

Research Article

Discrepancy between Reviewing Clinicians and Radiologists in Reporting of Chest Radiograph Signs of Coronavirus Disease 2019 (COVID-19)

Keir Ovington*, Rhys Metters

Royal United Hospital, Bath, UK

*Corresponding Author: Keir Ovington, Royal United Hospital, Bath, BA1 3NG, UK, Tel: +447796351814

Received: 12 February 2021; Accepted: 08 March 2021; Published: 30 March 2021

Citation: Keir Ovington, Rhys Metters. Discrepancy between Reviewing Clinicians and Radiologists in Reporting of Chest Radiograph Signs of Coronavirus Disease 2019 (COVID-19). Journal of Radiology and Clinical Imaging 4 (2021): 050-056.

Abstract

Introduction: Chest radiographs form an important aid to COVID-19 diagnosis, however their utility is limited by the reviewers ability to accurately assess for its radiological features, This study seeks to assess for any difference in reporting between radiologists and clinicians.

Methods: 135 admission chest radiographs of patients without a known COVID-19 diagnosis were gathered opportunistically. Radiologist reports, reviewing clinicians

and clinicians with no knowledge of the patient categorised radiographs as having either “no covid signs” (category 0), “indeterminate covid signs” (cat. 1) or “classic/probable covid signs” (cat. 2). Cohen’s Kappa was used to evaluate the inter-reporter reliability between these groups.

Results: Radiologists identified 69% as cat. 0, 29% as cat. 1, 5% as cat. 2. Reviewing clinicians agreed with 73% of these reports achieving a Kappa of 0.43 (95% CI 0.32 - 0.54). Consultants performed best with a kappa of 0.77

(0.56-0.98). Clinicians without knowledge of the patient agreed with 54% of reports, Kappa 0.17 (-0.16 -0.50).

Conclusion: There is a significant discrepancy between radiologist and non-radiologist reporting of chest radiographs in COVID-19 supporting the use of rapid radiologist reporting of chest radiographs to aid with diagnosis.

Advances in Knowledge: This is the first paper to our knowledge to assess the difference in reporting of COVID-19 between radiologists and reviewing clinicians, indicating that radiologist reporting of chest x-rays has a measurable advantage in detecting COVID-19 signs, compared to clinician reports alone.

Keywords: Chest radiograph; COVID-19; Reporting; Radiologist

1. Introduction

Chest radiographs are being widely used as part of diagnosing and evaluating the severity of COVID-19 in patients attending emergency departments or being admitted to hospital [1]. The reporting of chest radiographs however varies between hospitals with some using prompt or “hot” radiologist reporting [2, 3], while others rely only upon the reviewing clinician's interpretation. This paper seeks to evaluate the reliability between radiologist and clinician reporting in a large district general hospital during the first ‘wave’ of COVID-19 cases. Previous studies have shown significant discrepancies between reporting of chest radiographs by specialty clinicians versus radiologists. The degree of discrepancy varies depending on the subject, for example emergency medicine specialists show high levels

of agreement with radiologist reports for traumatic chest radiographs but not for pneumonia [4, 5, 6]. No prior study to our knowledge has looked at this discrepancy in COVID-19.

Evaluating whether specialty clinicians reports vary from those of radiologists would help to determine if rapid reporting by radiologists is useful in diagnosing and assessing the severity of COVID-19, ensuring that patients receive the optimum care and that the risk of transmission to staff or other patients can be minimized. The results may be of further importance in the systemic response to future emerging diseases with radiological features, helping to determine if early application of radiologist reporting of radiographs is beneficial. More senior clinicians tend to be more accurate in reporting radiographs and more confident in their decisions [7]. This is presumed to be due to increased training and experience, however no similar data has been collected assessing if this trend holds in a novel disease. This study also evaluates the levels of agreement between radiologist and clinician reports, depending on the grade of clinician, to assess if senior input on radiographs can achieve similar results to radiologist reporting.

A key difference in radiologist reporting from clinician reporting is the level of clinical details known. A clinician will generally interpret a chest radiograph with their pre-existing clinical impression in mind, contrasting from the minimal clinical details generally known to radiologists. Comparing the difference in reporting by clinicians with and without knowledge of the patients clinical details allows us to assess the extent to which their interpretation of the chest radiograph is biased towards their clinical impression rather than based solely on radiological findings.

Inter-rater reliability between two raters, as is being assessed in this study, is statistically evaluated using Cohen's Kappa. This measures the degree of concordance taking into account the probability of agreement occurring by chance given the distributions of ratings. A value of 1 indicates perfect agreement while a value of zero indicates agreement occurs at the rate you would expect from chance alone [8].

2. Methods

Data collection was undertaken at Royal United Hospital Bath, a large district general hospital in south-west England between April and August 2020 - the first peak of COVID-19 cases in the UK.

2.1 Initial clinician chest radiograph report

Patients were opportunistically selected on admission meeting the following criteria:

- New admission
- New chest radiograph performed
- Clinician clerking patient or providing senior decision making available to interpret chest radiograph
- Covid status not already known

Clinicians included junior and senior doctors from a foundation level up to consultant along with nurse practitioners who regularly review and independently interpret chest radiographs. Clinicians had either seen the patient themselves or were providing senior input/decision making after a junior clerking and were therefore aware of the pertinent clinical details.

Clinicians were asked to categorise the radiograph into a simplified scoring system based upon published recommendations of reporting language [9]:

- Category 0: No signs in keeping with COVID-19 infection.
- Category 1: Signs that could be in keeping with COVID-19 infection but do not follow classical pattern - indeterminate.
- Category 2: Classical signs of COVID-19 infection.

A radiograph in any category may have signs of a different pathology, this would not affect the categorisation. The scoring system was explained to clinicians in above terms but no further tutoring was given into identifying signs of COVID-19 infection or other pathologies. Clinicians were informed that they could look at previous imaging and current or past clinical information.

2.2 Blind clinician report

Clinicians (grades as above) were opportunistically asked to review five chest radiographs collected as above for patients whom they had not reviewed, without knowledge of the patients clinical information beyond age (as date of birth was visible on viewing software). Clinicians were informed that they could look at previous imaging but not previous reports, requests or other clinical information. They were asked to categorise the chest radiographs using the same criteria as above.

2.3 Radiologist report

Radiologist reports were taken from finalised radiology written reports by consultant and registrar radiologists. The majority of reports specified a classification based on published recommendations on reporting language [9]:

- No covid signs - marked category 0 for comparison
- Indeterminate for COVID-19 - marked category 1 for comparison
- Classic/Probable COVID-19- marked category 2 for comparison
- Alternative pathology, not in keeping with COVID-19 - marked category 0 for comparison

In reports where a classification was not specified one was determined from the body of the report. Reporting radiologists had access to clinical information supplied by the requesting clinician who is asked to specify (through computerised form) if the patient has a cough, is febrile, has raised CRP or lymphopenia. Other information may also have been added as free-text.

2.4 Data processing

Data processing was undertaken using Microsoft Excel 2002. Precision of data calculations was limited to the numeric precision of Excel (15 significant figures). Kappa statistics and their confidence intervals were calculated according to the methods set out by ML McHugh [8]. Kappa statistic is used in favour of simple percentage agreement in order to take into account the probability of agreement occurring by chance.

3. Results

Data was obtained on 135 radiographs of which 90 were re-reviewed by 'Blind' Clinicians. Additionally 5 radiographs were rejected due to not being findable on the image viewing system (presumably due to incorrect data entry). No radiographs required exclusion after data entry. All data

has been rounded to the nearest percentage point or 2 decimal places.

3.1 Overall distribution of radiograph reports

The Majority (69%) of reviewed chest radiographs had no covid signs as determined by radiologist report with 5% having classic covid signs and the rest being indeterminate. A similar distribution of reports is seen in clinician reports with a slightly higher rate of indeterminate and classic covid signs being reported, particularly in 'Blind' Clinician reports - see table 1.

3.2 Reviewing clinician reports

The Overall Kappa value of 0.43 shows a moderate, significantly better than chance, agreement between reviewing clinicians and radiologists however a statistically significant difference remains even among the best performing group (consultants) see table 2.

3.3 'Blind' clinician reports

Overall blind clinicians did not significantly differ from chance in their reports agreement with radiologist reports. This was true for all groups except consultants achieving a moderate to good agreement (Kappa 0.58) although a statistically significant difference in reports remained. A further breakdown by grade is seen in table 3.

3.4 Evaluation of discrepancies

Overall, as shown in table 4, when compared to radiologist reports a similar percentage of clinicians under-reported Covid signs as over-reported them. Consultants were less likely than SHO's to under-report Covid signs, with similar levels of over-reporting. 'Blind' Clinicians were much more likely to over-report covid signs.

	Category 0 ‘No covid’	Category 1	Category 2 ‘Classic covid’
Radiologist reports	93	35	7
	69%	26%	5%
Reviewing clinician reports	88	36	11
	65%	27%	8%
‘Blind’ clinician reports	46	32	12
	51%	36%	13%

Table 1: Overall distribution of radiograph reports.

Reviewer Grade	N	Kappa	95% CI	Percentage agreement
Overall	135	0.43	0.32, 0.54	73%
F1	1	-0.05	N/A*	0%
SHO	91	0.39	0.26, 0.53	73%
SPR	25	0.34	0.08, 0.60	64%
Consultant	18	0.77	0.56, 0.98	89%

*Confidence interval cannot be calculated due to lone data point.

Table 2: Agreement between reviewing clinician and radiologist reports by clinician grade.

Reviewer Grade	N	Kappa	95% CI	Percentage agreement
Overall	90	0.17	-0.16, 0.50	54%
F1	38	0.11	-0.11, 0.32	50%
SHO	28	0.06	-0.19, 0.32	50%
SPR	10	0.18	-0.21, 0.58	50%
Consultant	14	0.58	0.28, 0.88	79%

Table 3: Agreement between ‘Blind’ clinician and radiologist reports by clinician grade.

Comparison to radiologist report	Reviewing clinicians			'Blind' Clinicians
	Overall	Consultant	SHO	
In agreement	73%	89%	73%	54%
Covid signs under identified	16%	0%	15%	10%
Covid signs over identified	11%	11%	12%	36%

Table 4: Comparison of reporting discrepancies.

3.5 Other statistics

The statistical sensitivity of clinicians in reporting category 3 radiographs (taking radiologist reports as the ‘True positive’) was 71%. The statistical specificity of category 0 reports was 81%. We found 31% of radiographs reported by radiologists as category 1 & 2 radiographs were reported as category 0 by clinicians. Equivalent to clinicians missing probable or indeterminate covid signs in 10% of all radiographs. Additionally 19% of radiographs reported by radiologists as category 0 were reported as category 1 or 2 by clinicians. Equivalent to clinician’s inaccurately reporting indeterminate/probable covid signs in 13% of total radiographs.

4. Limitations and Recommendations for Further Research

Chest radiographs were opportunistically gathered when data collectors were on shift and may not represent a random sample, additionally radiograph and corresponding report collection would have favoured those clinicians who requested more chest radiographs. As the pandemic progresses both radiologists and other clinicians are likely to gain experience and improve their skills in detecting COVID-19 signs on chest radiographs which may change the discrepancies studied here. This review assesses the

agreement between radiologist and non-radiologist reporting however does not assess the accuracy of these reports with regards CT or PCR findings. Recommended topics for further research therefore include: comparison with CT and PCR findings, comparison with reports collected later in the course of the pandemic and on the benefit of targeted teaching sessions in reducing the discrepancies between radiologist and non-radiologist reporting.

5. Discussion and Conclusion

We have shown that there is a significant disparity between interpretation of chest radiograph signs of COVID-19 between radiologist reporting and reviewing clinicians (Cohen’s Kappa 0.48). Overall clinicians had similar levels of under and over reporting of COVID-19 signs (11% vs 16%). More consultant clinicians had much closer levels of agreement to radiologist reports than juniors (Kappa 0.77). Clinicians with knowledge of the patient's clinical presentation had closer levels of agreement to radiologist reporting than those without (Kappa 0.47 vs 0.17). These results suggest that a clinically significant difference exists between radiologist and clinician reporting of COVID-19 signs in chest radiographs. In our data 10% of total radiographs had COVID-19 signs missed by clinicians, 13%

had COVID-19 signs falsely reported; together this implies that 23% of chest radiographs are significantly miscategorised by clinician reporting with potential to result in erroneous clinical decision making regarding treatment and isolation precautions. This data supports the use of rapid chest radiograph reporting by radiologists to assess for COVID-19 signs, however consultant clinicians with knowledge of the patients clinical details corroborate reasonably well with radiologists. Given the strain on radiologist reporting research into improving clinician reporting with targeted teaching should be considered, artificial intelligence (AI) has also shown promising results that may soon form a potential alternative [10].

References

1. Rubin GD, Ryerson CJ, Haramati LB, et al. The Role of Chest Imaging in Patient Management during the COVID-19 Pandemic: A Multinational Consensus Statement from the Fleischner Society. *Radiology* 296 (2020): 172-180.
2. Cheng LT-E, Chan LP, Tan BH, et al. Déjà Vu or Jamais Vu? How the Severe Acute Respiratory Syndrome Experience Influenced a Singapore Radiology Department's Response to the Coronavirus Disease (COVID-19) Epidemic. *Am J Roentgenol.* 214 (2020): 1206-1210.
3. Glover T, Alwan S, Wessely K, et al. Radiology department preparedness for COVID-19 – experience of a central-London hospital. *Future Healthc J* 7 (2020): 174-176.
4. Atamna A, Shiber S, Yassin M, et al. The accuracy of a diagnosis of pneumonia in the emergency department. *Int J Infect Dis IJID Off Publ Int Soc Infect Dis* 89 (2019): 62-65.
5. Gatt M, Spectre G, Paltiel O, et al. Chest radiographs in the emergency department: is the radiologist really necessary?. *Postgrad Med J* 79 (2003): 214-217.
6. Safari S, Baratloo A, Negida AS, et al. Comparing the interpretation of traumatic chest x-ray by emergency medicine specialists and radiologists. *Arch Trauma Res* 3 (2014): e22189.
7. Satia I, Bashagha S, Bibi A, et al. Assessing the accuracy and certainty in interpreting chest X-rays in the medical division. *Clin Med* 13 (2013): 349.
8. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Medica* 22 (2012): 276-282.
9. Litmanovich DE, Chung M, Kirkbride R, et al. Review of Chest Radiograph Findings of COVID-19 Pneumonia and Suggested Reporting Language. *J Thorac Imaging* (2020).
10. Zhang R, Tie X, Qi Z, et al. Diagnosis of COVID-19 Pneumonia Using Chest Radiography: Value of Artificial Intelligence. *Radiology* (2020): 202944.



This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC-BY\) license 4.0](https://creativecommons.org/licenses/by/4.0/)