Anthracycline cardiotoxicity in childhood cancer. Frequency, risk factors, and early detection
Kremer, L.C.M.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Reproducibility of the assessment of indium-111 antimyosin scintigraphy in children

LCM Kremer\textsuperscript{1}
MC Ubbink\textsuperscript{1}
M Offringa\textsuperscript{2}
MC Tiel-van Buul\textsuperscript{3}
PMM Bossuyt\textsuperscript{4}
PA Voûte\textsuperscript{1}

Department of Pediatric Oncology\textsuperscript{1}
Department of Pediatrics\textsuperscript{2}
Emma Kinderziekenhuis, AMC, University of Amsterdam
Department of Nuclear Medicine\textsuperscript{3}
Department of Clinical Epidemiology and Biostatistics\textsuperscript{4}
Academic Medical Center, Amsterdam, The Netherlands

Submitted for publication
Introduction

In adults, 111-Indium-Antimyosin (111-In-AM) scintigraphy can detect myocardial damage in patients with myocardial infarction, myocarditis, cardiomyopathy, cardiac rejection after heart transplantation, and cardiac dysfunction due to anthracycline.\textsuperscript{1-4} Antimyosin binds to intracellular myosin only when damage to myocytes causes sarcolemmal disruption.\textsuperscript{5} The cardiac uptake of 111-In-AM can be determined qualitatively, by visually assessing 111-In-AM uptake in the myocardium, and quantitatively, by measuring the heart-to-lung ratio (HLR) of 111-In-AM uptake. In adults, an abnormal uptake of the tracer in the heart is defined as an HLR higher than 1.58.\textsuperscript{3,6}

Although two studies have reported on the intra- and inter-observer variability of the analysis of 111-In AM scintigrams in adults, a high correlation does not necessarily mean a high agreement.\textsuperscript{6,7,8} We do not know of studies in which the agreement between different measurements and observers was investigated. 111-In AM scintigraphy has been used in children to detect heart damage caused by anthracycline therapy, but little is known about the reproducibility of the assessment of 111-In AM scintigrams of children.\textsuperscript{9}

The present study was performed to evaluate the intra-and inter-observer agreement of the assessment of 111-In AM scintigrams in children and young adults, and to determine whether the magnitude of the standard deviation of the measurement is clinically relevant.

Methods

Patients

Twenty-seven 111-In AM scintigrams were recorded from 23 patients for four clinical indications: 1/ a baseline scintigraphy before anthracycline therapy in order to evaluate cardiotoxicity (4 scintigrams); 2/ during follow up of anthracycline therapy to evaluate cardiotoxicity (9 scintigrams); 3/ detection of myocarditis (6 scintigrams); 4/ detection of metastasis in rhabdomyosarcoma (8 scintigrams). In 4 patients two scintigrams were for two different indications. The mean age of the patients was 10.2 years (range: 1 month to 25 years). Two patients were older than 18 years. (Table 1)

Antimyosin scintigraphy

Antimyosin scintigraphy was performed using 111-Indium Chloride in 0.5 mg R11-D10-Fab-DTPA (Mallinckrodt Medical B.V., Petten, The Netherlands). The administered activity was 40 MBq in children younger than 10 years, 60 MBq in children aged between 10 and 14 years, and 80 MBq in children older than 14 years. Scintigraphy was performed after
Table 1 Data of the assessments of 27 111-Indium Antimyosin Scintigraphic Studies of 23 patients

<table>
<thead>
<tr>
<th>Patient No</th>
<th>Indication for AM scintigraphy*</th>
<th>Qualitative assessment (No)</th>
<th>Quantitative assessment Mean HLR³</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>1</td>
<td>1/2</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1/2</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>2/2</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>2/2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>4</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>17</td>
<td>4</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>4</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>22</td>
<td>4</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>4</td>
<td>9</td>
<td>1</td>
</tr>
</tbody>
</table>

* Indication for antimyosin scintigraphy: 1: baseline scan before anthracycline therapy, 2: evaluation cardiotoxicity during anthracycline therapy, 3: suspected myocarditis, 4: detection of metastasis of rhabdomyosarcoma.

¹ Number of judgements in the qualitative assessment. A: Normal uptake, B: suspected uptake, C: Increased uptake.

³ Mean heart lung ratio of 15 measurements

48 hours in both the anterior and left anterior oblique 45° views, using a large-field-of-view gamma camera (Diacam, Siemens), fitted with a medium energy all purpose collimator and a 20% window set around both photopeaks of 111-Indium. The images were made in a 128 by 128 matrix with a preset time of 10 minutes.

**Qualitative assessment**

Ten experienced nuclear medicine physicians from eight hospitals were asked to assess independently 25 scintigrams. The observers had to classify the scans into one of three categories: A/ normal uptake, B/ suspected increased uptake and C/ definite increased uptake in the heart. Two scans were used as examples: one showed a normal uptake of
Figure 1 Two 111-Indium-Antimyosin scintigrams of two different patients used as example for the qualitative assessment. A: Anterior view showing a normal tracer uptake in the heart. B: Anterior view showing an increased tracer uptake in the heart.
Figure 2 111-Indium-Antimyosin scintigram with drawn regions of interest used to calculate the heart-to-lung ratio for the quantitative assessment. 1: region of interest over the heart. 2: region of interest over the left lung. 3: region of interest over the right lung.

tracer (category A) and the other showed an extremely high uptake of tracer (category C) (Figure 1).

Quantitative assessment
To calculate the HLR, five observers (three nuclear medicine physicians, two research analysts) were asked to manually draw regions-of-interest (ROI) over the heart, the left lung, and the right lung. Each observer could adjust the screen intensity of the images, by changing the upper and lower threshold. All five observers were experienced in nuclear medicine and in drawing ROI. The drawn regions were not allowed to overlap with liver, kidneys, sternum, mediastinum, or mammae. The HLR was then calculated by dividing the mean counts per pixel of the heart by the mean counts per pixel of the average of the lungs (Figure 2). Each observer repeated the measurements three times at 1-week intervals. This provided a total of 405 HLRs (27 times 5 times 3).

Blinding of the observers
The scintigrams were numbered and the clinical information of the patient was removed. For the repeated measurements, the sequence of the scintigrams was changed.
Statistical analysis

Qualitative assessment
Kappa statistics (κ) was calculated for each pair of observers. The overall mean Kappa value was determined.

Quantitative assessment
In order to determine the intra- and inter-observer agreement of the quantitative assessment, different analyses were used. Analysis of variance was used to identify the amount of variation due to differences between the observers (inter-observer), differences in the repeated measurements (intra-observer), and differences between the patients. Intra-class correlation coefficients (ICC) were calculated to determine the true variance as a fraction of the total variance. The intra-observer ICC was calculated by dividing the patient variance by the sum of the patient variance and the intra-observer variance. The inter-observer ICC was calculated by dividing the patient variance by the sum of the patient variance and the intra- and the inter-observer variance.
To assess the error of a measurement apart from the patient variance, the error standard deviation (error sd) of one measurement was determined by taking the square root of the intra- and inter-observer variance. To assess the change of the error sd with repeated measurement of HLR, the error sd was divided by the square root of the number of measurements. The error sd for each observer was determined by the square root of intra-observer variance. To assess whether this error sd was of potential clinical relevance, we calculated the chance that the mean of two measured HLR values would differ by less than 0.10, 0.15, or 0.20 from the true value. In this calculation we assumed that the measured HLR values had a normal distribution. SPSS 7.5 for Windows was used for statistical calculations.

Results

Qualitative assessment
The overall kappa of the visual evaluation of the 25 scintigrams by the ten observers was 0.41 (confidence interval 0.22-0.60, range: 0.21-0.54).

Quantitative assessment
The mean HLR values of 15 measurements of each of the 27 111-In AM scintigrams ranged from 1.22 to 2.34 (Table 1). The inter-observer variance, patient variance, and intra-observer/error variance were 0.004, 0.061, and 0.013, respectively. The intra-observer ICC was 0.82 and the inter-observer ICC was 0.78. The error sd of 1 measurement was 0.13. The mean of 2 and 3 measurements of the HLR was 0.09 and
0.08, respectively. The error sd of the measurements was different for each observer (range: 0.09-0.17). The chance that the mean of two HLR measurements would differ by more than 0.10, 0.15, or 0.20 from the true value of the HLR was 27%, 10%, and 3%, respectively. In case of the best observer, the error sd of two repeated measurements will decrease from 0.09 to 0.06.

Discussion

Before 111-In AM scintigraphy can be useful in children, we need to know how reproducible the qualitative and quantitative assessments are. In this study we did not address the issue of accuracy, i.e. the relationship between the test results and the disease state or the severity of cardiac cell damage in individuals.

The inter-observer agreement of the ten different observers was moderate. This suggests that it is difficult to distinguish between normal, suspect, or increased uptake. A possible explanation for this low agreement is that while the observers were experienced nuclear medicine physicians, they had limited experience in the qualitative assessment of 111-In AM scintigrams of children. Moreover, tracer uptake was relatively low, even in pathological cases (Figure 1).

The intra- and inter-observer intra-class correlation coefficients (ICC) of the quantitative assessment, as a measurement of agreement, were reasonable. A high ICC indicates that there is a relatively small error variance regardless of the patient variation. The same error variance will give a higher ICC in a heterogeneous study group than in a homogeneous group. Our group was rather heterogeneous (four different indications), with the mean HLR ranging from 1.22 to 2.34. Because the ICC is a relative measure, we also used an alternative approach that is relevant for the clinical interpretation of the reproducibility, namely the error sd of a measurement. Calculation of the mean of two HLR measurements decreased the sd of a measurement from 0.13 to 0.09. A third measurement gave only a further improvement from 0.09 to 0.08. So, the HLR should be measured twice and the mean should be used as a measure of 111-In-AM uptake. The relation between the measured HLR value and the ‘true’ HLR value of one AM scintigram is clinically relevant. In 95% of the cases, the mean of two measurements of the HLR will be in the range of the true HLR value ± 0.18. This means that in 5% of the measurements the mean of two measured HLR values will differ by more than 0.18 units from the true value of the HLR, and in 10% of the cases by more than 0.15. As a consequence, in patients with an HLR that is less than 0.15 from the normal value, scintigraphy results may be classified incorrectly. However, for patients with measured HLR values of more than 0.20 from the normal value, the chances of misclassification are small (3%).

Furthermore, it is of clinical relevance to know whether there is a ‘true’ difference between the HLR measurements of two different AM scintigrams. Because the HLR value
of each scintigram can be in the range of the measured HLR value ± 0.18, a difference of more than 0.36 between two scintigrams will be a ‘true’ difference in 95% of the cases. In case of the best observer, a difference of more than 0.24 represents a ‘true’ difference in the tracer uptake.

The marked difference in the error sd between the observers could not be explained by experience. When we retrospectively looked at the ROIs drawn we found differences between the observers, despite the guidelines given. These differences may be due to precision in drawing the ROI, the size of the ROI, and the color intensity of the image when the ROI was drawn. Drawing a large ROI and drawing a ROI with a high intensity, so that anatomical lines can be distinguished better, will lead to a lower error sd. The reproducibility of the HLR assessment can probably be improved if observers try to reach consensus on the optimal ROIs. Further investigation is required. The present study reveals that the qualitative, visual assessment of 111-In AM scintigrams has a poor inter-observer reproducibility. The quantitative assessment, using the HLR, showed a reasonable intra- and inter-observer agreement and can be used for the detection of myocardial damage in children. For the best results, the HLR should be measured twice and the mean used as a measure of 111-In-AM uptake. The error sd of the mean of two measurements is acceptable. However caution has to be taken when a HLR value is less than 0.18 around the normal value.

Acknowledgement

The authors thank J Baas MD, RJ Bennink MD, RO Boer MD, K de Bruin MD, BLF van Eck-Smit MD Phd, AJM Rijnders MD, W Roolker PhD, F Smit MD, Prof GJJ Teule MD PhD, JF Verzijlbergen MD, DEA Zanin MD and A Zwijnenburg MD PhD for the assessments of the scintigraphic studies and J Habraken MSc for his assistance with the scintigraphy.
References


