ORIGINAL ARTICLE

Assessment of Interobserver Agreement in the Interpretation of ^{99m}Tc-PYP SPECT Imaging of Cardiac Amyloidosis

Aiganym Imakhanova, MD[1)], Reiko Ideguchi, MD, PhD[2)], Akiyo Chiba, MD[3)], and Takashi Kudo, MD, PhD[1), 2)]

Received: June 7, 2023/Revised manuscript received: August 2, 2023/Accepted: August 3, 2023
J-STAGE advance published: September 26, 2023

Abstract

***Background*: Technetium-99m pyrophosphate single photon emission computed tomography (^{99m}Tc-PYP SPECT) imaging is widely used to diagnose cardiac amyloidosis, a disease characterized by amyloid protein deposits in the myocardium. The effects of viewing perspectives on interobserver agreement in the interpretation of ^{99m}Tc-PYP SPECT images for the diagnosis of cardiac amyloidosis remain unclear.**

***Methods*: A retrospective analysis of 32 consecutive patients who underwent ^{99m}Tc-PYP imaging for the diagnosis of cardiac amyloidosis at Nagasaki University Hospital between October 2017 and February 2020 was performed. Four evaluators independently reviewed coronal, sagittal, and transaxial images and then all images together and made a categorical diagnosis based on predefined criteria. Interobserver agreement was analyzed using Cohen's Kappa analysis.**

***Results*: Kappa coefficient values in the four-grade grading system (grades 0–3) ranged between 0.31 and 0.95, while those in the binary grading system (positive/negative) ranged between 0.88 and 1. The sagittal view showed the highest value in the four-grade grading system (0.95) and the lowest in the binary grading system (0.88). The transaxial view was more likely to show a consistently high kappa value in both the four-grade and binary grading systems. The use of the multiplanar view reduced the number of subjects classified as grade 1.**

***Conclusion*: Our study demonstrates that the transaxial view provides the most consistent interpretation of ^{99m}Tc-PYP SPECT images for the diagnosis of cardiac amyloidosis. The use of the multiplanar view may also reduce equivocal interpretations, which are graded as grade 1. Further studies with larger sample sizes and a quantitative analysis are needed to confirm the present results.**

Keywords: ^{99m}Tc-PYP imaging, Cardiac amyloidosis, Diagnostic accuracy, Interobserver agreement, Viewing perspective

Ann Nucl Cardiol 2023; 9 (1): 48–53

Systemic amyloidosis is a group of rare diseases caused by the abnormal accumulation of amyloid proteins in tissues and organs throughout the body (1). In transthyretin cardiac amyloidosis (ATTR-CA), the transthyretin (TTR) protein is deposited in the myocardium, leading to progressive heart failure (2). Although cardiac amyloidosis was once considered a rare disease, it is now recognized as an important cause of heart failure, particularly in the elderly, with an increasing prevalence worldwide (3–5). However, due to the ambiguous clinical characteristics of ATTR-CA, a diagnosis may be challenging. Patients with unexplained heart failure, particularly those with a preserved ejection fraction (HFpEF), need to be examined for cardiac amyloidosis (6).

Radionuclide imaging is currently one of the leading modalities for the diagnosis of cardiac amyloidosis. The sensitivity and specificity of technetium-99m pyrophosphate (^{99m}Tc-PYP) imaging for the diagnosis of cardiac amyloidosis is high and, thus, it has been rapidly accepted by clinical cardiologists (7, 8). A diagnosis is performed by two approaches: a quantitative measurement of the heart-to-

DOI: 10.17996/anc.23-00004

1) Department of Radioisotope Medicine, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan
2) Department of Radioisotope Medicine, Atomic Bomb Disease Institute, Nagasaki University, Nagasaki, Japan
3) Department of Cardiovascular Medicine, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan

contralateral lung ratio on planar images and a visual assessment using the Perugini grade (9, 10). Accurate image interpretation is crucial not only to correctly identify ATTR-CA and assess the severity and extent of cardiac involvement, but also to guide treatment decisions and improve patient outcomes.

The ability of a radiologist to accurately interpret medical images depends on many parameters, including their experience, knowledge, and training. Depending on their level of expertise and experience, radiologists may interpret the same images differently, which leads to divergent diagnoses or evaluations (11). This is an issue not only in radiological imaging, but also in nuclear medicine imaging. In visual assessments of amyloidosis imaging with bone scan agents, it is sometimes difficult to distinguish blood activity in the ventricle/atrial cavity from myocardial uptake. Ventricle pool activity may sometimes be mis-interpreted as wall activity (12). The use of single photon emission computed tomography (SPECT) / computed tomography (CT) fusion images is recommended to simplify visual interpretations, but is not widely available. In perfusion imaging, the heart is typically sliced into three plain-view slices: a short axis, horizontal long axis, and vertical long axis. However, in PYP images with only faint or no myocardial uptake, it is challenging to slice the myocardium in the same manner because of the difficulties associated with visualizing the shape and axis of the left ventricle. In these cases, we generally slice PYP SPECT images into a transaxial image, similar to conventional CT, whole-body SPECT, and PET images. However, it remains unclear whether a transaxial image of PYP is the best approach for its interpretation. The practice guidelines from the American Society of Nuclear Cardiology (ASNC) and other published guidelines recommend the use of SPECT images for visual interpretations. However, recommendations on how to present slice images for interpretations have not yet been established. We hypothesized that using a multiplane image presentation (simultaneously viewing transaxial, coronal, and sagittal images) may avoid misinterpretations and reduce interobserver variability.

Materials and methods

Study population

Thirty-two consecutive patients who underwent ^{99m}Tc-PYP imaging for the diagnosis of cardiac amyloidosis at Nagasaki University Hospital between October 2017 and February 2020 were retrospectively analyzed. This was a retrospective study using a standard clinical examination and was approved by the Ethics Committee of Nagasaki University Hospital (Approval number, 20111631). All procedures conducted in studies involving human participants were in accordance with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Image acquisition

Patients were intravenously injected with 700–900 MBq of ^{99m}Tc-PYP and SPECT images were obtained 3 hours later. SPECT images fused with CT are used for a clinical diagnosis. Due to the purpose of the present study, CT fusion images were not interpreted. Planar images of anterior (for H/CL measurement) and lateral projections are also obtained 1 and 3 hours after an injection of ^{99m}Tc-PYP but were not used in the present study.

Planar and SPECT/CT images were acquired using a Symbia T Series SPECT/CT dual-head camera (Siemens, Germany). Only SPECT data without SPECT/CT fusion were used in this study. SPECT images were acquired using a low-energy high-resolution collimator with a 128 × 128 matrix, magnification = 1.0, 140 keV photon energy, 22% energy window. The detector configuration was 180 degrees × 2 camera, number of views was 36, time per stop was 15 seconds. The pixel size was 4.8 mm. Image was reconstructed using OSEM reconstruction with Gaussian filter. Attenuation correction using CT data and scatter correction were performed, but resolution correction was not performed.

Image interpretation

Four doctors, one expert in nuclear cardiology (TK), one radiologist with 10 years' experience (RI), one cardiologist with 10 years' experience (AC), and one researcher with 3 years' experience in cardiac amyloidosis imaging (AI) visually interpreted SPECT images. Four interpreters were given transaxial, coronal, and sagittal SPECT images with a 5 × 5 display (jpeg format) separately and read them. Interpreters then read three combined images. All images were numbered randomly to prevent them being matched to images used in previous interpretations. All interpretations (transaxial, coronal, sagittal, and multiplanar images) were performed at least two weeks apart to avoid any memory of previous interpretations. All images were classified according to the Perugini grade (0–3).

Statistical analysis

Statistical analyses were performed using JMP Pro15 software. Cohen's Kappa analysis was used to measure interobserver agreement between two observers when evaluating a categorical variable. Interobserver agreement was calculated 1) using the original grade 0–3 grading system and 2) with a binary classification using grades 0 and 1 as negative and 2 and 3 as positive.

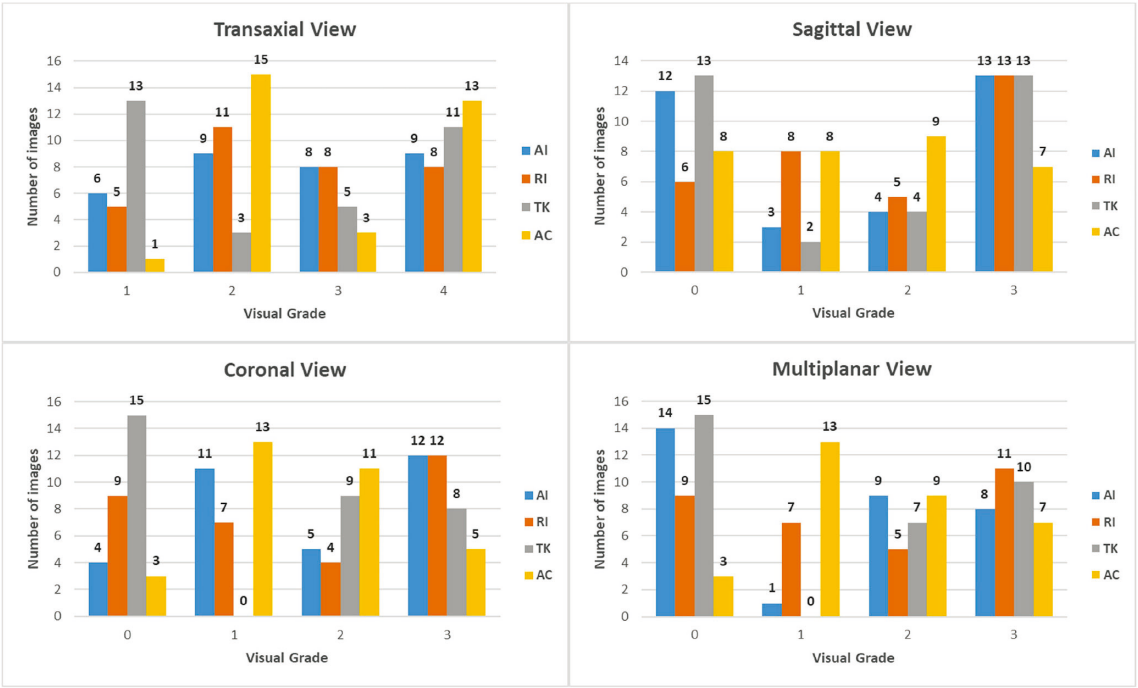

Figure 1　Visual grade distribution of ^{99m}Tc-PYP SPECT images in four different scan views.
The figure shows the results of the inter-observer agreement analysis by 4 observers and 32 images in four different views: transaxial, sagittal, coronal, and multiplanar, using the Perugini score. Each observer scored each image on a scale of 0 to 3 based on the degree of radiotracer uptake, with 0 representing no uptake and 3 representing intense uptake.

Table 1　Kappa coefficients of interobserver agreement

Grades	AI vs TK	AI vs RI	AI vs AC	TK vs RI	TK vs AC	RI vs AC
Transaxial	0.51	0.62	**0.57**	*0.48*	**0.45**	**0.60**
Coronal	*0.42*	*0.49*	0.48	0.54	0.39	*0.53*
Sagittal	**0.95**	**0.65**	0.38	**0.61**	0.34	0.59
Multiplanar	0.85	0.54	*0.31*	0.57	*0.31*	0.59

(The highest values among four interpretations are presented in **BOLD** and the lowest in *ITALICS*)

Results

Distribution of perugini grades

The results on the distribution of Perugini grades for ^{99m}Tc-PYP SPECT images in four different scan views - transaxial, sagittal, coronal, and multiplanar - are shown in Figure 1. Data presented in the figure are organized into four bar graphs, each corresponding to one of the four scan views. The figure shows the number of images graded by each observer in each scan view and the corresponding scores. No significant differences were observed between the four scan views; however, the use of the multiplanar view slightly reduced the number of images graded as grade 1, except for observer AC.

Interobserver agreement of the perugini grade

The agreement between two observers (6 pairs) for four views were analyzed. Agreement tables for each pair (6 pairs ×4 views = 24 tables) are shown in the Appendix.

Kappa coefficients represent the degree of agreement between the four observers for each view of ^{99m}Tc-PYP SPECT images according to the Perugini score, with values ranging between 0 and 1, as shown in Table 1. A kappa coefficient of 0 indicated no agreement beyond chance, while a coefficient of 1 indicated perfect agreement. In the transaxial view, kappa coefficients ranged between 0.45 and 0.62, indicating moderate interobserver agreement. Three of the 6 paired interpreters had the highest kappa value, and one had the lowest kappa value in the transaxial view interpretation. In the coronal view, kappa coefficients ranged between 0.39 and 0.54, indicating moderate agreement among observers. Three of the 6 paired interpreters had the lowest kappa value in the coronal view interpretation. No pair had the highest kappa value. In the sagittal view, kappa coefficients ranged between 0.34 and 0.95, indicating variable agreement among the observers. Three of the 6 paired interpreters had the highest kappa value in the sagittal view interpretation. None of the

Table 2 Kappa coefficients of interobserver agreement

Pos/Neg	AI vs TK	AI vs RI	AI vs AC	TK vs RI	TK vs AC	RI vs AC
Transaxial	0.94	0.94	0.94	**1**	**1**	**1**
Coronal	**1**	0.94	0.94	**1**	0.94	**1**
Sagittal	**1**	0.94	0.94	0.94	0.94	*0.88*
Multiplanar	**1**	0.94	0.94	0.94	0.94	**1**

(The highest values among four interpretations are presented in **BOLD** and the lowest in *ITALICS)*

pairs had the lowest kappa value. In the multiplanar view, kappa coefficients ranged between 0.31 and 0.85, indicating variable agreement among the observers. Two of the 6 paired interpreters had the lowest kappa value in the multiplanar view interpretation. None of the pairs had the highest kappa value. The average and standard deviation of kappa values in each of the four views were as follows: 0.54 ± 0.07, 0.48 ± 0.06, 0.59 ± 0.22, and 0.53 ± 0.20 for the transaxial, coronal, sagittal, and multiplanar views, respectively. These results suggest that the level of agreement among observers varied depending on the views of the ^{99m}Tc-PYP SPECT images analyzed, with the sagittal view showing slightly better agreement.

Table 2 shows kappa coefficients representing the degree of agreement between the four observers for positive and negative findings on ^{99m}Tc-PYP SPECT images, with values ranging between 0 and 1.

In the transaxial view, kappa coefficients ranged between 0.94 and 1, indicating high to perfect agreement among observers. Three of the 6 paired interpreters showed perfect agreement. In the coronal view, kappa coefficients ranged between 0.94 and 1, indicating high to perfect agreement among observers. Three of the 6 paired interpreters showed perfect agreement. In the sagittal view, kappa coefficients ranged between 0.88 and 1, indicating high to perfect agreement among observers. One of the 6 paired interpreters showed perfect agreement. In the multiplanar view, kappa coefficients ranged between 0.94 and 1, indicating high to perfect agreement among observers. Two of the 6 paired interpreters showed perfect agreement. The averages and standard deviations of kappa values in each of the four views were as follows: 0.97 ± 0.03, 0.97 ± 0.03, 0.94 ± 0.04, and 0.96 ± 0.03 for transaxial, coronal, sagittal, and multiplanar views, respectively. These results suggest that the level of agreement among observers was generally high to perfect for positive and negative findings on ^{99m}Tc-PYP SPECT images across all views. Furthermore, the grading system used to analyze these images appeared to have good consistency and reliability for identifying both positive and negative findings.

Discussion

In the present study, we assessed interobserver agreement in the visual interpretation of ^{99m}Tc-PYP SPECT images for the diagnosis of cardiac amyloidosis using four different scan views. The results obtained showed moderate to good interobserver agreement in the transaxial, coronal, and multiplanar views, with kappa coefficients ranging between 0.42 and 0.85 for the four-grade grading system and good to perfect agreement with kappa coefficients ranging between 0.88 and 1 for the binary (positive/negative) classification system. The highest agreement with a kappa coefficient of 0.95 for the four-grade grading system was found in the sagittal view interpretation, which was the lowest for the binary classification system (kappa 0.88). Interpretations of transaxial images consistently showed high agreement in both the four-grade grading and binary classification systems (highest kappa value in 3 of the 6 interobserver pairs, perfect agreement in 3 of the 6 pairs). However, the use of the multiplanar view slightly reduced the number of subjects graded as grade 1. These results indicate that when only a single view of ^{99m}Tc-PYP SPECT is allowed to be interpreted, the interpretation of the transaxial view may provide consistent results.

The reason why the transaxial view showed consistently high agreement in both the four-grade and binary classification systems is difficult to explain. One simple reason is familiarity. Two of our interpreters are non-radiologists/nuclear cardiologists. The use of the transaxial image in medical imaging adheres to a standardized orientation that is recognized and familiar to non-radiologists. The transaxial view facilitates recognition of characteristic patterns of amyloid deposition, including global or focal involvement and preferential deposition in certain segments. This familiarity promotes better interobserver agreement in image interpretation.

The present results agree with previous findings on interobserver agreement in the visual interpretation of ^{99m}Tc-PYP SPECT images for the diagnosis of cardiac amyloidosis. For example, a study by Caobelli et al (13) evaluated interobserver agreement in the visual interpretation of ^{99m}Tc-PYP SPECT images for the diagnosis of cardiac amyloidosis using a binary scoring system and found good interobserver agreement (kappa = 0.65) in the interpretation of ^{99m}Tc-PYP SPECT images, which was similar to interobserver agreement in the present study.

The use of the multiplanar view did not result in good agreement in the four-grade grading system. However, regarding the binary classification system (positive/negative), the use of the multiplanar view resulted in good agreement and slightly reduced the number of subjects graded as grade 1. Perugini grade 1 is an equivocal grade. In the American Heart Association statement (14), grade 1 is considered to be negative, whereas in the European Society of Cardiology recommendation (15), biopsy is recommended for patients classified as grade 1. The management of subjects classified as grade 1 is still considered controversial. Therefore, there may be some benefit to reducing the number of patients classified as grade 1 using the multiplanar view.

Our study has certain limitations that warrant acknowledgment. Firstly, the absence of a well-defined gold standard and the challenges associated with histological assessment in patients with cardiac amyloidosis hindered the establishment of a definitive standard. Myocardial biopsy, recommended for definite diagnosis, was performed in only 10 cases, with 6 cases confirmed as ATTR cardiac amyloidosis. However, due to its invasiveness, myocardial biopsy is infrequently conducted. Considering the rapid acceptance of PYP (and other bone tracers) in cardiac amyloidosis within the medical community, along with its use in general hospitals, standardization of image interpretation was deemed important. Hence, our focus was on assessing standardization of image interpretation and interobserver agreement rather than sensitivity or specificity. Secondly, our analysis did not include SPECT/CT. While we believe SPECT/CT should become the standard method, a significant number of installed two-head gamma cameras in Japan lack SPECT/CT capabilities (438 out of 1088) (16). Therefore, we concentrated on SPECT image interpretation without CT fusion. For the same reason, we did not use images sliced into short axis/long axis, which is hard to perform in case of poor PYP uptake without help of CT fused images. Thirdly, the imaging protocol was slightly suboptimal compared to the standard acquisition protocol (40 views, 20 seconds each) recommended by ASNC. Our acquisition protocol was determined as part of the bone scan protocol, and while the agreement between observers may improve using the ASNC standard protocol, we believe that the observed tendency will remain consistent across present and ASNC protocols, as the difference in image quality is expected to be minor.

Conclusion

The present study demonstrated that the transaxial view may provide the most consistent interpretation of ^{99m}Tc-PYP SPECT images for the diagnosis of cardiac amyloidosis. The use of the multiplanar view may also reduce equivocal interpretations, which are graded as grade 1.

Acknowledgments

None.

Sources of funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Conflicts of interest

The authors have no relevant financial or non-financial interests to disclose.

Reprint requests and correspondence:
Aiganym Imakhanova, MD
Department of Radioisotope Medicine, Graduate School of Biomedical Sciences, Nagasaki University, Sakamoto 1-12-4, Nagasaki 852-8523, Japan
E-mail: imakhanova.a@mail.ru

References

1. Wechalekar AD, Gillmore JD, Hawkins PN. Systemic amyloidosis. Lancet 2016; 387: 2641–54.
2. Dubrey SW, Hawkins PN, Falk RH. Amyloid diseases of the heart: Assessment, diagnosis, and referral. Heart 2011; 97: 75–84.
3. Bianco M, Parente A, Biolè C, et al. The prevalence of TTR cardiac amyloidosis among patients undergoing bone scintigraphy. J Nucl Cardiol 2021; 28: 825–30.
4. González-López E, Gallego-Delgado M, Guzzo-Merello G, et al. Wild-type transthyretin amyloidosis as a cause of heart failure with preserved ejection fraction. Eur Heart J 2015; 36: 2585–94.
5. Quock TP, Yan T, Chang E, Guthrie S, Broder MS. Epidemiology of AL amyloidosis: A real-world study using US claims data. Blood Adv 2018; 2: 1046–53.
6. Rubin J, Maurer MS. Cardiac amyloidosis: Overlooked, underappreciated, and treatable. Annu Rev Med 2020; 71: 203–19.
7. Bokhari S, Cerqueira MD. Tc-99m-PYP imaging for cardiac amyloidosis: Defining the best protocol before the flood gates burst. J Nucl Cardiol 2020; 27: 1816–9.
8. Castano A, Haq M, Narotsky DL, et al. Multicenter study of planar technetium 99m pyrophosphate cardiac imaging: Predicting survival for patients with ATTR cardiac amyloidosis. JAMA Cardiol 2016; 1: 880–9. 2019; 10032.
9. Dorbala S, Bokhari S, Miller EJ, Bullock-Palmer R, Soman P, Thompson R. ASNC Practice Points: 99mTechnetium-Pyrophosphate Imaging for Transthyretin Cardiac Amyloidosis (American Society of Nuclear Cardiology website). 2019. Available at https: //www. asnc. org/Files/Amyloid/ ASNC%20Practice%20Point-99mTechnetium-Pyrophosphate.2019.pdf.
10. Dorbala S, Cuddy S, Falk RH. How to image cardiac amyloidosis: A practical approach. JACC Cardiovasc Imaging 2020; 13: 1368–83.
11. Benchoufi M, Matzner-Lober E, Molinari N, Jannot A-S,

Soyer P. Interobserver agreement issues in radiology. Diagn Interv Imaging. 2020; 101: 639–41.
12. Singh V, Falk R, Di Carli MF, Kijewski M, Rapezzi C, Dorbala S. State-of-the-art radionuclide imaging in cardiac transthyretin amyloidosis. J Nucl Cardiol 2019; 26: 158–73.
13. Caobelli F, Braun M, Haaf P, Wild D, Zellweger MJ. Quantitative 99mTc-DPD SPECT/CT in patients with suspected ATTR cardiac amyloidosis: Feasibility and correlation with visual scores. J Nucl Cardiol 2020; 27: 1456–63.
14. Kittleson MM, Maurer MS, Ambardekar AV, et al. Cardiac amyloidosis: Evolving diagnosis and management: A scientific statement from the American Heart Association. Circulation 2020; 142: e7–22.
15. Garcia-Pavia P, Rapezzi C, Adler Y, et al. Diagnosis and treatment of cardiac amyloidosis: A position statement of the ESC Working Group on Myocardial and Pericardial Diseases. Eur Heart J 2021; 42: 1554–68.
16. Subcommittee on Survey of Nuclear Medicine Practice in Japan, Medical Science and Pharmaceutical Committee, Japan Radioisotope Association. The Present state of nuclear medicine practice in Japan —A report of the 9th nationwide survey in 2022—. Radioisotopes 2023; 72: 49–100. [Article in Japanese]