Copyright © 2021 The Korean Society of Emergency Medicine

Factors affecting incorrect

Jong-Hak Park

incorrectly interpreted scans.

tion; Emergency medicine

Ansan, Korea

by novice emergency physicians

Seong Geun Lee, Hanjin Cho, Joo Yeong Kim, Juhyun Song,

terpretation were calculated using multivariable logistic regression analysis.

focus on these two areas to improve abdominal CT interpretation accuracy.

tomography in non-traumatic patients

Department of Emergency Medicine, Korea University Ansan Hospital, Korea University College of Medicine,

Objective Accurate interpretation of computed tomography (CT) scans is critical for patient care

in the emergency department. We aimed to identify factors associated with an incorrect inter-

pretation of abdominal CT by novice emergency residents and to analyze the characteristics of

Methods This retrospective analysis of a prospective observational cohort was conducted at three urban emergency departments. Discrepancies between the interpretations by postgraduate year-1 (PGY-1) emergency residents and the final radiologists' reports were assessed by independent adjudicators. Potential factors associated with incorrect interpretation included patient age, sex, time of interpretation, and organ category. Adjusted odds ratios (aORs) for incorrect in-

Results Among 1,628 eligible cases, 270 (16.6%) were incorrect. The urinary system was the most correctly interpreted organ system (95.8%, 365/381), while the biliary tract was the most incorrectly interpreted (28.4%, 48/169). Normal CT images showed high false-positive rates of incorrect interpretation (28.2%, 96/340). Organ category was found to be a major determinant of incorrect interpretation. Using the urinary system as a reference, the aOR for incorrect interpretation of biliary tract disease was 9.20 (95% confidence interval, 5.0-16.90) and the aOR for incorrectly interpreting normal CT images was 8.47 (95% confidence interval, 4.85-14.78).

Conclusion Biliary tract disease is a major factor associated with incorrect preliminary interpretations of abdominal CT scans by PGY-1 emergency residents. PGY-1 residents also showed high false-positive interpretation rates for normal CT images. Emergency residents' training should

Keywords Acute abdomen; Computed tomography; Internship and residency; Medical educa-

interpretation of abdominal computed

Received: 31 August 2020 Revised: 6 November 2020 Accepted: 16 November 2020

Correspondence to: Jong-Hak Park Department of Emergency Medicine, Korea University Ansan Hospital, 123 Jeokgeum-ro, Danwon-gu, Ansan 15355, Korea E-mail: roadrunner@korea.ac.kr ORCID https://orcid.org/0000-0001-9481-6358

elSSN: 2383-4625

How to cite this article:

Lee SG, Cho H, Kim JY, Song J, Park JH. Factors affecting incorrect interpretation of abdominal computed tomography in nontraumatic patients by novice emergency physicians. Clin Exp Emerg Med 2021;8(3):207-215. https://doi. org/10.15441/ceem.20.118

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https:// creativecommons.org/licenses/by-nc/4.0/).

Original Article



Capsule Summary

What is already known

With increased usage of computed tomography (CT), it has become more important than ever to accurately interpret a computed tomography (CT) image. There have been some studies on discrepancies in interpretation between radiology specialists and other department residents, but there has been no research on a checklist-based CT learning program for residents.

What is new in the current study

This study showed a major factor associated with novice postgraduate year-1 (PGY-1) emergency residents' incorrect interpretation for abdominal CT scans is biliary tract disease, and also showed high false-positive interpretation rates for normal CT images in PGY-1 emergency residents.

INTRODUCTION

Acute abdominal pain is one of the most common patient complaints, and it accounts for 11% to 12% of all patients who visit the emergency department (ED) each year.¹ Performing a proper differential diagnosis is critical for the management of these patients, and abdominal computed tomography (CT) is often essential to this process. CT has been shown to be superior to other diagnostic methods for patients presenting to the ED with trauma, acute abdominal pain, or for geriatric patient evaluation.²⁻⁴ Several studies have since reported the increased utilization of CT in the adult ED.⁵⁻⁷

Today, it has become more important than ever to interpret a CT scan accurately given the clinical repercussions of its use; CT is a high-value diagnostic tool that is used with an increased frequency and has a potential impact on patient care and disposition. However, many hospitals do not always have a radiologist available. Therefore, the ability of emergency physicians to interpret CT findings appropriately is critical in the patient care process, and efforts to reduce interpretation errors are necessary. Identifying factors that affect incorrect interpretations of abdominal CT and addressing them in educational programs may help improve the interpretation capability of physicians in training.

Some studies have reported on discrepancies in interpretation between radiology specialists and residents of other departments, but there has been no research on a checklist-based CT learning program for residents. This study aimed to identify the factors that are mainly associated with the incorrect interpretation of abdominal CT by novice postgraduate year-1 (PGY-1) residents in a checklist-based learning program.^{8,9} We also analyzed the characteristics of incorrectly interpreted CT.

METHODS

Study design and setting

A convenience sample of patients who visited the ED and underwent abdominal CT from March 2016 to February 2018 was included in this retrospective analysis of a prospective observational cohort. During this study period, a checklist-based CT learning program for PGY-1 residents had been in place in the EDs of three tertiary academic hospitals (Appendix 1). The checklists were revised over time by resident educational staff during the study period^{8,9} based on the frequencies and diagnoses of the CT scans performed in the ED. The checklist was completed by the PGY-1 residents in the checklist-based CT learning program immediately after the CT scan. The three hospitals operate EDs with approximately 50,000 to 60,000 patient visits per year and run a 4-year resident training program, during which time the residents circulate between each hospital. This study conformed with the ethical norms and standards as set out in the Declaration of Helsinki. The institutional review board of Korea University Ansan Hospital approved the study (AS15030). The requirement for informed consent was waived by the board, as this study was performed as part of an education program.

Study protocol

We included patients who visited the ED and underwent an abdominal CT examination, as ordered by a PGY-1 resident, during their ED stay. CT images were excluded if interpretations from both a board-certified radiologist and an ED PGY-1 resident were not available. Patients who visited the ED for trauma, those who underwent imaging as part of a routine disease follow-up for the evaluation of postoperative complications, and patients under 15 years of age were also excluded from the analysis. Fig. 1 shows a schematic flow chart of the study. The CT protocols ranged from

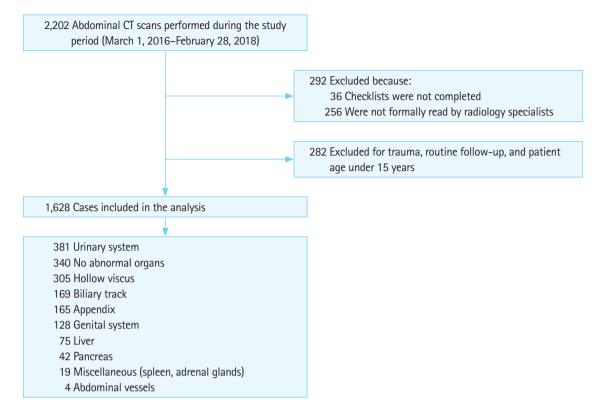


Fig. 1. Flow chart showing enrollment for the study. CT, computed tomography.

non-contrast to intravenous contrast-enhanced three-phase scans depending on the clinically suspected disease entity. CT scanners with 64 or 256 detector rows were used.

Outcome measures

A trained researcher (registered nurse) entered the results from the checklists recorded by the PGY-1 and the radiologists' reports into a database. Variables in the database included patients' age, sex, time of interpretation (on-duty vs. off-duty; on-duty time was defined as 8 a.m. to 5 p.m., weekday vs. weekend), reason for the visit, chief complaints, suspected disease, and interpretations by the PGY-1 resident and radiologist. Two board-certified emergency physicians independently assessed the discrepancies between the interpretations of PGY-1 residents and radiologists. The assessors also evaluated whether a resident's interpretation was clinically acceptable, despite not being identical to that of a radiologist. When the assessors were not in agreement, the decision was discussed among the assessors until a consensus was reached.

The assessors determined the interpretations as either correct or incorrect. Interpretations were judged to be correct if both were identical and clinically acceptable, while any other case was defined as incorrect. Incorrect interpretations were further divided into three subgroups: false negative, false positive, and misinterpretation. A false negative was defined as interpreting an abnormal CT image as normal, and a false positive was defined as interpreting a normal CT image as abnormal. Misinterpretation was defined as the PGY-1 resident correctly identifying that the CT was abnormal but incorrectly identifying the pathological lesion.

Data analysis

The distributions of categorical data are presented as percentages. Continuous variables that were not normally distributed were presented as medians with interquartile ranges. Categorical variables were compared using the chi-squared test, and continuous variables were evaluated using the Wilcoxon rank sum test. To assess inter-rater agreement, Cohen's kappa coefficient was calculated; its value ranged from 0 (perfect discordance) to 1 (perfect accordance). Multivariable logistic regression analysis was performed to evaluate the factors associated with incorrect interpretation; the model was adjusted for potential confounders such as patient age, sex, time of interpretation, and organ systems. Data analyses were performed using RStudio ver. 3.5.1 (RStudio Inc., Boston, MA, USA), and P-values < 0.05 were considered statistically significant.

Baseline characteristics of enrolled patients

During the study period, 2,202 abdominal CT scans were performed in the ED. Of these, several were excluded because checklists were not completed (n = 36), scans were not formally read by radiology specialists (n = 256), or they involved cases of patients with trauma, patients who underwent routine disease follow-up, or patients who were under 15 years of age (total, n = 282). Ultimately, 1,628 checklists completed by PGY-1 emergency residents were included in the analysis (Fig. 1).

Among the 1,628 CT scans, 1,358 (83.4%) were assessed as correct and 270 (16.6%) were incorrect. The incorrect interpretation rates for female and male patients were 19.6% (158/808) and 13.7% (112/820), respectively. Interpretation accuracy varied widely across organ categories (Fig. 2). The biliary tract is the organ system that was most frequently incorrectly interpreted (28.4%, 48/169). The second highest incorrect interpretation rate occurred in cases of no abnormal organs (28.2%, 96/340). The urinary system was the most correctly interpreted organ system (95.8%, 365/381).

There were no significant differences between the two groups regarding characteristics (age, duty time [on-duty vs. off-duty], and day of the week [weekday vs. weekend or holiday]). PGY-1 residents also displayed non-significant heterogeneity in their CT interpretation ability (Table 1).

Characteristics of incorrect interpretation

Table 2 shows the characteristics of incorrect interpretation ac-

cording to organ categories. In the misinterpretation subgroup, the organ system incorrectly interpreted most frequently was the biliary tract (28.9%, 43/149). Hollow viscus disease had the highest proportion of incorrect interpretations in both the false negative (24.0%, 6/25) and false positive (32.3%, 31/96) groups.

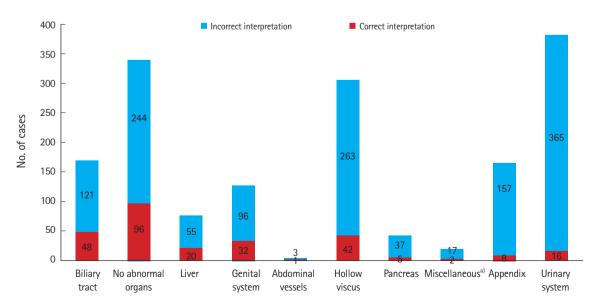
Factors affecting the accuracy of novice emergency residents' CT interpretations

Multivariable logistic regression analysis, which was used to analyze the factors that affected incorrect interpretations, showed that organ categories were the major determinants of incorrect interpretation (Fig. 3). Patient age, sex, and the time of interpretation did not affect the interpretation status.

In descending order, the statistically significant adjusted odds ratios for incorrect interpretation by organ systems are as follows: biliary tract, no abnormal organs, liver, genital system, hollow viscus, and pancreas, with the urinary system as the reference.

DISCUSSION

Biliary tract disease was found to be a major factor associated with the incorrect preliminary interpretation of abdominal CTs, among various confounders such as patient age, sex, and time of interpretation. We did not find any association between incorrect interpretations and duty time or weekend duty. An accurate and confident interpretation of normal CT images must be ensured for high-quality emergency care. However, ED residents showed high false-positive interpretation rates of normal CT images.



Although the use of CT as a diagnostic tool has increased be-

Fig. 2. Preliminary interpretation accuracy across organ categories. ^{a)}Includes spleen, adrenal glands, etc.

Table 1. Baseline characteristics of enrolled patients

	•			
Characteristics	Total (n = 1,628)	Correct interpretation (n = 1,358)	Incorrect interpretation (n = 270)	
Age (yr)	47 (32.0–60.5)	47 (34.0–62.0)	47 (31.0–60.0)	
< 60	1,191 (73.2)	995 (73.3)	196 (72.6)	
≥ 60	437 (26.8)	363 (26.7)	74 (27.4)	
Sex				
Male	820 (50.4)	708 (52.1)	112 (41.5)	
Female	808 (49.6)	650 (47.9)	158 (58.5)	
Location of pain				
Lower abdomen	518 (31.8)	444 (32.7)	74 (27.4)	
Upper abdomen	312 (19.2)	237 (17.5)	75 (27.8)	
Flank				
Right	167 (10.2)	150 (11.0)	17 (6.3)	
Left	148 (9.1)	143 (10.5)	5 (1.8)	
Whole abdomen	164 (10.1)	132 (9.7)	32 (11.9)	
Others (nonspecific pain, etc.)	319 (19.6)	252 (18.6)	67 (24.8)	
Organ category				
Urinary system	381 (23.4)	365 (26.9)	16 (5.9)	
Hollow viscus	305 (18.7)	263 (19.4)	42 (15.6)	
Biliary tract	169 (10.4)	121 (8.9)	48 (17.8)	
Appendix	165 (10.1)	157 (11.6)	8 (3.0)	
Genital system	128 (7.9)	96 (7.1)	32 (11.8)	
Liver	75 (4.6)	55 (4.0)	20 (7.4)	
Pancreas	42 (2.6)	37 (2.7)	5 (1.8)	
Miscellaneous ^{a)}	19 (1.2)	17 (1.2)	2 (0.7)	
Abdominal vessels	4 (0.2)	3 (0.2)	1 (0.4)	
No abnormal organs	340 (20.9)	244 (18.0)	96 (35.6)	
Time of CT interpretation				
On-duty time ^{b)}	632 (38.8)	536 (39.5)	96 (35.6)	
Off-duty time	996 (61.2)	822 (60.5)	174 (64.4)	
Weekdays	1,035 (63.6)	877 (64.6)	158 (58.5)	
Weekend (or holidays)	593 (36.4)	481 (35.4)	112 (41.5)	
Reader				
PGY-A	256 (15.7)	216 (15.9)	40 (14.8)	
PGY-B	244 (15.0)	209 (15.4)	35 (12.9)	
PGY-C	243 (14.9)	201 (14.8)	42 (15.6)	
PGY-D	224 (13.8)	183 (13.5)	41 (15.2)	
PGY-E	185 (11.4)	151 (11.1)	34 (12.6)	
PGY-F	180 (11.1)	153 (11.3)	27 (10.0)	
PGY-G	167 (10.2)	140 (10.3)	27 (10.0)	
PGY-H	129 (7.9)	105 (7.7)	24 (8.9)	

Values are presented as median (interquartile range) or number (%).

CT, computed tomography; PGY, postgraduate year-1 resident.

^{a)}Includes the spleen, adrenal glands, etc. ^{b)}On-duty time was defined as 8 a.m. to 5 p.m.

cause of its accessibility and convenience within the ED, there are few educational programs that help novice emergency physicians improve their interpretation capability. We developed a checklistbased learning program, revised over time, to enable novice emergency residents to interpret CT images easily.^{8,9}

Diagnostic errors are among the main causes of medical malpractice, and they can lead to adverse outcomes such as treatment delay, maltreatment, unnecessary use of medical resources, and patient disability or death.^{10,11} As diagnostic technologies evolve, the CT scan has been shown to have great advantages in the ED, such as an increased level of diagnostic certainty and use in early decision-making and formulation of treatment plans.¹² The ability to interpret abdominal CT images correctly is important for accurate and rapid decision-making in the ED, where highly heterogeneous patients, including mild to critically ill, trauma, pediatric, and geriatric patients, visit simultaneously.

Most previous studies examined the discrepancies in interpretation between radiology specialists and radiology residents. The overall discrepancy rate varied from 1% to 10% depending on the type of modality (CT or magnetic resonance imaging) and body parts scanned.¹³⁻¹⁷ In our study, the discrepancy rate between pre-

 Table 2. Subgroup analysis of incorrect interpretations according to organ category

Organ astagorias	Radio abnor	No radiologic abnormality		
Organ categories	Misinterpreta- tion ^{a)} (n = 149)	False negative ^{b)} (n = 25)	False positive ^{c)} (n=96)	
Urinary system	13 (8.7)	3 (12.0)	13 (13.5)	
Appendix	6 (4.0)	2 (8.0)	9 (9.4)	
Miscellaneous ^{d)}	1 (0.7)	1 (4.0)	4 (4.2)	
Pancreas	4 (2.7)	1 (4.0)	2 (2.1)	
Hollow viscus	36 (24.1)	6 (24.0)	31 (32.3)	
Genital system	27 (18.1)	5 (20.0)	13 (13.5)	
Abdominal vessels	0 (0)	1 (4.0)	1 (1.0)	
Liver	19 (12.8)	1 (4.0)	11 (11.5)	
Biliary tract	43 (28.9)	5 (20.0)	12 (12.5)	

Values are presented as number (%).

^{a)}Misinterpretation: computed tomography was correctly identified as abnormal, but residents did not correctly identify the pathologic lesion. ^{b)}False negative interpretation: abnormal computed tomography, interpreted as normal. ^{cl}False positive interpretation: normal computed tomography, interpreted as abnormal. ^{dl}Includes the spleen, adrenal glands, etc.

Unadjusted

liminary interpretations by novice emergency residents and the final radiologists' reports was 16.6%, which was higher than that of previous studies. This is most probably due to differences in the experience and level of education of the residents included in the study with respect to CT interpretation.

False-positive interpretation rates were also higher in our study than in previous studies.¹⁷ This could be explained by several possible reasons, one of which might be the residents' experience levels. Resident learner feedback revealed that PGY-1 residents are often not fully confident about their interpretations and show a tendency to fill in the checklist with any findings related to the chief complaint. Another possible reason is that emergency physicians are clinicians who encounter patients firsthand. Therefore, they are prone to interpreting diagnostic images with the underlying assumption that there must be a pathologic lesion to be found. Radiologists, on the other hand, often take an objective approach to image interpretation. This result shows that the ability to identify normal CT images, an aspect that is often overlooked in emergency care, requires improvement and is a potential training goal for novice emergency residents. False-positive interpretations can lead to unnecessary laboratory tests, incorrect disposition, and increased ED stay.

Ruutiainen et al. found that major discrepancies increased significantly during the final two hours of consecutive overnight call

	Una	adjusted /		djusted®				
	OR	95% Cl	OR	95% Cl	-			
Organ categories					_			
Urinary system	ref		ref		1			
Appendix	1.16	0.49-2.77	1.08	0.45-2.59	⊢	1		
Miscellaneous ^{b)}	2.68	0.57-12.62	2.67	0.57-12.59	F	•		
Pancreas	3.08	1.07-8.89	3.09	1.07-8.94	ŀ	•	-	
Hollow viscus	3.64	2.00-6.62	3.49	1.92-6.36		- - I		
Genital system	7.60	4.01-14.43	6.16	3.16-12.02		⊢		
Abdominal vessel	7.60	0.75-77.20	7.03	0.68–72.29	H	•		
Liver	8.30	4.05-16.97	8.23	4.02-16.87		—	→	
Biliary tract	9.05	4.96-16.52	9.20	5.00-16.90		—	→	
No abnormal organs	8.98	5.16-15.61	8.47	4.85–14.78		—	♦ ——I	
Age								
< 60	ref		ref					
≥ 60	1.03	0.77-1.38	0.91	0.66-1.25	⊢ ∳ ⊣			
Sex								
Male	ref		ref					
Female	1.54	1.17-2.00	1.32	0.99–1.78	↓			
Time of interpretation								
On-duty time ^{c)}	ref		ref					
Off-duty time	1.18	0.90-1.55	1.10	0.83-1.46	i i ∳−i			
					1 (aOR)	4	16	64

A divicto da)

Fig. 3. Results of the multivariable logistic regression analysis for incorrect interpretation. OR, odds ratio; Cl, confidence interval; ref, reference; aOR, adjusted odds ratio. ^{a)}The logistic regression model was adjusted for age, sex, and time of interpretation. ^{b)}Includes the spleen, adrenal glands, etc. ^{c)}On-duty time is defined as 8 a.m. to 5 p.m.

shifts and that this finding could be related to either fatigue or circadian desynchronization.¹⁸ However, in our study, we did not find any statistically significant relationship between incorrect interpretations and shift time or holiday and weekend duty.

This study has several limitations. First, the degree of subjectivity in the assessment of checklists cannot be ruled out. One of the most important prerequisites of this study is consistency among assessors. Although a high agreement rate of judgment supports the objectivity of our outcome measure, we cannot completely discount the subjective nature of human assessment. To mitigate this, we analyzed 100 pilot test cases and conducted extensive discussions. There was a 91.5% agreement in the assessment with substantial inter-rater reliability (kappa=0.75; 95% confidence interval, 0.70–0.80).¹⁹ Second, the checklist alone may not always accurately reflect the residents' interpretation abilities. The checklist was developed to have a simple and focused format for user convenience, compliance, and rapid completion. Therefore, it may not be appropriate to judge PGY-1 residents' image interpretation skills based on checklist accuracy alone. However, it is a guick, objective assessment tool, and its data can be coded and compared against radiologists' reports. Third, for abdominal CT interpretations, there may be differences in interpretation results depending on the residents' training periods.⁸ However, the monthly analysis could not be performed for the residents in this study. Fourth, whether the CT scan was enhanced with contrast could affect residents' preliminary interpretation. However, the number of cases of non-enhanced CT in our study was very small, so the presence of contrast was not included as a study variable. Finally, ED patient characteristics may vary by region, culture, insurance, institute, and country, thereby reducing the generalizability of our findings.

In conclusion, the presence of biliary tract disease is a major factor associated with the incorrect preliminary interpretation of abdominal CT scans by PGY-1 ED residents. Furthermore, no association was found between incorrect interpretations and the time of interpretation. Additionally, PGY-1 residents showed high false-positive interpretation rates for normal CT images. These are two areas on which emergency residents' training should focus to improve abdominal CT interpretation accuracy and the quality of emergency care.

CONFLICT OF INTEREST

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

REFERENCES

- 1. Bhuiya FA, Pitts SR, McCaig LF. Emergency department visits for chest pain and abdominal pain: United States, 1999–2008. NCHS Data Brief 2010:1–8.
- Rosen MP, Sands DZ, Longmaid HE 3rd, Reynolds KF, Wagner M, Raptopoulos V. Impact of abdominal CT on the management of patients presenting to the emergency department with acute abdominal pain. AJR Am J Roentgenol 2000;174: 1391-6.
- 3. Abujudeh HH, Kaewlai R, McMahon PM, et al. Abdominopelvic CT increases diagnostic certainty and guides management decisions: a prospective investigation of 584 patients in a large academic medical center. AJR Am J Roentgenol 2011;196: 238-43.
- 4. Gardner CS, Jaffe TA, Nelson RC. Impact of CT in elderly patients presenting to the emergency department with acute abdominal pain. Abdom Imaging 2015;40:2877-82.
- 5. Boone JM, Brunberg JA. Computed tomography use in a tertiary care university hospital. J Am Coll Radiol 2008;5:132-8.
- Broder J, Warshauer DM. Increasing utilization of computed tomography in the adult emergency department, 2000–2005. Emerg Radiol 2006;13:25–30.
- 7. Oh HY, Kim EY, Cho J, et al. Trends of CT use in the adult emergency department in a tertiary academic hospital of Korea during 2001-2010. Korean J Radiol 2012;13:536-40.
- Song JH, Cho H, Park JH, et al. Learning curve and period of experience required for the competent diagnosis of acute appendicitis using abdominal computed tomography: a prospective observational study. Clin Exp Emerg Med 2017;4:222-31.
- 9. Suh JY, Song JH, Moon SW, et al. Current state of abdominal computed tomography performed in emergency department of a tertiary university hospital and development of a preliminary interpretation checklist. J Korean Soc Emerg Med 2016; 27:336-44.
- Walsh JN, Knight M, Lee AJ. Diagnostic errors: impact of an educational intervention on pediatric primary care. J Pediatr Health Care 2018;32:53–62.
- Singh H, Sittig DF. Advancing the science of measurement of diagnostic errors in healthcare: the Safer Dx framework. BMJ Qual Saf 2015;24:103–10.
- 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.
- 13. Strub WM, Vagal AA, Tomsick T, Moulton JS. Overnight resident preliminary interpretations on CT examinations: should the process continue? Emerg Radiol 2006;13:19-23.

- 14. 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.
- Ruchman RB, Jaeger J, Wiggins EF 3rd, et al. 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.
- Filippi CG, Meyer RE, Cauley K, et al. The misinterpretation rates of radiology residents on emergent neuroradiology magnetic resonance (MR) angiogram studies: correlation with level of residency training. Emerg Radiol 2010;17:45-50.
- Kang MJ, Sim MS, Shin TG, et al. Evaluating the accuracy of emergency medicine resident interpretations of abdominal CTs in patients with non-traumatic abdominal pain. J Korean Med Sci 2012;27:1255-60.
- Ruutiainen AT, Durand DJ, Scanlon MH, Itri JN. Increased error rates in preliminary reports issued by radiology residents working more than 10 consecutive hours overnight. Acad Radiol 2013;20:305-11.
- 19. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22:276-82.



Appendix 1. Checklist for preliminary abdominal computed tomography interpretation

Abdomen CT checklist

Patient Baseline Data: 1. ID: 2. Name: 3. Age/Gender:	
1. Disease Trauma chief complaint: Epigastric Vpper RUQ Periumblical Lower RLQ LLQ CRt. flank Lt. flank other: describe 3. physical related impressions (clinical impressions)	Patients Barcode
Abdominal CT finding: check abnormal findings √	
1. Liver - parenchyma Cyst mass abscess periparenchyma enhancement color change (□ hepatitis fatty liver) surface (□ regular irregular) - hepatic duct □ dilatation □ dilatation stone	9. Uterus and adnexa - Uterus mass abscess - Ovary and tubule abnormal cyst mass abscess peri-fluid collection infiltration
2. Biliary tract - GB sac mass stone polyp - GB wall thickening enhancement peri wall infiltration - Common Bile Duct duct dilatation stone mass duct enhancement	10. Etc. Peritoneum, Mesentery and Abdominal wall mass peritoneum thickening mesentery or fat infiltration LN enlargement : site () fluid collection Prostate mass alargement
3. Pancreas -	- 🗆 mass 🔲 abscess 🗌 enlargement Adrenal gland - 🗆 cyst 🔹 mass
 4. Kidney, Ureter and Bladder kidney size (normal abnormal) renal cyst mass abscess wedge multiple patch (APN) peri-wall infiltration fluid collection hydronephrosis (renal calyces renal pelvis) ureter 	11. Vessel check list Aorta IVC femoral artery femoral vein SMA IMA renal artery renal vein splenic artery splenic vein celiac trunk portal vein
□ mass □ stone □ wall thickening □ enhancement □ hydro ureter - bladder □ mass □ wall thickening □ enhancement	- 🗆 aneurism 🛛 thrombus 🗌 dissection
5. Spleen – size (normal abnormal) - infiltration	Any other description →
6. Stomach - mass wall thickening perforation diverticulum herniation Duodenum	Impression at ED: 1.
- mass wall thickening perforation	2. 3.
 mass wall thickening perforation bowel dilatation bowel wall enhancement ischemia diverticulum internal hernia inguinal hernia obstruction 	
8. Appendix - □ dilatation (size mm) - □ not found - □ wall enhancement □ wall thickening □ appendicolith	□ R1 □ R2 □ R3 □ R4

RUQ, right upper quadrant; LUQ, left upper quadrant; RLQ, right lower quadrant; LLQ, left lower quadrant; GB, gall bladder; LN, lymph node; IVC, inferior vena cava; SMA, superior mesenteric artery; IMA, inferior mesenteric artery.