Anatomical and/or pathological predictors for the "incorrect" classification of red dot markers on wrist radiographs taken following trauma

R KRANZ, BSc (Hons), MSc

Radiology Department, Queen Elizabeth Hospital, Gateshead, Tyne and Wear, UK

P COSSON, DCR(R), D.Prof.

Medical Imaging, Teesside University, Middlesbrough, UK

Address correspondence to: Dr Philip Cosson

E-mail: p.cosson@tees.ac.uk

ABSTRACT

Objective: To establish the prevalence of red dot markers in a sample of wrist radiographs and to identify any anatomical and/or pathological characteristics that predict "incorrect" red dot classification.

Methods: Accident and emergency (A&E) wrist cases from a digital imaging and communications in medicine/ digital teaching library were examined for red dot prevalence and for the presence of several anatomical and pathological features. Binary logistic regression analyses were run to establish if any of these features were predictors of incorrect red dot classification.

Results: 398 cases were analysed. Red dot was "incorrectly" classified in 8.5% of cases; 6.3% were "false negatives" ("FNs") and 2.3% false positives (FPs) (one decimal place). Old fractures [odds ratio (OR), 5.070 (1.256–20.471)] and reported degenerative change [OR, 9.870 (2.300–42.359)] were found to predict FPs. Frykman V [OR, 9.500 (1.954–46.179)], Frykman VI [OR, 6.333 (1.205–33.283)] and non-Frykman positive abnormalities [OR, 4.597 (1.264–16.711)] predict "FNs". Old fractures and Frykman VI were predictive of error at 90% confidence interval (CI); the rest at 95% CI.

Conclusion: The five predictors of incorrect red dot classification may inform the image interpretation training of radiographers and other professionals to reduce diagnostic error. Verification with larger samples would reinforce these findings.

Advances in knowledge: All healthcare providers strive to eradicate diagnostic error. By examining specific anatomical and pathological predictors on radiographs for such error, as well as extrinsic factors that may affect reporting accuracy, image interpretation training can focus on these "problem" areas and influence which radiographic abnormality detection schemes are appropriate to implement in A&E departments.

BACKGROUND

Diagnostic error in accident and emergency (A&E) department medical imaging may negatively impact on patient care and the reputation of the hospital, and potentially result in litigation—phenomena that all healthcare providers strive to eradicate.[1] In a district general A&E, Guly[1] reported 779 "false negative" ("FN") diagnoses over 4 years—patients discharged with unidentified fractures or dislocations. 77.8% of these were owing to radiograph misinterpretation, and such misdiagnoses may become disabling or life threatening if not immediately recognized. Furthermore, the distress that false-positive (FP) results may cause is highlighted in mammography and may reduce service use[2]—something that A&E departments must also avoid. Three potential strategies to prevent such errors—employed individually or in unison—are apparent: improve image interpretation training for A&E staff, 24-h reporting services and radiographer abnormality detection schemes (RADS).[1] Immediate image interpretation to inform diagnostic decisions is frequently undertaken by A&E staff, before the radiological report is available. However, it has been claimed that junior doctors (whose responsibility this often is) are insufficiently experienced, trained or supervised.[3] Reported retrospective error rates vary between 2.5% and 6.8%,[4,5] with up to 68% reported in prospective studies.[6] Reliable comparison of such studies is challenging with the inability to control significant variables such as FP exclusions, nontrauma inclusions and length of data collection.[7,8] Nevertheless, the variability in image interpretation skills is striking and not conducive to a reliable 24-h service. The image interpretation gold standard is the full radiological report.[9] National Confidential Enquiry into Patient Outcome and Death[9] stated that all emergency patients should have immediate reporting of projection radiography and CT examinations to influence A&E treatment decisions. The Royal College of Radiologists[10] issued standards for a 24-h "hot" reporting service, but the UK shortage of radiologists per million population is a barrier—47 radiologists per million compared with 67 per million in much of Europe and Australia.[10,11] Also, significant geographical variation makes the actual provision of 24-h reporting difficult to establish. The use of reporting radiographers in projection radiography in the UK may alleviate the deficit in radiologist numbers (indeed, radiographer reports have been shown to be equivalent to those of consultant radiologists[12]), but it is suggested they are underutilized with regard to working hours (typically, 9 am–5 pm, Monday–Friday).[13] Future development of this service in line with 7-day working and/or increases in radiologist numbers could improve the report turnaround time, but at present, immediate reporting is not always a reality.[13] RADS are the third option to minimize diagnostic error, by alerting A&E staff to the presence of fracture/abnormality.[14] 92.8% of UK hospitals use RADS; 77.8% of these using red dot— a method initiated in 1985 involving radiographers attaching asterisks to radiographs considered to demonstrate abnormality.[15,16] There are moves towards refining this system to provide a written comment [preliminary clinical evaluation (PCE)] rather than a simple indication of abnormality, which is considered less ambiguous than red dot.[17,18] Numerous studies have examined the diagnostic accuracy of red dot systems.[19–21] The variability of the reported sensitivities (45.8–81.3%) and specificities (96.4–98.0%) from these studies shows it is not yet clear what the benefits are of using these tests. Without a stated or clearly defined benchmark, it is impossible to tell how "hard" the radiographs were to diagnose. This weakens the use of sensitivity and specificity as markers of radiographer ability, as conclusions are hard to draw across articles if the difficulty of the "tests" is unclear. Therefore, it is hard to evaluate effectively the merit of the red dot in relation to other RADS in this way.[20] Studies to investigate predictive factors for diagnostic errors in radiology have focused on extrinsic factors affecting reporters. Night-shift fatigue [odds ratio (OR), 1.94; confidence interval (CI), 1.18–3.21], increased workload and inexperience [OR, 1.6 (1.5–1.7) and OR, 1.3 (1.2–1.4), respectively] have been shown to predict reporter error by evaluation of "major" interpretation discrepancies between initial reports (junior staff) and senior review.[22,23] Ruutiainen et al[22] defined this as a discrepancy influencing patient management, whereas Davenport et al (p. 923)[23] used a scale of interobserver agreement that changed during the study, with categories including "difficult diagnosis, not ordinarily expected to be made" and "change in report may be clinically significant", making it difficult to dissect what specifically constituted error. The subjective nature of this decreases internal validity, but if validated by other groups, these conditions extrinsic to the interpreted images may have wider implications for radiographers and red dot performance. Kim et al[24] identified patient age as a predictor for discordant radiograph interpretation and was the only study to propose a pathological factor (presence of degenerative change) as a reason for error. However, this was suggested alongside difficulties in history taking and physical examination and warrants further investigation. Importantly, follow-up review of reports may identify errors, but too late to influence treatment or discharge.[22] Red dot (at the time of imaging) may impact on patient management. Given the limited consideration of intrinsic predictive factors for diagnostic error, in particular those that are directly applicable to radiographic practice, the aims of this study were to establish the prevalence of red dot in a sample of radiographs taken in A&E departments following trauma and to identify anatomical and/or pathological predictors for "incorrect" red dot classification.

METHODOLOGY

A retrospective, quantitative approach was used to examine existing radiographic cases using a digital imaging and communications in medicine/digital teaching library (DTL). The DTL was constructed from a radiology information system (RIS)/picture archiving and communication system migration from clinical sites and therefore represented clinical practice at the time (2009/2010). Previous ethical approval was obtained from the National Information Governance Board for this migration, confirming all patient, staff and site data were extracted and given pseudonyms, so these details are completely anonymous to users.[25] Therefore, demographics of the reporters cannot be identified, although reporting radiographers are shown to be equivalent to radiologists in plain-film emergency imaging.[12] Similarly, the demographics of the radiographers performing the examination are also unknown. Institutional ethical release for this study was obtained prior to data collection to ensure the safety of the researcher, and that the confidentiality was protected.

Sample

A single examination type was selected for review to reduce confounding factors associated with analysing multiple body areas and classification systems in one study. The wrist has been stated as the most commonly fractured area (maximizing sample size),[26,27] with delayed diagnosis causing continued pain and disability.[28,29] This clinical justification was reinforced by the availability of objective-recognized classifications of wrist fractures. The Frykman classification has been found to have better interobserver reliability than does Arbeitsgemeinschaft fur Osetosynthesefragen/Orthopaedic Trauma Association, and so the former was selected for use in this study.[30–33] Figure 1 illustrates the classification as used. Odd-numbered classifications depict fracture location, and even-numbered classifications include ulnar styloid involvement. Types I and II are extra-articular, Types III and IV are intra-articular at the radiocarpal joint (RCJ), Types V and VI are intra-articular at the distal radioulnar joint (DRUJ) and Types VII and VIII are intraarticular at both the RCJ and DRUJ.[30] A non-probability, purposive sample was obtained from the DTL and subjected to inclusion and exclusion criteria.

The inclusion criteria were wrist examinations, A&E referrals (first attendance for each injury) and all age groups (to reflect a typical A&E workload). The red dot scheme is implemented in A&E environments for highlighting fractures.[16] It is not utilized for patients with known fractures such as following a cast fitting (postreduction), surgery or in musculoskeletal clinics, so these were excluded. Where a report was unavailable for a particular case, it was also excluded for the absence of a gold-standard diagnosis. Following a pilot study of five cases, it was established that each case would take approximately 2min to assess. Therefore, all 452 A&E wrist cases in the DTL could be examined in the time available.

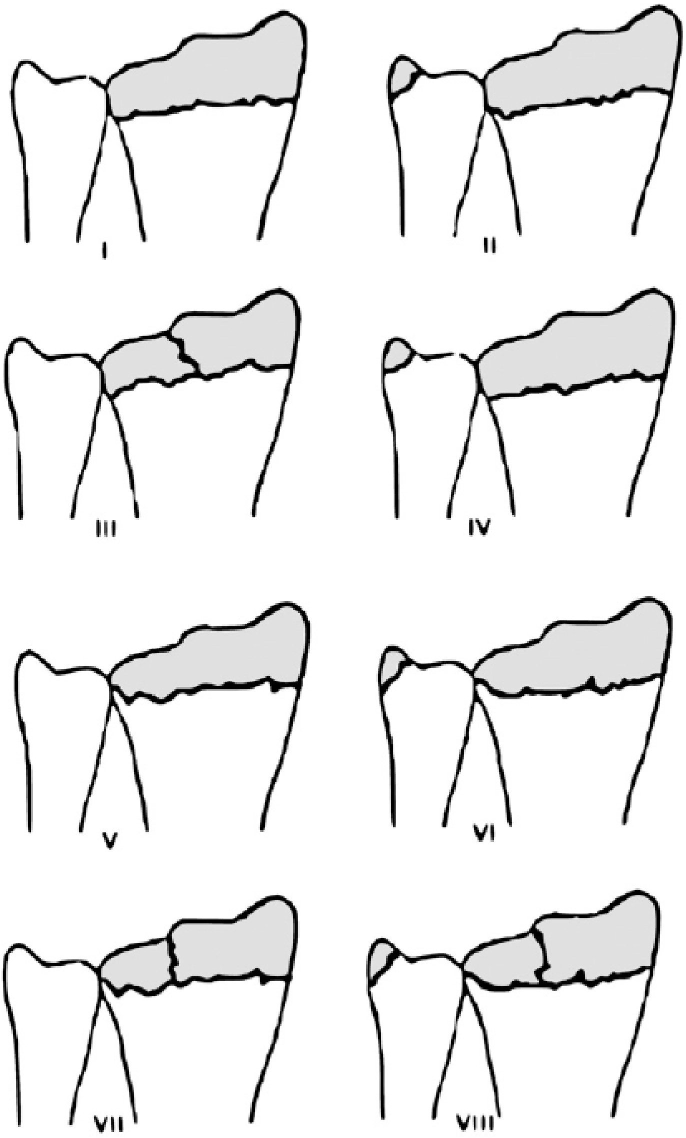
Predictors

Given the lack of previous research, three other anatomical/ pathological characteristics were selected as potential predictors of error in image interpretation, measured by correct or incorrect red dot classification. Degenerative change has been linked to radiograph misinterpretation,[24] whilst epiphyseal growth plates may obscure paediatric fracture detection.[28,35] Old fractures also present susceptibility to error, as they may be misinterpreted as acute injuries or confused with tumours, warranting inclusion in this study.[36,37] These predictors were identified following a systematic search of the literature and by subsequent hand searches, given the lack of focus on anatomical and pathological reasons for errors in the literature.

Procedure

All cases in the DTL were available for sampling.[25] “Wrist” and “accident and emergency” terms were inputted into the RIS advanced search tool to obtain the sample. The reports were read, inclusion and exclusion criteria were applied and, if included, the radiographs were examined by the researcher (RK). The accession numbers, red dot prevalence, abnormality status (present or not) and the presence of the stated characteristics (including Frykman classifications and non-Frykman positive group for abnormal cases) were recorded in a pro forma. The non-Frykman positive group was included for radiographs with an abnormality, but which did not fit any of the Frykman classification groups. Degenerative change and old fracture presence were recorded separately, whether they were mentioned in the report or observed but not reported. Not all anatomical/ pathological features (such as growth plates) are described in reports but may be significant. Report terminology was utilized to aid Frykman classification if this detail was included (i.e. “intra-articular”, “comminuted”, “ulnar styloid involvement”). Otherwise, this was performed by eye, primarily with the dorsopalmar (DP) projection as in Figure 1, but with the lateral projection used to assess the extent of fracture lines, for example, to establish if the fracture extended into a joint space, thus affecting classification. An expert reviewer (reporting radiographer with 13 years’ experience) was consulted if there was uncertainty regarding Frykman classification.

Figure 1. The Frykman classification[34] for fractures of the distal radius and ulna.



Data analysis

Data analysis was performed using SPSS® Statistics (IBM Corporation, Armonk, NY).[38] The four outcome groups were collated to allow "yes/no" (correct/incorrect red dot classification) input required by SPSS but were still discernible as demonstrated in Table 1. Crosstab analyses were performed for each factor against the outcome to see if any associations existed. This was performed in a split field, to observe if "abnormal" or "normal" diagnosis also affected whether an anatomical/pathological factor was significant. The level of significance was set at 10% (p=0.10) rather than 5% (p=0.05) to maximize the opportunity for any further trends to be identified, given the limited sample size.

Table 1. Discerning "incorrect" cases by splitting the field

Red dot classification	No acute abnormality	Abnormality present
Correct	True negative—no red dot and no abnormality present	True positive—red dot and abnormality present
Incorrect	False positive—red dot but no abnormality present	False negative—no red dot but abnormality present

Binary logistic regression models were used to identify which factors were predictive of red dot error and to test their independence from each other. FP error and "FN" error were regarded as dependent variables, and Frykman I–VIII classification, non-Frykman positives, degenerative change (reported), degenerative change (unreported) epiphyseal growth plate(s), old fracture (reported) and old fracture (unreported) were regarded as independent variables. One model was run for cases where no abnormality was present, to establish predictors for FP error, and another for fractured/abnormal cases for FN error. All potential factors were inputted into each model nonhierarchically, and those with the highest p-value (i.e. nonsignificant) removed following each run until only the significant (p=0.10) remained. This step-wise technique is suitable for exploratory studies, as the presence or absence of other variables can affect which are significant.[39,40] True diagnostic accuracy was incalculable owing to the potential for radiographer non-participation, affecting the classification of cases as "FNs". It cannot be established whether an abnormality was erroneously missed or the radiographer opted out of red dot. The red dot outcomes (Table 1) were referred to as true positive (TP), "FN", FP and true negative (TN), with inverted commas around "FN" reflecting the provisory nature of this label.

RESULTS

Decisions and exclusions

There were 54 exclusions from the 452 A&E wrist cases; therefore, 398 cases were eligible for full analysis. 45 cases were excluded as post-reduction projections, 7 where no report was available, 1 for a repeat entry of a case and 1 case that was performed following insertion of metalware abroad. The annotation "?*" (query fracture/abnormality) was observed in eight cases. It was considered TP if the report confirmed a fracture (five cases) and FP where "no fracture" was observed (three cases). Four TPs were paediatric, whilst the FPs had unreported degenerative change, an old fracture and growth plates, respectively. There was one examination of "bilateral wrists", so projections for both wrists were found under the same accession number.

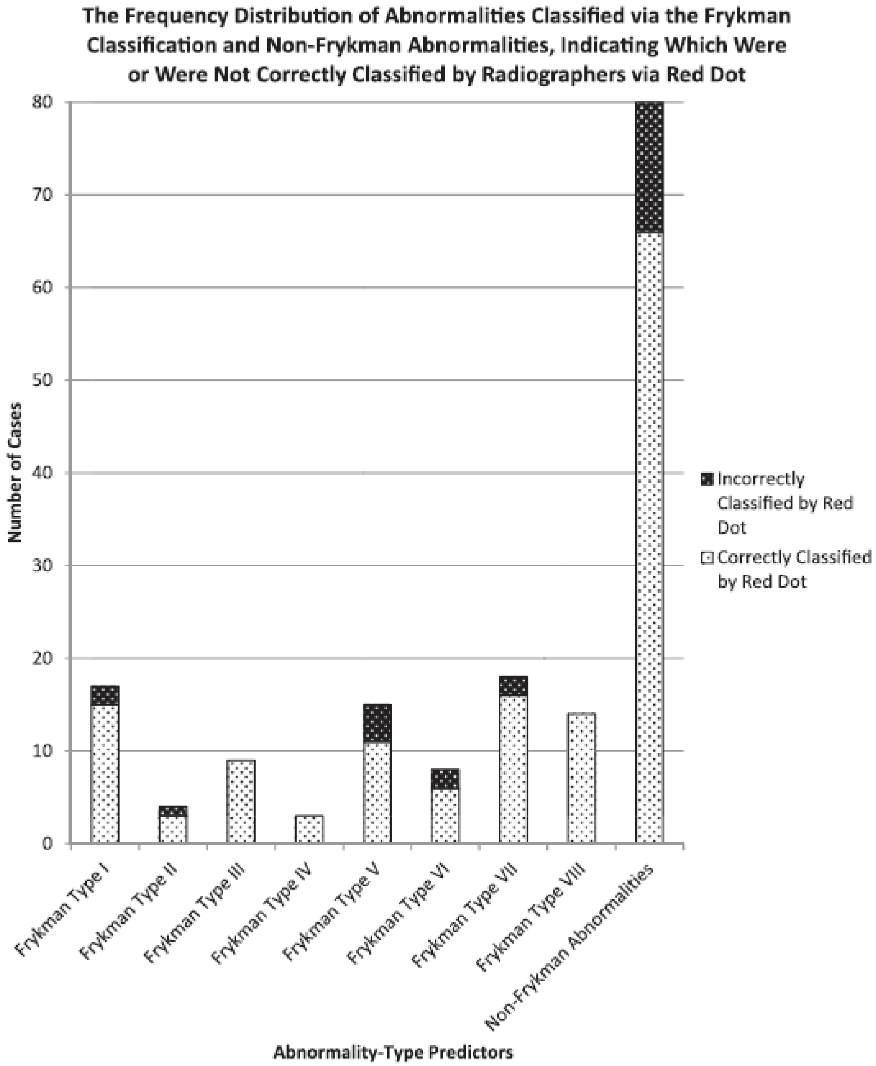
The left wrist was TP, whereas the right was "FN". These were recorded separately in the pro forma, as it was the first attendance for each wrist, with different red dot outcomes. Another case had two separate accession numbers—one for the initial wrist projections and another for an additional, modified (oblique) projection. A fracture was queried in the second visit, and it may have been a different examining radiographer. Indeed, both examinations were reported separately, not necessarily by the same reporter. Therefore, two entries met the inclusion criteria and were entered individually.

An expert reviewer (reporting radiographer) was consulted to clarify the Frykman classification of three cases, with agreement settled at Frykman I (extra-articular), V (DRUJ articulation) and VIII (RCJ and DRUJ articulation with ulnar styloid involvement). In a fourth case, the report did not mention an ulnar styloid fracture, but the expert reviewer agreed one was present. As the report was the specified gold standard, it was classified as Frykman VII (no ulnar styloid involvement), highlighting the weakness of a single, "human" gold standard. The Frykman V was FP; the others TP.

Prevalence

Of the 398 cases, 364 (91.5%) had correct red dot classification— 135 (33.9%) TPs and 229 (57.5%) TNs. 25 (6.3%) of the samples were "FNs" and 9 (2.3%) FPs (1 decimal place). This prevalence appears to be similar to other studies examining red dot (although these went on to calculate diagnostic accuracy). Hargreaves and Mackay[19] in their assessment of radiographer red dot performance prior to a training programme and the retrospective Brown and Leschke[21] study, both also had the lowest to highest prevalence groups as FP, "FN", TP and TN, respectively, despite differences in the overall number of cases analysed.

Figure 2. Frequency distribution of abnormalities.



The frequency distribution of abnormal radiographs is given in Figure 2. Of these cases, 88 (55%) had fractures that were classified via Frykman. No existing data were found for comparison. Frykman defines intra-articular fractures individually at the RCJ (III and IV) and DRUJ (V and VI), and in combination (VII and VIII) but only gives DP projections. However, two projections (DP and lateral) are standard practice as the wrist is three dimensional.[41] The most difficult to classify were therefore fractures that were not obvious on DP but evident on lateral projections. However, lateral projections were useful to analyse the extent of fracture lines and thus whether they were intraarticular (affecting classification). Making the distinction between extra- and intra-articular fractures was sometimes challenging owing to suboptimal DP positioning and natural variation, where the extent of distal radioulnar articulation differs between individuals. Thus distinguishing between Types I and V (extra- vs intra-articular) and Types II and IV (extra- vs intra-articular with ulnar styloid involvement) was problematic, seen with expert reviewer consultation.

There were 168 individual abnormalities overall in the 160 cases: 88 Frykman fractures and 80 abnormalities that could not be classified by the Frykman classification (the latter included nonfractures: radio-opaque foreign bodies, acute scapholunate joint widening and carpometacarpal joint dislocation). In three cases, two different non-Frykman abnormalities occurred in the same patient and in five other cases, a non-Frykman abnormality was observed alongside a Frykman fracture (therefore, 72 cases had just 1 non-Frykman abnormality only). If a red dot was observed on the radiographs with more than one abnormality, it was considered a TP. The abnormalities are detailed in Table 2, indicating the TPs and “FNs”.

Table 2. A breakdown of all abnormalities, indicating true positives and false negatives ("FNs"), and percentage missed

Description	*n*	TP	"FN"	Percentage missed
Scapholunate joint widening	2	1	1	50
Metacarpal fracture	5	3	2	40
Scaphoid fracture	6	4	2	33
Frykman V	15	11	4	27
Frykman VI	8	6	2	25
Frykman II	4	3	1	25
Radial styloid-only fracture	6	5	1	17
Frykman I	17	15	2	12
Frykman VII	18	16	2	11
Distal radial fracture too proximal for Frykman	36[a]	32	4[a]	11
Frykman VIII	14	14	0	0
Radius and ulna fractures—paediatric	14	14	0	0
Frykman III	9	9	0	0
Frykman IV	3	3	0	0
Ulna fracture only	3	3	0	0
Radio-opaque foreign body	2	2	0	0
Radius and ulna fractures—adult	2	2	0	0
Triquetral fracture	2	2	0	0
Carpometacarpal joint dislocation	1	1	0	0
Perilunate fracture dislocation	1	1	0	0

[a]All but one were paediatric.

Three of the FPs had degenerative change (two of which had this mentioned in the report), one had an old fracture, three had both reported degenerative change and an old fracture and two had neither of these characteristics (one of which had growth plates).

Figure 3. Frequency distribution of additional anatomical/ pathological factors. No cases had "old fracture—not reported" therefore was not considered further.

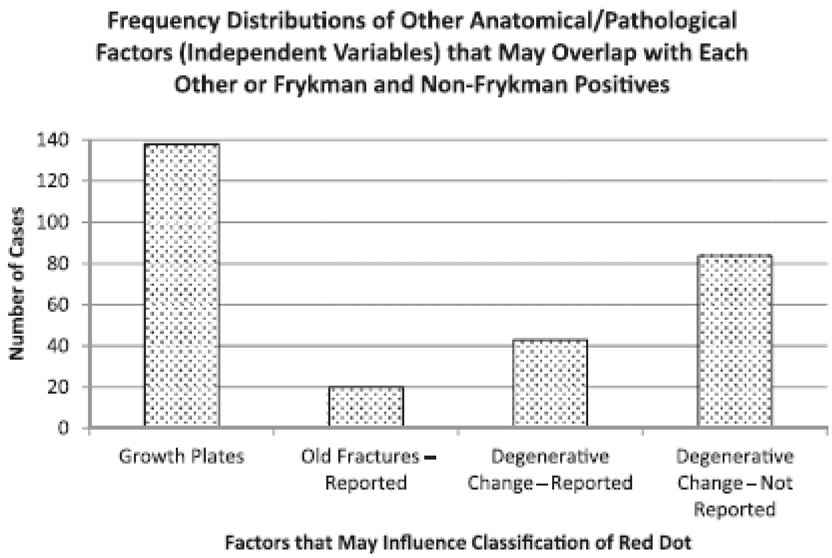


Figure 3 illustrates the distribution of the remaining factors that could influence red dot, as previously described in the literature. These were not exclusive and were present in some fracture cases (for example, a classified Frykman fracture may also demonstrate degenerative change). Some degenerative change was reported, and some not mentioned in the reports but observed by the researcher (RK). These two groups were considered separately in the analysis. Crosstabulations Crosstabulations (crosstabs) assessed each anatomical/ pathological factor against red dot classification, in nonfractured and fractured/abnormal instances. Old fractures, reported degenerative change and Frykman V fractures had higher than expected minimum counts and were significant (p=0.014, p=0.001 and p=0.062, respectively). The p-value is that of Fisher's exact test, as the assumption of expected frequencies was violated.[42] As the groups of interest are incorrect red dot classification, correct classification is shaded out (Tables 3–5).

Table 3. Crosstabulations: red dot classification * old fractures—reported ($p = 0.014$; minimum expected count = 0.57)

Non-fractured	Old fractures—reported	
	No	Yes
Red dot classification		
Incorrect	6	3
Correct	217	12

Table 4. Crosstabulations: red dot classification * degenerative change—reported ($p = 0.001$; minimum expected count = 0.95)

Non-fractured	Degenerative change—reported	
	No	Yes
Red dot classification		
Incorrect	4	5
Correct	209	20

Table 5. Crosstabulation: red dot classification * Frykman V ($p = 0.062$; minimum expected count = 2.34)

Fractured	Frykman V	
	No	Yes
Red dot classification		
Incorrect	20	5
Correct	125	10

Logistic regression analyses

Table 6 indicates the significant predictors for incorrect red dot classification, obtained by binary logistic regression analyses. Each model was statistically significant [χ^2 (14, n=398)=15.059 (non-fractured); χ^2 (14, n=398)=10.883 (fractured/ abnormal)], indicating they could distinguish between correct and incorrect red dot classifications. Again, separating fractured and non-fractured groups yielded different significant factors, which vary for FPs and "FNs". Two predictors of FP error were identified, with the corresponding logistic regression equation as follows:

$$\mathrm{Ln}\left(\frac{p}{1-p}\right) = 4.122 - 1.623 \times [\text{old fracture(reported)}] + 2.289 \times [\text{degenerative change(reported)}]$$

where p is the probability of FP error $[x^2\ (14,\ n = 398) = 10.883\ (14,\ n = 398)]$.

The ORs of old fractures—reported and degenerative change— reported were 5.070 and 9.870, respectively (Table 6); therefore, the risk of FP red dot error with old fracture—reported was 5.070 times that without this characteristic; the risk of a FP red dot error with degenerative change—reported was 9.870 times that without degenerative change (reported), controlling for other factors in the model. Thus, having degenerative change that is mentioned in the radiological report is an important predictive feature of FP image interpretation error.

Three predictors of “FN” error were identified, with the logistic regression equation as follows:

$$\mathrm{Ln}\left(\frac{p}{1-p}\right) = 2.944 - 2.251 \times (\text{Frykman V fracture presence}) + 1.846 \times (\text{Frykman VI fracture presence}) + 1.525 \times (\text{non} - \text{Frykman fracture presence})$$

where p is the probability of “FN” error $[x^2\ (14,\ n = 398) = 15.059(\text{non} - \text{fractured})]$.

The ORs of Frykman V, Frykman VI and non-Frykman positives were 9.500, 6.333 and 4.597, respectively (Table 6); thus, the risk of missing a Frykman V fracture was 9.5000 times greater than without this fracture type present and was the most important predictor of “FN” image interpretation error. Neither Frykman VI nor non-Frykman positives were significant individually in crosstabs but were identified as predictors of “FN” red dot classification when considered with all the other variables in the logistic regressions, thus the inclusion of all predictors in the model was justified.[40]

Table 6. Significant predictors for "incorrect" red dot classification from binary logistic regression models

Logistic regression model	*p*-value	Odds ratio	CI (95%)		CI (90%)	
			Lower	Upper	Lower	Upper
No acute abnormality cases (false-positive predictors)						
Old fractures—reported	0.056	5.070	0.961	26.746[a]	1.256	20.471
Degenerative change—reported	0.002	9.870	2.300	42.359	2.906	33.514
Fractured/abnormal cases (false-negative predictors)						
Frykman V	0.005	9.500	1.954	46.179	2.520	35.813
Frykman VI	0.067	6.333	0.877	45.737[a]	1.205	33.283
Non-Frykman positives	0.021	4.597	1.264	16.711	1.556	13.579

CI, confidence interval.

[a]Not significant at 95% CI as contains one between upper and lower values, but significant at 90% CI.

DISCUSSION

The red dot RADS has been contested, despite the wide use in practice.[18] The present study established red dot prevalence in a sample of A&E wrist radiographs (Figure 2), with 91.5% correct classification. This was not dissimilar to previous studies; Brown and Leschke[21] had 93.5% with ten times the sample size. This strengthens the external validity and generalizability of the study, as radiographer practice in the DTL appears to correlate with other groups. Full verification is difficult, however, as the red dot protocols in the DTL and Brown and Leschke[21] were unknown, whereas Hargreaves and Mackay[19] defined their system. Consequently, diagnostic accuracy calculations were precluded. If red dot was voluntary, it would reflect the accuracy of a non-mandatory practice and if enforced would give a true reflection of radiographer ability ("FNs" could reduce, but FPs may increase). Neither could be confirmed, highlighting a drawback of retrospective studies in this area and reflecting the provisory use of inverted commas for "FNs", incorrect and "error" throughout. Published data of enforced PCE would aid the direct comparison and evaluation of these RADS.

There was no comparable Frykman prevalence data; therefore, this study may be the first assessment. The absence of existing data means this work cannot be verified and weakens the generalizability of findings until analogous figures are compiled. Focusing on one anatomical area allowed thorough exploration of fracture type via this established classification, strengthening internal validity. Classification was assigned on assessment of each radiograph; as anatomical description rather than eponym systems such as Frykman is generally used to describe fractures in radiological reports. This researcher assessment may be subjective, but use of a recognized classification system was favoured over approximations of fracture severity to evaluate fracture type as a predictor of error. The researcher (RK) was a Masters student at the time of data collection, thus inexperience may have affected the assessment. The reporters in the DTL include reporting radiographers and radiologists of varied experience and seniority; these factors that may affect report accuracy could not be taken into account as these details were extracted to protect identity. A study with broader scope could consider such alongside anatomical and pathological factors.

In the context of RADS, this study represents the first exploration of anatomical/pathological predictors for "error". Five predictors for incorrect red dot classification were identified: "old fractures—reported" and "degenerative change—reported" for FPs, "Frykman V", "Frykman VI" and "non-Frykman positives" for "FNs". Crosstabs indicated significant predictor– outcome relationships existed within the field (Tables 3–5), however, the subsequent emergence of the latter two predictors validated the inclusion of all factors (not just those significant via crosstabs) in the logistic regression models. "Old fractures— reported" and "Frykman VI" were significant at 90% CI, whereas the others were significant at 95% CI. This threshold relaxation uncovered these further predictors, however, verification with a larger sample at 95% CI would be needed to ensure Type I error did not occur (non-significant results reported as significant) and reinforce internal validity.

Although previous research is sparse, reported degenerative change and the presence of old fractures have been associated with interpretation discrepancies.[24,37] This study supports these findings with radiographers; therefore, image interpretation training may need to focus on such areas to improve FP and "FN" rates. Given the lower frequency of reported degenerative change (Figure 3), its predictive effect may have been masked if all cases of degeneration had been considered together. The omission of unreported degenerative change from reports may reflect clinical insignificance as well as insignificance on interpretation "error". There were no unreported old fractures, thus separating this characteristic was inconsequential. Frykman V and VI are intra-articular at the DRUJ, suggesting this fracture type may be harder to spot than are extra-articular or those with RCJ articulation. Of the non-Frykman positives, metacarpal fractures had the highest "FN" percentage, followed by radial fractures too proximal for Frykman classification (Table 2). Such "misses" may be attributable to radiographers not assessing the whole radiograph, although this cannot be established without further research.

Limitations of the sample size may have prevented other predictors of error being detected, with only four and three cases in Frykman II and IV categories, respectively.[43] A larger sample may be able to collect enough paediatric cases to allow classification using the Salter and Harris[44] system. As a single group in this study, paediatric fractures contributed to 45% of non- Frykman positives (Table 2), warranting further investigation. Additionally, a larger sample may allow for individual analysis of non-Frykman positives as potential predictors for error. Despite the described limitations, the study demonstrated that fracture type affects red dot "error".

The present study has expanded the knowledge surrounding red dot, whilst stimulating several recommendations for future research. Red dot appears to be effective (91.5% correct); however, the prevalence statistics would be complemented by qualitative research to understand fully the motivation (or lack of) for radiographer participation. This would also allow insight into the "?*" annotation; whether it is a crisis of confidence or whether radiographers would rather risk FP diagnosis than miss an abnormality. Conversely, if all "FNs" were errors, it would be interesting to see whether this is due to lack of time for full image assessment, or poor image interpretation skills. This could be quantitatively evaluated with eye tracking, as used to show experience improves reporting accuracy and speed of abnormality detection.[45,46] Together with the present study, such insights may provide a strong basis to inform image interpretation training and to establish the most appropriate RADS to aid timely detection of abnormalities in the A&E setting. Extrinsic predictors such as radiographer experience, radiographer confidence, time of day and previous education could influence red dot error. These were not considered in this study and could be the focus of future work.

CONCLUSION

This study has established red dot prevalence and identified predictors for incorrect classification. The unknown protocol leaves questions regarding participation, but these predictors may inform image interpretation training to reduce both "FN" and FP errors and allow radiographers and other health professionals to impact more positively on patient diagnoses. There are numerous additional avenues for future study, which together would expand the understanding of "error" to evaluate the hotly debated RADS

REFERENCES

1. Guly H. Diagnostic errors in an accident and emergency department. Emerg Med J 2001; 18: 263–9.

2. Bond M, Pavey T, Welch K, Cooper C, Garside R, Dean S, et al. Systematic review of the psychological consequences of falsepositive screening mammograms. Health Technol Assess 2013; 17: 1–170, v–vi. doi: 10.3310/hta17130

3. Touquet R, Driscoll P, Nicholson D. Teaching in accident and emergency medicine: 10 commandments of accident and emergency radiology. BMJ 1995; 310: 642–5.

4. Berman L, de Lacey G, Craig O. Survey of accident and emergency reporting: results and implications. Clin Radiol 1985; 36: 483–4.

5. Vincent CA, Driscoll PA, Audley RJ, Grant DS. Accuracy of detection of radiographic abnormalities by junior doctors. Arch Emerg Med 1988;5: 101–9.

6. McLauchlan CA, Jones K, Guly HR. Interpretation of trauma radiographs by junior doctors in accident and emergency departments: a cause for concern? J Accid Emerg Med 1997; 14: 295–8.

7. Gleadhill DN, Thomson JY, Simms P. Can more efficient use be made of x ray examinations in the accident and emergency department? Br Med J (Clin Res Ed) 1987; 294: 943–7.

8. Petinaux B, Bhat R, Boniface K, Aristizabal J. Accuracy of radiographic readings in the emergency department. Am J Emerg Med 2011; 29: 18–25. doi: 10.1016/j. ajem.2009.07.011

9. National Confidential Enquiry into Patient Outcome and Death (NCEPOD). Emergency admissions: a journey in the right direction? Executive summary. London, UK: National Confidential Enquiry into Patient Outcome and Death; 2007.

10. Royal College of Radiologists (RCR). Standards for providing a 24-hour diagnostic radiology service. London, UK: Royal College of Radiologists; 2009.

11. Nicholson I. Changing out-of-hours radiology practice: a response by the Royal College of Radiologists to five NCEPOD reports. London, UK: Royal College of Radiologists; 2009.

12. Brealey S, Scally A, Hahn S, Thomas N, Godfrey C, Coomarasamy A. Accuracy of radiographer plain radiograph reporting in clinical practice: a meta-analysis. Clin Radiol 2005; 60: 232–41.

13. Hardy M, Hutton J, Snaith B. Is a radiographer led immediate reporting service for emergency department referrals a costeffective initiative? Radiography 2013; 19: 23–7.

14. Swinburne K. Pattern recognition for radiographers. Lancet 1971; 1: 589–90.

15. Snaith B, Hardy M. Radiographer abnormality detection schemes in the trauma environment—an assessment of current practice. Radiography 2008; 14: 277–81.

16. Berman L, de Lacey G, Twomey E, Twomey B, Welch T, Eban R. Reducing errors in the accident department: a simple method using radiographers. Br Med J (Clin Res Ed) 1985; 290: 421–2.

17. Kelly BS, Rainford LA, Gray J, McEntee MF. Collaboration between radiological technologists (radiographers) and junior doctors during image interpretation improves the accuracy of diagnostic decisions. Radiography 2012; 18: 90–5.

18. Society and College of Radiographers (SCoR). Preliminary clinical evaluation and clinical reporting by radiographers: policy and practice guidance. London, UK: Society and College of Radiographers; 2013.

19. Hargreaves J, Mackay S. The accuracy of the red dot system: can it improve with training? Radiography 2003; 9: 283–9.

20. Brealey S, Scally A, Hahn S, Thomas N, Godfrey C, Crane S. Accuracy of radiographers red dot or triage of accident and emergency radiographs in clinical practice: a systematic review. Clin Radiol 2006; 61: 604–15.

21. Brown N, Leschke P. Evaluating the true clinical utility of the red dot system in radiograph interpretation. J Med Imaging Radiat Oncol 2012; 56: 510–13. doi: 10.1111/ j.1754-9485.2012.02398.x

22. Ruutiainen AT, Durand DJ, Scanlon MH, Itri JN. Increased error rates in preliminary reports issued by radiology residents working more than 10 consecutive hours overnight. Acad Radiol 2013; 20: 305–11. doi: 10.1016/ j.acra.2012.09.028

23. Davenport MS, Ellis JH, Khalatbari SH, Myles JD, Klein KA. Effect of work hours, caseload, shift type and experience on resident call performance. Acad Radiol 2010; 17: 921–7. doi: 10.1016/j.acra.2010.03.006

24. Kim SJ, Lee SW, Hong YS, Kim DH. Radiological misinterpretations by emergency physicians in discharged minor trauma patients. Emerg Med J 2012; 29: 635–9. doi: 10.1136/emj.2011.111385

25. Cosson P, Willis N. Digital teaching library (DTL) development for radiography education. Radiography 2012; 18: 112–16.

26. Diaz-Garcia RJ, Oda T, Shauver MJ, Chung KC. A systematic review of outcomes and complications of treating unstable distal radius fractures in the elderly. J Hand Surg Am 2011; 36: 824–35.e2. doi: 10.1016/j. jhsa.2011.02.005

27. Leversedge FJ, Srinivasan RC. Management of soft-tissue injuries in distal radius fractures. Hand Clin 2012; 28: 225–33. doi: 10.1016/j.hcl.2012.03.005

28. Ryan LM, DePiero AD, Sadow KD, Warmink CA, Chamberlain JM, Teach SJ, et al. Recognition and management of pediatric fractures by pediatric residents. Paediatrics 2004; 114: 1530–3.

29. Moore CM, Leonardi-Bee J. The prevalence of pain and disability one-year post-fracture of the distal radius in a UK population: a crosssectional survey. BMC Musculoskelet Disord 2008; 9: 129. doi: 10.1186/1471-2474-9-129

30. Frykman G. Fractures of the distal radius including sequelae-shoulder–handfinger syndrome, disturbance in the distal radio-ulnar joint and impairment of nerve function. A clinical and experimental study. Acta Orthop Scand 1967; 38(Suppl. 108): 1–61.

31. AO (Arbeitsgemeinschaft fur Osteosynthesefragen) Foundation. Muller AO classification of fractures—long bones (AOE-E1-018.5). Davos, Switzerland: AO Foundation; 2010.

32. Illarramendi A, Gonzalez Della Valle A, Segal E, De Carli P, Maignon G, Gallucci G. Evaluation of simplified Frykman and AO classifications of fractures of the distal radius. Assessment of interobserver and intraobserver agreement. Int Orthop 1998; 22: 111–15.

33. Belloti JC, Tamaoki MJ, Franciazi CE, Santos JB, Balbachevsky D, Chap Chap E, et al. Are distal radius fracture classifications reproducible? Intra and interobserver agreement. Sao Paulo Med J 2008; 126: 180–5.

34. Barton T, Chambers C, Bannister G. A comparison between subjective outcome score and moderate radial shortening following a fractured distal radius in patients mean age 69 years. J Hand Surg Eur Vol 2007; 32: 165–9.

35. Gufler H, Schulze CG, Wagner S, Baumbach L. MRI for occult physeal fracture detection in children and adolescents. Acta Radiol 2013; 54: 467–72. doi: 10.1177/ 0284185113475606

36. Verhaven E, De Boeck H, Opdecam P. Osteosarcoma appearing as a pathologic fracture. Acta Orthop Belg 1991; 57: 437–41.

37. Maheshwari J. Approach to a patient with limb injury. In: Maheshwari J, ed. Essential orthopaedics. 4th edn. New Delhi, India: JP Medical Ltd; 2011. pp. 35–40.

38. IBM Corporation. IBM SPSS statistics for windows, version 21.0. Armonk, NY: IBM Corporation; 2012.

39. Zeigler-Hill V. Logistic regression. [Updated 11 February 2013; cited 9 July 2014.] Available from: http://www.zeigler-hill.com/uploads/ 7/7/3/2/7732402/psy_512_logistic_regression. Pdf

40. Menard S. Model specification, variable selection, and model building. In: Menard S, ed. Logistic regression—from introductory to advanced concepts and applications. Thousand Oaks, CA: Sage Publications Inc.; 2010. pp. 105–24.

41. Johnson N. Upper limb. In: Bontrager KL, Lampignano JP, eds. Textbook of radiographic positioning and related anatomy. 7th edn. St Louis, MO: Mosby Elsevier; 2010. pp. 123–70.

42. Field A. Categorical data. In: Field A, ed. Discovering statistics using IBM SPSS statistics. 3rd edn. London, UK: Sage Publications Inc.; 2009. pp. 686–724.

43. Batterham AH, Atkinson G. How big does my sample need to be? A primer on the murky world of sample size estimation. Phys Ther Sport 2005; 6: 153–63.

44. Salter RB, Harris WR. Injuries involving the epiphyseal plate. J Bone Joint Surg Am 1963;

45: 587–622. 45. Reed WM, Ryan JT, McEntee MF, Evanoff MG, Brennan PC. The effect of abnormality— prevalence expectation on expert observer performance and visual search. Radiology 2011; 258: 938–43. doi: 10.1148/ radiol.10101090

46. Wood G, Knapp KM, Rock B, Cousens C, Roobottom C, Wilson MR. Visual expertise in detecting and diagnosing skeletal fractures. Skeletal Radiol 2013; 42: 165–72. doi: 10.1007/s00256-012-1503-5