Double Reading of Outsourced CT/MR Radiology Reports: Retrospective Analysis

Jean-François Vendrell, MD, PhD, Julien Frandon, MD,* Bastien Boussat, MD, PhD,† François Cotton, MD, PhD,‡ Gilbert Ferretti, MD, PhD,§ Nicolas Sans, MD, PhD,|| Jean-Pierre Tasu, MD, PhD,¶ Jean-Paul Beregi, MD, PhD,* and Ahmed Larbi, MD**

Objectives: Our objective was to determine disagreement rates in radiological reports provided by using a double-reading protocol in a national teleradiology company.

Methods: From January 2015 to July 2016, 134169 radiological exams from 36 French centers, benefited outsourced interpretations by certified radiologists, in both regular and after-hours activities. Of these, 2040 CT and MR-scans (1.5%) were subjected to a second opinion by other radiologists in the field of their anatomical specialty (cerebral, thoracic, abdominal-pelvic, and osteoarticular). A five-point agreement scale graded from 0 to 4 was assigned for each exam. Disagreements were considered as minor if no clinical consequence for patient (scores 1 and 2) and major if potential clinical consequence (score 3 and 4). Independent radiologists performed a retrospective analysis and a stratified statistical analysis.

Results: Double reading was performed on CT-scans (n 934/2040, 45.8%) and MR-scans (n 1106/2040, 54.2%) performed in regular (80.1%) and after-hours activities (19.9%). Disagreement scores occurred in 437 exams (21.4%), including major disagreements in 59 (2.9%). Among these, 48/754 were assigned by the thoracic second reader (6.4%), 6/70 by the abdominal-pelvic second reader (8.6%), 3/901 by the osteoarticular second reader (0.3%), and 2/315 by the cerebral second reader (0.6%), with statistical significant difference. No additional disagreement rate was observed in regular and after-hours activities (*P* 0.63).

Conclusions: Double-reading of outsourced CT and MRI interpretations yielded 21.4% disagreement rate, with potential clinical consequence for patient in 2,9% of the cases. These results are in accordance with those previously reported and suggests that quality assurance of outsourced interpretations is needed.

Key Words: teleradiology, quality assurance, CT-scan, MR-scan, disagreement

*From the department of Radiology, Nîmes University Hospital, Nîmes, France; †Quality of care unit Grenoble Alpes University Hospital, TIMC UMR 5525 CNRS, Grenoble Alpes University, France; ‡Department of Radiology, Université de Lyon 1, Hospices Civils de Lyon, Centre Hospitalier Lyon Sud, 69495 Pierre Bénite Cedex, CREATIS INSA 502, 69621 Villeurbanne Cedex, France; §Department of Medical Informatics, Centre Hospitalier et Universitaire de Grenoble, Hôpital Nord, Boulevard de la Chantourne, 38700 La Tronche, France; ||Department of Radiology, Centre Hospitalier et Universitaire de Toulouse, Hôpital Pierre-Paul Riquet, Place du Docteur Baylac - TSA 40031, 31059 Toulouse cedex 9; and ¶Department of radiology, Centre Hospitalier et Universitaire de Poitiers, Hôpital de la Milétrie, 2 Rue de la Milétrie, 86021 Poitiers cedex, France.

Correspondence: Jean-François Vendrell, MD, PhD, Centre Hospitalier et Universitaire de Nîmes, Hôpital Carémeau, Place du Professeur Robert Debré, 30029 Nîmes, France (e mail: jfvendrell@yahoo.fr).

Funding: No funding source.

Competing Interests: Drs. Vendrell Jean-François, Frandon Julien, Cotton François, Ferretti Gilbert, Sans Nicolas, and Tasu Jean-Pierre are shareholders in TeleDiag.

Ethics approval: local ethics committee, University Hospital Center of Nîmes, France.

Misdiagnosis and discrepancies in radiological reports are common in daily practice, involving radiologists in regular and after hours activities.[1,2] Many contributing factors have long been recognized, such as excessive workload, cognitive overload, imperfect information processing, poor communication, and flawed decision making.[3–7] Recent developments in computed tomography (CT) and magnetic resonance (MR) imaging have slowly induced a simultaneous increase of healthcare consumer demands on providers for access to radiological exams, and an increase in radiologist productivity. This excess of workload is in part responsible for fatigue, which can result in errors, especially when combined with inadequate clinical information and poor communication with patients.[3,8,9] The current workload means that radiologists have never been so at risk of making a misdiagnosis with potential patient harm.[9] Fortunately, most of these errors are minor with no consequence, or quickly rectified following clinical patient evolution.[4,10]

Considering the general tight management of the radiological environment, with x ray dose for example being stringently controlled, lack of exam interpretation evaluation represents an overlooked potential flaw in patient management in France. By comparison, A systematic peer review of 5% of radiological exams is required in USA and recommended in UK.[11,12] Recent changes in radiologist demography, practice and organization have led to rural depopulation of radiologists, consequently necessitating teleradiology services for regular and after hours activities.[13] Thus, the emergence of teleradiology companies without evaluation of their reports is concerning.[14] The purpose of this study was to independently analyze a double reading protocol of current CT and MR radiological exams through a teleradiology company.[6,15] A percentage of all interpreted exams were subjected to a second analysis by senior radiologists hired by the teleradiology company, with a particular focus on diagnostic errors that potentially induced morbidity or life threatening prognosis for patients.

METHODS

Institutional review board approval of the University Hospital Center of Nîmes was obtained for this retrospective study and informed consent from patients was waived. A teleradiology company (TeleDiag, Lyon, France) offering its services and support to French medical centers without adequate number of radiologists was investigated. Members of this company have a validated radiological certificate for interpreting radiographies, CT scans, and MR scans proposed via a secure internet link (Rxeye, SECTRA group, Stockholm, Sweden). Most of these radiologists also have a routine activity in public and/or private establishments, with a subspecialty in the field of radiology. They trained in French university hospital centers. For each patient, previous examinations and clinical details were available to the radiologist through electronic communication or phone conversation with the radiologic technologist or the prescribing doctor. Emergencies and after hours

TABLE 1. Five Point Agreement Scale

Conformity Score	Criteria
0	No discordance between reports from radiologists
1	Writing error in the report without any consequence
2	Minor discrepancy without potential for morbidity or life threatening outcome
3	Writing error or misdiagnosis potentially inducing morbidity without potential for life threatening
4	Writing error or misdiagnosis potentially inducing life threatening outcome

reports had to be interpreted within 1 hour, and within 24 hours in regular hours practice. A pool of 200 radiologists performing ra diological activities in the setting of TeleDiag, interpreted all exams. The radiologist supervising the session interpreted all exams in their own field of competence and offered the others to specific radiologists via the TeleDiag platform. Four "senior ra diologists (university professors >10 years of experience)" spe cializing in cerebral imaging (F.C., radiologist n°1), thoracic imaging (G.F., radiologist n°2), osteo articular imaging (N.S., ra diologist n°3), and abdominal pelvic imaging (J.P.T., radiologist n°4) were requested to give a second opinion as many of the pri mary radiologists' imaging reports of CT and MR scan examina tions as they could find time for. From 1st January 2015 to 30th June 2016, 134169 radiological exams from 36 public French centers were interpreted: 1) 72190 radiographies (53.8%), 2) 42225 CT scans (31.5%), and 3) 19754 MR scans (14.7%). The analysis addressed a total of 2040 radiological exams that included 934 CT scans and 1106 MR scans for which a second read was per formed. Second reads were independently chosen and interpreted by second radiologists from the total list of exams in the worklist viewer, which displays basic information relating to anatomical area of scan, patient age and geographical location. The second ra diologists chose exams from their field specialty, with knowledge of the report provided by the first reader, without a fixed number of controls to perform. These exams were acquired on working days from 8 a.m. to 6 P.M. (n 1635; 80.1%) and in after hours from 6 p.m. to 8 A.M., weekend and public holidays (n 405; 19.9%). Depending on the anatomical area exploration and the exam indication, they were classified as cerebral, thorax, osteo articular, abdominal, and thoracic abdominal pelvic sub specialties by authors (JFV and JF). Thoracic abdominal pelvic anatomical area was designed for the classification of oncologic thoracic abdominal pelvic CT scans, these exams had a second intepretation provided either by radiologist n°2 or n°4. A five point agreement scale was attributed by the second readers as presented in Table 1. The disagreement frequency was analyzed with regards to the sec ond read, and graded from score 1 to 4. Disagreements were considered as minor if no clinical consequence for patient (scores 1 and 2) and major if potential clinical consequences (score 3 and 4). In cases of "major disagreement", a systematic phone conversa tion between the two radiologists was immediately organized. Af ter obtaining a consensus, prescriber doctors were informed about the final reports with the modified interpretation. Radiologists from the University Hospital Center of Nimes independently col lected and analyzed all the data (JPB and AL).

Minor and major disagreements were compared to identify risk factors using the Chi square test. When the expected number in any cell was less than five, the Fisher exact test was used for 2×2 tables and the Freeman Halton extension test was used for tables larger than 2×2. We then stratified our comparison of no or minor error versus major error by technical exams, across working hours and organ specialty. Two sided P values <0.05 were considered signifi cant. All statistical analyses were performed using Stata Special Edition version 14 (Stata Corporation, College Station, TX, USA).

RESULTS

Among the 2040 included radiological exams, 315 (15.4%) were interpreted by radiologist n°1 (cerebral specialty), 754 exams (37.0%) by radiologist n°2 (thorax specialty), 901 exams (44.2%) by radiologist n°3 (osteo articular specialty), and 70 exams (3.4%) by radiologist n°4 (abdominal pelvic specialty) (Table 2). A total of 62 Thoracic abdominal pelvic CT scans were interpreted either by radiologist n°2 (n 41) or n°4 (n 21). After second read, a score 0 was attributed in 1603 cases (78.6%) (CT scans, n 696 and MR scans, n 907), a score 1 was attributed in 271 cases (13.3%) (CT scans, n 107 and MR scans n 164), a score 2 was attributed in 107 cases (5.2%) (CT scans, n 78 and MR scans, n 29), a score 3 was attributed in 42 cases (2.1%) (CT scans, n 36 and MR scans, n 6), and a score 4 was attributed in 17 cases (0.8%) (CT scans, n 17) (Table 3). A total of 1,635 exams (80.1%) were interpreted in regular hours and 405 (19.9%) in after hours. Major disagree ments were observed in 34 regular hours exams and in 25 after hours exams (n 59, 2.9%). Disagreement scores occurred in

TABLE 2. Distribution of All Exams With Double Reading

	CT		MR		Total	
Examination	**n**	**%**	**n**	**%**	**n**	**%**
Cerebral	113	12.1	202	18.2	315	15.4
Thoracic	713	76.4	0	0	713	35.0
Abdominal pelvic	45	4.8	4	0.4	49	2.5
Osteoarticular	1	0.1	900	81.4	901	44.1
Thoracic abdominal pelvic	62	6.6	0	0	62	3.0
Regular hours	541	57.9	1094	98.9	1635	80.1
After Hours	393	42.1	12	1.1	405	19.9
Total	934	100	1106	100	2040	100

TABLE 3. Conformity Score Presentation According to Technical Exam, Working Hours, and Anatomical Examination

Disagreement	No		Minor				Major					
Score	0		1		2		3		4		Total	
Total	n	%	n	%	n	%	n	%	n	%	n	%
Exam												
CT	696	74.5	107	11.5	78	8.4	36	3.8	17	1.8	934	100
MR	907	82	164	14.8	29	2.6	6	0.6	0	0	1106	100
Activity												
Regular hours	1305	79.8	222	13.6	74	4.5	25	1.5	9	0.6	1635	100
After hours	298	73.6	49	12.1	33	8.1	17	4.2	8	2	405	100
Radiologist												
N°1 (cerebral)	172	54.6	115	36.5	26	8.3	2	0.6	0	0	315	100
N°2 (thoracic)	596	79.1	53	7	57	7.6	32	4.2	16	2.1	754	100
N°3 (osteoarticular)	799	88.7	86	9.6	13	1.4	3	0.3	0	0	901	100
N°4 (abdominal pelvic)	36	51.4	17	24.3	11	15.7	5	7.2	1	1.4	70	100

437/2040 exams (21.4%), including minor (grade 1 and 2) dis agreements in 378 cases (18.5%) and major (grade 3 and 4) dis agreements in 59 cases (2.9%). Of the major disagreements (n 59), 48/754 were assigned by the thoracic second reader (6.4%), 6/70 by the abdomen pelvic second reader (8.6%), 3/901 by the osteoarticular second reader (0.3%), and 2 by the cerebral second reader (0.6%), with statistical significant difference.

Minor and major disagreements observed in the series are pre sented in Table 4.Stratified analysis underlined no statistically signif icant difference in term of disagreement rate was observed between radiological interpretations performed in regular hours and in after hours for both CT and MRI (P 0.63 and P 0.06, respectively) (Table 5). Interpretations of thoracic, abdominal pelvic and thoracic abdominal pelvic CT scans presented increased major disagreement rates as compared to cerebral and osteoarticular MR scans. Interpreta tion of osteoarticular MR scans provided the lowest disagreement rates, particularly without graded score 4. High number of thorax CT scans (n 713) and osteoarticular MR scans (n 901) were included as compared to other anatomical locations.

DISCUSSION

The double reading of teleradiologist exams reported a dis agreement rate of 21.4% (437/2040). Major disagreement rate with potential for clinical implication was 2.9% (59/2040) includ ing potential for life threatening in 0.8% of the cases (17/2040). These data underlined the interest of a double reading protocol for patient therapeutic management, and suggest quality assurance of outsourced radiological interpretations.

Errors in medical diagnosis, and most particularly in radiology, have long been acknowledged; before 2001, error rates ranging from 2 to 30% were reported.[5,6,16,17] These results cannot be di rectly compared against ours because of the large diversity of in cluded radiological exams such as mammography, obstetrical

TABLE 4. Disagreement Description

Minor Disagreements	Grade 1	Grade 2	Major Disagreements	Grade 3	Grade 4
Report anomaly (ex: unstructured, incomplete, typo error, dose lengh product anomaly)	n = 65	.	Pulmonary embolism	n = 2	n = 9
			Pleural plaque	n = 6	.
Incomplete acquisition protocol (ex: absence of a CT acquisition or a MR sequence)	n = 14	.	Infectious pneumonia	n = 4	n = 1
			Diffuse interstitial pulmonary fibrosis	n = 1	.
Cerebral finding (ex: leucopathy, atrophy, begnin lesion, developement venous anomaly, sequella)	n = 61	n = 18	Mediastinal adenopathy	n = 4	n = 1
			Acute pulmonary oedema	n = 2	.
			Lymphangioleiomyomatosis	n = 1	.
Osteoarticular finding (ex: chondropathy, meniscopathy, begnin lesion, popliteal cyst)	n = 78	n = 18	Pulmonary node	.	n = 2
			Hepatic node	n = 1	n = 1
			Kidney node	.	n = 1
Thoracic finding (ex: cardiomegalia, small isolated lung node <6 mm, mediastinal lymphnode, limited pleural effusion)	n = 43	n = 51	Bone lesion	.	n = 1
			Anterior cruciate ligament sprain	n = 1	.
			Bone fracture	n = 1	.
Abdominal finding (ex: steatosis, adrenal adenoma, begnin hepatic or renal lesion, limited peritoneal effusion, atheromatous disease)	n = 10	n = 20	Supraspinatus tendon tear	n = 1	.
			Stroke	.	n = 1
			Inappropriate CT scan Protocol	n = 6	.
			Incomplete report	n = 12	.
Total	n = 271	n = 107	Total	n = 42	n = 17

TABLE 5. Stratified Analysis to Determine Potential Risk Factor for Major Disagreement

Disagreement	No or Minor n (%)	Major n (%)	*P* Value
MRI			0.06
Regular hours	1089 (99.0)	5 (83.3)	
After hours	11 (1.0)	1 (16.7)	
CT			0.63
Regular hours	512 (58.1)	29 (54.7)	
After hours	369 (41.9)	24 (45.3)	
MRI			0.01
Cerebral	200 (18.2)	2 (33.3)	
Abdominal pelvic	3 (0.3)	1 (16.7)	
Thoracic	0 (0.0)	0 (0.0)	
Thoracic abdominal pelvic	0 (0.0)	0 (0.0)	
Osteoarticular	897 (81.5)	3 (50.0)	
CT			0.008
Cerebral	113 (12.8)	0 (0.0)	
Abdominal pelvic	43 (4.9)	2 (3.8)	
Thoracic	668 (75.8)	45 (84.9)	
Thoracic abdominal pelvic	56 (6.4)	6 (11.3)	
Osteoarticular	1 (0.1)	0 (0.0)	

ultrasound, and arteriography, and the lack of CT scan and MRI evaluation due to the emergence of these imaging techniques at this time. Since 2001, with the CT scan and teleradiology developments in North America, United Kingdom, and Australia, some previous reports have proposed a disagreement rate evaluation of outsourced teleradiology based on a second opinion performed by in house radiologists, with major discrepancy rates or clinically significant errors of teleradiologist <1.5%.[18–21] Other studies have evaluated in house radiologists, registrars or residents during after hours for CT scan interpretations with discrepancy rates <5% and for MR scan interpretations with a major discrepancy rate of 4.2%.[22–28] Although there were slight differences in the methodology of these studies, major disagreement rate observed in our study is in accordance with those presented above.

All included exams in the study were remotely interpreted by first and second readers, with potential increased risks of disagreement than in house radiological interpretations. We suggest that high level of experience and training of the first line teleradiologist in an organ specialty, and their awareness about a potential expert second read calling for careful interpretations, contribute to limit disagreement rates in regular and after hours. High association between thoracic, abdominal pelvic and thoracic abdominal pelvic CT scan and major disagreements was also observed. However, this requires a cautious interpretation, as it could have arisen due to the heterogeneity of the series, with the inclusion of a high proportion of thorax CT scans and osteoarticular MR scans as compared to other anatomical locations. For comparison, osteoarticular MR scan misinterpretations generally did not lead to significant errors inducing life threatening potential, unlike thorax or abdominal CT scans. Second readers (Radiologists n° 2 and 3) interpreted many more CT and MR exams as compared with the 2 others, with a tendency to seek fortuitous pulmonary embolism observed for radiologist n°2.

Control quality in teleradiology depends on several factors under the responsibility of three main operators including radiologist, radiologic technologist, and application engineer. Firstly, getting accurate information about patient symptomatology and past medical history using efficient telephone and internet communication is needed to determine the appropriate exam. Secondly, exam acquisition must be in accordance with, and relies on, the in house radiologic technologist expertise in the patient management, and the physician engineer in the image treatment. Thirdly, ergonomics of the viewer and accessibility of radiological anteriority for comparisons must also be taken into consideration to facilitate interpretations. The above factors are normally accounted for successfully in current practice, however, radiologist interpretation evaluation, which represents one of the most important points in quality interpretation, is often overlooked.

Our study presented some limitations, including the heterogeneous distribution of exams. This arose as second readers were not given a fixed number of controls to perform and instead were invited to perform as many as they wished, with some second readers being more active than others (n°2 and n° 3). In addition, use of the disagreement classification could be interpreted differently by the different radiologists, thus even after thorough training of the classification, it was still vulnerable to subjective interpretation. Overestimation of disagreement scores might have arisen in some cases, probably in the patient interest. Another limitation was the use of only one second radiologist, the assumption being that their evaluation was irrefutable and always better than the primary radiologists, but there is no basis for such an assumption. In cases of disagreement between first and second radiologists, a systematic phone conversation between the two readers was immediately organized to obtain a consensus. We also cannot exclude that second readers might have focused on exams considered most interesting to them, potentially introducing a bias of which exams were chosen for second evaluation. Finally, the observed disagreement rates using teleradiology require a careful interpretation due to the lack of gold standard for comparison. In theory, the gold standard might be the evaluation of radiologists working in radiology department in a hospital or clinic, although this is very rarely reported.

CONCLUSION

We conclude that double reading protocol of outsourced radiology exams, in regular and after hours activities, yielded a 21.4% of disagreement rate with 2.9% considered as major. Our results are in accordance with those previously reported and suggests that quality assurance of outsourced CT and MR interpretations is needed.

ACKNOWLEDGMENTS

We thank Sarah Kabani for help in editing the manuscript.

REFERENCES

1. Diagnostic accuracy in radiology: defining a literature-based benchmarkad. Riology quality institute. 2012. Available at: http://www.radisphereradiology.com.
2. Brady A. Error and discrepancy in radiology: inevitable or avoidable? *Insights Imaging*. 2017;8:171 182.
3. Reiner B, Krupinski E. The insidious problem of fatigue in medical imaging practice. *J Digit Imaging*. 2012;25:3 6.
4. Brady A, Laoide RÓ, McCarthy P, et al. Discrepancy and error in radiology: concepts, causes and consequences. *Ulster Med J*. 2012;81:3 9.
5. Fitzgerald R. Errors in radiology. *Clin Radiol*. 2001;56:938 946.
6. Goddard P, Leslie A, Jones A, et al. Error in radiology. *Br J Radiol*. 2001;74:949 951.

7. Balogh E, Miller BT, Ball J. *Improving Diagnosis in Health Care*. Washington DC: Institute of Medicine (U.S.). Committee on Diagnostic Error in Health Care; 2015.
8. Berlin L. Liability of interpreting too many radiographs. *AJR Am J Roentgenol*. 2000;175:17 22.
9. Krupinski EA, Berbaum KS, Caldwell RT, et al. Long radiology workdays reduce detection and accommodation accuracy. *J Am Coll Radiol*. 2010;7: 698 704.
10. Carancil F, Tedeschi E, Leone G, et al. Errors in neuroradiology. *Radiol Med*. 2015;120:795 801.
11. Mahgerefteh S, Kruskal JB, Yam CS, et al. Peer review in diagnostic radiology: current state and a vision for the future. *Radiographics*. 2009;29:1221 1231.
12. The Royal College of Radiologists. *Quality Assurance in Radiology Reporting: Peer Feedback*. London, UK: The Royal College of Radiologists, 2014:1 18.
13. Conseil National de l'Ordre des Médecins (2016) Atlas de la démographie médicale en France. Available at: https://www.conseil-national.medecin.fr/sites/default/files/atlas de la demographie medicale 2016.pdf.
14. European Society of Radiology (ESR). ESR teleradiology survey: results. *Insights Imaging*. 2016;7:463 479.
15. Lauritzen PM, Andersen JG, Stokke MV, et al. Radiologist-initiated double reading of abdominal CT: retrospective analysis of the clinical importance of changes to radiology reports. *BMJ Qual Saf*. 2016;25: 595 603.
16. Brenner RJ, Lucey LL, Smith JJ, et al. Radiology and medical malpractice claims: a report of the practice standards claims survey of the Physician Insurers Association of America and the American College of Radiology. *Am J Roentgenol*. 1998; 171:19 22.
17. Soffa DJ, Lewis RS, Sunshine JH, et al. Disagreement in interpretation: a method for the development of benchmarks for quality assurance in imaging. *J Am Coll Radiol*. 2004;1:212 217.
18. Wong WS, Roubal I, Jackson DB, et al. Outsourced teleradiology imaging services: an analysis of discordant interpretation in 124,870 cases. *J Am Coll Radiol*. 2005;2:478 484.
19. Hohmanna J, de Villiers P, Urigo C, et al. Quality assessment of out sourced after-hours computed tomography teleradiology reports in a Central London University Hospital. *Eur J Radiol*. 2012;81:e875 e879.
20. Platts-Mills TF, Hendey GW, Ferguson B. Teleradiology interpretations of emergency department computed tomography scans. *J Emerg Med*. 2010; 38:188 195.
21. Agrawal A, Agrawal A, Pandit M, et al. Systematic survey of discrepancy rates in an international teleradiology service. *Emergency Radiology*. 2011; 18:23 29.
22. Strub WM, Vagal AA, Tomsick T, et al. Overnight resident preliminary interpretations on CT examinations: should the process continue? *Emerg Radiol*. 2006;13:19 23.
23. Stevens KJ, Griffiths KL, Rosenberg J, et al. Discordance rates between preliminary and final radiology reports on cross-sectional imaging studies at a level 1 trauma center. *Academic Radiol*. 2008;15:1217 1226.
24. 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.
25. Howlett DC, Drinkwater K, Frost C, et al. The accuracy of interpretation of emergency abdominal CT in adult patients who present with non-traumatic abdominal pain: results of a UK national audit. *Clin Radiol*. 2017;72:41 51.
26. Hanna TN, Loehfelm T, Khosa F, et al. Overnight shift work: factors contributing to diagnostic discrepancies. *Emerg Radiol*. 2016;23:41 47.
27. Issa G, Taslakian B, Itani M, et al. The discrepancy rate between preliminary and official reports of emergency radiology studies: a performance indicator and quality improvement method. *Acta Radiol*. 2015;56 598 604.
28. Filippi CG, Schneider B, Burbank HN, et al. Discrepancy rates of radiology resident interpretations of on-call neuroradiology MR imaging studies. *Radiology*. 2008;249:972 979.