

# Evaluating the Accuracy of Emergency Medicine Resident Interpretations of Abdominal CTs in Patients with Non-Traumatic Abdominal Pain

Mun Ju Kang, Min Seob Sim,  
Tae Gun Shin, Ik Joon Jo,  
Hyoung Gon Song, Keun Jeong Song,  
and Yeon Kwon Jeong

Department of Emergency Medicine, Samsung  
Medical Center, Sungkyunkwan University School of  
Medicine, Seoul, Korea

Received: 19 March 2012  
Accepted: 27 July 2012

Address for Correspondence:  
Min Seob Sim, MD  
Department of Emergency Medicine, Samsung Medical Center,  
Sungkyunkwan University School of Medicine, 81 Irwon-ro,  
Gangnam-gu, Seoul 135-710, Korea  
Tel: +82-2-3410-2053, Fax: +82-2-3410-0012  
E-mail: minsub01.sim@samsung.com

Abdominal computed tomography (CT) is widely used as a diagnostic tool in emergency medicine (EM) to accurately diagnose abdominal pain. EM residents must be able to offer preliminary interpretations of CT imaging. In this study, we evaluated the preliminary interpretation ability of a sample of emergency residents presented with adult abdominal CT images, and compared their results with those of radiology residents. We conducted a prospective observational study from November 16, 2008 to June 30, 2009. During this time, we gathered preliminary interpretations of consecutive abdominal CT made by emergency and radiology residents. We assessed the discrepancy rates of both samples by comparing their findings to the final reports from attending radiologists. A total of 884 cases were enrolled in the present study. The discrepancy rates of emergency and radiology residents were 16.7% and 12.2%, respectively. When female genital organs, peritoneum, adrenal glands, or the musculoskeletal system were abnormal, we found that emergency residents' preliminary interpretations of CT images were insufficient compared to those of radiology residents. Therefore more formal education is needed to emergency residents. If possible, the preliminary interpretations of radiology attending physicians are ideal until improving the ability of interpretations of emergency residents in abdomen CT.

**Key Words:** Abdominal Pain; CT Scan; Emergency Medicine; Radiographic Image Interpretation

## INTRODUCTION

Acute abdominal pain is a common chief complaint in clinical practice, particularly in the emergency department (ED) (1). In the ED, prompt differential diagnosis of acute abdominal pain is essential, since delayed or missed diagnoses are associated with high morbidity and mortality (1-4). In order to obtain an accurate diagnosis, abdominal computed tomography (CT) has been widely used as a diagnostic tool (1-4). Therefore, the ability offer accurate preliminary interpretations of CT images is required for an emergency resident (5).

The ability of radiology residents to interpret radiologic images, including CT images, has been well established in previous studies (6-15). However, only a few such studies have been conducted for clinical department residents, focusing on simple radiograph or CT images of traumatic patients (5, 16). Therefore, in this study, we evaluated the ability of emergency residents to accurately interpret abdominal CT images for non-traumatic patients.

We conducted a prospective observational study to investigate the discrepancy rates of preliminary reports of emergency

abdominal CT imaging of emergency residents and to compare them to the results of similar studies of radiology residents. We also analyzed the factors-including major abnormal organs, time of interpretation, and grades of residents that affect the accuracy of emergency resident reports. To evaluate the relationship between discrepancies and clinical outcomes, we compared ED stay time, admission rate, admission length, operation rate, revisit rate, and mortality rate of patients whose abdominal CT were inaccurately preliminary interpreted with patients whose CT were accurately interpreted.

## MATERIALS AND METHODS

### Subjects and study period

We enrolled all CT scan protocols involving the entire abdomen and pelvic cavity, e.g., abdominal pelvic scan, liver scan, biliary-gall bladder scan, pancreas scan, abdominal CT angiography, that were ordered by the ED at the Samsung Medical Center, a tertiary educational hospital, from November 16, 2008 to June 30, 2009. A Centricity Web-Fax system (G.E., Fairfield, CT, USA) was used for all radiologic studies. Exclusion criteria included

trauma patients, pediatric patients (under 15 yr), and patients whose chief complaint was not abdominal pain. Cases where first preliminary interpretations were performed by attending faculty in emergency, radiology, or radiology residents were excluded. Cases where the primary pathognomonic findings on the final formal report were not for abdominal organs, but for lungs that were visualized by the abdominal CT, were excluded.

### Data collection

In order to collect preliminary interpretation results from emergency and radiology residents, data collection sheets were created. First, emergency residents recorded their preliminary interpretations on the data collection sheets. Second, the residents requested and received preliminary interpretation results from the radiology department by telephone, and entered this information on the data collection sheets. The grades assigned by emergency and radiology residents who performed the preliminary interpretations, and the time of the interpretation, were also collected. The time of interpretation was classified as either on-duty time (8 a.m. and 5 p.m. Monday to Friday) or off-duty time.

After requesting and receiving preliminary interpretations from radiology, the emergency residents supplied patient information, including chief complaint and laboratory findings, to the radiology residents. Discussions about the CT images were not permitted.

### Abnormal organ classification and discrepancies

In this study, the final formal reports that were interpreted by radiology attending physicians were used as the gold standard to interpret the CT scans. The abnormal organs identified by the final formal reports were classified into nine categories: liver and spleen, gall bladder and biliary tract, pancreas, hollow viscera except appendix, appendix, urinary system (kidney, ureter, bladder), female genital organ (uterus, ovary, pelvic cavity), blood vessels, and miscellaneous (peritoneum, adrenal gland, musculoskeletal system). If there were more than two abnormal organs in the final formal report, only one organ was selected based on the most pathognomonic findings of the report. If there were no radiological abnormal findings on the final formal reports, the abnormal organs were classified as non-specific abdominal pain.

A discrepancy was defined as a change in diagnosis, treatment, or disposition. Minor differences between final formal reports and preliminary interpretations that did not result in a change were not considered discrepancies. Discrepancies were classified into three categories, with the first being false negative in which the most important diagnosis in the final reports was not made in preliminary reports. The second category represented a misdiagnosis, in which preliminary interpretations and final interpretations described the same organs, but yield-

ed different diagnoses. The final category was false positive, in which the final reports were of normal findings, but abnormal findings were recorded in the preliminary reports.

To classify abnormal organs and resolve discrepancies, two emergency medicine faculty members independently compared the final formal reports with preliminary interpretations from emergency medicine and radiology residents. If the opinions of both of the faculty members did not coincide, the opinions were discussed until an agreement was reached.

### Clinical outcomes of patients

Patient outcomes were evaluated by retrospective chart review. We investigated each patient's ED stay time and whether or not the patient was admitted. If the patient was admitted, his or her admission length, whether or not he or she underwent an emergency operation, and whether or not he or she survived after discharge, were investigated. If the patient was not admitted, we investigated whether or not the same patient revisited the ED within 24 hr with the same chief complaint.

### Statistical analysis

Statistical analyses were performed using STATA 11.0 (Stata-Corp, Texas, USA). To assess the differences between the discrepancy rates of emergency and radiology residents, we performed chi-square tests with  $P$  values  $< 0.05$  considered to be statistically significant. To determine the factors-abnormal organs, grade of residents, time of interpretations-that affect the accuracy of CT interpretations by emergency residents, multivariate analyses using a logistic regression test model were applied with robust variance estimates across the clusters of the emergency residents' individual identification. The regression model was adjusted with patient's sex, age, whether admitted or not, mortality and emergency department stay time. Hosmer and Lomeshow tests were used to evaluate modeling fitness.

To evaluate the relationship between discrepancies and clinical outcomes, the admission rate, emergency operation rate, revisit rate, and mortality rate were compared by chi-square test for correct and incorrect interpretations by emergency residents. Admission length and ED stay time were compared by Student's  $t$ -test.

### Ethics statement

This study was approved by the institutional review board of the Samsung Medical Center (IRB File No. 2010-01-091). The informed consent from the patients and residents who participated in this study was waived by the board.

## RESULTS

A total of 1,153 abdominal CT scan examinations were ordered by the ED during the study period. We excluded 140 cases be-

cause they were trauma patients or pediatric patients, and 110 cases because the chief complaint was not abdominal pain. Eight cases were excluded because the first preliminary interpretations were performed by emergency medicine faculty members. Two cases in which data regarding preliminary interpretations were not complete and nine cases in which the primary pathognomonic organ was the lung were excluded. After exclusions, 884 cases were enrolled.

In 884 cases, 741 cases were examined by the abdominal pelvic protocol, 14 by liver protocol, 109 by biliary gall bladder protocol, 11 by pancreas protocol, and 9 by abdominal CT angiography. The mean age of the enrolled patients was 46.83 yr; 493 (55.8%) patients were female, and 391 (44.2%) patients were male (Table 1).

In the ED, 404 (45.7%) scans were preliminarily interpreted by first year residents, 203 (22.9%) by second year residents, 231 (26.1%) by third year residents, and 46 (5.2%) by fourth year residents. In the radiology department, 421 (47.6%) scans were

preliminarily interpreted by second year residents and 463 (52.4%) by third year residents.

**Discrepancy rates**

In emergency resident preliminary reports, there were 148 discrepant cases (16.7%). In radiology resident preliminary reports, there were 108 discrepant cases (12.2%) (Table 2). The discrepancy rate of emergency residents was significantly higher than that of radiology residents ( $P = 0.01$ ). The discrepancy rate of emergency residents was significantly higher than that of radiology residents for images of female genital organs ( $P < 0.01$ ) and miscellaneous organs ( $P = 0.01$ ) (Fig. 1).

**Factors affecting the accuracy of emergency resident CT interpretations**

In the regression test used to analyze the factors that affected the discrepancy, the grades of residents and times of interpretations did not affect discrepancy rates. Except hollow viscera, all other cases of classified abnormal organs were associated with significant discrepancies when compared to cases of non-spe-

**Table 1.** Baseline characteristics of 884 enrolled patients

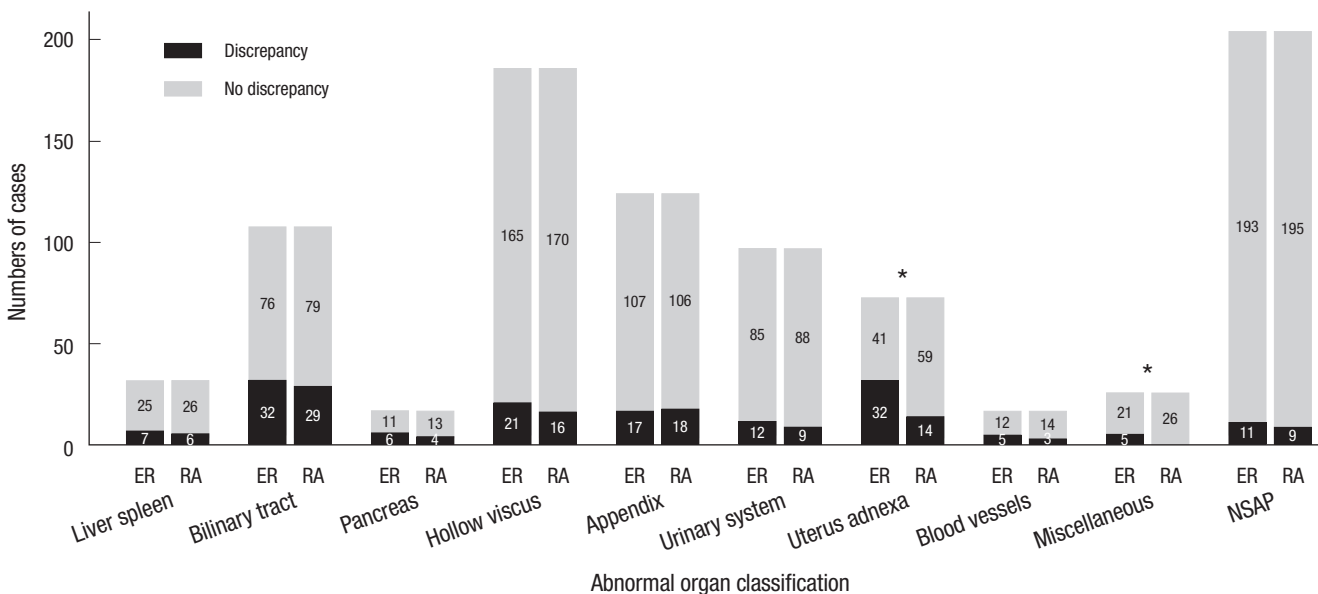
Characteristic	Value
Age (yr)	48.83 ± 18.50
Sex (%)	
Male	391 (44.23)
Female	493 (55.76)
ED stay time (hr)	14.12 ± 21.00
ED revisit within 48 hr (%)	64 (7.24)
Admission (%)	336 (38.01)
Emergency operation (%)	142 (16.06)
Mortality (%)	11 (1.25)

ED, Emergency department.

**Table 2.** The discrepancy rates of emergency and radiology residents

	EM residents	Radiology residents
Total discrepancies (%)	148 (16.7)	108 (12.2)
False negative report (%)	135 (91.2)	97 (89.8)
False positive report (%)	11 (7.4)	9 (8.3)
Misinterpretation of pathology presents (%)	2 (1.3)	2 (1.8)
No discrepancies (%)	736 (83.3)	776 (87.8)

EM, Emergency medicine.



**Fig. 1.** A comparison of the discrepancy rate between emergency medicine and radiology residents according to the abnormal organ. \*There were statistically significant differences for female genital organs and miscellaneous organs. ER, Emergency resident; RA, Radiology resident; Uterus adnexa, Female genital organs (uterus, adnexa, ovary, pelvic cavity); Hollow viscera, Hollow viscera except appendix; Miscellaneous, Peritoneum, adrenal gland, musculoskeletal system; NSAP, Non-specific abdominal pain (no radiological abnormality).

**Table 3.** The results of multivariate logistic regression tests of discrepancy after adjusted patient's factors\*

Factors	Odds ratio (95% Confidence Interval)	P value
Grade of residents		
First	Reference	
Second	0.87 (0.56-1.32)	0.55
Third	1.16 (0.86-1.56)	0.33
Fourth	0.59 (0.21-1.67)	0.32
Time of interpretation		
On-duty time	Reference	
Off-duty time	1.25 (0.80-1.95)	0.32
The classified abnormal organs		
NSAP (normal findings)	Reference	
Liver, spleen	4.87 (1.76-13.44)	< 0.01
Biliary tract	7.52 (3.56-15.87)	< 0.01
Pancreas	11.47 (4.53-29.04)	< 0.01
Hollow viscera	2.27 (0.92-5.58)	0.07
Appendix	2.90 (1.19-7.04)	0.02
Urinary system	2.69 (1.51-4.81)	< 0.01
Female genital system	12.50 (6.59-23.73)	< 0.01
Blood vessels	8.14 (2.39-27.76)	< 0.01
Miscellaneous	4.14 (1.34-12.76)	0.01

\*The regression model was adjusted with patient's sex, age, whether admitted or not, mortality and emergency department stay time. On-duty time: 8 a.m. and 5 p.m. Monday to Friday. Off-duty time: any time other than on-duty time. NSAP: Non-specific abdominal pain (No radiologic abnormality); Hollow viscera: Hollow viscera except appendix; Female genital organs: Uterus, adnexa, ovary, pelvic cavity; Miscellaneous: Peritoneum, adrenal gland, musculoskeletal system.

cific abdominal pain (Table 3).

### The relationship between clinical outcomes and accuracy of emergency resident interpretations

There were no significant differences in admission rates between inaccurately diagnosed patients and accurately diagnosed patients (37.3% vs 41.2%, respectively,  $P = 0.37$ ) or ED stay time (14.3 hr vs 13.1 hr, respectively,  $P = 0.55$ ). There were no significant differences in admission length (9.2 days vs 11.0 days, respectively,  $P = 0.36$ ), emergency operation rate (43.2% vs 34.6%, respectively,  $P = 0.20$ ), mortality rate (1.3% vs 0.7%, respectively,  $P = 0.49$ ) of the admitted patients, and ED revisit rate (5.6% vs 5.7%, respectively,  $P = 0.96$ ) of the discharged patients.

## DISCUSSION

We investigated and compared discrepancy rates of abdominal CT interpretations made by emergency residents and radiology residents. Several previous studies have assessed discrepancy rates of radiology resident interpretations of CT scans (5-8, 10, 12, 13, 15, 17-19). In these studies, the discrepancy rate varied from 1.2% to 10% (5-7, 10, 12, 13, 15, 18). In the present study, the discrepancy rates of emergency and radiology residents were 16.7% and 12.2%, respectively. These values were higher than those reported in previous studies. However, precise comparisons between studies are not possible because of many differ-

ences, including resident specialty, the part of the body examined by CT scans, and the disease scope that was enrolled. The criteria for defining discrepancies were also different because there are no standard definitions of discrepancy. In previous studies, both major and minor discrepancies were assessed, with definitions of "major" and "minor" differing between studies (5-7, 10, 12, 13, 18). In the present study, discrepancies were recognized when they resulted in a change in diagnosis, treatment, or disposition, leading to high discrepancy rates.

In the present study, the discrepancy rate of emergency residents (16.7%) was higher than that of radiology residents (12.2%) when the same CT scans and criteria were used. This may be explained by the different experiences and levels of education of the residents included in the study regarding CT interpretation. At our institute, radiology residents are trained systematically in radiology of the abdomen, chest, and musculoskeletal system during the first year. In addition the residents had participated in teaching conferences before their emergency department rotation. In Samsung Medical Center, radiology residents are on call during their 2nd, 3rd years of radiology training. On the other hand emergency residents are only trained for a few months in an informal training program, and separately focused on common and serious diseases. The residents had participated in teaching conferences in emergency radiology before their emergency department rotation. Usually this training program has conducted during 1st year of emergency resident training. In addition all emergency residents attend a radiology conference once a week. This training difference may explain why CT scans of female genital organs and miscellaneous organs (the peritoneum, adrenal gland and musculoskeletal system) were more likely to be misinterpreted by emergency residents, as these organs are difficult to interpret in general, and are often the source of errors in interpretation (20). Abnormalities of these organs are relatively uncommon compared to other organs (21). The discrepancy rates for the appendix and biliary track, which are the most common regions of abdominal pain in the ED, were similar between ED and radiology residents. Therefore we could suggest that education of emergency residents is important to decrease the discrepancy rates. Especially the education for unfamiliar organs that were more likely to be misinterpreted by emergency residents may be needed to increase the capacity of interpretation of unfamiliar organs in emergency residents. And this type of education may be able to decrease the discrepancy rates of emergency residents. In the future, we need further study of the relationship between formal radiology education of emergency residents and the discrepancy rates.

Tieng et al. (7) reported that the discrepancy of radiology residents were common. And they emphasized that ED physicians should exercise caution when relying on radiology residents' interpretation of body CT scans. We found that the discrepancy

of emergency residents were higher than that of radiology residents. Therefore emergency residents may need to discuss with radiology physicians about their preliminary interpretation of body CT scans when emergency residents make decisions based on their interpretation of body CT scans. If possible, the preliminary interpretations of radiology physicians are needed to make up for the interpretation of emergency residents. Some studies (22, 23) reported that the importance of systematic support to reduce the discrepancy in radiology imaging. Therefore we think that we could reduce the discrepancy of emergency residents by systematic support. More investigations are required for evaluating the effect of systematic supports on the discrepancy.

Previous studies showed that discrepancy rates were associated with the training level of the residents (6, 18). The discrepancy rates of senior residents are lower than those of junior residents (6, 18). In the present study, we detected no significant difference in discrepancy rates associated with level of training. We did additional analysis between the training level of the residents and the disease severity. There were no significant difference between the training level of the residents and the disease severity; the admission rate, the emergency operation rate and the mortality rate. Therefore this may be because the imaging education program for emergency residents is concentrated during the first few months of the first year of residency for emergency physicians at our institute. Also lack of a formal education about radiology contributed to induce this result.

It is difficult to assess whether discrepancies in preliminary interpretation affect clinical outcomes, since there is no standard method for making such comparisons. As a result, various methods have been used in previous studies (9, 12, 18, 24). In this study, the same parameters were used; e.g., ED stay time, admission rate, admission length, emergency operation rate, ED revisit rate, and mortality rate. There were no significant differences for these parameters between patients who were diagnosed accurately or inaccurately by emergency residents. Our data of clinical effect are consistent with other study (7, 11, 12, 23). It seems that EM residents are relying on other available data (patient's history, physical examination, observation and consultation) to prevent clinical mistakes. However the interpretation of that result was very limited because the discussion with radiology residents or notification to ED faculties after the preliminary interpretation data collection was not restricted due to ethical problems.

This study had several limitations. First, the composition of the patient sample, and the educational levels of the residents, varied due to national origin, regional origin, and institutes. Although we adjusted the results for individual identifications of emergency residents, variation in individual ability and experience could have affected the results. Therefore, the results of this study cannot be generalized. The results of this study may be more appropriate for use as basic data to establish CT imag-

ing education guidelines for emergency residents. Second, like similar previous studies, we considered final reports of radiology faculty members to represent the gold standard. However, a previous study showed that 2.3% of major body CT scans were incorrectly interpreted by attending radiology faculty members (24). Third, because radiology residents started ED rotation at 2nd year of resident, we could not compare the same training period between emergency residents and radiology residents. Finally, we were unable to clearly evaluate the influence of discrepant interpretations on clinical outcomes due to ethical problems.

Emergency residents were more likely to inaccurately interpret preliminary CT scans than radiology residents. When the female genital organs, peritoneum, adrenal glands, or the musculoskeletal system were abnormal, the preliminary interpretation abilities of emergency residents proved to be especially insufficient compared to those of radiology residents. Therefore more formal education is needed to emergency residents. EM physicians should pay attention to clinical decision when relying on emergency residents' interpretation of body CT. If possible, the preliminary interpretations of radiology attending physicians are ideal until improving the ability of interpretations of emergency residents in abdomen CT.

## REFERENCES

1. Rosen MP, Siewert B, Sands DZ, Bromberg R, Edlow J, Raptopoulos V. *Value of abdominal CT in the emergency department for patients with abdominal pain.* *Eur Radiol* 2003; 13: 418-24.
2. Urban BA, Fishman EK. *Tailored helical CT evaluation of acute abdomen.* *Radiographics* 2000; 20: 725-49.
3. Tsushima Y, Yamada S, Aoki J, Motojima T, Endo K. *Effect of contrast-enhanced computed tomography on diagnosis and management of acute abdomen in adults.* *Clin Radiol* 2002; 57: 507-13.
4. Stromberg C, Johansson G, Adolfsson A. *Acute abdominal pain: diagnostic impact of immediate CT scanning.* *World J Surg* 2007; 31: 2347-54.
5. Aydyn SA, Bulut M, Topal NB, Akgoz S, Koksall O, Orcan S, Turan M, Aydyn T, Gultekin E, Oncu MR, et al. *Performance of emergency medicine residents in the interpretation of radiographs in patients with trauma.* *Emerg Med J* 2008; 25: 482-5.
6. Walls J, Hunter N, Brasher PM, Ho SG. *The DePICTORS Study: discrepancies in preliminary interpretation of CT scans between on-call residents and staff.* *Emerg Radiol* 2009; 16: 303-8.
7. Tieng N, Grinberg D, Li SF. *Discrepancies in interpretation of ED body computed tomographic scans by radiology residents.* *Am J Emerg Med* 2007; 25: 45-8.
8. Wysoki MG, Nassar CJ, Koenigsberg RA, Novelline RA, Faro SH, Faerber EN. *Head trauma: CT scan interpretation by radiology residents versus staff radiologists.* *Radiology* 1998; 208: 125-8.
9. Chung JH, Strigel RM, Chew AR, Albrecht E, Gunn ML. *Overnight resident interpretation of torso CT at a level 1 trauma center an analysis and review of the literature.* *Acad Radiol* 2009; 16: 1155-60.
10. Strub WM, Vagal AA, Tomsick T, Moulton JS. *Overnight resident prelim-*

- inary interpretations on CT examinations: should the process continue? *Emerg Radiol* 2006; 13: 19-23.
11. Carney E, Kempf J, DeCarvalho V, Yudd A, Noshier J. *Preliminary interpretations of after-hours CT and sonography by radiology residents versus final interpretations by body imaging radiologists at a level 1 trauma center. AJR Am J Roentgenol* 2003; 181: 367-73.
  12. Ruchman RB, Jaeger J, Wiggins EF 3rd, Seinfeld S, Thakral V, Bolla S, Wallach S. *Preliminary radiology resident interpretations versus final attending radiologist interpretations and the impact on patient care in a community hospital. AJR Am J Roentgenol* 2007; 189: 523-6.
  13. Cooper VF, Goodhart LA, Nemcek AA Jr, Ryu RK. *Radiology resident interpretations of on-call imaging studies: the incidence of major discrepancies. Acad Radiol* 2008; 15: 1198-204.
  14. Roszler MH, McCarroll KA, Rashid T, Donovan KR, Kling GA. *Resident interpretation of emergency computed tomographic scans. Invest Radiol* 1991; 26: 374-6.
  15. Jin KN, Jae HJ, Shin CI, Chai JW, Chun SR, Shin SD, Kang YJ, Na DK. *Overnight preliminary interpretations of CT and MR images by radiology residents in ER: How accurate are they? J Korean Soc Emerg Med* 2008; 19: 205-10.
  16. Park SO, Song KJ, Sim MS. *Appropriation evaluation of simple Radiography interpretation by EM resident. J Korean Soc Emerg Med* 2004; 15: 30-5.
  17. Arentz C, Griswold JA, Halldorsson A, Quattromani F, Dissanaika S. *Best poster award: accuracy of surgery residents' interpretation of computed tomography scans in trauma. Am J Surg* 2008; 196: 809-12.
  18. Wechsler RJ, Spettell CM, Kurtz AB, Lev-Toaff AS, Halpern EJ, Nazarian LN, Feld RI, Needleman L, Alexander AA. *Effects of training and experience in interpretation of emergency body CT scans. Radiology* 1996; 199: 717-20.
  19. Wong C, Taubman K, Hoehne F, Malhotra S, Naidu M, Parent D, Taylor EW. *The quality of surgical resident interpretation of after-hour CT scans of the head and abdomen in trauma. Am Surg* 2005; 71: 772-5.
  20. Bechtold RE, Chen MY, Ott DJ, Zagoria RJ, Scharling ES, Wolfman NT, Vining DJ. *Interpretation of abdominal CT: analysis of errors and their causes. J Comput Assist Tomogr* 1997; 21: 681-5.
  21. O'Brien MC. Acute abdominal pain. In: *Tintinalli M, editor, Tintinalli's emergency medicine. 6th ed. New York: The McGraw-Hill Companies Inc., 2004, p487-510.*
  22. Espinosa JA, Nolan TW. *Reducing errors made by emergency physicians in interpreting radiographs: longitudinal study. BMJ* 2000; 320: 737-40.
  23. Kim SJ, Lee SW, Hong YS, Kim DH. *Radiological misinterpretations by emergency physicians in discharged minor trauma patients. Emerg Med J* 2011; 29: 635-9.
  24. Yoon LS, Haims AH, Brink JA, Rabinovici R, Forman HP. *Evaluation of an emergency radiology quality assurance program at a level I trauma center: abdominal and pelvic CT studies. Radiology* 2002; 224: 42-6.