Can emergency physicians reliably interpret cardiac CT images? A prospective observational study

Joon-myoung Kwon*, Joonghee Kim*, Kyuseok Kim, Taeyun Kim, You Hwan Jo, Jin Hee Lee, Jae Hyuk Lee, Yu Jin Kim, Jae Yun Jung
Department of Emergency Medicine, Seoul National University Bundang Hospital, Seongnam, Korea

Objective Cardiac computed tomography (CCT) is useful for evaluation of acute chest pain in the emergency department (ED). Though the test needs proper interpretation by someone with expertise in cardiovascular imaging, the critical nature of the information the test provides frequently lead emergency physicians (EPs) to act on their own interpretation. We performed this study to assess how often EPs' interpretations are in agreement with radiologists'.

Methods This study is a prospective observational study. The target population was patients assessed with CCT for acute chest pain or discomfort. EPs with at least one year CCT experience underwent a one-hour training session before study participation. The most significant lesion, if any, in each arterial segment was assessed for coronary stenosis and plaque calcification. The agreement between EPs' and radiologists' interpretation was assessed with Cohen's kappa and Gwet's AC1.

Results One hundred and three patients were enrolled and 412 segments were analyzed. Stenosis grading was identical in 363 segments (88.1%) and the interrater agreement was good (kappa=0.6439, AC1=0.8810). Similarly, the plaque calcification grading was identical in 354 segments (86.6%) and the kappa and AC1 values were 0.5660 and 0.8501, respectively. EPs classified 6 of the 17 arterial segments with significant stenosis reported by radiologists as non-significant stenosis (n = 5) or clear (n = 2), all of which were proved to be significant by following subsequent invasive coronary angiography.

Conclusion There was substantial discordance of CCT interpretation between EPs and radiologists. For now, EPs need more education prior to independent CCT reading.

Keywords Multidetector computed tomography; Chest pain; Cardiac imaging techniques; Atherosclerotic plaque

What is already known
Cardiac CT has become a critical diagnostic test in the arsenal of emergency physicians. However, it is unknown whether they can properly interpret the test.

What is new in the current study
There was substantial discordance of cardiac CT interpretation between emergency physicians and radiologists. For now, emergency physicians need further education prior to independent cardiac CT reading.
INTRODUCTION
Cardiac computed tomography (CT) is useful for evaluation of undifferentiated acute chest pain in the emergency department (ED). It provides a comprehensive description of the anatomy of a heart and coronary arteries within 10 to 20 minutes non-invasively. It has high negative predictive value for acute coronary syndrome (ACS) and can also detect serious non-coronary etiologies of chest pain such as pulmonary embolism or aortic dissection. These characteristics have appeal to the needs of busy emergency physicians (EPs) and the test is rapidly gaining popularity. Currently, the most up-to-date guideline supports its use in low to intermediate risk patients without obvious ischemic electrocardiographic (ECG) changes or biomarker elevation.

Because of its inherent difficulty, the test needs proper interpretation by someone with expertise in cardiovascular imaging. However, the critical nature of the information the test provides frequently leads EPs to act on their own interpretation before the formal report is available despite their questionable image interpreting skills. Therefore it is necessary to evaluate how often the average EPs' interpretation agrees with the radiologists' and to determine whether such behavior can be regarded as safe for patients. The objective of this study was to measure the interrater agreement between EPs and radiologists in the interpretation of the cardiac CT images of ED patients with acute chest pain.

METHODS
This study was a prospective observational study assessing the interrater agreement of cardiac CT image interpretation between ED physicians and radiologists. The Institutional Review Board of Seoul National University Bundang Hospital approved the study (IRB number: B-1110/137-001).

Study setting
The study facility was a urban teaching hospital with an annual ED census of 65,000 visits per year. The study was conducted in 2008 when the facility used a 64-channel multi-detector CT for cardiac imaging. At the beginning of the study, the cardiac CT imaging had been actively used for more than a year in the study ED for evaluation of patients with acute chest pain.

Study participants and patient characteristics
A heterogeneous group of emergency medicine (EM) residents including two postgraduate year (PGY)-2 residents, two PGY-3 residents, and two PGY-4 residents participated in this study. As cardiac CT imaging had been actively used in the study ED for more than a year, they had basic knowledge about how to read a cardiac CT image. At the beginning of the study, they had undergone a one-hour training session (didactic, image review) provided by a radiologist specialized in cardiovascular imaging. The participating EPs were not blinded to the clinical information of enrolled patients which included initial examination findings, ECG and biomarker levels.

Patients presenting with acute chest pain who were determined to require a cardiac CT imaging after initial evaluation that included both ECG and cardiac biomarker tests were prospectively enrolled by the participating EPs unless the following exclusion criteria were met (1) age less than 18 years, (2) pregnancy, (3) contraindication to iodinated contrast or β-blocking agents, (4) atrial fibrillation or markedly irregular rhythm, (5) renal insufficiency (creatinine > 1.4 mg/dL), (6) high risk features or previous coronary artery bypass grafting or recent (within 6 months) percutaneous coronary intervention. High risk features included (1) ST-segment elevation, new onset LBBB or pathologic Q-wave in more than two consecutive leads, (2) positive cardiac biomarkers, (3) typical chest pain with ischemic ST–T change (ST–segment depression > 1 mm or T-wave inversion > 3 mm in more than two consecutive leads that is not proven to be old), (4) any clinical feature of decompensated heart failure, (5) recurrent ventricular arrhythmia or high-degree atrioventricular block, (6) ongoing or recurrent angina at rest or minimal effort, (7) history of recent revascularization within 6 months, and (8) previously documented high-degree stenosis without history of revascularization. Patients who developed any of the high risk features after cardiac CT imaging were not excluded. Patients with suboptimal image result were also not excluded.

Interventions
All enrolled patients were provided with the usual standard care which was consistent with 2005 American Heart Association guidelines for acute coronary syndrome. For heart rate control, 100 mg of metoprolol was administered orally to patients with resting heart rate above 70 bpm. If further rate control was indicated, intravenous esmolol bolus injection was administered just before the test. Participating EPs read cardiac CT images right after the imaging process. Relevant clinical information and the CT findings observed by the participating physicians were recorded on a study registry form. Although the participating physicians were allowed to make clinical decisions based on their own readings, it was recommended that they wait for a formal report from the radiology department before making any new significant decision. Patients without significant coronary and/or non-coronary lesions were discharged from the ED if no other conflicting clini-
Can emergency physicians reliably interpret cardiac CT?

We calculated and reported both of the coefficients in addition to the proportionate agreement. The data were analyzed using IBM SPSS ver. 19.0 (IBM Co., Armonk, NY, USA) and Agreestat (Advanced Analytics, Gaithersburg, MD, USA; http://www.agreestat.com).

RESULTS

A total of 104 patients with primary complaints of acute chest pain or discomfort were prospectively enrolled from May 2008 to June 2008. One patient with previous history of coronary artery bypass graft was excluded due to protocol violation. Baseline characteristics of the remaining 103 patients are summarized in Table 1. The mean age was 56 ± 2 years, with 47.6% being men. Heart rate control was required in 84 patients (81.6%). Average heart rate just before CT imaging was 65 ± 8. EPs described 22 cardiac CT (21.4%) as difficult to read. Relatively frequent complaints were beam-hardening/blooming effect due to dense calcification (n = 11), motion artifact (n = 6), and hypoplastic vessel (n = 4). Following CT acquisition, two patients’ serial troponin level increased above reference range and six patients’ serial ECGs showed significant dynamic ST-T change.

The detection of stenosis on each of the four anatomic locations including LM, LAD, LCx, and RCA by EPs was compared to
Table 2. Interrater agreement of coronary artery stenosis between emergency physicians and radiologists

| Agreement | Whole segments | “Not difficult” subgroup | LM | LAD | LCx | RCA |
|-----------|----------------|--------------------------|----|-----|-----|-----|
| Percent agreement (%) | 88.1 | 94.1 | 87.4 | 86.4 | 91.3 | 87.4 |
| AC1 | 0.88 | 0.95 | 0.87 | 0.85 | 0.93 | 0.88 |
| Kappa | 0.64 | 0.76 | 0.14 | 0.74 | 0.69 | 0.64 |
| Presence of significant lesion (%) | 96.1 | 99.1 | 95.1 | 95.1 | 98.1 | 96.1 |

LM, left main coronary artery; LAD, left anterior descending coronary artery; LCx, left circumflex coronary artery; RCA, right coronary artery; AC1, first-order agreement coefficient.

that of radiologists’ formal reports (Table 2). Among the 412 arterial segments that were analyzed by the radiologists, 17 arterial segments (4.1%) showed significant stenosis (> 50% stenosis) and 53 arterial segments (12.9%) showed non-significant stenosis. The grading of stenosis was identical in 363 segments (88.1%) and the interrater agreement was good (kappa = 0.6439, AC1 = 0.8810). Cases rated as ‘not difficult to read’ (n = 324, 78.6%) were analyzed separately. In 305 segments (94.1%), the grading of stenosis was identical, and the calculated kappa and AC1 values were 0.7552 and 0.9306, respectively. Separate analysis of each coronary artery segment showed a similar rate of agreement. Stenosis grading of the LM, LAD, LCx, and RCA was identical in 87.4%, 86.4%, 91.3%, and 87.4%, respectively. Each segments’ kappa values varied significantly (0.1421, 0.7400, 0.6929, 0.6384) while their AC1 values did not (0.8667, 0.8499, 0.9266, 0.8813).

Of the seventeen arterial segments reported as having significant stenosis by radiologists, EPs classified six as non-significant (Fig. 1, Table 3). EPs also classified 9 of 395 segments that were reported as non significant by radiologists as having significant stenosis of which two were proved to be significant by subsequent invasive coronary angiography. All other discrepancies were regarding the presence of any minor (non-significant) stenosis.

Four hundred and nine arterial segments were analyzed for interrater agreement of plaque characterization. Three hundred fifty-four segments (86.6%) were identical and the kappa and AC1 values were 0.5660 and 0.8501, respectively. Among 323 (80.0%) ‘not difficult to read’ segments 300 segments (92.9%) were identically classified. Overall kappa and AC1 value were 0.6626, 0.9234, respectively. Separate analysis of the LM, LAD, LCx, and RCA showed kappa values of 0.3036, 0.6338, 0.4297, and 0.4297, respectively. AC1 value of each of them was 0.8982, 0.7784, 0.8408, and 0.8408 (Table 4).

Among the 103 study patients, 19 patients (18.5%) were admitted for further evaluation and management. Seventeen patients were admitted to cardiology and two to other departments. Fifteen of the 17 patients underwent coronary angiography and 10 were found to have significant stenosis. The two patients who were admitted to the non-cardiology wards were diagnosed as common bile duct stone and newly-diagnosed lymphoma, respectively. Among the 84 patients who were discharged, two patients were admitted at a follow-up visit to the cardiology outpatient department because of recurrent symptoms. Both of them underwent invasive coronary angiography which found no significant fixed stenosis. Each of them were given a presumptive diagnosis of vasospasm and myocardial bridging, respectively. Sixty-eight patients without significant stenosis were followed more than a year after ED discharge. None of them experienced MACE during the first year.

DISCUSSION

There have been several studies measuring the interrater agreement of cardiac CT interpretation between radiologists. Their study population and the method of measurement (visual or machine
Plaque characterization can help EPs to localize culprit lesions. The three CT characteristics of culprit lesions (positive vascular remodeling, non-calcified plaque with low HU [< 30], spotty calcification) were previously reported. In the present study, we could test the calcification density of plaques only. The overall interrater agreement of plaque characterization using this method was even worse than that of stenosis grading (kappa = 0.5660, AC1 = 0.8501). However, they were better in the ‘not difficult to read’ subgroup (kappa = 0.6626, AC1 = 0.9234).

Limitations of this study are as follows. (1) Even the formal reports from radiologists specialized in cardiovascular imaging are not the gold-standard. And this study we measured only the interrater agreement of coronary CT interpretation between radiologists and EPs. (2) The agreement coefficients such as kappa should be interpreted in the context of population characteristics. (3) More detailed description of plaque and stenosis localization using an 18 or 28 segment system has been recommended for formal reporting of coronary CT reading. However we think it is not practical as EPs who participated in this study have limited knowledge and such fine segmentation could cause spuriously low interrater agreement due to discordant localizations. (4) Ancillary findings such as coronary calcium score, wall-motion abnormalities, systolic function and extra-cardiac abnormalities were not analyzed, because of the scarcity of such findings. (5) Only a limited number of EM residents from a single institution participated in this study and the results of this study cannot be generalized to all EM physicians.

In conclusion, the overall interrater agreement for stenosis grading and plaque characterization between EPs and radiologists was not good enough to justify EPs’ independent interpretation and further management based on it. Therefore, for now, EPs need further education before being allowed to independently read cardiac C Ts.

**CONFLICT OF INTEREST**

No potential conflict of interest relevant to this article was reported.

---

### Table 3. Missed significant lesions (>50% stenosis) by EPs

| Age | Sex | Location | EP interpretation | Difficulty | Radiologist interpretation | Coronary angiography |
|-----|-----|----------|-------------------|------------|---------------------------|----------------------|
| 72  | F   | LM       | < 50%, mixed      | Difficult  | 50%–60%, mixed            | 50%–60%              |
| 42  | F   | LAD      | < 50%, mixed      | Difficult  | 60%, mixed                | > 90%                |
| 65  | M   | LAD      | No lesion         | Easy       | > 90%, non-calcified       | 90%                  |
| 70  | F   | LAD      | No lesion         | Difficult  | 80%, mixed                | 90%                  |
| 65  | F   | LAD      | < 50%, mixed      | Difficult  | 95%–99%, non-calcified     | 99%                  |

EP, emergency physician; LM, left main coronary artery; LAD, left anterior descending coronary artery.

### Table 4. Interrater agreement of plaque characterization between EPs and radiologists

| Agreement | Whole segments | "Not difficult" subgroup | LM | LAD | LCx | RCA |
|-----------|----------------|--------------------------|----|-----|-----|-----|
| Percent agreement (%) | 86.6 | 92.9 | 90.3 | 81.6 | 85.4 | 86.4 |
| AC1       | 0.85 | 0.92 | 0.90 | 0.78 | 0.84 | 0.84 |
| Kappa     | 0.57 | 0.66 | 0.30 | 0.63 | 0.43 | 0.43 |

EP, emergency physician; LM, left main coronary artery; LAD, left anterior descending coronary artery; LCx, left circumflex coronary artery; RCA, right coronary artery; AC1, first-order agreement coefficient.
REFERENCES

1. Hoffmann U, Bamberg F. Is computed tomography coronary angiography the most accurate and effective noninvasive imaging tool to evaluate patients with acute chest pain in the emergency department? CT coronary angiography is the most accurate and effective noninvasive imaging tool for evaluating patients presenting with chest pain to the emergency department. Circ Cardiovasc Imaging 2009;2:251-63.

2. Hollander JE, Litt HI, Chase M, Brown AM, Kim W, Baxt WG. Computed tomography coronary angiography for rapid disposition of low-risk emergency department patients with chest pain syndromes. Acad Emerg Med 2007;14:112-6.

3. Broder J, Warshauer DM. Increasing utilization of computed tomography in the adult emergency department, 2000-2005. Emerg Radiol 2006;13:25-30.

4. Taylor AJ, Cerqueira M, Hodgson JM, et al. ACCF/SCCT/ACR/AHA/ASE/ASNC/NASCI/SCAI/SCMR 2010 appropriate use criteria for cardiac computed tomography: a report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, the Society of Cardiovascular Computed Tomography, the American College of Radiology, the American Heart Association, the American Society of Echocardiography, the American Society of Nuclear Cardiology, the North American Society for Cardiovascular Imaging, the Society for Cardiovascular Angiography and Interventions, and the Society for Cardiovascular Magnetic Resonance. J Am Coll Cardiol 2010;56:1864-94.

5. Thomas JD, Zoghbi WA, Beller GA, et al. ACCF 2008 Training Statement on Multimodality Noninvasive Cardiovascular Imaging A Report of the American College of Cardiology Foundation/American Heart Association/American College of Physicians Task Force on Clinical Competence and Training Developed in Collaboration With the American Society of Echo-cardiography, the American Society of Nuclear Cardiology, the Society of Cardiovascular Computed Tomography, the Society for Cardiovascular Magnetic Resonance, and the Society for Vascular Medicine. J Am Coll Cardiol 2009;53:125-46.

6. Lowe RA, Abbaugh SB, Baumrinner A, et al. Radiology services in emergency medicine residency programs: a national survey. Acad Emerg Med 2002;9:587-94.

7. Eng J, Mysko WK, Weller GE, et al. Interpretation of Emergency Department radiographs: a comparison of emergency medicine physicians with radiologists, residents with faculty, and film with digital display. AJR Am J Roentgenol 2000;175:1233-8.

8. ECC Committee, Subcommittees and Task Forces of the American Heart Association. 2005 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation 2005;112(24 Suppl):IV1-203.

9. Motoyama S, Kondo T, Sarai M, et al. Multislice computed tomographic characteristics of coronary lesions in acute coronary syndromes. J Am Coll Cardiol 2007;50:319-26.

10. Leber AW, Knez A, Becker A, et al. Accuracy of multidetector spiral computed tomography in identifying and differentiating the composition of coronary atherosclerotic plaques: a comparative study with intracoronary ultrasound. J Am Coll Cardiol 2004;43:1241-7.

11. Schroeder S, Flohr T, Kopp AF, et al. Accuracy of density measurements within plaques located in artificial coronary arteries by X-ray multislice CT: results of a phantom study. J Comput Assist Tomogr 2001;25:900-6.

12. Becker CR, Nikolaou K, Maders M, et al. Ex vivo coronary atherosclerotic plaque characterization with multi-detector-row CT. Eur Radiol 2003;13:2094-8.

13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.

14. Uebersax JS. Diversity of decision-making models and the measurement of interrater agreement. Psychol Bull 1987;101:140-6.

15. Feinstein AR, Cicchetti DV. High agreement but low kappa: II. The problems of two paradoxes. J Clin Epidemiol 1990;43:551-8.

16. Cicchetti DV, Feinstein AR. High agreement but low kappa: I. The measurement of interrater agreement. Psychol Bull 1987;101:140-6.

17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.

18. Uebersax JS. Diversity of decision-making models and the measurement of interrater agreement. Psychol Bull 1987;101:140-6.

19. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol 1990;43:543-9.

20. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. J Clin Epidemiol 1990;43:551-8.

21. Gwet KL. Handbook of inter-rater reliability. Gaithersburg, MD: STATAxis Publishing Co.; 2002.

22. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. Br J Math Stat Psychol 2008;61(Pt 1):29-48.

23. Gouya H, Varenne O, Trinquart L, et al. Coronary artery stenosis in high-risk patients: 64-section CT and coronary angiography—prospective study and analysis of discordance. Radiology 2009;252:377-85.

24. Higashi M. Noninvasive assessment of coronary plaque using multidetector row computed tomography: does MDCT accurately estimate plaque vulnerability? (Con). Circ J 2011;75:1522-8.

25. Raff GL, Abidov A, Achenbach S, et al. SCCT guidelines for the interpretation and reporting of coronary computed tomographic angiography. J Cardiovasc Comput Tomogr 2009;3:122-36.