



Peer Review of Teleradiology at a Teleradiology Clinic: Comparison of Unacceptable Diagnosis and Clinically Significant Discrepancy between Radiology Sections and Imaging Modalities

원격판독 의료기관에서 시행한 원격판독의 동료평가:
세부전공분야와 촬영장비에 따른 판독 불일치 정도와
불일치 판독의 임상적 의의 비교

Hyung Suk Seo, MD¹ , Jai Soung Park, MD^{2*} , Yu-Whan Oh, MD³ ,
Dongwook Sung, MD⁴ , A Leum Lee, MD²

¹Department of Radiology, Eulji University Uijeongbu Hospital, Uijeongbu, Korea

²Department of Radiology, Soonchunhyang University Bucheon Hospital, Bucheon, Korea

³Department of Radiology, Korea University Anam Hospital, Seoul, Korea

⁴Department of Radiology, Kyung Hee University Hospital, Seoul, Korea

Purpose The purpose of this study was to evaluate the rates of unacceptable diagnosis and clinically significant diagnostic discrepancy in radiology sections and imaging modalities through a peer review of teleradiology.

Materials and Methods Teleradiology peer reviews in a Korean teleradiology clinic in 2018 and 2019 were included. The peer review scores were classified as acceptable and unacceptable diagnoses and clinically insignificant and significant diagnostic discrepancy. The diagnostic discrepancy rates and clinical significance were compared among radiology sections and imaging modalities using the chi-square test.

Results Of 1312 peer reviews, 117 (8.9%) cases had unacceptable diagnoses. Of 462 diagnostic discrepancies, the clinically significant discrepancy was observed in 104 (21.6%) cases. In radiology sections, the unacceptable diagnosis was highest in the musculoskeletal section (21.4%) ($p < 0.05$), followed by the abdominal section (7.3%) and neuro section (1.3%) ($p < 0.05$). The proportion of significant discrepancy was higher in the chest section (32.7%) than in the musculoskeletal (19.5%) and abdominal sections (17.1%) ($p < 0.05$). Regarding modalities, the number

Received November 3, 2020

Revised January 18, 2021

Accepted February 3, 2021

*Corresponding author

Jai Soung Park, MD
Department of Radiology,
Soonchunhyang University
Bucheon Hospital,
170 Jomaru-ro,
Bucheon 14584, Korea.

Tel 82-2-2267-4912

Fax 82-2-2269-5164

E-mail jspark@schmc.ac.kr

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>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ORCID iDs

Hyung Suk Seo
<https://orcid.org/0000-0003-2303-7415>
Jai Soung Park
<https://orcid.org/0000-0003-0382-0605>
Yu-Whan Oh
<https://orcid.org/0000-0003-2646-0497>
Dongwook Sung
<https://orcid.org/0000-0001-8598-6670>
A Leum Lee
<https://orcid.org/0000-0001-7662-4264>

of unacceptable diagnoses was higher with MRI (16.2%) than plain radiology (7.8%) ($p < 0.05$). There was no significant difference in significant discrepancy.

Conclusion Peer review provides the rates of unacceptable diagnosis and clinically significant discrepancy in teleradiology. These rates also differ with subspecialty and modality.

Index terms Teleradiology; Peer Review; Diagnostic Errors; Diagnostic Equipment

서론

21세기로 접어들면서 사회적으로 인터넷의 기술과 보급이 빠르게 발달하고, 의료영상은 Picture Archiving and Communication System을 이용한 디지털화가 이루어져 원격판독에 대한 기술적 인 조건은 성립되었다. 의료법 34조에 ‘의료인은 컴퓨터와 화상통신 등 정보통신 기술을 활용하여 먼 곳에 있는 의료인에게 의료지식이나 기술을 지원하는 원격의료를 할 수 있다.’고 규정되어 있고, 2011년 2월 복지부가 국회에 전달한 의견서에서 “원격판독은 영상의학과 전문의가 의료지식을 활용하여 환자와 대면하는 의사의 최종진단 판정을 지원해 주는 것으로 볼 때 의료법에 위반된다고 보기 어렵다”고 밝힘으로써 의료영상 원격판독의 합법성에 대한 근거도 마련되었다(1).

한국에서 원격판독은 2008년에 시작되어 현재 대표적인 원격판독 의료기관은 약 30여 개 정도로 파악되고, 소규모 기관까지 포함하면 100여 개까지 추정되기도 한다(2, 3). 2017년에 국내 원격판독은 300–400억 원 매출 규모를 나타내었고 그 규모는 점점 커지고 있다(4). 그러나 원격판독 의뢰기관에서는 원활하지 않은 의사소통과 판독의 오류 등으로 인해 원격판독의 정확성을 문제점으로 제기하였다(2).

현재 한국에서 널리 이루어지는 원격판독의 정확도에 대한 객관적 측정과 분석이 있다면 원격판독과 관련된 정책 방향을 설정하는데 도움을 줄 수 있을 것이다. 그러나 지금까지 원격판독의 정확도를 체계적으로 분석한 연구는 없었다. 이 연구는 국내의 한 원격판독 의료기관에서 자체적으로 시행한 체계적 동료평가를 정리하고 분석하고, 이를 통해서 영상의학 세부전공분야와 촬영 장비별로 원격판독의 부적격 진단율과 임상적으로 유의한 불일치율을 평가하여 현재 이루어지고 있는 원격판독의 정확도를 알아보는 것을 목적으로 하였다.

대상과 방법

연구 대상

2018년 1월부터 2019년 12월까지 국내 한 원격판독 의료기관에서 원격판독을 시행했던 영상 중에서 무작위로 선택하여 분기당 1회씩 총 7회의 동료평가(peer review)를 자체적으로 시행하였다. 신경두경부, 흉부, 복부, 근골격으로 나누어서 네 개의 세부전공분야에서 시행하였고, 각 세부전공분야의 평가 영상은 일반촬영, CT 영상, MR 영상으로 구성되었다. 다만, 첫 회에 평가된 신경두경부 분야의 결과를 분실하여 연구 대상에서 제외하였다. 또 흉부 분야는 MR 영상을 포함하지 않았다.

전체적으로 일반촬영 816건, CT 영상 397건의 MR 영상 99건에 대해서 동료평가를 시행하여 총 1312건의 영상을 평가하였다. 이는 해당 기간에 원격판독 의료기관에서 이루어진 일반촬영영상 991235건의 0.08%, CT 영상 50432건의 0.79%, MR 영상 13025건의 0.76%에서 동료평가를 시행한 것이다. 전체적으로 원격판독 영상 1054692건의 0.12%를 대상으로 동료평가를 시행하였다.

각 영상의학 세부전공분야에서 촬영장비별 동료평가 영상 숫자는 달랐다. 신경두경부는 일반촬영 영상 185건, CT 영상 84건, 그리고 MR 영상 34건에서 동료평가를 시행하였다. 흉부는 일반촬영 영상 211건, CT 영상 139건, 그리고 MR 영상 0건에서 시행하였다. 복부는 일반촬영영상 200건, CT 영상 104건, 그리고 MR 영상 24건에서 시행하였다. 근골격은 일반촬영영상 220건, CT 영상 70건, 그리고 MR 영상 47건에서 시행하였다.

동료평가 방법

네 세부전공분야에서 15년 이상 3차 의료기관에서 근무한 경력을 가진 각 분야전공의 영상의학과 전문의 한 명이 2016 revised RADPEER scoring system (Table 1)를 이용하여 동료평가를 시행하였다(5). 동료평가를 시행할 때에 각 세부전공분야 평가 전문의들에게 환자의 임상정보나 기존 영상 검사는 제공되지 않았다. 매 분기당 각 분야의 평가자들이 모여 동료평가 결과를 발표하고 토론을 거쳐 평가점수를 조정하였다.

통계 분석을 위해 RADPEER 1점과 2점을 '적격 진단(acceptable diagnosis)'으로, RADPEER 3점을 '부적격 진단(unacceptable diagnosis)'으로 분류하여 평가하였다. 또 RADPEER 2a와 3a를 '임상적으로 무의한 불일치(clinically insignificant diagnostic discrepancy)'로, 2b와 3b를 '임상적으로 유의한 불일치(clinically significant diagnostic discrepancy)'로 분류하여 임상적 의의가 있는 판독 불일치의 빈도를 분석하였다. RADPEER 3b점으로 평가된 경우에는 즉시 판독을 의뢰한 해당 의료기관에 통보하여 수정된 판독 의견을 전달하였다.

통계 분석

동료평가에 대한 통계는 카이제곱 검증을 이용하여 시행하였다. 세부전공분야와 촬영장비에 따른 판독 불일치율과 임상적으로 유의한 판독 불일치율의 차이는 p 값 < 0.05 를 기준으로 통계적 의의를 평가하였다. 전체 영상에서 통계적 의의가 있는 경우에는 Bonferroni correction을 적용하여 각 세부전공분야와 촬영장비 간의 다중 비교를 시행하였다. 따라서 각각의 세부전공분야에 대한 차이는 p 값 < 0.0083 ($= 0.05/6$)을 기준으로 통계적 의의를 평가하였고, 각각의 장비에 대한

Table 1. RADPEER Scoring System (Effective May 2016)

Score	Meaning	Optional
1	Concur with interpretation	
2	Discrepancy in interpretation/not ordinarily expected to be made (understandable miss)	Unlikely to be clinically significant Likely to be clinically significant
3	Discrepancy in interpretation/should be made most of the time	Unlikely to be clinically significant Likely to be clinically significant

차이는 p 값 < 0.017 ($= 0.05/3$)을 기준으로 통계적 의의를 평가하였다.

결과

판독 불일치

전체 동료평가에서 RADPEER score 1점의 판독 일치는 830건으로 63.3%였고, RADPEER score 2점과 3점을 합친 판독 불일치는 482건으로 36.7%를 차지하였다. 그러나 적격과 부적격 진단으로 나누었을 때 적격 진단은 1195건으로 91.1%였고, 부적격 진단은 117건으로 8.9%였다. 세부전공분야에서 부적격 진단은 근골격(21.4%), 복부(7.3%), 흉부(4.9%), 신경두경부(1.3%) 순서로 높았고(Table 2, Fig. 1), 통계적으로 의미 있는 차이를 보였다($p < 0.001$).

각각의 세부전공분야를 비교하였을 때에 근골격 분야의 부적격 진단율은 다른 세부전공분야에 비해 의미 있게 높았다(각각 $p < 0.001$). 복부의 부적격 진단율이 신경두경부에 비해 높았다($p < 0.001$).

촬영장비 간에 부적격 진단율을 비교하였을 때에 MR 영상(16.2%), CT 영상(9.3%), 일반촬영(7.8%) 순서로 부적격 진단율이 높게 나타났다(Table 2). 촬영장비별 부적격 진단율은 의미 있는 통계적 차이를 나타냈다($p = 0.022$). 각각의 촬영장비의 부적격 진단율을 비교하였을 때 MR 영상이 일반촬영에 비해 부적격 진단율이 높았고($p = 0.006$). 나머지 촬영장비 간에는 의미 있는 차이가 없었다.

임상적으로 유의한 판독 불일치

전체 동료평가 1312건 중에서 임상적으로 유의한 불일치는 104건으로 7.9%를 차지하였다. 또 동료평가에서 판독 불일치였던 482건 중에서 21.6%가 임상적으로 유의한 불일치였다. RADPEER score 2점인 365건 중에서 임상적으로 무의한 2a (Fig. 2)가 311건, 유의한 2b (Fig. 3)가 54

Table 2. Acceptable and Unacceptable Diagnosis in Teleradiology

Sub-Specialty	Acceptance of Diagnosis	CR (%)	CT (%)	MR (%)	Total (%)
Neuro	Acceptable*	183 (98.9)	83 (98.8)	33 (97.1)	299 (98.7)
	Unacceptable†	2 (1.1)	1 (1.2)	1 (2.9)	4 (1.3)
Chest	Acceptable	204 (96.7)	129 (92.8)		333 (95.1)
	Unacceptable	7 (3.3)	10 (7.2)		17 (4.9)
Abdomen	Acceptable	192 (96.0)	89 (85.6)	23 (95.8)	304 (92.7)
	Unacceptable	8 (4.0)	15 (14.4)	1 (4.2)	24 (7.3)
Musculoskeletal	Acceptable	173 (78.6)	59 (84.3)	33 (70.2)	265 (78.6)
	Unacceptable	47 (21.4)	11 (15.7)	14 (29.8)	72 (21.4)
Total	Acceptable	752 (92.2)	360 (90.7)	83 (83.8)	1195 (91.1)
	Unacceptable	64 (7.8)	37 (9.3)	16 (16.2)	117 (8.9)

*Acceptable diagnosis: RADPEER score 1 & 2.

†Unacceptable diagnosis: RADPEER score 3.

CR = computed radiography

Fig. 1. Results of teleradiology peer review between radiology sections and imaging modalities.

The lower brown portion in the graph shows RADPEER 1, which means 'concurs with interpretation'. The middle light brown portion in the graph shows RADPEER 2, which means 'discrepancy in interpretation/not ordinarily expected to be made (understandable miss)'. The upper white portion shows RADPEER 3, which means 'discrepancy in interpretation/should be made most of the time'.

CR = computed radiography

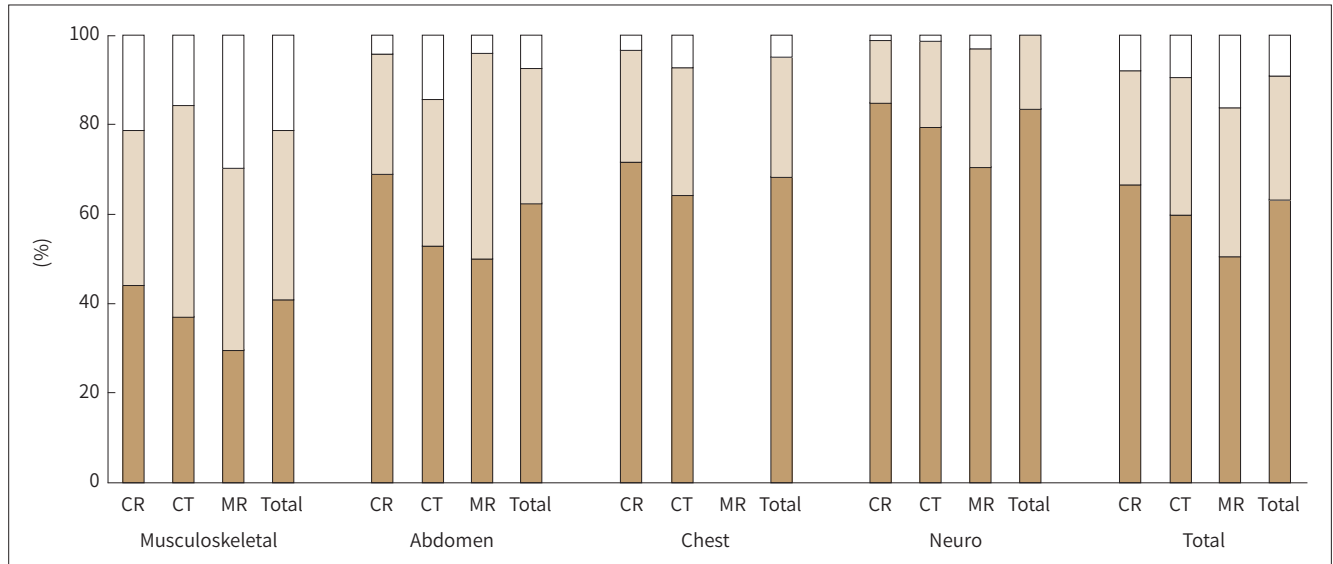


Fig. 2. Representative case of RADPEER 2a.

Radiologist observed fecal material in the large bowel, suggestive of constipation on plain abdominal radiography. The patient was diagnosed with mild paralytic ileus on peer review.



건이었다. RADPEER score 3점인 117건 중에서 임상적으로 무의한 3a (Fig. 4)가 67건, 유의한 3b (Fig. 5)가 50건이었다.

네 세부전공분야에서 임상적으로 유의한 불일치율은 흉부(32.7%), 신경두경부(25.5%), 근골격 (19.5%), 복부(17.1%) 순으로 높았다(Table 3). 전체 영상에서 임상적으로 유의한 불일치율은 의

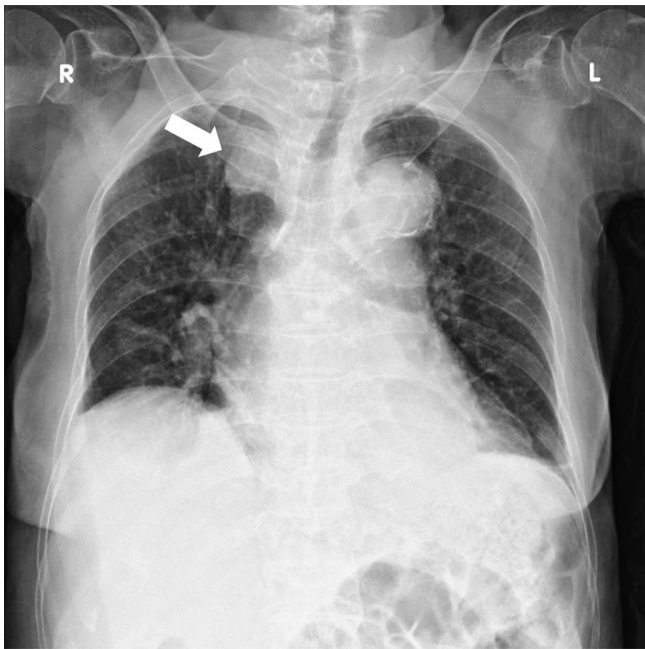


Fig. 3. Representative case of RAD-PEER 2b.

Radiologist observed nodule or mass in the right apical lung (arrow) on chest plain radiography. This finding was interpreted as the buckling of innominate vessels on peer review.

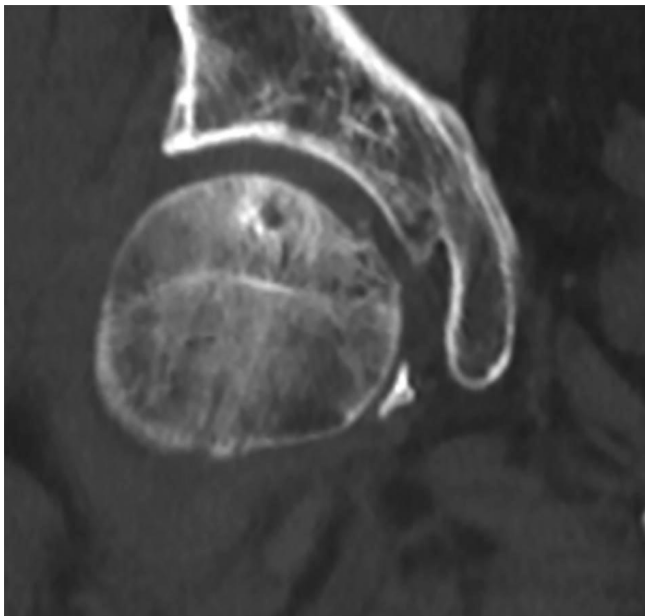


Fig. 4. Representative case of RAD-PEER 3a.

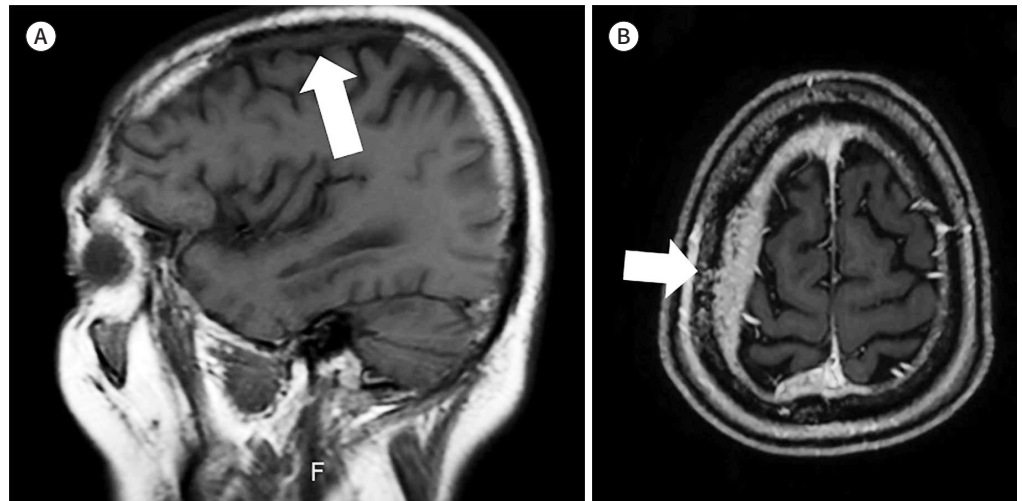
Radiologist observed mild osteoarthritis on hip joint CT. Hip joint CT scans performed by a radiologist revealing mild osteoarthritis. It was interpreted as avascular necrosis of the right femoral head on peer review.

미 있는 통계적 차이를 나타내었다($p=0.036$). 각 세부전공분야별로 비교하였을 때에 흉부에서 유의한 불일치율이 복부($p=0.006$)와 근골격($p=0.009$) 분야에 비해 높게 나타났고, 다른 분야 간에는 유의한 통계적 차이가 없었다.

촬영장비에 따른 임상적으로 유의한 불일치율은 CT 영상(26.9%), 일반촬영(19.4%), MR 영상(16.3%) 순서로 높았다(Table 3). 그러나 촬영장비 간에 의미 있는 차이는 나타나지 않았다($p=0.122$).

Fig. 5. Representative case of RADPEER 3b.

A, B. The radiologist observed meningeal thickening on brain MRI in a lung cancer patient. On peer review: T1 hypointensity (arrow) (**A**) and T1 enhancement (arrow) (**B**) in the right parietal lobe were interpreted as suggestive of metastasis.

**Table 3.** Clinically Insignificant and Significant Diagnostic Discrepancies in Teleradiology

Sub-Specialty	Clinical Significance	CR (%)	CT (%)	MR (%)	Total (%)
Neuro	Insignificant*	22 (78.6)	13 (76.5)	6 (60.0)	41 (74.5)
	Significant†	6 (21.4)	4 (23.5)	4 (40.0)	14 (25.5)
Chest	Insignificant	38 (63.3)	36 (72.0)		74 (67.3)
	Significant	22 (36.7)	14 (28.0)		36 (32.7)
Abdomen	Insignificant	59 (95.2)	34 (69.4)	9 (75.0)	102 (82.9)
	Significant	3 (4.8)	15 (30.6)	3 (25.0)	21 (17.1)
Musculoskeletal	Insignificant	101 (82.1)	34 (77.3)	26 (78.8)	161 (80.5)
	Significant	22 (17.9)	10 (22.7)	7 (21.2)	39 (19.5)
Total	Insignificant	220 (80.6)	117 (73.1)	41 (83.7)	378 (78.4)
	Significant	53 (19.4)	43 (26.9)	8 (16.3)	104 (21.6)

*Clinically insignificant diagnostic discrepancy: RADPEER score: 2a & 3a.

†Clinically significant diagnostic discrepancy: RADPEER score: 2b & 3b.

CR = computed radiography

고찰

이 연구는 우리나라의 원격판독 기관에서 자체적으로 시행한 첫 번째 동료평가로서 원격판독의 정확도를 평가했다는 점에서 의의를 갖는다. 또 원격판독의 불일치를 세부전공분야와 촬영장비별로 분석하여 특성을 파악하였다. 이 연구의 결과에 의하면 RADPEER 1점인 판독에 다른 의견 없이 동의하는 비율(RADPEER 1)은 63.3%에 불과하였다. 그러나 RADPEER 1점과 2점을 포함한 적격 진단의 비율은 91.1%로 증가하였다. 세부전공분야별 비교 결과에서 근골격 분야는 부적격 진단율이 가장 높았고, 다른 분야들과 통계적 의미 있는 차이를 나타냈다. 그러나 임상적으로 유의한 불일치율은 흉부 분야에 가장 높았고 복부나 근골격 분야와 비교해 통계적으로 의미 있는 차

이가 있었다. 이번 연구에서 부적격 진단율이 높은 분야와 임상적으로 유의한 불일치율이 높은 분야가 서로 달랐다. 즉 부적격 진단의 빈도가 높은 복부와 근골격 분야는 임상적으로 유의한 불일치율이 낮았다. 반면에 부적격 진단의 빈도가 낮은 신경두경부와 흉부 분야는 임상적으로 유의한 불일치율이 높게 나타났다. 세부전공분야별로 부적격 진단의 빈도나 임상적 의의를 비교한 이전 연구는 찾을 수 없었다.

촬영장비별 비교에서는 MR 영상에서 부적격 진단율이 일반촬영에 비해 높게 나타났으나 CT 영상은 다른 촬영장비와 부적격 진단율에서 차이가 없었다. 이 결과는 CT 영상에서 임상적으로 유의한 불일치율이 높았던 이전 연구 결과와 차이를 나타내고 있다(6, 7). MR 영상과 다른 촬영장비영상의 동료평가 결과를 비교한 이전 연구는 없었다. MR 영상에는 영상의 개수가 많고, 병원별로 조건을 달리하는 다양한 시퀀스 영상이 포함되어 있다. 이러한 이유로 MR 영상에서 부적격 진단율이 높게 나타난 것으로 생각된다(8). 촬영장비별 비교에서 임상적으로 유의한 불일치율의 의미 있는 차이는 없었다.

의료 영상의 디지털화와 초고속 정보통신망의 보급이라는 기술의 발전을 통해서 영상학과 의사의 판독이나 자문을 시간과 공간의 제약 없이 받을 수 있게 되면서 원격판독의 이용은 급격히 늘어나고 있다. 그러나 원격판독의 정확성에 대한 의문이 제기되고 있다. 원격판독에 대한 동료평가를 시행하는 원격판독 의뢰기관이 있었으나 동료평가 결과는 공개되지 않았다(2).

이 연구 결과 중 전체 영상에서 나타난 36.7%의 판독 불일치율은 야간 당직 전공의와 전문의를 비교한 이전 연구의 7.8%에 비교해 상대적으로 매우 높다(6). 그러나 이전 전공의와 전문의의 비교 연구에서 비교한 영상은 대부분 일반촬영영상이었고 CT와 MR 영상이 26.1%를 이루고 있었다. 이 연구에서는 CT와 MR 영상이 전체 영상의 37.8%였던 것이 불일치율이 높은 이유 중의 하나일 것이다. 또 이전 연구는 응급실에서 촬영한 영상에 국한되어 있고 환자의 임상 정보와 이전 영상을 판독에 활용할 수 있는 반면에 이 연구에서는 다양한 환경에서 촬영한 영상이 연구 대상에 포함되고 환자의 임상 정보와 이전 영상의 이용이 제한적이었다는 차이가 있다.

최근에는 단순히 판독 불일치 여부를 판단하기보다 임상적 의의를 고려한 판독 불일치를 평가하여 교육적 효과 등에 활용한다(6). 이전 RADPEER score는 1-4점으로 나누어져 있었고 불일치 판독의 임상적 의의에 대한 분류가 없었다(7). 그러나 지금 사용되고 있는 2016 revised RADPEER scoring system은 ‘거의 항상 진단되어야 할 소견의 오류’인 4점을 제거하고, 1-3점으로 분류하였다 그뿐만 아니라 임상적 의의 유무에 대한 분류가 추가되어 임상적 의의를 중시하는 경향이 반영되었다(Table 1) (5). 이 연구에서는 RADPEER score 2점과 3점을 합친 판독 불일치로 평가된 원격판독 중에서 21.6%가 임상적으로 유의한 불일치였다.

원격판독의 영상의 질이나 보안에 대한 문제를 해결하기 위해서 미국은 원격판독의 가이드라인을 1994년에 처음 제정하였고, 최근까지 여러 차례 개정을 시행하였다(9). American College of Radiology에서는 원격판독자나 환경에 대한 기술적 가이드라인을 일찍부터 마련하였다(10, 11). 국내에서는 2005년에 원격판독 표준화의 가이드라인을 제안하였고 영상 전송방식, 압축과 보안, 그리고 영상 입출력 환경을 최적화하기 위해 노력해왔다(12). 그러나 원격판독 의료기관에서는 의뢰받은 촬영 영상의 질을 통제할 수 없고, 환자 정보가 제한되는 경우가 많다. 또 다섯 국내 원격판

독기관의 인터뷰를 통한 연구에서 원격판독을 수행하는 판독의의 90%가 비전속 전문의로, 전속된 기관의 업무 외에 부가적으로 원격판독을 수행하고 있었다(2). 이러한 원격판독의 특성들이 높은 판독 불일치율의 원인으로 추정된다.

영상 진단 오류는 크게 인식(perception)의 오류와 해석(interpretation)의 오류로 나눌 수 있다(13). 이전 연구 결과에 따르면 전체 오류 중에서 해석의 오류가 약 15%–28%를 차지하고(14–16), 나머지 약 2/3를 인식의 오류가 차지한다(17). 영상의 질, 판독 환경, 환자 정보, 판독자의 피로도, 그리고 판독자의 추론 오류 등이 영상 진단 오류를 일으키는 요소로 지적되고 있다(13).

일반적으로 진단 오류를 줄이기 위해서 여러 방안들이 제시되고 있다. 그중에서 동료평가와 구조적 판독은 원격판독의 불일치 또는 진단 오류를 줄이는데에 적용 가능하다(13). 구조적 판독은 병변의 발견율을 높여 인식의 오류를 줄일 수 있지만 해석의 오류를 해결하지는 못한다(18). 동료평가를 통한 피드백은 교육적 효과가 크고, 인식 오류와 해석 오류를 줄이고 진단의 정확성을 높이는 데 도움이 된다(16, 19). 특히 진단 오류를 일으킬 요인을 많이 내포하고 있는 원격판독에서는 적극적으로 동료평가를 시행해서 문제점을 분석하고 진단 오류를 감소시키려는 노력이 필요하다.

이 연구의 제한점은 짧은 연구 기간으로 인해 동료평가에 의한 판독 불일치의 개선 정도를 평가할 수 없었다는 점이다. 또 동료평가 영상이 의뢰된 전체 원격판독 영상 중에서 0.12%로 매우 적어서 판독 불일치의 개선 효과를 나타내기에는 제한적이었을 것으로 생각된다. 따라서 더 많은 영상에 대해 동료평가를 꾸준히 시행한 후 판독 불일치 개선 효과를 측정하는 연구가 필요하다. 또 다른 연구의 제한점은 각 영상의학 세부전공분야 별로 한 명의 전문가가 독립적으로 평가하는 방법을 택하여 평가 결과의 충분한 보편성을 얻지 못하였다는 것이다. 그러나 각 분야의 평가자들이 모여 서로의 동료평가 결과를 논의하고 조정함으로 객관성을 확보하려고 노력하였다. 마지막 연구의 제한점은 적격 또는 부적격 진단, 그리고 유의한 또는 무의한 불일치 진단의 근거로 객관적 병리나 임상적 진단을 제시하지 못한 것이다.

결론적으로 원격판독의 동료평가에서 판독 불일치의 빈도가 다소 높으나 적격 진단은 90% 이상을 나타내고 있다. 복부와 근골격 분야에서 부적격 진단율이 높은 반면에 신경두경부와 흉부 분야에서 임상적으로 유의한 판독 불일치율이 높았다. 촬영장비별로는 MR에서 부적격 진단율이 높았다. 원격판독 의료기관이 자체적으로 시행하는 동료평가를 통해 판독 불일치를 분석하고 판독 시행한 영상의학과 의사에게 지속적으로 피드백을 주면 원격판독의 정확성을 높일 수 있을 것으로 기대된다.

Author Contributions

Conceptualization, S.H.S., P.J.S., O.Y., S.D.; data curation, S.H.S.; formal analysis, S.H.S.; investigation, S.H.S., O.Y.; methodology, S.H.S.; supervision, P.J.S., S.D.; writing—original draft, S.H.S., L.A.L.; and writing—review & editing, all authors.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

Funding

None

REFERENCES

1. Medical Today. Available at: <http://www.mdtoday.co.kr/mdtoday/index.html?no=150028>. Published 2011. Accessed Oct 10, 2020
2. Woo H, Choi MH, Eo H, Jung SE, Do K, Lee JS, et al. Teleradiology of Korea in 2017: survey and interview of training hospitals and teleradiology center. *J Korean Soc Radiol* 2019;80:490-502
3. Maeil Business News. Available at: <https://mk.co.kr/news/it/view/2018/01/38640/>. Published 2018. Accessed Oct 10, 2020
4. Chosun Biz. Available at: https://biz.chosun.com/site/data/html_dir/2018/02/06/2018020600920.html. Published 2018. Accessed Oct 10, 2020
5. Goldberg-Stein S, Frigini LA, Long S, Metwalli Z, Nguyen XV, Parker M, et al. ACR RADPEER committee white paper with 2016 updates: revised scoring system, new classifications, self-review, and subspecialized reports. *J Am Coll Radiol* 2017;14:1080-1086
6. Maloney E, Lomasney LM, Schomer L. Application of the RADPEER™ scoring language to interpretation discrepancies between diagnostic radiology residents and faculty radiologists. *J Am Coll Radiol* 2012;9:264-269
7. Borgstede JP, Lewis RS, Bhargavan M, Sunshine JH. RADPEER quality assurance program: a multifacility study of interpretive disagreement rates. *J Am Coll Radiol* 2004;1:59-65
8. Mangrum W, Christianson K, Duncan S, Hoang P, Merkle E, Song A. *Duke review of MRI principles*. St. Louis: Mosby 2012
9. American College of Radiology. ACR-AAPM-SIIM technical standard for electronic practice of medical imaging. Revised 2017. Available at: <https://www.acr.org/-/media/ACR/Files/Practice-Parameters/elec-practice-medimag.pdf>. Accessed Dec 11, 2020
10. Silva E 3rd, Breslau J, Barr RM, Liebscher LA, Bohl M, Hoffman T, et al. ACR white paper on teleradiology practice: a report from the Task Force on Teleradiology Practice. *J Am Coll Radiol* 2013;10:575-585
11. Van Moore A, Allen B Jr, Campbell SC, Carlson RA, Dunnick NR, Fletcher TB, et al. Report of the ACR task force on international teleradiology. *J Am Coll Radiol* 2005;2:121-125
12. Ha DH. Standard guideline for teleradiography. *The Korean Journal of PACS* 2005;11:129-133
13. Degnan AJ, Ghobadi EH, Hardy P, Krupinski E, Scali EP, Stratchko L, et al. Perceptual and interpretive error in diagnostic radiology—causes and potential solutions. *Acad Radiol* 2019;26:833-845
14. Donald JJ, Barnard SA. Common patterns in 558 diagnostic radiology errors. *J Med Imaging Radiat Oncol* 2012;56:173-178
15. Kielar AZ, McInnes M, Quan M, O'Sullivan J. Introduction of QUIP (quality information program) as a semi-automated quality assessment endeavor allowing retrospective review of errors in cross-sectional abdominal imaging. *Acad Radiol* 2011;18:1358-1364
16. Mankad K, Hoey ET, Jones JB, Tirukonda P, Smith JT. Radiology errors: are we learning from our mistakes? *Clin Radiol* 2009;64:988-993
17. Kim YW, Mansfield LT. Fool me twice: delayed diagnoses in radiology with emphasis on perpetuated errors. *AJR Am J Roentgenol* 2014;202:465-470
18. Lee CS, Nagy PG, Weaver SJ, Newman-Toker DE. Cognitive and system factors contributing to diagnostic errors in radiology. *AJR Am J Roentgenol* 2013;201:611-617
19. Tudor GR, Finlay DB. Error review: can this improve reporting performance? *Clin Radiol* 2001;56:751-754

원격판독 의료기관에서 시행한 원격판독의 동료평가: 세부전공분야와 촬영장비에 따른 판독 불일치 정도와 불일치 판독의 임상적 의의 비교

서형석¹ · 박재성^{2*} · 오유환³ · 성동욱⁴ · 이아름²

목적 동료평가를 시행하여 원격판독 불일치의 정도와 불일치 판독의 임상적 의의를 세부전공분야와 촬영장비별로 평가하고자 하였다.

대상과 방법 한 국내 원격판독 의료기관에서 2018년과 2019년에 시행한 동료평가를 대상으로 세부전공분야와 촬영장비별로 판독 불일치를 분석하였다. 동료평가 점수를 적격 진단과 부적격 진단으로 분류하였고, 판독 불일치는 임상적 의의에 따라 무의한 불일치와 유의한 불일치로 분류하였다. 카이제곱 검정으로 통계를 시행하였다.

결과 총 1312건의 동료평가 중 부적격 진단은 117건(8.9%)이고, 총 462건의 진단 불일치 중 유의한 불일치는 104건(21.6%)이었다. 세부전공분야에서 부적격 진단율은 근골격(21.4%)이 가장 높았고(각각 $p < 0.05$), 복부(7.3%)는 신경두경부(1.3%)에 비해 높았다($p < 0.05$). 유의한 불일치는 흉부(32.7%)가 복부(17.1%)와 근골격(19.5%)에 비해 높았다(각각 $p < 0.05$). 촬영장비에서 부적격 진단은 MR (16.2%)이 일반촬영(7.8%)에 비해 높았고($p < 0.05$), 유의한 불일치는 차이가 없었다.

결론 동료평가를 통해서 원격판독의 부적격 진단과 임상적 유의한 불일치의 정도, 세부전공분야와 촬영장비에 따른 차이를 알 수 있었다.

¹울지대학교 의정부병원 영상의학과,

²순천향대학교 부천병원 영상의학과,

³고려대학교 안암병원 영상의학과,

⁴경희대학교병원 영상의학과