JAMA Surgery | Original Investigation

Diagnostic Accuracy of Unenhanced Computed Tomography for Evaluation of Acute Abdominal Pain in the Emergency Department

Hiram Shaish, MD; Justin Ream, MD; Chenchan Huang, MD; Jonathan Troost, PhD; Sonia Gaur, MD; Ryan Chung, MD; Sooah Kim, MD; Hanisha Patel, MD; Jeffrey H. Newhouse, MD; Shokoufeh Khalatbari, MS; Matthew S. Davenport, MD

IMPORTANCE Intravenous (IV) contrast medium is sometimes withheld due to risk of complication or lack of availability in patients undergoing computed tomography (CT) for abdominal pain. The risk from withholding contrast medium is understudied.

OBJECTIVE To determine the diagnostic accuracy of unenhanced abdominopelvic CT using contemporaneous contrast-enhanced CT as the reference standard in emergency department (ED) patients with acute abdominal pain.

DESIGN, SETTING, AND PARTICIPANTS This was an institutional review board–approved, multicenter retrospective diagnostic accuracy study of 201 consecutive adult ED patients who underwent dual-energy contrast-enhanced CT for the evaluation of acute abdominal pain from April 1, 2017, through April 22, 2017. Three blinded radiologists interpreted these scans to establish the reference standard by majority rule. IV and oral contrast media were then digitally subtracted using dual-energy techniques. Six different blinded radiologists from 3 institutions (3 specialist faculty and 3 residents) interpreted the resulting unenhanced CT examinations. Participants included a consecutive sample of ED patients with abdominal pain who underwent dual-energy CT.

EXPOSURE Contrast-enhanced and virtual unenhanced CT derived from dual-energy CT.

MAIN OUTCOME Diagnostic accuracy of unenhanced CT for primary (ie, principal cause[s] of pain) and actionable secondary (ie, incidental findings requiring management) diagnoses. The Gwet interrater agreement coefficient was calculated.

RESULTS There were 201 included patients (female, 108; male, 93) with a mean age of 50.1 (SD, 20.9) years and mean BMI of 25.5 (SD, 5.4). Overall accuracy of unenhanced CT was 70% (faculty, 68% to 74%; residents, 69% to 70%). Faculty had higher accuracy than residents for primary diagnoses (82% vs 76%; adjusted odds ratio [OR], 1.83; 95% CI, 1.26-2.67; $P = .002$) but lower accuracy for actionable secondary diagnoses (87% vs 90%; OR, 0.57; 95% CI, 0.35-0.93; $P < .001$). This was because faculty made fewer false-negative primary diagnoses (38% vs 62%; OR, 0.23; 95% CI, 0.13-0.41; $P < .001$) but more false-positive actionable secondary diagnoses (63% vs 37%; OR, 2.11, 95% CI, 1.26-3.54; $P = .01$). False-negative (19%) and false-positive (14%) results were common. Interrater agreement for overall accuracy was moderate (Gwet agreement coefficient, 0.58).

CONCLUSION Unenhanced CT was approximately 30% less accurate than contrast-enhanced CT for evaluating abdominal pain in the ED. This should be balanced with the risk of administering contrast material to patients with risk factors for kidney injury or hypersensitivity reaction.

JAMA Surg. doi:10.1001/jamasurg.2023.1112
Published online May 3, 2023.

Invited Commentary

Supplemental content

Author Affiliations: Department of Radiology, Columbia University, New York, New York (Shaish, Patel, Newhouse); Department of Radiology, Cleveland Clinic, Cleveland, Ohio (Ream); Department of Radiology, New York University Langone, New York (Huang, Kim); Michigan Institute for Clinical and Translational Research, Ann Arbor (Troost, Khalatbari); Department of Radiology, Michigan Medicine, Ann Arbor (Gaur); Department of Radiology, Massachusetts General Hospital, Boston (Chung); Department of Radiology and Urology, Michigan Medicine, Ann Arbor (Davenport); Michigan Radiology Quality Collaborative, Michigan Medicine, Ann Arbor (Davenport).

Corresponding Author: Matthew S. Davenport, MD, 1500 E Medical Center Dr, A2-B209A, Ann Arbor, MI 48109 (matdaven@med.umich.edu).

Contrast-enhanced computed tomography (CT) of the abdomen and pelvis is the most common and most appropriate imaging test for the evaluation of acute abdominal pain in adult nonpregnant patients in the emergency department (ED).[1-3] Intravenous (IV) contrast medium is used in this setting to improve diagnostic accuracy. For some diagnoses, contrast medium is essential for diagnosis (eg, active bleeding), for others it is helpful (eg, liver mass), and for others it is not required (eg, ureteral stone).[1-4] However, at the time of imaging, the diagnosis often is uncertain and so unless certain diagnoses that do not require contrast medium are strongly favored (eg, urolithiasis), contrast medium is generally administered to cover the range of possible pathologies that might be identified.[1-4] Due to ethical considerations, randomized trials comparing the diagnostic accuracy of contrast-enhanced CT and unenhanced CT do not exist for the general population of patients with acute abdominal pain in the ED.

Contrast medium sometimes is withheld due to risk of complication (eg, prior hypersensitivity reaction to iodinated contrast medium, severe kidney disease) or lack of availability (eg, the 2022 acute shortage arising from Shanghai, China).[5-7] When contrast medium is withheld, it is not always clear what loss of diagnostic accuracy is imparted, due in part to an insufficient evidence base but also to the wide range of diagnoses detectable at abdominopelvic CT. Diagnoses can include the principal cause (s) of pain (eg, bowel ischemia, abscess), as well as clinically important incidental findings (eg, kidney cancer, new liver metastasis). False-negative results at unenhanced CT may occur due to misdiagnosis or underdiagnosis and false-positive results may occur from impaired radiologist confidence. These errors can harm patients, delay care, and result in additional unneeded testing and intervention.[8] It is important to know the risk of diagnostic error caused by withholding contrast medium to better inform the risk-benefit analysis in patients for whom contrast medium may be withheld.

Existing studies that have attempted to answer this question in general have focused either on a narrow set of diagnoses (eg, appendicitis, diverticulitis) or lacked a robust reference standard.[9-12] In the present work, we used dual-energy techniques to digitally subtract IV contrast medium from a consecutive cohort of patients undergoing contrast-enhanced CT for evaluation of abdominal pain in the ED. We then evaluated the diagnostic accuracy of unenhanced CT with respect to the principal cause(s) of pain (ie, primary diagnoses) and incidental findings likely to require additional imaging or clinical management (ie, actionable secondary diagnoses). This study design has the potential to minimize selection, longitudinal, and verification bias that might result in cohorts without a simultaneous reference standard. The purpose of our study was to determine the diagnostic accuracy of unenhanced abdominopelvic CT in ED patients with acute abdominal pain using contemporaneous contrast-enhanced CT as the reference standard.

Key Points

Question What is the diagnostic accuracy of unenhanced computed tomography (CT) in patients admitted to an emergency department with abdominal pain?

Findings In this multicenter diagnostic accuracy study, unenhanced CT was approximately 30 percentage points less accurate than contrast-enhanced CT for diagnosing the cause of pain and identifying actionable secondary diagnoses.

Meaning In a general population of emergency department patients with abdominal pain, using unenhanced CT to avoid risks of intravenous contrast medium administration was associated with a large diagnostic penalty.

Methods

This Health Insurance Portability and Accountability Act-compliant (HIPAA) multicenter retrospective diagnostic accuracy study of single-center CT data was approved by the institutional review boards of the 3 participating institutions (institutions A, B, and C). Standards for Reporting of Diagnostic Accuracy (STARD) reporting guidelines were used in the conduct and reporting of this study. Patient informed consent was determined to be not required by the institutional review boards due to the retrospective nature of this investigation and anonymity of the data set.

Study Population

The study population was initially composed of 202 consecutive ED patients who underwent dual-energy contrast-enhanced CT at institution A from April 1, 2017 through April 22, 2017, for the evaluation of acute abdominal pain. Each of the 3 institutions involved in this study was a quaternary care academic institution with a resident training program. Institution A provided the imaging data, and institutions A, B, and C provided radiology residents and specialist faculty radiologists to interpret those data. Inclusion criteria were (1) adult (18 years of age or older) ED patient, (2) contrast-enhanced dual-energy CT performed at institution A with institutional standard portal venous phase technique, (3) CT performed to evaluate acute abdominal pain, and (4) virtual unenhanced CT data set could be derived from the contrast-enhanced data set using dual energy techniques. All imaging studies were acquired from institution A to ensure a consistent imaging protocol and to minimize risk of selection bias; dual-energy contrast-enhanced CT was initiated before the study period at institution A (instead of single-energy CT) for routine evaluation of acute abdominal pain. Exclusion criterion was missing imaging or clinical data (N = 1). The final study population was composed of 201 patients (eFigure in Supplement 1).

A Priori Sample Size Calculation

Prior to study initiation, a pilot sample of 30 consecutive adult patients who underwent contrast-enhanced CT with institutional standard portal venous phase dual-energy technique for acute abdominal pain at institution A were identified. Two fellowship-trained subspecialist abdominal radiologists who provide clinical care for ED patients independently reviewed the images and in consensus determined the prevalence of primary diagnoses and actionable secondary diagnoses to inform a power analysis. Based on that observed prevalence

(40%), we performed a power analysis with the following assumptions: α = .05, prevalence of actionable findings of 40%, width of CI of 15%, assumed sensitivity of 90%, and assumed specificity of 80%. This resulted in a target sample size of 202. The patients in the pilot sample were not included in the study cohort.

Dual-Energy CT Technique and Creation of Unenhanced Imaging Data

All CT imaging examinations were performed on a dual-source dual-energy CT scanner (Force; Siemens Healthineers) following the IV administration of 0.68 mL × (patient weight in pounds) (maximum, 150 mL) iopamidol, 300 mgI/mL at 3 mL/s. Iodine-containing oral contrast medium (Gastrografin, diatrizoate meglumine) was administered. Imaging was performed in the portal venous phase approximately 80 seconds after contrast material administration. Scan parameters were 90 kV peak, variable milliamperes, and field of view to include diaphragm to symphysis pubis. Virtual unenhanced CT data were derived by removing iodine from the images using established dual-energy techniques.[13] The resulting unenhanced CT data formed the study cohort and had neither IV nor oral contrast media. The unaltered contrast-enhanced CT data formed the reference standard.

Reference Standard

The reference standard was derived from 3 independent reviews of the contrast-enhanced CT examinations: 2 retrospective reviews by members of the study team (J.R. and C.H.) and the prospective read rendered as part of clinical care. The reference standard was composed of a list of primary and actionable secondary diagnoses and related recommendation(s) for each contrast-enhanced CT, like a list of impressions in a typical radiology report. The primary diagnoses responsible for abdominal pain were identified (if present) in addition to any actionable secondary findings (ie, incidental findings requiring additional imaging or clinical management). At least 2 of 3 of the independent reviews were required to independently agree on the presence of an actionable finding or recommendation for it to be included in the reference standard. Primary diagnosis was defined as the diagnoses most likely to explain the patient's abdominal pain. Actionable secondary diagnosis was defined as an incidental finding not necessarily connected to the abdominal pain but that was likely to affect patient management, such as imaging or clinical follow-up or medical or surgical intervention.

The reference standard was chosen to minimize the risks of selection, longitudinal, and verification bias that might have resulted had we included patients who received unenhanced and contrast-enhanced CT at separate times. All patients in our consecutive cohort had available contemporaneous unenhanced and contrast-enhanced CT data. Prospective study designs requiring back-to-back acquisitions with and without contrast medium would be ethically challenging to implement due to additional radiation exposure without clear patient benefit. Alternative retrospective study designs restricted to patients who happened to undergo both unenhanced CT and contrast-enhanced CT as part of clinical care would be biased by the clinical situations requiring such scanning, as well as any clinical events that may have occurred between scans separated in time. We used contrast-enhanced CT as the reference standard rather than a composite clinical outcome (recognizing that contrast-enhanced CT does not have perfect accuracy) because unenhanced CT was being tested as a potential replacement for contrast-enhanced CT, the current standard of care.

Blinded Review of Unenhanced CT Data

Six different radiologists from 3 institutions (1 fellowship-trained subspecialist abdominal radiologist who provides clinical care for ED patients and with 20 [institution A], 7 [institution B], or 12 years' [institution C] experience, and 1 senior diagnostic radiology resident from each institution) who knew the clinical indication for the CT examination (ie, acute abdominal pain), but were blinded to other clinical and imaging data, reviewed the unenhanced CT scans and recorded findings and recommendations. The primary diagnoses responsible for abdominal pain were identified (if present) in addition to any actionable secondary findings.

Primary and Secondary Outcomes

The primary outcome was the diagnostic accuracy of unenhanced CT for the primary cause(s) of abdominal pain (ie, primary diagnoses), using contrast-enhanced CT as the reference standard. The diagnostic accuracy of unenhanced CT for important incidental findings (ie, actionable secondary diagnoses) was evaluated as a secondary outcome.

Data Collection and Analysis

Demographic data (age, gender, BMI) are expressed as counts and percentages and mean SDs. Diagnostic accuracy of unenhanced CT was assessed in 2 ways and reported by radiologist with ranges where appropriate. Simple diagnostic accuracy was calculated by evaluating sensitivity, specificity, negative predictive value, and positive predictive value at the level of the examination by asking if any primary or actionable secondary diagnosis was correctly identified. For simple diagnostic accuracy calculations, if any correct primary or actionable secondary diagnosis was identified in a patient with more than one such diagnosis, it was considered a true positive. Detailed diagnostic accuracy was calculated by evaluating overall accuracy, false-positive rate, and false-negative rate at the level of diagnosis, considering all primary and actionable secondary diagnoses. Given that each CT could have multiple diagnoses (both in the reference standard and in the unenhanced CT interpretations), it was possible for an unenhanced CT to simultaneously be a false positive and a false negative. For example, if the reference standard and unenhanced CT had different primary diagnoses, it would result in both a false positive (ie, diagnosis assigned at unenhanced CT was not in reference standard) and a false negative (ie, primary diagnosis in reference standard was not identified). Hence, overall accuracy was defined as the proportion of CT interpretations from a given rater in which all primary and actionable secondary diagnoses were correctly identified and no other diagnoses were made. False-positive rate was defined as the proportion of CT

interpretations with at least 1 false-positive primary diagnosis or false-positive actionable secondary diagnosis. Similarly, false-negative rate was defined as the proportion of CT interpretations in which at least 1 primary diagnosis or actionable secondary diagnosis was not identified by the rater. Results for primary diagnoses and actionable secondary diagnoses are presented in combination and separately. 95% CIs were calculated. Additional details are provided in the eMethods and eResults in Supplement 1.

Multivariable generalized linear mixed models were performed to determine whether the likelihood of a diagnostic error at unenhanced CT changes with radiologist experience, patient BMI, patient age, or patient gender. *P* value adjustment was performed to correct for multiple comparisons (false discovery rate-adjusted *P* values).[14,15] BMI was incorporated because smaller patients have less intraperitoneal and retroperitoneal fat and therefore less intrinsic image contrast. Sankey diagrams were drawn comparing the diagnoses in the reference standard with the diagnoses made by the radiologists. The Gwet interrater agreement coefficient (AC) was calculated. Statistical analyses were performed using SAS software version 9.4 (SAS Institute).

Results

Study Population and Reference Standard

There were 201 included patients (female, 108; male, 93) with a mean age of 50.1 (SD, 20.9) years and mean BMI of 25.5 (SD, 5.4) (**Table 1**). Most patients did not have a specific location of abdominal pain (103 of 201 [51%]) or a suggested diagnosis stated on the CT order (152 of 201 [76%]). Five percent of patients had an abdominal operation in the last 30 days (10 of 201). There were 104 primary diagnoses in 98 patients (49% of patients; 6 patients had 2 primary diagnoses) and 17 secondary diagnoses in 17 patients (8% of patients). Ninety-two patients had no primary or actionable secondary diagnosis (46%). **Figures 1** and **2** illustrate reference standard diagnoses at contrast-enhanced CT and assigned diagnoses at unenhanced CT for all radiologists. There were more than 30 primary and secondary diagnoses in the reference standard.

Simple Diagnostic Accuracy (by Examination)

Simple diagnostic accuracy by radiologist at the level of the CT examination (ie, multiple diagnoses in a single patient are disregarded) (eTable in Supplement 1) was calculated as sensitivity (faculty, 78% to 89%; residents, 67% to 87%), specificity (faculty, 73% to 86%; residents, 75% to 92%), positive predictive value (faculty, 79% to 88%; residents, 81% to 92%), and negative predictive value (faculty, 77% to 87%; residents, 70% to 83%). These diagnostic accuracy data are higher than the detailed diagnostic accuracy due to treating the test result as a binary outcome.

Detailed Diagnostic Accuracy (by Diagnosis)

Unenhanced CT was approximately 30 percentage points less accurate than contrast-enhanced CT for identification of all primary and actionable secondary diagnoses (**Table 2**). The reduction in diagnostic accuracy was similar for faculty and residents (overall accuracy for primary and secondary diagnoses was 70% [pooled], 69% to 74% [faculty], and 69% to 70% [residents]) (Table 2). However, overall accuracy was higher for faculty when only primary diagnoses were considered (faculty, 85% to 90%; residents, 76% to 79%) and higher for residents when only secondary diagnoses were considered (faculty, 81% to 89%, residents, 88% to 92%) (Table 2). In general,

Table 1. Study Population Characteristics[a]

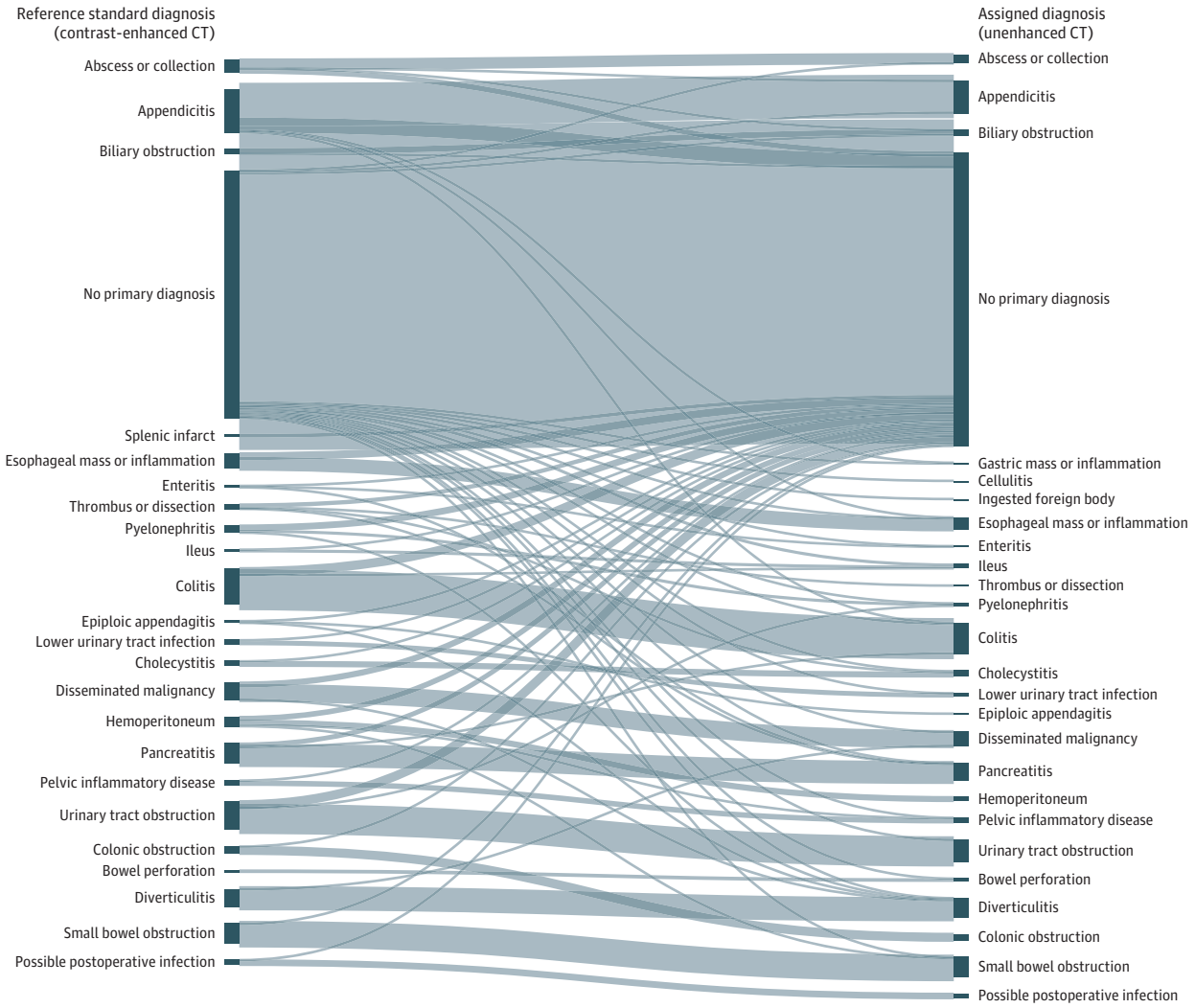
Patient characteristic	Data
No.	201
Sex, No. (%)	
Female	108 (54)
Male	93 (46)
Age, y, mean (SD)	50.1 (20.9)
BMI,[b] mean (SD)	25.5 (5.4)
Actionable diagnosis at CT, No. (%)	109 (54%)
Primary diagnosis	104 in 98 Patients (49)
Actionable secondary diagnosis	17 in 17 Patients (8)
No primary or actionable secondary diagnosis	92 (46)
Abdominal pain location, No. (%)	
Diffuse	3 (1)
Epigastric	5 (2)
Left lower quadrant	20 (10)
Left upper quadrant	2 (1)
Left abdomen	1 (0.5)
Lower abdomen	8 (4)
Periumbilical	2 (1)
Right lower quadrant	50 (25)
Right upper quadrant	4 (2)
Right abdomen	3 (1)
Pain location not specified	103 (51)
Operation within last 30 d, No. (%)	10 (5)
Suspected diagnosis by ordering health care professional, No. (%)	
Abscess	3 (1)
Appendicitis	9 (4)
Colitis	1 (0.5)
Crohn disease	4 (2)
Diverticulitis	9 (4)
Hemorrhage	1 (0.5)
Ischemic colitis	1 (0.5)
Neoplasm	2 (1)
Obstruction	10 (5)
Pancreatitis	3 (1)
Perforation	3 (1)
Pyelonephritis	1 (0.5)
Sepsis	1 (0.5)
Trauma	1 (0.5)
No suspected diagnosis provided	152 (76)

Abbreviation: CT, computed tomography

[a] Patients presented to the emergency department with abdominal pain. Listed clinical information represents clinical history provided on the CT requisition in addition to abdominal pain.

[b] Calculated as weight in kilograms divided by height in meters squared.

Figure 1. Diagnostic Accuracy of Unenhanced Computed Tomography (CT) for the Primary Cause(s) of Pain



Sankey diagram illustrating the reference standard diagnosis (on left, contrast-enhanced CT) and the assigned diagnoses (on right, unenhanced CT) for the primary diagnosis (ie, primary cause(s) of acute abdominal pain).

faculty were less likely to miss a finding and more likely to identify a false positive (Table 2 and **Table 3**). The false-negative rates were 13% to 19% (faculty) and 15% to 27% (residents), and the false-positive rates were 10% to 21% (faculty) and 8% to 19% (residents).

Multivariable Regression

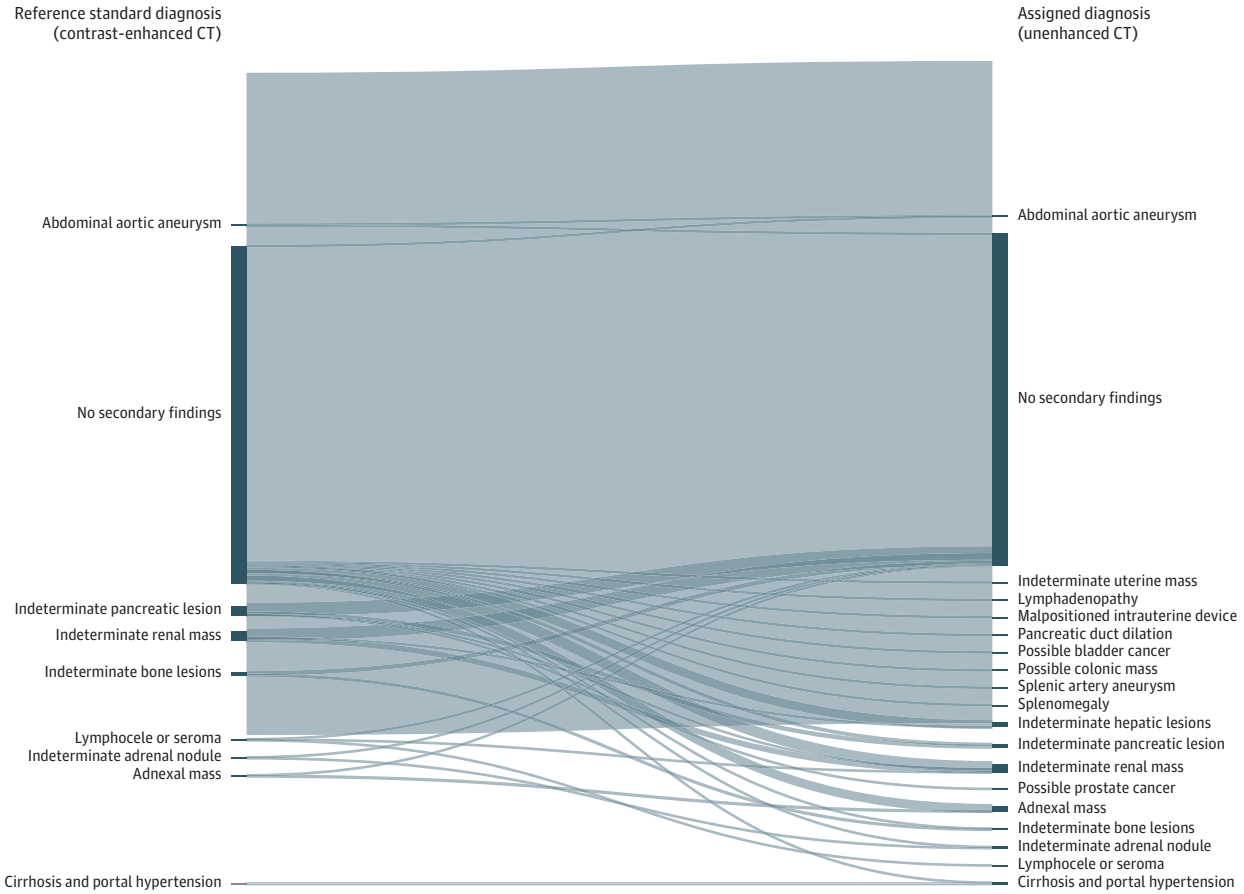
Faculty radiologists had higher overall diagnostic accuracy than radiology residents for identifying primary diagnoses (82% vs 76%; adjusted odds ratio [OR] 1.83; 95% CI, 1.26-2.67; $P = .002$), but lower overall diagnostic accuracy for identifying secondary diagnoses (87% vs 90%; adjusted OR, 0.57; 95% CI, 0.35-0.93; $P = .02$) (Table 3). This was because faculty had a lower false-negative rate for primary diagnoses (38% vs 62%; adjusted OR, 0.23; 95% CI, 0.13-0.41; $P < .001$) and a higher false-positive rate for secondary diagnoses (63% vs 37%; adjusted OR, 2.11; 95% CI, 1.26-3.54; $P = .01$).

Older age was associated with a minor reduction in overall diagnostic accuracy due to mild increases in the false-positive and false-negative rates for actionable secondary diagnoses in older patients (Table 3). Neither patient BMI nor patient gender predicted diagnostic error, either independently or as a statistical interaction with radiologist experience (data not shown but available upon request) (Table 3).

Interrater Agreement

Interrater agreement for overall accuracy was moderate (Gwet AC, 0.58). Agreement was higher when there was no primary diagnosis (Gwet AC, 0.67) compared with when there was a primary diagnosis (Gwet AC, 0.49). Sixteen of 201 examinations were misinterpreted by all 6 radiologists (8%) and 74 of 201 examinations were diagnosed correctly by all 6 radiologists (37%). Of the 16 examinations that no radiologist correctly interpreted, 94% had a primary diagnosis and 38% had a secondary

Figure 2. Diagnostic Accuracy of Unenhanced Computed Tomography (CT) for Actionable Secondary Diagnoses



Sankey diagram illustrating the reference standard diagnosis (on left, contrast-enhanced CT) and the assigned diagnoses (on right, unenhanced CT) for actionable secondary diagnoses.

diagnosis. Of the 74 examinations that all radiologists correctly interpreted, only 31% had a primary diagnosis and only 1% had a secondary diagnosis.

Discussion

Contrast-enhanced abdominopelvic CT is the most common and most accurate test for the evaluation of acute abdominal pain in adult nonpregnant ED patients. However, contrast medium is sometimes withheld in this setting due to risk of complication or lack of availability. It is important to understand the risk of withholding contrast medium so informed risk-benefit analyses can be made. In this consecutive cohort of ED patients presenting with abdominal pain, unenhanced CT was consistently approximately 30 percentage points less accurate than contrast-enhanced CT for primary and secondary actionable findings (overall accuracy, faculty, 68% to 74%; residents, 69% to 70%). Faculty had higher accuracy than residents for primary diagnoses (OR, 1.83; 95% CI, 1.26-2.67) but lower accuracy for actionable secondary diagnoses (OR, 0.57; 95% CI, 0.35-0.93). This was because faculty made fewer false-negative primary diagnoses (OR, 0.23; 95% CI, 0.13-0.41) and more false-positive actionable secondary diagnoses (false-positive OR, 2.11; 95% CI, 1.26-3.54). Prior studies have reported higher diagnostic accuracy of unenhanced CT for the evaluation of acute abdominal pain, but in general have either focused on a limited number of diagnoses or lacked a robust reference standard.[9-12,16-18]

False-negative (faculty, 13% to 19%; residents, 15% to 27%) and false-positive (faculty, 10% to 21%; residents, 8% to 19%) results were common at unenhanced CT for all radiologists. This likely was because reduced image contrast reduces accuracy and radiologist confidence. Example false-positive results included pancreatitis, bowel perforation, diverticulitis, pyelonephritis, and neoplasm. Example false-negative results included vascular dissection, hemoperitoneum, infection, and neoplasm. The commonality of both false-positive and false-negative results challenges efforts to adjust reading style to reduce error. In other words, the diagnostic penalty resulting from the elimination of contrast medium is not easily fixable by simply raising or lowering the threshold to report a diagnosis.

Table 2. Detailed Diagnostic Accuracy of Unenhanced Computed Tomography at the Findings Level in Emergency Department Patients With Abdominal Pain, Using Contrast-Enhanced Computed Tomography as Reference Standard[a]

Diagnostic accuracy methods	Radiologist					
	Faculty A	Faculty B	Faculty C	Resident A	Resident B	Resident C
Overall accuracy, No. (%) [95% CI]						
Primary and secondary diagnoses	139/201 (69) [62-75]	146/201 (73) [66-79]	148/201 (74) [67-80]	141/201 (70) [63-76]	139/201 (69) [62-75]	138/201 (69) [62-75]
Primary diagnoses only	163/201 (81) [75-86]	165/201 (82) [76-87]	169/201 (84) [78-89]	158/201 (79) [72-84]	151/201 (75) [69-81]	152/201 (76) [69-81]
Secondary diagnoses only	170/201 (85) [79-89]	180/201 (90) [84-93]	(173/201) (86) [80-91]	177/201 (88) [83-92]	184/201 (92) [87-95]	181/201 (90) [85-94]
False-positive rate, No. (%) [95% CI]						
Primary and secondary diagnoses	43/201 (21) [16-28]	20/201 (10) [6-15]	33/201 (16) [11-22]	38/201 (19) [14-25]	17/201 (8) [5-13]	22/201 (11) [7-16]
Primary diagnoses only	20/201 (10) [6-15]	10/201 (5) [2-9]	16/201 (8) [5-13]	24/201 (12) [8-17]	12/201 (6) [3-10]	12/201 (6) [3-10]
Secondary diagnoses only	24/201 (12) [8-17]	11/201 (5) [3-10]	20/201 (10) [6-15]	16/201 (8) [5-13]	5/201 (2) [1-6]	11/201 (5) [3-10]
False-negative rate, No. (%) [95% CI]						
Primary and secondary diagnoses	27/201 (13) [9-19]	38/201 (19) [14-25]	28/201 (14) [9-20]	30/201 (15) [10-21]	52/201 (26) [20-33]	54/201 (27) [21-34]
Primary diagnoses only	21/201 (10) [7-16]	27/201 (13) [9-19]	19/201 (9) [6-14]	22/201 (11) [7-16]	43/201 (21) [16-28]	46/201 (23) [17-29]
Secondary diagnoses only	7/201 (3) [1-7]	12/201 (6) [3-10]	10/201 (5) [2-9]	8/201 (4) [2-8]	12/201 (6) [3-10]	11/201 (5) [3-10]

[a] Overall accuracy is (true positive with false negative) divided by (true positive with true negative with false positive with false negative). To be considered accurate, all true-positive diagnoses had to be identified without any false positives. False-positive rate is false positive divided by (true positive with true negative with false positive with false negative). Any false positive in an examination qualifies. False-negative rate is false negative divided by (true positive with true negative with false positive with false negative). Any false negative in an examination qualifies.

Table 3. Multivariable Generalized Linear Mixed Models (Binary Logit) of Diagnostic Accuracy With Random Effects for Case

Diagnostic accuracy methods	Faculty vs resident	*P* value	BMI[a] of patient	*P* value	Age, y	*P* value	Female vs male patient	*P* value
Overall accuracy, OR (95% CI)[b]								
Primary and secondary diagnoses	1.19 (0.86-1.65)	.29	1.02 (0.96-1.08)	.54	0.97 (0.96-0.99)[c]	<.001	0.62 (0.31-1.25)	.18
Primary diagnoses only	1.83 (1.26-2.67)[c]	.002	1.05 (0.98-1.14)	.19	0.99 (0.97-1.01)	.34	0.78 (0.34-1.79)	.56
Secondary diagnoses only	0.57 (0.35-0.93)[c]	.02	0.98 (0.90-1.06)	.56	0.93 (0.91-0.96)[c]	<.001	0.61 (0.23-1.65)	.33
False-positive rate, OR (95% CI)								
Primary and secondary diagnoses	1.37 (0.96-1.97)	.08	1.03 (0.99-1.08)	.16	1.02 (1.01-1.03)[c]	.001	1.73 (1.00-2.99)	.05
Primary diagnoses only	0.94 (0.59-1.51)	.81	1.03 (0.96-1.10)	.42	1.01 (0.99-1.02)	.53	1.28 (0.60-2.76)	.52
Secondary diagnoses only	2.11 (1.26-3.54)[c]	.005	1.04 (0.98-1.11)	.20	1.04 (1.02-1.06)[c]	<.001	2.29 (1.04-5.03)	.04
False-negative rate, OR (95% CI)								
Primary and secondary diagnoses	0.31 (0.19-0.51)[c]	<.001	0.88 (0.76-1.00)	.06	1.05 (1.02-1.09)[c]	.002	1.04 (0.26-4.20)	.95
Primary diagnoses only	0.23 (0.13-0.41)[c]	<.001	0.88 (0.76-1.01)	.08	1.02 (0.99-1.06)	.23	1.76 (0.37-8.52)	.48
Secondary diagnoses only	0.76 (0.27-2.13)	.60	0.77 (0.51-1.17)	.23	1.28 (1.09-1.50)[c]	.002	0.12 (0.00-10.88)	.36

Abbreviation: OR, odds ratio.

[a] Calculated as weight in kilograms divided by height in meters squared.

[b] For overall accuracy, ORs more than 1 indicate higher accuracy and ORs less than 1 indicate lower accuracy. For false positive rate and false negative rate, ORs more than 1 indicate greater odds of false positives (or false negatives) and ORs less than 1 indicate lower odds of false positives (or false negatives).

[c] These data are statistically significant after adjustment for multiple comparisons (false discovery rate–adjusted *P* values[14,15]).

The diagnostic risk of withholding contrast medium should be part of an informed risk-benefit analysis. For example, patients without acute kidney injury (AKI) or an estimated glomerular filtration rate of 30 mL/min or more per 1.73 m^2 have a risk of contrast-induced AKI close to 0%.[6] Similarly, patients who have had a prior mild hypersensitivity reaction to iodinated contrast medium are at very low risk (less than 1%) of a life-threatening reaction on repeat administration, especially using standard-of-care risk-mitigation strategies.[5,19-21] Therefore, accepting a 30-percentage-point decrement in diagnostic accuracy in these settings likely causes harm. Patients who do have AKI or an estimated glomerular filtration rate less than 30 mL/min per 1.73 m^2 not receiving maintenance dialysis have a more uncertain risk of contrast-

induced AKI (0% to 17%).[6] Patients with a prior moderate or severe hypersensitivity reaction to iodinated contrast medium have an unclear risk that is complicated to mitigate.[5,19-21] Therefore, the risk benefit-analysis in these settings is more complicated and probably depends on the pretest suspected diagnoses.

We observed some interesting secondary results. A priori, we predicted that diagnostic error would be more common in smaller patients because patients with less body fat have less intrinsic tissue contrast. However, BMI was not predictive of diagnostic error in our study. This may have been because the study was not powered to detect that association or because BMI is not an accurate surrogate for visceral adiposity.[22] Diagnostic error was slightly more common in older patients. That is likely because older patients are more likely to have primary and actionable secondary diagnoses, and diagnostic error was more common in patients with those findings.

Limitations

Our study has limitations. It was retrospective and susceptible to related biases. We minimized the risk of selection bias, longitudinal bias, and verification bias by using a consecutive cohort at an institution in which dual-energy CT was routinely used to image patients with abdominal pain and through use of a contemporaneous reference standard that is the current standard of care. We minimized risk of reader bias by blinding our radiologists to the reference standard and clinical data outside the CT order, and by using radiologists with a range of experience at 3 separate institutions. We intentionally sampled a general population of ED patients with abdominal pain because dozens of diagnoses are commonly made with contrast-enhanced CT in that setting. Restricting the cohort to a specific diagnosis would have inflated the diagnostic accuracy. We subtracted oral, as well as IV, contrast medium from the reference standard to generate the unenhanced CT data. Diagnostic accuracy might have been higher had oral contrast medium been visible, or different had a single-energy unenhanced CT been used.

Conclusion

In summary, unenhanced CT was approximately 30 percentage points less accurate than contrast-enhanced CT for the evaluation of abdominal pain in the ED. Prior studies evaluating unenhanced CT in this population likely have overstated its accuracy due to focus on 1 or few diagnoses or lack of a robust reference standard. The consistent results we observed across 3 centers suggest that the substantial diagnostic penalty we observed is likely to be related to the removal of contrast medium rather than to radiologist idiosyncrasy. For patients with risk factors for receiving iodinated contrast medium (eg, prior hypersensitivity reaction, severe kidney disease) or for patients receiving care in locations where contrast media is in short supply, the diagnostic risk of withholding contrast medium should be considered in the risk-benefit analysis. In many patients, the risk of withholding iodinated contrast medium may be higher than the risk of administering it.

ARTICLE INFORMATION

Accepted for Publication: January 4, 2023.

Published Online: May 3, 2023. doi:10.1001/jamasurg.2023.1112

Author Contributions: Drs Ream and Davenport had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Shaish, Ream, Kim, Newhouse, Davenport.
Acquisition, analysis, or interpretation of data: Shaish, Ream, Huang, Troost, Gaur, Chung, Kim, Patel, Khalatbari, Davenport.
Drafting of the manuscript: Shaish, Troost, Kim, Patel, Khalatbari, Davenport.
Critical revision of the manuscript for important intellectual content: Shaish, Ream, Huang, Troost, Gaur, Chung, Newhouse, Khalatbari, Davenport.
Statistical analysis: Shaish, Troost, Khalatbari.
Administrative, technical, or material support: Shaish, Huang, Troost, Kim, Davenport.
Supervision: Shaish, Huang, Kim, Davenport.

Conflict of Interest Disclosures: Dr Troost reported grants from the National Center for Advancing Translational Sciences/National Institutes of Health (UL1TR002240) outside the submitted work. Dr Gaur reported a patent for artificial intelligence technology in prostate imaging with royalties paid from the National Institutes of Health. Dr Chung reported personal fees from BotImage outside the submitted work. Dr Davenport reported book royalties from Wolters Kluwer outside the submitted work. No other disclosures were reported.

Funding/Support: This work was supported in part by the National Center for Advancing Translational Sciences for the Michigan Institute for Clinical and Health Research (UL1TR002240).

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Sharing Statement: See Supplement 2.

REFERENCES

1. American College of Radiology. ACR Appropriateness criteria. Accessed March 24, 2023. https://www.acr.org/Clinical-Resources/ACR-Appropriateness-Criteria

2. Pandharipande PV, Reisner AT, Binder WD, et al. CT in the emergency department: a real-time study of changes in physician decision making. *Radiology*. 2016;278(3):812-821. doi:10.1148/radiol.2015150473

3. Eisenberg JD, Reisner AT, Binder WD, et al. Role of CT in the diagnosis of nonspecific abdominal pain: a multicenter analysis. *AJR Am J Roentgenol*. 2017;208(3):570-576. doi:10.2214/AJR.16.16669

4. Bae KT. Intravenous contrast medium administration and scan timing at CT: considerations and approaches. *Radiology*. 2010;256(1):32-61. doi:10.1148/radiol.10090908

5. American College of Radiology. Manual on contrast media. Accessed March 24, 2023. https://www.acr.org/Clinical-Resources/Contrast-Manual

6. Davenport MS, Perazella MA, Yee J, et al. Use of intravenous iodinated contrast media in patients with kidney disease: consensus statements from the American College of Radiology and the National Kidney Foundation. *Radiology*. 2020;294(3):660-668. doi:10.1148/radiol.2019192094

7. Davenport MS, Chu P, Szczykutowicz TP, Smith-Bindman R. Comparison of strategies to conserve iodinated intravascular contrast media for computed tomography during a shortage. *JAMA*. 2022;328(5):476-478. doi:10.1001/jama.2022.9879

8. Balogh EP, Miller BT, Ball JR, eds. *Improving diagnosis in health care*. The National Academy of Sciences, The National Academies Press; 2015. doi:10.17226/21794

9. Lamoureux C, Weber S, Hanna T, Grabiel AJ, Clark RH. Effect of intravenous contrast for CT abdomen and pelvis on detection of urgent and non-urgent pathology: can repeat CT within 72 hours be avoided? *Emerg Radiol*. 2019;26(6):601-608. doi:10.1007/s10140-019-01704-0

10. Hill BC, Johnson SC, Owens EK, Gerber JL, Senagore AJ. CT scan for suspected acute abdominal process: impact of combinations of IV, oral, and rectal contrast. *World J Surg*. 2010;34(4):699-703. doi:10.1007/s00268-009-0379-6

11. Lane MJ, Katz DS, Ross BA, Clautice-Engle TL, Mindelzun RE, Jeffrey RB Jr. Unenhanced helical CT for suspected acute appendicitis. *AJR Am J*

Roentgenol. 1997;168(2):405-409. doi:10.2214/ajr.168.2.9016216

12. Tack D, Bohy P, Perlot I, et al. Suspected acute colon diverticulitis: imaging with low-dose unenhanced multi-detector row CT. *Radiology*. 2005;237(1):189-196. doi:10.1148/radiol.2371041432

13. Coursey CA, Nelson RC, Boll DT, et al. Dual-energy multidetector CT: how does it work, what can it tell us, and when can we use it in abdominopelvic imaging? *Radiographics*. 2010;30(4):1037-1055. doi:10.1148/rg.304095175

14. Glickman ME, Rao SR, Schultz MR. False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. *J Clin Epidemiol*. 2014;67(8):850-857. doi:10.1016/j.jclinepi.2014.03.012

15. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B*. 1995;57:289-300. doi:10.1111/j.2517-6161.1995.tb02031.x

16. Basak S, Nazarian LN, Wechsler RJ, et al. Is unenhanced CT sufficient for evaluation of acute abdominal pain? *Clin Imaging*. 2002;26(6):405-407. doi:10.1016/S0899-7071(02)00535-1

17. Millet I, Sebbane M, Molinari N, et al. Systematic unenhanced CT for acute abdominal symptoms in the elderly patients improves both emergency department diagnosis and prompt clinical management. *Eur Radiol*. 2017;27(2):868-877. doi:10.1007/s00330-016-4425-0

18. Barat M, Paisant A, Calame P, et al. Unenhanced CT for clinical triage of elderly patients presenting to the emergency department with acute abdominal pain. *Diagn Interv Imaging*. 2019;100(11):709-719. doi:10.1016/j.diii.2019.05.004

19. Umakoshi H, Nihashi T, Takada A, et al. Iodinated contrast media substitution to prevent recurrent hypersensitivity reactions: a systematic review and meta-analysis. *Radiology*. 2022;305(2):341-349. doi:10.1148/radiol.220370

20. Mervak BM, Davenport MS, Ellis JH, Cohan RH. Rates of breakthrough reactions in inpatients at high risk receiving premedication before contrast-enhanced CT. *AJR Am J Roentgenol*. 2015;205(1):77-84. doi:10.2214/AJR.14.13810

21. Davenport MS, Cohan RH, Caoili EM, Ellis JH. Repeat contrast medium reactions in premedicated patients: frequency and severity. *Radiology*. 2009;253(2):372-379. doi:10.1148/radiol.2532090465

22. Shuster A, Patlas M, Pinthus JH, Mourtzakis M. The clinical importance of visceral adiposity: a critical review of methods for visceral adipose tissue analysis. *Br J Radiol*. 2012;85(1009):1-10. doi:10.1259/bjr/38447238