Current Problems in Diagnostic Radiology 000 (2020) 1–4

Current Problems in Diagnostic Radiology

journal homepage: www.cpdrjournal.com

Appropriateness of Computed Tomography and Ultrasound for Abdominal Complaints in the Emergency Department

Martina Zaguini Francisco, MD[a], Stephan Altmayer, MD[b], Nupur Verma, MD[c], Guilherme Watte, PhD[a,b], Marina de Campos Brandao, MD[a], Marcelo Cardoso Barros, MD[b], Tan-Lucien Mohammed, MD[c], Bruno Hochhegger, MD, PhD[a,b,*]

[a] *Universidade Federal de Ciencias da Saude de Porto Alegre, Porto Alegre, Rio Grande do Sul, Brazil*
[b] *Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil*
[c] *Department of Radiology, University of Florida, Gainesville, FL*

A B S T R A C T

Objective: The purpose of this study was to evaluate the appropriateness of ultrasound (US) and computed tomography (CT) examinations ordered in the emergency department (ED) for abdominal complaints.
Materials and Methods: We reviewed 154 CTs and 154 US orders for appropriateness using evidence-based recommendations by the American College of Radiology. The sample was powered to show a prevalence of inappropriate orders of 25% with a margin of error of 7.5%. Findings in the final reports were compared to the initial clinical diagnosis classified in 4 categories: normal, compatible with initial diagnosis, alternative diagnosis, and inconclusive. We also evaluated the frequency in which a second imaging modality was ordered on the same visit.
Results: A total of 135 CT and 143 US examinations had complete clinical information to allow evaluation of order appropriateness. The rate of inappropriate orders was 36.3% for CT and 84.4% for US. The final report of appropriate orders was significantly more likely to demonstrate findings compatible with the initial diagnosis for both CT (76.7% vs 20.4%, $P < 0.0001$) and US (38.9% vs 14.4%, $P = 0.0093$). Inappropriately ordered CT scans were more likely to show no abnormalities (46.9 vs 16.3%, $P = 0.0001$). An additional imaging order with a secondary modality was requested in 20% of the inappropriate US orders, and 8.2% of the inappropriate CT orders.
Conclusion: The prevalence of inappropriate examinations in the ED was 36.3% for CT and 84.4% for US. Appropriately ordered exams were more likely to yield imaging findings compatible with the initial diagnosis for both modalities.

Introduction

Overutilization of imaging tests is defined as the performance of imaging procedures where clinical circumstances indicate that they are unlikely to improve patient outcomes.[1] Over 85% of the emergency physicians agree they request too many diagnostic tests.[2] Abdominal complaints, which are often caused by gastrointestinal or urological disease, are among the principal reasons for emergency department (ED) visits.[3-4,5] Imaging studies of the abdomen and pelvis, such as computed tomography (CT) and ultrasound (US), are often ordered in the workup of abdominal complaints to assist in diagnosis and managing potential etiologies.

Imaging exams orders in the ED may be considered inappropriate for several reasons, such as not increasing the post-test probability of a diagnosis, being chosen over a more appropriate first exam modality, or when it does not change the therapeutic management. Choosing the wrong imaging modality can also be considered inappropriate ordering. The American College of Radiology (ACR) has developed and continually revised series of evidence-based guidelines called "ACR Appropriateness Criteria" to assist physicians in making the most appropriate imaging decision in many clinical contexts.[6] Despite the open availability of these guidelines and considered work in introducing them to nonradiology providers, approximately 25% of all US and CT scans ordered in the ED are considered inappropriate orders.[7]

In light of the prevalence of abdominal complaints in the ED, the purpose of this study was to evaluate the appropriateness of US and CT scans ordered in the ED for this common complaint.

Methods

Population, Imaging Protocol, and Study Design

We retrospectively identified consecutive adult patients who had undergone either US or CT imaging for abdominal complaints in the ED of a nontrauma tertiary care urban hospital from January to March 2019. There is no formal imaging protocol for requesting CT and US for patients presenting with abdominal complaints in our institution. Ordering physicians are educated and encouraged to follow the

*Reprint requests: Bruno Hochhegger, MD, PhD, Medical Imaging Research Lab, LABIMED, Department of Radiology, Pavilhão Pereira Filho Hospital, Irmandade Santa Casa de Misericórdia de Porto Alegre - Av. Independência, 75, Porto Alegre, Brazil. 90020-160.
E-mail address: brunoho@ufcspa.edu.br (B. Hochhegger).

https://doi.org/10.1067/j.cpradiol.2020.11.004
0363-0188/

recommendations from the "ACR Appropriateness Criteria." However, clinicians still preserve final discretion on ordering tests. No decision-support software is used when ordering imaging studies.

Data on demographics, entry date, clinical indication for imaging referral, the first choice of imaging modality, relevant findings, additional imaging requests were collected from the patient's charts and radiology reports. Patients with documented iodine allergy or estimated glomerular filtration rate < 30 mL/m^2 were not included in the sample as the choice of test by the ordering physician could be influenced by these factors. If a patient had undergone both imaging modalities during the same visit, only the first entry was considered for the main analysis. Ordering physicians from the ED were not actively involved in the study. The collection of data for this study was approved by the institutional review board.

Appropriateness of imaging order and outcomes

We primarily utilized the "ACR Appropriateness criteria" guidelines for assessing first ordered exam appropriateness.[6] We defined as appropriate imaging requests those classified by the ACR as "usually appropriate"—that is, those with a favorable risk-benefit ratio for patients. For the basis of our analysis, imaging was considered inappropriate when classified by the ACR as "may be appropriate" (ie, risk-benefit ratio is equivocal, or alternative modality in specific clinical scenarios, such as pregnancy or children) or "usually not appropriate"——those with possibly unfavorable risk-benefit ratio for patients.

Patients were classified into clinical variants (eg, suspected small-bowel obstruction) and scenarios (eg, acute presentation vs indolent presentation) according to the ACR appropriateness criteria topics simultaneously by 2 independent radiologists based on the primary assessment of the referring physician and available clinical information at the time of imaging request. The reviewers were blinded to the existing imaging report for classifying patients into the clinical scenarios. Any disagreements were resolved by consensus. If any patient with a specific clinical scenario could not be classified into the available ACR guidelines (eg, gastroenteritis), the 2 radiologists judged the request's appropriateness independently, and disagreements were solved by consensus.

Radiology reports were classified as "normal" when they did not show any abnormal findings that relate to the clinical question by the ordering physician. Positive reports were classified as "compatible with the initial diagnosis," "alternative diagnosis," or "inconclusive" in comparison to the primary clinical assessment leading to the imaging referral. For instance, if a patient referred for CT due to suspected appendicitis after the initial assessment had imaging findings of appendicitis, it was considered "compatible with initial diagnosis." On the other hand, if the appendix was normal and the CT revealed an obstructive stone in the right ureter, it would be classified as "alternative diagnosis." "Inconclusive" findings were defined as any unspecific findings related to the clinical question, such as perirenal fat stranding on CT for suspected pyelonephritis, ureteral dilation on US for suspected urolithiasis, etc. Incidental findings not related to the clinical scenario were considered within the "normal" reports. Two independent radiologists performed the classification of the radiology reports. Disagreements were solved by consensus. It was also recorded if a patient had undergone a second imaging modality during the same visit (eg, CT scan after an inconclusive US, or vice versa).

Statistics

Sample size was calculated for an estimated proportion of inappropriate tests of 25% based on the results of previous studies.[7,8] Using a margin of error of 0.075, a 95% confidence interval (95% CI), and a 20% expected rate of subjects with incomplete clinical information for the appropriateness of a test to be determined, the final sample size required was 154 patients for each imaging modality. Continuous variables are expressed as mean ± SD and comparisons were performed using two-sample, two-tailed Student's *t*-test. Categorical variables are reported as frequency (percentages) and were compared with the Pearson's chi-squared test or Fisher exact test when the value of any cell was equal to zero. A 2-tailed *P*-value of less than 0.05 was considered significant. Subgroup comparison tests regarding the verdict (normal, compatible, alternative, inconclusive) of the final report were calculated using adjusted residuals and Bonferroni correction to reduce the probability of a Type I error ($P < 0.0125$ for significance).[9] Statistical analysis was performed using IBM SPSS v.23 (IBM Corp, Armonk, NY).

Results

A total of 154 CT and 154 US examinations performed for abdominal complaints in the ED were reviewed. Out of the 154 exams reviewed, 19 CT scans and 11 US were excluded due to incomplete clinical information to determine test's clinical appropriateness. Thus, 135 CTs and 143 US were included in the final revision for analysis (Table 1). Overall, only 37.4% of all examinations were considered appropriate. The rate of inappropriate tests was 36.3% for CT scans and 84.4% of US. Approximately half of all inappropriate exams (55.1% for CT, and 44.8% for US) were classified under the ACR category of "usually not appropriate." The rate of inappropriate tests (CT and US) were 66.6% for females and 47.8% for males, which yielded no significant association between gender and appropriateness of imaging tests ($P = 0.10$). There was a statistically significant difference in the mean age between patients within the appropriate and inappropriate categories (mean age 51.3 ± 16.0 years vs 46.8 ± 17.8 years, respectively; $P = 0.03$). There was only 1 pregnant patient in the sample.

The analysis of all CT orders is shown in Table 2. The most common reasons for CT orders were urolithiasis ($n = 36$, 26.7%), acute abdominal pain ($n = 18$, 13.3%), pyelonephritis ($n = 12$, 8.9%), small bowel obstruction ($n = 11$, 8.1%), diverticulitis ($n = 11$, 8.1%), and appendicitis ($n = 10$, 7.4%). Gastroenteritis ($n = 10$, 7.4%) was the most frequent diagnosis not contemplated by the ACR criteria leading to a CT. Regarding the final report, appropriate orders were significantly more likely to show imaging findings compatible with the initial clinical diagnosis that lead to the imaging referral (76.7% vs 20.4%, $P < 0.0001$). Inappropriate CT scans were more likely to show no abnormalities (46.9 vs 16.3%, $P = 0.0001$), suggest an alternative diagnosis (20.4% vs 4.6%, $P = 0.0037$), or be considered inconclusive (12.2 vs 2.3%, $P = 0.0214$) (Fig 1). There was no significant difference in the rate of inappropriate exams between contrast-enhanced CT and unenhanced CT (33.3% vs 41.2%, respectively; $P = 0.358$)

The analysis of the US orders is shown in Table 3. Only 18 (12.6%) out of the 143 US requests were considered appropriate. The most common reasons for US were urolithiasis ($n = 29$, 20.3%), acute abdominal pain ($n = 23$, 16.1%), pyelonephritis ($n = 15$, 10.4%), biliary disease ($n = 9$, 6.3%), diverticulitis ($n = 6$, 4.2%), and appendicitis ($n = 3$, 2.1%). Gastroenteritis ($n = 26$, 18.2%) and uncomplicated urinary tract

TABLE 1
Study characteristics

	CT ($n = 135$)	US ($n = 143$)	CT + US ($n = 278$)
Age (years)	51.4 ± 17.5	45.7 ± 16.6	48.4 ± 17.2
Female (%)	77 (57.0)	85 (59.4)	162 (58.3)
Appropriate (%)	86 (63.7)	18 (12.6)	104 (37.4)
Inappropriate (%)	49 (36.3)	125 (84.4)	174 (62.6)
"Maybe appropriate" (%)	22 (16.3)	69 (48.3)	91 (32.7)
"Usually not appropriate" (%)	27 (20.0)	56 (39.2)	83 (29.9)

CT, computed tomography; US, ultrasound

TABLE 2
Analysis of CT orders regarding appropriateness, diagnosis, and outcomes

	Appropriate ($n = 86$)	Inappropriate ($n = 49$)	*P*-value
ACR diagnosis			
Urolithiasis	30/36	6/36	-
Acute abdominal pain	17/18	1/18	-
Pyelonephritis	5/12	7/12	-
Small bowel obstruction	10/11	1/11	-
1Diverticulitis	6/11	5/11	-
Appendicitis	9/10	1/10	-
Biliary disease	0/4	4/4	-
Pancreatitis	0/4	4/4	-
Renal failure	0/3	3/3	-
Other	5/9	4/9	-
Non-ACR diagnosis			
Gastroenteritis	0/10	10/10	-
Other	4/7	3/7	-
Final report			
Normal	14/86 (16.3)	23/49 (46.9)	.0001*
Compatible w/ initial dx	66/86 (76.7)	10/49 (20.4)	<.0001*
Alternative dx	4/86 (4.6)	10/49 (20.4)	.0037*
Inconclusive	2/86 (2.3)	6/49 (12.2)	.0214
Secondary imaging modality	2/86 (2.3)	4/49 (8.2)	.189

ACR, American College of Radiology; dx, diagnosis.
*Statistically significant at the level of $P = 0.0125$.

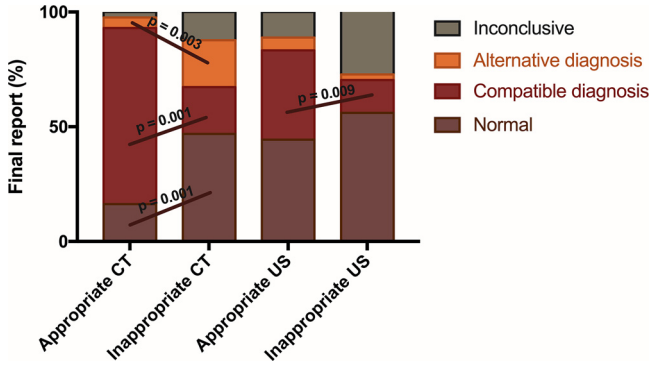

FIG 1. Final report findings for CT and US stratified by appropriateness of the order. Only the *P*-values of statistically significant differences between appropriate and inappropriate groups are shown. If the p-value is not provided, differences between groups was not statistically significant.

TABLE 3
Analysis of US orders regarding appropriateness, diagnosis, and outcomes

	Appropriate ($n = 18$)	Inappropriate ($n = 125$)	*P*-value
ACR diagnosis			
Urolithiasis	1/29	28/29	-
Acute abdominal pain	0/23	23/23	-
Pyelonephritis	0/15	15/15	-
Biliary disease	9/9	0/9	-
Diverticulitis	0/6	6/6	-
Appendicitis	0/3	3/3	-
Other	4/7	3/7	-
Non-ACR diagnosis			
Gastroenteritis	0/26	26/26	-
Lower UTI	0/13	13/13	-
Dyspepsia	0/4	4/4	-
Other	3/8	5/8	-
Final report			
Normal	8/18 (44.4)	70/125 (56.0)	.3681
Compatible w/ initial dx	7/18 (38.9)	18/125 (14.4)	.0093*
Alternative dx	1/18 (5.6)	3/125 (2.4)	.4237
Inconclusive	2/18 (11.1)	34/125 (27.2)	.1336
Secondary imaging modality	0 (0.0)	25/125 (20.0)	.043†

ACR, American College of Radiology; UTI, urinary tract infection; dx, diagnosis.
*Statistically significant at the level of $P = 0.0125$.
†Statistically significant at the level of $P = 0.05$

infection ($n = 13$, 9.1%) were the most common reasons for US not contemplated by the ACR criteria. Appropriate orders were significantly more likely to demonstrate findings compatible with the initial diagnosis (38.9 vs 14.4%, $P = 0.0093$) (Fig 1). An additional imaging order with a secondary modality was requested in 20% of the inappropriate US orders and none of the appropriate ($P = 0.043$). There were no significant differences in the rates of normal, alternative, and inconclusive reports between the 2 groups. A more comprehensive table with all diagnoses leading to US and CT orders is available in Supplementary Table 1.

Discussion

This study of ED ordered CT and US exams shows that the rate of inappropriate orders for abdominal complaints was 36.3% for CT scans and 84.4% of US. Appropriate exams were significantly more likely to report findings compatible with the initial clinical diagnosis for both CT and US. This highlights the high impact that correct exam selection has on finding confirmative or actionable results on imaging. Inappropriate CT scans were more likely to show no abnormalities, suggest an alternative diagnosis, or be considered inconclusive. Furthermore, inappropriate US orders resulted in an additional imaging modality being performed in 20% of cases, which exposes the patient to the risks of unnecessary exams, extension of ED length of stay, and costs. Although there was a statistically significant difference between age of patients undergoing inappropriate imaging compared to the appropriate group, we believe the divergence was small (mean age 46.8 vs 51.3 years, respectively) and has no clinical significance.

Using the ACR Appropriateness Criteria as our reference, the most frequent inappropriate uses of CT were evaluation of biliary disease, pancreatitis, renal failure, and uncomplicated pyelonephritis. Other less common inappropriate orders were related to the lack of intravenous (IV) contrast when it is usually indicated (acute abdominal pain, SBO, diverticulitis, appendicitis), or using IV contrast when it is not indicated (eg, urolithiasis). Inappropriate US orders were more commonly requested for patients with acute abdominal pain, uncomplicated pyelonephritis, diverticulitis, and appendicitis. Except for uncomplicated pyelonephritis, where no imaging testing is required, CT is the most appropriate test for all of the remaining.[6]

A fraction of the imaging orders (24.4%) included in this study were not clearly defined by the ACR Appropriateness Criteria. Among these orders, 89.7% were considered inappropriate, which demonstrates that the ACR guidelines are very comprehensive and that abdominal complaints not listed in the ACR guidelines usually do not require an imaging test. The proportion of inappropriate US orders were considerably high in our study. There are several possible explanations for this observation. First, this modality has low cost and no exposure to ionizing radiation; therefore, ordering physicians may be less diligent requesting US compared to CT, which would yield a higher rate of inappropriate orders for the former. The other explanation would be the profile of clinical indications that led patients to this imaging modality. There was only 1 study in the literature that investigated the clinical indications leading to US orders in patients with nontraumatic acute abdominal pain in the ED.[10] Although the authors did not analyze the appropriateness of these orders, according to the clinical indications provided in the study, the rate of inappropriate orders would probably lay between 55% and 60% (compared to an 84% rate in our study).[10] In this study by Raman and colleagues, biliary pathology and acute pancreatitis, both of which being appropriate indications for abdominal US, were 2 of the top indications for US referral. On the other hand, our study had mostly nonspecific acute abdominal pain, gastroenteritis, uncomplicated pyelonephritis, and lower urinary tract infection, as clinical indications for US, which were all deemed inappropriate. This raises concern that in addition to work on reminding ordering providers of

the guidelines for the indication of abdominal US, greater attention may be needed to teach specific modality limitations.

Also, there was a considerable number of CT and US orders for patients with acute gastroenteritis. Imaging for initial acute gastroenteritis was considered inappropriate in our analysis.[11] Most patients presented with a typical history of acute new diarrheal disease, nausea or vomiting, and abdominal pain with or without fever. Virtually all patients with gastroenteritis who underwent CT or US had normal or inconclusive results due to unspecific findings, such as bowel wall thickening or distension. Only 1 patient had an "alternative diagnosis" due to the imaging findings suggestive of cholelithiasis on US, which most likely could be considered an incidental finding due to nonacuity. Therefore, imaging for new onset gastroenteritis in the ED is unlikely to be beneficial.

Multiple evidence-based guidelines exist to direct the appropriate use of medical imaging and prevent overuse. Although guidelines are not perfect, their implementation may help diminish defensive ordering of imaging and improve quality, cost, and patient outcomes.[12] Several studies prior to ours performed in the emergency care setting have shown poor adherence to standard-of-care imaging recommendations. Martins et al. reported a 23.8% rate of inappropriate CTs and US (for all causes) in the ED, using the ACR guidelines as standard-of-care.[7] In that study, only 33.9% of all tests performed showed relevant findings. Similarly, overuse of head CT for minor head injury is also well reported in the literature, with inappropriate rates ranging from 30%-70% of cases.[12-14,13] Our study was the first to our knowledge to investigate imaging overuse exclusively for abdominal complaints, which are one of the main reasons for ED visits and deserves further evaluation and mediating action.

The reasons behind the high number of inappropriate imaging orders in the ED are many and multifactorial. Concerns of missing a low-probability diagnosis and malpractice deserve acknowledgment as probably leading reasons for this behavior.[2,15,16] Other factors include availability (convenience of performing imaging exams), volume of patients and workload, and educational gaps in understanding modality limitations. Although there is wide and ready dissemination of ACR tools, the lack of awareness of existing guidelines remains a major problem.[8,17,18] This results not only in imaging overuse but also in wrong modalities being requested, leading to additional imaging orders during the same visit. Considering almost all imaging methods involve some risk or exposure to radiation, it is essential that patients receive the most appropriate modality the first time and every time. Lack of adherence to recommendations despite knowledge of the guidelines is also a significant issue. In the primary care setting, physicians followed decision-support advice for inappropriate imaging orders in only 25% of cases. Thus, there is a long run from education to implementation of evidence-based guidelines.

This study has some limitations in addition to those inherent to its retrospective design. The generalizability of the results may be limited as it was performed at a single institution. There was only 1 pregnant woman in our sample, and therefore our findings may not be generalizable to this population. The sample was powered for the prevalence of inappropriate exams, and therefore some subgroup analysis of the final report may not be powered to show significant differences. The appropriateness of some of the orders (24%) were not contemplated by the evidence-based guidelines but rather based on the consensus of 2 experienced radiologists.

In summary, 36.3% of CT and 84.4% of US orders for abdominal complaints in the ED were not considered appropriate. Appropriate exams were more likely to yield findings compatible with the initial clinical diagnosis for both CT and US.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgments

Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior – Brazil (CAPES)

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1067/j.cpradiol.2020.11.004.

References

1. Hendee WR, Becker GJ, Borgstede JP, et al. Addressing Overutilization in Medical Imaging. Radiology 2010;257:240–5.
2. Kanzaria HK, Hoffman JR, Probst MA, et al. Emergency physician perceptions of medically unnecessary advanced diagnostic imaging. Acad Emerg Med 2015:390–8.
3. Rui P, Kang K. National Hospital Ambulatory Medical Care Survey: 2017 emergency department summary tables. National Center for Health Statistics. Available from: https://www.cdc.gov/nchs/data/nhamcs/web_tables/2017_ed_web_tables-508.pdf.
4. Caporale N, Morselli-Labate AM, Nardi E, et al. Acute abdominal pain in the emergency department of a university hospital in Italy. United Eur Gastroenterol J 2016;4:297–304.
5. Thijssen WAMH, van Miero E, Willekens M, et al. Complaints and diagnoses of emergency department patients in the Netherlands: a comparative study of integrated primary and emergency care. PLoS One 2015;10:e0129739.
6. American College of Radiology. ACR appropriateness criteria ®. Available at https://www.acr.org/Clinical-Resources/ACR-Appropriateness-Criteria. Assessed in June, 2020.
7. Martins R, Raimundo P, Alves P, et al. Appropriateness of radiology test requests by an emergency department: a retrospective Study. Acta Med Port 2020;33:7.
8. Lehnert BE, Bree RL. Analysis of appropriateness of outpatient CT and MRI referred from primary care clinics at an academic medical center : how critical is the need for improved decision support ? JACR 2010;7:192–7.
9. MacDonald PL, Gardner RC. Type I error rate comparisons of post hoc procedures for I j Chi-Square tables. Educ Psychol Meas 2000;60:735–54.
10. Raman S, Somasekar K, Winter R, Lewis M. Are we overusing ultrasound in non-traumatic acute abdominal pain? Postgrad Med J 2004;80:177–9.
11. Cartwright SL, Knudson MP. Evaluation of acute abdominal pain in adults. Am Fam Physician 2008;77:971–8.
12. Cellina M, Panzeri M, Floridi C, et al. Overuse of computed tomography for minor head injury in young patients : an analysis of promoting factors. Radiol Med 2018;123:507–14.
13. Melnick ER, Szlezak CM, Bentley SK, Dziura JD. CT overuse for mild traumatic brain injury. Jt Comm J Qual Patient Saf. The Joint Commission 2012;38:483–9.
14. Klang E, Beytelman A, Greenberg D, et al. Overuse of head CT examinations for the investigation of minor head trauma: analysis of contributing factors. J Am Coll Radiol 2017;14:171–6.
15. Kuhn GJ, Wyer PC, Cordell WH, Rowe BH. A survey to determine the prevalence and characteristics of training in Evidence-Based Medicine in emergency medicine residency programs. J Emerg Med 2005;28:353–9.
16. Baskerville JR, Herrick J. Head multidetector computed tomography: emergency medicine physicians overestimate the pretest probability and legal risk of significant findings. Am J Emerg Med 2012;30:367–70.
17. Bautista AB, Burgos A, Nickel BJ, Yoon JJ, Tilara AA, Amorosa JK. Do clinicians use the American College of Radiology Appropriateness Criteria in the management of their patients? Am J Roentgenol 2009;192:1581–5.
18. Curry L, Reed MH. Electronic decision support for diagnostic imaging in a primary care setting. J Am Med Informatics Assoc 2011;18:267–70.