The American Journal of Surgery 213 (2017) 498–501

Contents lists available at ScienceDirect

The American Journal of Surgery

journal homepage: www.americanjournalofsurgery.com

Assessing surgical residents' imaging interpretation skills

Joseph J. Eid [a,*], Francisco Igor B. Macedo [a], Edsa Negussie [b], Vijay K. Mittal [a]

[a] Department of Surgery, Providence Hospital and Medical Centers, Michigan State University College of Human Medicine, Southfield, MI, USA
[b] Department of Radiology, Providence Hospital and Medical Centers, Michigan State University College of Human Medicine, Southfield, MI, USA

ARTICLE INFO

Article history:
Received 26 July 2016
Received in revised form
10 November 2016
Accepted 18 November 2016
Podium Presentation at the Midwest Surgical Association 2016 Annual Meeting, Mackinanc Island, MI, USA.

Keywords:
Radiology interpretation
General surgery residency
Imaging interpretation
Surgical education

ABSTRACT

Purpose: During general surgery (GS) training, residents are expected to accurately interpret radiologic images. Objective evidence evaluating residents' ability to provide accurate interpretation of imaging studies is currently lacking.
Methods: A 30-item web-based test was developed using images from different radiologic modalities. Residents from 6 ACGME accredited GS programs participated. Residents from 1 radiology program served as a control group.
Results: 74 GS residents (GSR) enrolled in the online test. The mean score for GSR was 75% (±9) and 83% (±6) for RR ($p = 0.006$). Residents correctly answered 63% x-rays, 74%, CT(head), 84% CT(body), 69% ultrasound, and 88% tube/line localization questions. Senior residents were more proficient than junior residents at interpreting CT (body) and ultrasound images.
Conclusion: GS residents were able to accurately interpret 75% of basic radiology images. In an effort to improve patient care, programs should consider integrating radiological education during surgical training.

1. Introduction

As surgeons-in-training, residents are expected to use multiple modalities to accurately diagnose surgical diseases. With technological advancement and the integration of digitally enhanced and reconstructed images, radiology is now an essential tool used to manage various pathologies. In general surgery (GS) specifically, resident and attending physicians have incorporated radiographic analysis as part of patients' perioperative care. Such skills have also become an integral part of critical and trauma care where surgeons need to provide rapid interventions. Concurrently, the American Board of Surgery has emphasized the importance of training surgeons to critically assess, interpret, and integrate clinical imaging in their patient care.[1] Despite the lack of formalized curricula during training, GS programs have met the requirements for board certification by providing residents radiology education via conferences, online modules, review of patient encounters under the guidance and supervision of attending physicians.

Although there is an ongoing effort by academic programs to provide radiology education to GS, and other specialty residents, variable results have been published when investigators assessed trainees' proficiency in diagnostic imaging evaluation. These investigational studies have limited resident assessment to one radiologic modality, focused clinical scenarios such as trauma radiographs, or to a single postgraduate year level.[2–4] Similar variability in trainees' imaging interpretation accuracy is seen among different specialties including radiology residents.[5–8] Given the discordance in the reported accuracy rates and the ongoing expectation of GSR to provide preliminary interpretation of clinical radiographs, we sought to assess their ability to correctly interpret basic radiology images. We also sought to identify areas of deficiency in surgical education that may ultimately improve patient care.

2. Methods

A 30-item web-based test was created on Proprofs.com® that included five different categories of imaging modalities: (1) x-rays (chest/abdomen), (2) computer tomography of the Head (CT head), (3) CT (body), (4) ultrasound (solid organs/vascular), and (5) invasive lines/tube localizing images. After reviewing a previously established radiology education database, de-identified images that were considered common conditions encountered by GSR's

* Corresponding author. Department of Surgery, Providence Hospital and Medical Centers, Michigan State University College of Human Medicine, 16001 W 9 Mile Rd, Southfield, MI, 48075, USA.
E-mail address: joseph.eid@stjohn.org (J.J. Eid).

http://dx.doi.org/10.1016/j.amjsurg.2016.11.029
0002-9610/

were selected by an expert panel. The panel included clinical radiologists, general surgeons and both GS and radiology program directors. The questions were formatted to include a brief clinical vignette and a corresponding static radiographic image. True/False and multiple choice questions which allowed for single or multiple answers were utilized.

Seven allopathic general surgery programs from the Southeast Michigan Center of Medical Education consortium were invited to participate. Radiology residents (RR) from our institution served as the control group. Resident participation was on voluntary basis. An initial invitation letter and 4 additional reminders were sent to improve participation rates. Institutional Review Board approval was obtained prior to the initiation of the study.

Results were collected between September and November 2015. GS postgraduate year (PGY)1 and PGY2 trainees were considered junior residents (GS-JR) and PGY3, PGY4, and PGY5 trainees were considered senior residents (GS-SR). PGY2/R1 and PGY3/R2 RR were considered RR-JR), and PGY4/R3 and PGY5/R4 were considered senior residents (RR-SR). GSR's competency in interpreting images was established by the percentage of correct answers. Performance among GSR was compared at different levels of training, and overall performance was compared to RR.

2.1. Statistical analysis

Student's t-test and analysis of variance (ANOVA) F-test followed by post-hoc Tukey-Kramer test were used to analyze mean differences in resident accuracy rates between programs, different levels of training, and specialties. Categorical proportions were analyzed using Chi-square Test (χ^2) or Fisher's exact test as appropriate. Data is presented as mean (±standard deviation) and counts (percent, %). $P < 0.05$ was considered statistically significant. Statistical analysis was performed using Prism© 7.0 software (GraphPad Inc., CA, USA).

3. Results

Of the 173 eligible subjects, 74 (43%) GS residents enrolled from 6 programs of which 3 were community-based with university affiliations, 2 community-based without university affiliations, and 1 university-based. One program was excluded from the analysis due to lack of participation. Examinees were predominantly PGY1 residents (n = 27) compared to PGY2 (n = 14), PGY3 (n = 10), PGY4 (n = 12), and PGY5 (n = 11) residents. Nine (75%) of 12 eligible RR participated. GSR accurately interpreted 75% (±9) of the queried images compared to 83% (±6) for RR (p = 0.006). GS residents from different programs performed similarly (p = 0.9) (Table 1).

Table 1 highlights difference in performance among GSR across various PGY levels (p < 0.0001). This discrepancy in scores is seen on post-hoc analysis when comparing GS PGY1 residents to PGY3, 4, and 5 residents (Fig. 1). GS-SR scored higher than their junior colleagues (p = 0.0007). RR-JR (n = 5) scored 81% (±6) compared to their GS-JR counterparts [72% (±9), p = 0.037]. Similar results were seen for SR-RR (n = 4) who interpreted more radiology images correctly than GS-SR [78% (±6) vs. 86% (±4), p = 0.026]. No difference in mean scores was seen between male [74% (±9)] and female residents [74% (±8), p = 0.88].

GSR required 13 min less to complete the test compared to RR (p = 0.019), and no difference in duration was seen between GS-JR and GS-SR (p = 0.12) (Table 1). A correlation between performance and duration of the quiz was not seen (r = −0.11, p = 0.34). Residents who completed the test during working hours answered 76% (±8) of questions correctly, compared to those that completed the quiz after working hours [71% (±9), p = 0.04].

RR were more competent the interpretation of images on CT (head) and x-rays (chest/abdomen) compared to GSR (Table 1). No difference in performance was seen between GSR and RR in the assessment of CT (body), ultrasound, and invasive line/tube localizing images. Subgroup analysis revealed no difference in accuracy rate between GS-SR and RR in the evaluation of CT (body) [90% (±13) vs. 92% (±10), p = 0.81], however RR were more proficient than GS-JR [79% (±16) vs. 92% (±10), p = 0.019].

The level of training appeared to influence the interpretation of several diagnostic imaging modalities. GS-SR answered 10% more questions correctly involving images on CT (body) compared to their JR colleagues (p = 0.001). Similar results were observed for the ultrasound modality, where GS-SR were more precise than their JR counterparts [65% (±17) vs. 74% (±15), p = 0.03]. Seniority in GS residency did not appear to impact interpretation scores for x-rays, invasive tube/lines localizing, CT (head) images. Level of training in RR did not impact performance on any of the queried modalities.

4. Discussion

GSR as well other physicians continue to find themselves providing preliminary reads of clinical images. Such instances occur

Table 1
Participant demographics and performance.

	RR	p-value	GSR	p-value
Participants N (rate %)	9 (75%)		74 (43%)	
Gender				ns
Male	3 (6%)		44 (94%)	
Female	6 (17%)		30 (83%)	
Duration of exam, mean minutes (±SD)	32 (±31)		19 (±13)	0.019*
Score, mean %(±SD)	83 (±6)		75 (±9)	0.006*
GS programs				ns
Program A			74 (±11)	
Program B			75 (±9)	
Program C			73 (±6)	
Program D			76 (±5)	
Program E			70 (±3)	
Program F			76 (±7)	
Modalities				
X-rays (Chest/Abdomen)	77 (±12)		63 (±14)	0.019*
CT (Head)	100 (±0)		74 (±22)	0.001*
CT (Body)	92 (±10)		84 (±15)	ns
Ultrasound	69 (±16)		69 (±15)	ns
Invasive Lines/Tubes	90 (±15)		88 (±13)	ns
GS: PGY-Level				<0.0001
PGY1			70 (±10)	
PGY2			75 (±6)	
PGY3			80 (±7)	
PGY4			77 (±3)	
PGY5			79 (±7)	

	RR-JR	RR-SR	p-value	GS-JR	GS-SR	p-value
Participants, N (%)	5 (55%)	4 (45%)	ns	41 (55%)	33 (45%)	ns
Duration of exam, mean minutes (±SD)	37 (±35)	26 (±27)	ns	21 (±14)	17 (±11)	ns
Score, mean %(±SD)	79 (±6)	86 (±4)	ns	72 (±9)	78 (±6)	0.0007*
X-rays (Chest/Abdomen)	73 (±15)	81 (±6)	ns	60 (±15)	67 (±12)	ns
CT (Head)	100 (±0)	100 (±0)	ns	72 (±14)	76 (±23)	ns
CT (Body)	89 (±12)	96 (±7)	ns	79 (±16)	90 (±13)	0.001*
Ultrasound	70 (±7)	67 (±24)	ns	65 (±19)	74 (±15)	0.03*
Invasive Lines/Tubes	84 (±17)	95 (±10)	ns	89 (±14)	90 (±12)	ns

*p < 0.05: statistically significant; ns: not statistically significant.