on the Prognostic Value of Myocardial Fatty Acid Imaging in Asymptomatic Hemodialysis Patients in a Multicenter Trial in Japan

Tomonari Kiriyama, MD; Shin-ichiro Kumita, MD; Masao Moroi, MD; Tsunehiko Nishimura, MD; Nagara Tamaki, MD; Naoyuki Hasebe, MD; Kenjiro Kikuchi, MD

**Background:** The severity of impaired fatty acid utilization in the myocardium can predict cardiac death in asymptomatic patients on hemodialysis. However, interpretive variability and its impact on the prognostic value of myocardial fatty acid imaging are unknown.

**Methods and Results:** A total of 677 patients who received hemodialysis for ≥20 years and had one or more cardiovascular risk factors underwent 123I-labeled β-methyl iodophenyl-pentadecanoic acid (BMIPP) single-photon emission computed tomography (SPECT) at 48 hospitals across Japan. SPECT images were interpreted by experts at the nuclear core laboratory and by readers with varying skill levels at clinical centers, based on the standard 17-segment model and 5-point scoring systems, independently. The κ values only reached fair agreement both for overall impression (κ=0.298, normal vs. abnormal) and for categorical impression (κ=0.244, normal vs. mildly abnormal vs. severely abnormal). The normalcy rate was lower in readers at the clinical centers (60.9%) than in experts (69.9%). In contrast to the results assessed by experts, a Kaplan-Meier analysis based on the interpretation by readers at the clinical centers failed to distinguish the risk of events in patients with normal scans from that of patients with mildly abnormal scans.

**Conclusions:** Considerable variability and its impact on prognostic value were observed in the visual interpretation of BMIPP SPECT images between experts and readers at the clinical centers. (Circ J 2015; 79: 153–160)

**Key Words:** β-methyl iodophenyl-pentadecanoic acid; Hemodialysis; Interpretive variability; Prognostic value; Single-photon emission computed tomography

Myocardial perfusion single-photon emission computed tomography (SPECT) is a well-established method and is widely used to evaluate the condition and function of the myocardium in patients with ischemic heart disease. However, stress myocardial SPECT is not readily performed in patients receiving hemodialysis, who have a 10–20-fold greater risk for cardiovascular mortality because of weakened physical strength, especially due to weakened leg muscles, renal anemia, peripheral artery disease (PAD), or other complications of underlying severe coronary artery atherosclerosis. In contrast, 123I-labeled β-methyl iodophenyl-pentadecanoic acid (BMIPP) imaging can detect a history of myocardial ischemia, even at rest, by visualizing depressed fatty acid utilization of the myocardium after severe and/or repetitive ischemia.
experts in image interpretation of BMIPP SPECT has not been investigated in contrast to that of myocardial perfusion SPECT (MPS), although the interpretation of BMIPP SPECT is not always undertaken by experts in actual clinical practice. Moreover, it is unknown how the prognostic value of BMIPP SPECT is influenced by the difference in image interpretation between experts and non-experts. In the main B-SAFE study, all BMIPP results were based on image interpretation by well-experienced readers at the nuclear core laboratory, not by readers with varying skill levels at clinical centers. The interpretative variability of BMIPP SPECT between experts and general readers was not assessed. Accordingly, the results of the B-SAFE study may not be directly applicable to daily clinical practice.

Thus, this sub-study aimed to clarify the extent of interpretative variability of BMIPP SPECT and the influence of the variability on its prognostic value, by examining the difference in interpretation between well-experienced readers at the core laboratory and readers with varying skill levels at the clinical centers in the B-SAFE study.

Methods

Study Design

In the population of the B-SAFE study, all BMIPP SPECT studies were interpreted by readers at each clinical center that consisted of physicians or radiologists with varying skill levels in interpreting cardiac nuclear images, and by well-experienced readers at the nuclear core laboratory. The decreased uptake score per segment and per coronary territory, summed BMIPP score, and agreements for overall impression were compared to determine the extent of interpretative variability. Kaplan-Meier survival curves were described based on the results from experts at the core laboratory as well as readers at the clinical centers and compared to evaluate the impact of variability in BMIPP SPECT image interpretation on its prognostic value.

Study Population

Between 1 July 2006 and 30 November 2007, a total of 683 patients were registered with the B-SAFE study at 48 hospitals across Japan. All patients had received hemodialysis for at least 20 years and had one or more cardiovascular risk factors. Fifteen readers were familiar with MPS but had only several years of experience interpreting BMIPP SPECT images. Forty-eight readers involved with image interpretation in the clinical centers.

Cardiovascular risk factors were as follows: hypertension, diabetes mellitus, hypercholesterolemia, PAD, smoking habit, family history of juvenile coronary artery disease (CAD), history of ischemic stroke, history of heart failure within 3 months of starting hemodialysis therapy, and dialysis hypotension. Each cardiovascular risk factor was defined as previously reported by Moroi et al. Patients with peritoneal dialysis therapy, severe valvular disorders requiring treatment, a diagnosis of dilated or hypertrophic cardiomyopathy before starting hemodialysis, a history of coronary revascularization, or a prior diagnosis of myocardial infarction or overt CAD were excluded from the study.

Age, sex, body mass index, and blood pressure at the beginning and end of hemodialysis were recorded within 2 months from the time of registration. Information recorded at registration included frequency and duration of hemodialysis during the week before registration. Consequently, 677 patients without chest symptoms who underwent hemodialysis were included in this study and followed up until 30 November 2010, or death.

Compulsory approval was obtained from the review boards of all institutions involved in this study. All patients provided written informed consent to participate in all study protocols before enrolment. All data in the present study were recorded and analyzed at the Translational Research Informatics Data Center (Kobe, Japan).

BMIPP SPECT Protocol

A total of 677 registered patients underwent BMIPP SPECT, on a day when the patient was not undergoing hemodialysis, using a variety of gamma cameras at each hospital. At least 6 h after fasting, patients were injected intravenously with 111–148 MBq of BMIPP (Nihon Medi-Physics Co Ltd), and SPECT data were acquired 10–30 min following the injection. MPS was also performed simultaneously using 74–111 MBq of 201TlCl in 154 patients. Dual isotope SPECT can simultaneously assess rest perfusion and fatty acid metabolism. All SPECT datasets were sent from each local imaging site to the nuclear core laboratory at Toho University Ohashi Medical Center (Tokyo, Japan).

Image Interpretation

BMIPP SPECT images were interpreted at the nuclear core laboratory and at each clinical center, independently. The standard 17-segment model and 5-point scoring systems (0, normal; 1, mildly reduced; 2, moderately reduced; 3, severely reduced; and 4, no uptake) were used for visual semi-quantitative analysis. MPS images were simultaneously interpreted in the same manner in 154 patients. Scored BMIPP uptake of each myocardial segment and summed BMIPP scores (range 0–68) were recorded. The overall impression (normal, 0–3; abnormal, ≥4) the BMIPP SPECT study was validated based on the summed BMIPP score. BMIPP SPECT results were divided into 3 categories based on the summed BMIPP score (normal, 0–3; mildly abnormal, 4–8; severely abnormal, ≥9).

At clinical centers, SPECT images were interpreted using routine analysis methodology for the purpose of clinical management of the patients, and the readers were allowed to use clinical information when interpreting. The environment of the image interpretation was not standardized at the clinical centers. Readers at the clinical centers consisted of physicians or radiologists with varying skill levels in interpreting cardiac nuclear imaging; this reflects the actual situation of interpreting BMIPP SPECT in daily clinical practice. Fourty-eight readers involved with image interpretation in the clinical centers.

At clinical centers, SPECT images were interpreted using routine analysis methodology for the purpose of clinical management of the patients, and the readers were allowed to use clinical information when interpreting. The environment of the image interpretation was not standardized at the clinical centers. Readers at the clinical centers consisted of physicians or radiologists with varying skill levels in interpreting cardiac nuclear imaging; this reflects the actual situation of interpreting BMIPP SPECT in daily clinical practice. Fourty-eight readers involved with image interpretation in the clinical centers.

Consequently, 677 patients without chest symptoms who underwent hemodialysis were included in this study and followed up until 30 November 2010, or death.

Compulsory approval was obtained from the review boards of all institutions involved in this study. All patients provided written informed consent to participate in all study protocols before enrolment. All data in the present study were recorded and analyzed at the Translational Research Informatics Data Center (Kobe, Japan).

BMIPP SPECT Protocol

A total of 677 registered patients underwent BMIPP SPECT, on a day when the patient was not undergoing hemodialysis, using a variety of gamma cameras at each hospital. At least 6 h after fasting, patients were injected intravenously with 111–148 MBq of BMIPP (Nihon Medi-Physics Co Ltd), and SPECT data were acquired 10–30 min following the injection. MPS was also performed simultaneously using 74–111 MBq of 201TlCl in 154 patients. Dual isotope SPECT can simultaneously assess rest perfusion and fatty acid metabolism. All SPECT datasets were sent from each local imaging site to the nuclear core laboratory at Toho University Ohashi Medical Center (Tokyo, Japan).

At clinical centers, SPECT images were interpreted using routine analysis methodology for the purpose of clinical management of the patients, and the readers were allowed to use clinical information when interpreting. The environment of the image interpretation was not standardized at the clinical centers. Readers at the clinical centers consisted of physicians or radiologists with varying skill levels in interpreting cardiac nuclear imaging; this reflects the actual situation of interpreting BMIPP SPECT in daily clinical practice. Fourty-eight readers involved with image interpretation in the clinical centers.

At clinical centers, SPECT images were interpreted using routine analysis methodology for the purpose of clinical management of the patients, and the readers were allowed to use clinical information when interpreting. The environment of the image interpretation was not standardized at the clinical centers. Readers at the clinical centers consisted of physicians or radiologists with varying skill levels in interpreting cardiac nuclear imaging; this reflects the actual situation of interpreting BMIPP SPECT in daily clinical practice. Fourty-eight readers involved with image interpretation in the clinical centers.

At clinical centers, SPECT images were interpreted using routine analysis methodology for the purpose of clinical management of the patients, and the readers were allowed to use clinical information when interpreting. The environment of the image interpretation was not standardized at the clinical centers. Readers at the clinical centers consisted of physicians or radiologists with varying skill levels in interpreting cardiac nuclear imaging; this reflects the actual situation of interpreting BMIPP SPECT in daily clinical practice. Fourty-eight readers involved with image interpretation in the clinical centers.

At clinical centers, SPECT images were interpreted using routine analysis methodology for the purpose of clinical management of the patients, and the readers were allowed to use clinical information when interpreting. The environment of the image interpretation was not standardized at the clinical centers. Readers at the clinical centers consisted of physicians or radiologists with varying skill levels in interpreting cardiac nuclear imaging; this reflects the actual situation of interpreting BMIPP SPECT in daily clinical practice. Fourty-eight readers involved with image interpretation in the clinical centers.

At clinical centers, SPECT images were interpreted using routine analysis methodology for the purpose of clinical management of the patients, and the readers were allowed to use clinical information when interpreting. The environment of the image interpretation was not standardized at the clinical centers. Readers at the clinical centers consisted of physicians or radiologists with varying skill levels in interpreting cardiac nuclear imaging; this reflects the actual situation of interpreting BMIPP SPECT in daily clinical practice. Fourty-eight readers involved with image interpretation in the clinical centers.

At clinical centers, SPECT images were interpreted using routine analysis methodology for the purpose of clinical management of the patients, and the readers were allowed to use clinical information when interpreting. The environment of the image interpretation was not standardized at the clinical centers. Readers at the clinical centers consisted of physicians or radiologists with varying skill levels in interpreting cardiac nuclear imaging; this reflects the actual situation of interpreting BMIPP SPECT in daily clinical practice. Fourty-eight readers involved with image interpretation in the clinical centers.

At clinical centers, SPECT images were interpreted using routine analysis methodology for the purpose of clinical management of the patients, and the readers were allowed to use clinical information when interpreting. The environment of the image interpretation was not standardized at the clinical centers. Readers at the clinical centers consisted of physicians or radiologists with varying skill levels in interpreting cardiac nuclear imaging; this reflects the actual situation of interpreting BMIPP SPECT in daily clinical practice. Fourty-eight readers involved with image interpretation in the clinical centers.

At clinical centers, SPECT images were interpreted using routine analysis methodology for the purpose of clinical management of the patients, and the readers were allowed to use clinical information when interpreting. The environment of the image interpretation was not standardized at the clinical centers. Readers at the clinical centers consisted of physicians or radiologists with varying skill levels in interpreting cardiac nuclear imaging; this reflects the actual situation of interpreting BMIPP SPECT in daily clinical practice. Fourty-eight readers involved with image interpretation in the clinical centers.

At clinical centers, SPECT images were interpreted using routine analysis methodology for the purpose of clinical management of the patients, and the readers were allowed to use clinical information when interpreting. The environment of the image interpretation was not standardized at the clinical centers. Readers at the clinical centers consisted of physicians or radiologists with varying skill levels in interpreting cardiac nuclear imaging; this reflects the actual situation of interpreting BMIPP SPECT in daily clinical practice. Fourty-eight readers involved with image interpretation in the clinical centers.
Interpretive Variability in BMIPP SPECT

Table 1. Summed $^{123}$I-Labeled BMIPP Score and Summed Scores of Coronary Territories at Core Laboratory and Clinical Centers

<table>
<thead>
<tr>
<th>Summed BMIPP score</th>
<th>Coronary territory</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAD</td>
<td>RCA</td>
</tr>
<tr>
<td>Core laboratory</td>
<td>3.3±5.3</td>
</tr>
<tr>
<td>Clinical centers</td>
<td>4.9±7.7</td>
</tr>
<tr>
<td>P value</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Results are presented as Mean±SD. LAD territory consists of segments 1, 2, 7, 8, 13, 14, 17; RCA territory consists of segments 3, 4, 9, 10, 15; LCX territory consists of segments 5, 6, 11, 12, 16. BMIPP, $\beta$-methyl iodophenyl-pentadecanoic acid; LAD, left anterior descending artery; LCX, left circumflex artery; RCA, right coronary artery; SD, standard deviation.

Table 2. Scores of Myocardial Segments at a Core Laboratory and Clinical Centers

<table>
<thead>
<tr>
<th>Segment</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core laboratory</td>
<td>0.06±0.31</td>
<td>0.06±0.31</td>
<td>0.13±0.48</td>
<td>0.46±0.89</td>
<td>0.14±0.53</td>
<td>0.05±0.31</td>
<td>0.14±0.53</td>
<td>0.08±0.38</td>
<td>0.06±0.34</td>
</tr>
<tr>
<td>Clinical centers</td>
<td>0.28±0.66</td>
<td>0.31±0.76</td>
<td>0.36±0.84</td>
<td>0.66±0.99</td>
<td>0.26±0.70</td>
<td>0.15±0.54</td>
<td>0.26±0.65</td>
<td>0.16±0.53</td>
<td>0.14±0.49</td>
</tr>
<tr>
<td>P value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Segment</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core laboratory</td>
<td>0.47±0.86</td>
<td>0.13±0.55</td>
<td>0.05±0.29</td>
<td>0.25±0.76</td>
<td>0.11±0.51</td>
<td>0.58±0.99</td>
<td>0.20±0.66</td>
<td>0.13±0.76</td>
</tr>
<tr>
<td>Clinical centers</td>
<td>0.53±0.88</td>
<td>0.19±0.63</td>
<td>0.12±0.44</td>
<td>0.26±0.69</td>
<td>0.12±0.54</td>
<td>0.52±0.93</td>
<td>0.19±0.83</td>
<td>0.39±0.85</td>
</tr>
<tr>
<td>P value</td>
<td>0.067</td>
<td>0.011</td>
<td>&lt;0.001</td>
<td>0.474</td>
<td>0.445</td>
<td>0.125</td>
<td>0.547</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Results are presented as Mean±SD. SD, standard deviation.

Figure 1. Discrepancy in the segmental uptake score between the core laboratory and the clinical centers. Readers at the clinical centers were inclined to give higher segmental uptake scores, particularly for basal and septal segments and the apex. SEM, standard error of the mean.

between teams 1 and 2, 0.624 between teams 1 and 3, and 0.454 between teams 2 and 3. After simultaneously consulting the 3 teams regarding the 50 patients, the remaining 627 SPECT images were randomly divided into 3 groups and assigned to each of the 3 teams for evaluation.

End-Point and Prognostic Value of Summed BMIPP Score

All 677 patients were followed up for at least 3 years after
compared using the Wilcoxon signed-rank test. Agreement between the core laboratory and clinical centers for their overall impression of the BMIPP SPECT scan (normal, 0–3 vs. abnormal, ≥4) and the category of the summed BMIPP score (normal, 0–3 vs. mildly abnormal, 4–8 vs. severely abnormal, ≥9) was assessed by calculating the $\kappa$ value. $\kappa$ values were classified as follows: <0.20=poor agreement; 0.21–0.40=fair agreement; 0.41–0.60=moderate agreement; 0.61–0.80=good agreement; and >0.81=excellent agreement. Kaplan-Meier survival curves were compared by using a 2-sided log-rank test. All data were statistically analyzed using SPSS version 19 (IBM Corporation, New York, NY, USA). P values of <0.05 were considered significant.

**Results**

Expert readers at the nuclear core laboratory diagnosed 473 (69.9%) out of 677 cases as normal, compared with 412 (60.9%) out of 677 cases diagnosed as normal by readers at the registration. The primary end-points were cardiac death or sudden death of an unknown cause. Details of the criteria have been described previously. Briefly, cardiac death was judged by physicians as being caused by heart failure, acute myocardial infarction, or other cardiac disorders. Of 677 study participants, 26 (3.8%) died of cardiac events (acute myocardial infarction, 10; congestive heart failure, 13; arrhythmia, 2; valvular heart disease, 1), and 20 (3.0%) died suddenly of unknown causes during a median follow-up period of 1,152 days. None were lost to follow up. We considered sudden death of an unknown cause to have been cardiac derived. The annual rate of cardiac-derived death in hemodialysis patients was 2.2%. Kaplan-Meier survival curves were described based on the results of interpretation by expert readers at the nuclear core laboratory and by readers at the clinical centers.

**Statistical Analysis**

Differences in the score of each myocardial segment, the score of each coronary territory, and the summed BMIPP score were compared using the Wilcoxon signed-rank test. Agreement between the core laboratory and clinical centers for their overall impression of the BMIPP SPECT scan (normal, 0–3 vs. abnormal, ≥4) and the category of the summed BMIPP score (normal, 0–3 vs. mildly abnormal, 4–8 vs. severely abnormal, ≥9) was assessed by calculating the $\kappa$ value. $\kappa$ values were classified as follows: <0.20=poor agreement; 0.21–0.40=fair agreement; 0.41–0.60=moderate agreement; 0.61–0.80=good agreement; and >0.81=excellent agreement. Kaplan-Meier survival curves were compared by using a 2-sided log-rank test. All data were statistically analyzed using SPSS version 19 (IBM Corporation, New York, NY, USA). P values of <0.05 were considered significant.

**Results**

Expert readers at the nuclear core laboratory diagnosed 473 (69.9%) out of 677 cases as normal, compared with 412 (60.9%) out of 677 cases diagnosed as normal by readers at the
Table 3. Causes of Death in Relation to BMIPP SPECT Results (Normal vs. Abnormal)

<table>
<thead>
<tr>
<th>BMIPP result</th>
<th>Core laboratory</th>
<th>Clinical centers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>Abnormal</td>
</tr>
<tr>
<td>Cardiac event</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMI</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>CHF</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Arrhythmia</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Valvular</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Sudden death</td>
<td>8</td>
<td>12</td>
</tr>
</tbody>
</table>

Results are presented as the number of cases.
AMI, acute myocardial infarction; CHF, congestive heart failure; SPECT, single-photon emission computed tomography. Other abbreviation as in Table 1.

Figure 3. Prognostic value for normal and abnormal $^{123}$I-labeled $\beta$-methyl iodo-phenyl-pentadecanoic acid single-photon emission computed tomography scans based on image interpretation by experts at the nuclear core laboratory (A), and by readers of varying skill level at clinical centers (B). In both cases, event-free survival was significantly lower in patients with abnormal results.
clinical centers (P<0.001). Among abnormal cases, 124 (18.3%) and 80 (11.8%) out of 204 cases were classified as mildly abnormal and severely abnormal at the core laboratory, respectively, and 132 (19.5%) and 133 (19.6%) out of 265 cases were classified as mildly abnormal and severely abnormal at the clinical centers, respectively. The summed BMIPP score and every score for coronary territories at the clinical centers were significantly higher than those at the core laboratory (Table 1). The segmental uptake score was also significantly higher in 12 out of 17 segments at the clinical centers (Table 2). The readers at the clinical centers were inclined to give higher segmental uptake scores, particularly in basal and septal segments and the apex (Figure 1). A representative case is shown in Figure 2.

The κ value indicated fair agreement (κ=0.244) for the categorical impression (normal, mildly abnormal, and severely abnormal) in all populations. Even for images interpreted as normal vs. abnormal, the κ value indicated fair agreement (κ=0.298). For 154 out of 677 patients in whom MPS images were available, the κ value for the categorical impression (κ=0.323) was a little higher than that for the overall population. The κ value for the overall impression reached moderate agreement (κ=0.469) and was modestly higher than that for the overall population. Among the rest of the 523 patients, κ
values for both the categorical impression (κ=0.220) and overall impression (κ=0.247) indicated fair agreement and were slightly lower than those for the overall population.

Based on the interpretation by experts at the core laboratory, an end-point was seen in 19 out of 473 patients with normal BMIPP SPECT scans and 27 out of 204 patients with abnormal BMIPP SPECT scans. Based on the interpretation by readers at the clinical centers, 16 out of 412 patients with normal BMIPP SPECT scans and 30 out of 235 patients with abnormal BMIPP scans reached an end-point. The detailed causes of the events are described in Table 3. Based on the results of the core laboratory, 11 (acute myocardial infarction [AMI], 2; congestive heart failure [CHF], 7; arrhythmia, 1; valvular disease, 1) cardiac events and 8 sudden deaths occurred in spite of normal BMIPP results. Based on the results of the clinical centers, 7 (AMI, 3; CHF 4) cardiac events and 9 sudden deaths occurred in spite of normal BMIPP results. Kaplan-Meier survival estimates for the overall impression of the experts and readers at the clinical centers are shown in Figure 3. Event-free survival was significantly different between patients with normal BMIPP SPECT scans and patients with abnormal scans. In both cases, event-free survival curves were based on interpretations at the core laboratory and clinical centers. When event-free survival curves for categories of the summed BMIPP score were plotted based on image interpretation by readers at the clinical centers, event-free survival rates were not significantly different between patients with normal BMIPP SPECT results and patients with mildly abnormal BMIPP SPECT results. However, significant differences were seen between patients with severely abnormal and each of the other 2 categories (Figure 4A). The 3-year cardiac-derived death-free rates were 96.0%, 95.0%, and 80.3% in patients with BMIPP summed scores of 0–3, 4–8, and ≥9, respectively. However, when the event-free survival curve was plotted based on the results of image interpretation by readers at the nuclear core laboratory, significant differences were seen among any 2 categories, which indicated 3-year cardiac-derived death-free rates of 95.7%, 90.6%, and 78.8% in patients with BMIPP summed scores of ≤3, 4–8, and ≥9, respectively (Figure 4B).

**Discussion**

This is the first study investigating the interpretive variability and its influence on the prognostic value of BMIPP SPECT by comparing the difference in interpretation between experienced readers at a core laboratory and readers of varying skill levels at clinical centers in a multicenter prospective cohort study. As a sub-study of the main B-SAFE study, it is demonstrated that the variability in interpretation of BMIPP SPECT between experts and readers or varying skill levels reached only fair agreement, which diminished the prognostic value of BMIPP SPECT considerably. By contrast, the main B-SAFE study reported valuable prognostic information of BMIPP SPECT for cardiac-derived death in patients receiving hemodialysis.

Readers at the clinical centers, who reflected the variety of the skill levels of physicians and radiologists who read BMIPP SPECT images in daily clinical practice, interpreted the BMIPP SPECT images more stringently than the experts at the core laboratory. This led to a lower normalcy rate and assignment of some of the BMIPP SPECT results to a more severe category. We assume that the reason why the Kaplan-Meier analysis, based on the interpretations of clinical center readers, failed to separate cardiac-derived death rates of patients with normal BMIPP SPECT results from those of patients with mildly abnormal BMIPP SPECT results. Although no previous studies have reported the variability in interpretation of BMIPP SPECT, false-positive results from less-experienced readers could be due to artifacts of SPECT. Indeed, readers at clinical centers were inclined to give higher segmental uptake scores, particularly in basal and septal segments and the apex. It might be explained by normal thinning of the basal membranous septum and apex, and attenuation in basal segments.

Results of previous studies assessing interobserver variability in the interpretation of myocardial perfusion imaging were in line with our results. Suleiman et al investigated the interpretative variability of MPS by comparing 2 trainees with 2 experienced readers and found that trainees tended to report more abnormal results than experienced readers. Most previous studies only evaluated interobserver reproducibility in the interpretation of myocardial perfusion imaging by assessing small numbers of readers and studies in a single clinical center or a specialized core laboratory. Only Wackers et al assessed interobserver agreement in the visual interpretation of 201TI planar myocardial perfusion imaging in a large multicenter trial. They reported that the normalcy rate in clinical centers was lower than that in the core laboratory, and that the κ value reached only fair agreement for the overall impression (normal vs. abnormal MPS).

To generalize the results from a large clinical research study in specialized laboratories to all laboratories performing nuclear cardiology imaging, the images must be highly reproducible, and the variation of interpretation should be minimal. Regarding MPS, the American Society of Nuclear Cardiology guidelines recommend image interpretation by experts to ensure reproducibility. However, BMIPP SPECT is less common than MPS and is not necessarily performed routinely in clinical practice at general, community hospitals. Moreover, because of a shortage of experts, the interpretation of BMIPP SPECT images is often performed by cardiologists or other types of physicians rather than nuclear cardiologists.

To improve the interpretive variability of BMIPP SPECT, several resolutions may be considered. First, as reported in previous studies of MPS, the use of automatic quantitative analysis compared with a normal database may improve the reproducibility of image interpretation. Some investigators have reported computed quantitative analysis of MPS and BMIPP SPECT based on normal databases for Japanese populations. Accordingly, the construction of a normal BMIPP uptake database is expected based on the BMIPP datasets in this study. Second, training in interpretation and an increase in the number of the experts in BMIPP SPECT image interpretation are necessary. In regard to MPS, Suleiman et al investigated the interpretive variability between 2 experts and 2 trainees, and reported that sufficient agreement was not only observed after 2 months of intense training, but also even after 2 weeks’ training. And third, the usage of 201TI and 123I-BMIPP dual isotope SPECT might increase reproducibility of image interpretation by referring MPS images. In our study, agreement in interpretation of BMIPP SPECT images was substantially higher in 154 patients who underwent dual isotope SPECT, where both SPECT images were simultaneously assessed, compared with the rest of the 523 patients who underwent BMIPP SPECT only.

Several limitations were considered. First, this study was not designed to evaluate the diagnostic accuracy of BMIPP SPECT. We examined only agreement between readers to assess variability in image interpretation, not the sensitivity and specificity of BMIPP SPECT for the detection of CAD. Second, in clinical centers, the environment of image interpretation was not...
necessarily standardized properly. The method of image display and interpretation varied in different centers. For example, the image display could be a designated picture archiving and communication system (PACS) monitor, workstation monitor, polaroid prints, or radiographs. Display color also could vary from a gray or monotone color scale to a multi-color or combined scale. The cut-off value of SPECT images could also be different. These non-uniform conditions might have affected the results of this study. Moreover, the availability of clinical information might have biased interpretations in the clinical centers. However, we believe clinical information had little effect on interpretation. This is supported by the study conducted by Wackers et al, who reported that clinical information had little effect on the interpretation in the clinical centers enrolled in a multicenter study using planar thallium imaging. Finally, radiation burden must be considered, especially in the dual isotope SPECT method. This dual isotope SPECT method is only available in combination with thallium, but not 99mTc tracers, because of the crosstalk between tracers.

Conclusions

Considerable variability in the visual interpretation of BMIPP SPECT images between experts at the nuclear core laboratory and non-experts at clinical centers was observed, and it had substantial impact on the prognostic value of BMIPP SPECT in asymptomatic hemodialysis patients in the multicenter prospective cohort B-SAFE study.

Acknowledgments

We thank nuclear cardiologists, Tomoaki Nakata, MD (Sapporo Medical University School of Medicine, Sapporo), Akioyoshi Hashimoto, MD (Sapporo Medical University School of Medicine, Sapporo), Keiichiro Yoshinaga, MD (Hokkaido University, Sapporo), and Masahiro Toba, MD (Nippon Medical School, Tokyo), for support with image interpretation; Kaori Kuronaka (Translational Research Informatics Center) for data management; Eiji Nakatani (Translational Research Informatics Center) for statistical analyses; and the technologists and clinical research coordinators at all participating institutions for valuable collaboration and cooperation with the B-SAFE study. This study was supported by the Foundation for Biomedical Research and Innovation (Kobe, Japan). Dr Moroi has received speaker fees from Nihon Medipysics (Kobe, Japan). The other authors declare that they have no relevant financial interests.

References
