Introduction

Acute abdominal pain, out of the trauma setting, is a common presenting symptom in the emergency department, with a wide spectrum of underlying causes[1].Abdominopelvic computed tomography (CT) has assumed an increasingly important role in the evaluation and diagnosis of these patients and is widely used as an integral part of surgical triage [2, 3, 4]. Abdominopelvic CT, although highly accurate in the assessment of the acute abdomen, can be challenging to report particularly in patients who are acutely unwell. The rapid increase in utilisation of CT, particularly out of hours, has created reporting pressures within United Kingdom (UK) radiology departments and this has led to the development of different reporting models. Provisional (initial) CT reports may be issued by trainee radiologists (registrars) with subsequent review by senior onsite Consultant radiologists, or reports may be issued by Consultants themselves. Alternatively in many departments reporting may be carried out by radiologists working offsite with no affiliation to the department where the imaging occurs. Offsite reporters are typically of Consultant level or equivalent, but may not be trained or working within the United Kingdom. Offsite reporting is particularly utilised out of hours, a practice that is well recognised across Europe and North America. Both registrar and offsite reports may be supplemented by an addendum report provided later by an onsite Consultant. Alongside these changes in practice has been recognition of the concept of radiological "error", more often referred to as "discrepancy" and the relationship of a discrepant report to potential or actual harm to the patient[5, 6].

Emergency abdominal or abdominopelvic CT performed out of hours in acutely ill patients is a complex investigation with the potential to impact positively or negatively on patient outcomes depending on the accuracy and timeliness of the report. Current UK reporting models involve radiologists of varying expertise and experience, some of whom are offsite and remote to both the patient and clinical interaction.

The aims of this national, UK-wide audit on acute non-traumatic abdominopelvic CT reporting in surgical and non-surgical groups included:

- Assessment of major/minor discrepancy rates for provisional (initial) and also addendum (supplementary) reports in unselected patients across a wide range of institutions.
- Examine factors affecting major discrepancy rate at the level of the provisional report.
- Examine reporting factors affecting cases of major discrepancy where patients came to harm and also to assess the nature of the harm.
- Obtain sensitivity and specificity of CT in the more common pathologies in both surgical and non-surgical patient groups
- Document any added value of a Consultant addendum report and to evaluate the availability of provisional and addendum reports pre-operatively in the surgical group.

Materials and Methods

The Royal College of Radiologists (RCR) works closely with individual radiology departments across the four countries within the United Kingdom, nominated individuals/fellows within the departments are responsible for co-ordinating both local audits and national RCR audit projects. As part of this emergency CT abdominal reporting audit all departmental audit leads were contacted by email and invited to participate and submit audit data to the RCR on behalf of their departments. Formal ethical approval for this type of study is not required in the UK as all submitted data is anonymised and only used to promote best medical practice.

Departments were requested to submit patient data in both non-surgical and surgical cohorts. Those departments with no onsite general surgery only submitted data in the non-surgical group. Access to relevant patient data on PACS (Picture Archiving and Communication System), RIS (Radiology Information System) and the patient record was necessary for inclusion.

Non-surgical group

A retrospective search was undertaken to identify 25 consecutive non-traumatic adult (> 16 years) emergency patients who underwent abdominopelvic CT from 1st January 2013 onwards from the radiological department database. The patients all had out of hours (6pm-8am weekdays or anytime at the weekend) emergency abdominal or abdominopelvic CT but no subsequent laparotomy. Patients who had another intervention during this admission e.g. colonic/JJ stent, percutaneous drainage, laparoscopy (to include laparoscopic surgical interventions) were included in this category. Patients who underwent non-contrast CT for suspected renal calculus were excluded from the audit.

Surgical group

Retrospective identification from 1st January 2013 onwards of 25 consecutive non-traumatic adult patients who had out of hours abdominal/abdominopelvic CT and had subsequent laparotomy. For the purposes of the audit it was expected that for the majority of patients CT would have been

performed within 24-48 hours pre-laparotomy. However patients could still be included if the time interval was greater than 48 hours but the CT deemed pertinent to that episode of care.

Data Collection

Data were entered into the Microsoft Office Excel 2007 spreadsheets – "Institutional", "non-surgical" and "surgical" group questionnaires. Auditors were able to toggle between the three questionnaires and also to access a drop down glossary of expanded terms (diagnoses) for truncated items in the drop down lists. These three questionnaires would provide a range of contextual data which would then be used to explore potential relationships to the chosen audit standards. Details of the three questionnaires are included in Appendix A.

The institutional questionnaire was used to assess more generic aspects relating to CT reporting out of hours, including the use of radiology trainees/registrars and offsite reporters in the provision of on call reports as well as onsite hospital based Consultants. The institutional questionnaire also explored availability of more specialised gastrointestinal (GI) radiology onsite, either as primary or supplementary/addendum reporter. A GI interest was classified as a radiologist with formal GI reporting sessions and involvement in GI multi-disciplinary team meetings (MDT); GI subspecialty interest was defined as a minimum of 5 sessions of GI reporting per week.

The "non-surgical" and "surgical" questionnaires looked more specifically at the patient journey, examining the diagnosis of the provisional report, nature of provisional (initial) reporter, presence/absence of an addendum report and its concordance with the provisional report. Correlation of provisional report with laparotomy findings was assessed in the surgical group together with the presence/absence of a pre-operative provisional and/or addendum report either documented in the patient notes or validated on the RIS system.

For any given case the questionnaires only allowed the auditor to select a single and representative major/minor diagnosis. The presence of additional secondary diagnoses could be selected but not itemised.

Auditors were instructed to select the provisional +/- addendum report diagnosis, recording their own auditor diagnosis if non-concordant and then also record the laparotomy diagnosis. This process would allow recording of major/minor discrepancy between reports by the auditor, also the type of discrepancy and using patient/radiology records to assess any harm that may have come to the patient. Correlation with surgical findings would also be undertaken.

Drop-down lists were widely used to facilitate data entry, data validation configuration restricted data entry to valid responses. Cell references in formulae enabled summary responses to be displayed and updated automatically. The questionnaires were initially piloted amongst members of the RCR audit committee to evaluate content and to confirm ease of use. The identity of respondents in terms of, a) teaching or district general hospital b) region of the UK, was used to evaluate potential bias between respondents and non-respondents.

Responses were incomplete in some parts of the questionnaires, with such data recorded as “no response”.

The CT Auditor

For the purposes of the audit it was proposed that the auditor evaluating provisional and addendum reports and the CT findings would be a substantive Consultant working onsite in the auditing institution. This individual should have experience in reporting abdominal CT, in cases of potential major discrepancy it was recommended that there should be case review with another onsite Consultant colleague, preferably with an interest in GI/abdominal radiology and a consensus reached.

It was specified within the audit proforma that the CT auditor should review the CT images blinded to original report content/reporter identity +/- surgical findings and then review the CT reports (provisional +/- addendum) and record concordance/discrepancy and their own diagnosis in cases of discrepancy. The CT auditor would then review the patient notes/RIS in surgical patients to determine presence/timing of a record of the provisional/addendum report and would also review provisional/addendum report findings compared to laparotomy findings in surgical patients.

Discrepancies

A major discrepancy comprised a change, or potential change in diagnosis or treatment as a result of either addendum report or CT auditor review. A minor discrepancy occurred where there were minor issues in provisional/addendum reports unlikely to result in harm or change in management. Major discrepancies were coded as false positive (provisional report diagnosis positive findings, negative on auditor review); false negative (provisional report negative diagnosis, positive findings on auditor review); misdiagnosis (incorrect provisional diagnosis); or indeterminate report (an indeterminate report defined as an inappropriately wide range of differential diagnoses, containing the correct diagnosis(es) but with no attempt at triaging the diagnoses or guiding the clinician to the most likely explanation for CT findings).

A dropdown menu also allowed grading of each case into;

- Major discrepancy patient came to harm (harm might include death, unnecessary intervention (e.g. colonoscopy, endoscopy, drainage), delay in diagnosis or treatment.
- Major discrepancy patient did not come to harm.
- Major discrepancy, outcome uncertain
- Minor discrepancy
- Concordance with reports, no issues of concern

Audit Standards

The derivation of audit standards followed similar practice previously outlined for RCR national audits[7]. It is established practice within the RCR to review all available literature and to adopt a standard/set of standards that is considered by the RCR audit committee to be both practical and achievable in everyday clinical practice.

The selected audit standards are included in Table 1 (compliance with standards is also documented in this table). The standards were derived following careful evaluation of relevant, current published literature taking into account the differing clinical scenarios and definitions of discrepancy included in these publications[8-20]. A comparison of the national audit findings against these standards were expressed as counts and percentages. A search of all available published literature (from 1950 onwards) was undertaken using the MEDLINE and National Health Service evidence (including the Cochrane library of systematic reviews and the National library of guidelines) to establish supporting literature and confirm/derive figures for the audit standards and made available during audit committee deliberations.

For the purposes of the audit there were three main groups providing provisional (initial) CT reports-

1) Registrar (trainee radiologist)
2) Offsiter (radiologist working remotely for an outsourcing agency at Consultant level or equivalent)
3) Consultant radiologist onsite (may or may not have GI radiology expertise)

Addendum (supplementary) reports to initial, provisional reports are provided by hospital-based onsite Consultant radiologists with varying degrees of GI radiology expertise.

Statistical analysis

Exploratory analyses of all variables in the three questionnaires (institutional, non-surgical and surgical) were performed to identify any significant variables that might predict discrepancy of the

provisional report with the auditor review. The auditor was used as the reference standard. The variables investigated included:

- Nature of provisional reporter (registrar, onsite Consultant or offsiter)
- Effect of registrar discussion of case with onsite Consultant (if documented)
- Effect of type of onsite Consultant (General vs GI radiologist) and also presence/absence of GI radiologist onsite.
- Effect of district general hospital vs teaching hospital
- Effect of availability of on-call registrar reporting of CT (present/absent)
- Effect of availability of on-call CT reporting by on site Consultants
- Effect of availability of on-call reporting of CT by offsite radiologists

For each of these variables major discrepancy risk ratios (95% CI) were estimated from generalised linear models with a binary outcome and log link, with robust standard errors to allow for non-independence of results from the same hospital. Separate models were first fitted to the surgical and non-surgical group data. A model was then fitted to the combined data: this allowed risks to differ in the surgical and non-surgical groups, as well as according to the variable being investigated. The model was also extended to allow for interactions (i.e. allowing the discrepancy rate ratios to differ between the surgical and non-surgical groups). For institutional comparisons (i.e. district general hospitals vs teaching hospital) further models were fitted adjusting for differences in the proportions of registrar, Consultant and offsite reports by including appropriate indicator variables as covariates in the models. An analogous series of models was used to analyse risk ratios for any discrepancy (major and minor combined).

For those subjects where an addendum report was available conditional logistic regression modes (with robust standard errors that allowed for non-independence of results from the same hospital) were used to investigate the value of the addendum report. The paired outcomes compared by the model were i) whether or not there was a major discrepancy between the provisional report and the

auditor ii) whether or not there was a major discrepancy between the addendum report and the auditor. Analogous analyses were performed for any discrepancy (major and minor combined). Sensitivity and specificity calculations were undertaken in relation to the ten most commonly occurring diagnosed pathologies in both the surgical and non-surgical groups. To allow for non-independence of results from the same hospital in Table 1 (compliance with standards) 95% confidence intervals for percentages were computed using the bootstrap: specifically, non-parametric, bias corrected and accelerated 95% confidence intervals were calculated from 100,000 bootstrap samples clustered by hospital.

~~Finally, sensitivity/specificity calculations were undertaken in relation to the most common pathologies in the surgical and non-surgical groups (pathology identified from provisional report if concordant with auditor, if not concordant then derived from the auditor or laparotomy diagnosis). Definitions for true positive, true negative, false positive and false negative are included in Appendix B. Two additional terms are used (see result tables and appendix B). Non-concurrence with indication of diagnosis (NCID) – the provisional CT report contains the diagnosis in question when compared to auditor/laparotomy findings, but the provisional report diagnosis is part of an indeterminate report and thereby recorded as non-concurrence. The second term is non-concurrence with no indication of diagnosis (NCNID) – in these cases neither the provisional nor auditor/laparotomy diagnoses contain the diagnosis in question, but there is also non-agreement between provisional and auditor/laparotomy findings. So, for example in NCNID, looking at cases negative for appendicitis, the provisional report and auditor/laparotomy would contain a diagnosis other than appendicitis but differing also from one another, so not true negatives for appendicitis for the purposes of the audit. NCID and NCNID cases were excluded from calculations. Bootstrap 95% confidence intervals (non-parametric, bias corrected and accelerated) for sensitivities and specificities were computed from 100,000 bootstrap samples clustered by hospital.~~

Results

The complete responses to the three audit questionnaires together are included in Appendix A. A total of 109/188 eligible departments responded to the audit (58%). Summary results of the institutional questionnaire/departmental demographics are included in Table 2. Case demographics are included in Table 3 – note 4931 patients were included in the audit (2568 non-surgical group, 2363 surgical group; 48% male, 52% female). Table 3 also includes information on the source of the CT request and the location and seniority of the provisional reporters.

In 179/887 (20.2%) provisional registrar reports there was evidence of discussion with an onsite Consultant radiologist documented in the provisional report.

Ninety- five departments submitted 25 cases in the non-surgical group, the remainder submitted 24. In the surgical group one department submitted 26 cases, 78 departments submitted 25 cases and the remainder between 4 and 24.

The identity of responding departments and hospitals were reviewed. The percentage of departments participating from teaching hospitals in England differed from district general hospitals by <1%. The geographic distribution of departments differed by 10.8% when respondents were compared with non-respondents in England. In Northern Ireland, the difference was 3.5%, in Wales 1.3% and in Scotland 8.5% however this was not statistically significant.

Overview of CT report Concordance

A detailed overview of these data is included in Appendix B.

- Non-surgical group

1947 patients had a provisional CT report with no evidence of addendum and of these there was concordance with the auditor in 1782 patients.

621 patients had evidence of an addendum report with provisional, addendum and auditor reports concordant in 472 patients. Varying levels of discordance were noted in the remaining patients (Appendix B) with the most prominent category being auditor concordance with addendum and not with provisional (75 patients).

- Surgical Group

1728 patients had a provisional CT report with no evidence of an addendum and of these the provisional report was concordant with the auditor in 1557 patients. In 1423/1557 there was also agreement with laparotomy.

635 patients had evidence of an addendum report with provisional, addendum and auditor reports concordant in 510 of these patients. Varying levels of discordance were noted in the remaining patients (Appendix B) with again the most prominent category being auditor concordance with addendum and not with provisional (72 patients). In the 510 patients with concordance of all 3 reports there was also agreement with laparotomy findings in 471 (39 disagreed).

Nature of discrepancies and patient harm

Summary characteristics and analysis by type of major discrepancies are included in Table 4. The number of additional incorrect secondary major diagnoses was greater in provisional (6 non-surgical, 10 surgical) than addendum reports (1 in each group). There were single incidents of additional indeterminate reporting in non-surgical provisional and addendum reports and in surgical provisional reports.

- Non-surgical group

In 47 patients there was evidence on notes/imaging review of subsequent additional procedures that may have been unnecessary following a major discrepancy. These were predominantly additional imaging procedures, but also included CT/ultrasound guided drainage (3 patients), laparoscopy (3 patients) and endoscopy (3 patients).

15/72 patients with provisional report major discrepancy were considered by the auditor to have come to harm as a result of the report: Delay in diagnosis (7 patients), delay in treatment (7 patients), unnecessary investigations (2 patients) and unspecified (1 patient).

- Surgical group

36/132 patients with provisional report major discrepancy were considered by the auditor to have come to harm as a result of the report and were detailed as follows: Delay in diagnosis (3 patients), delay in surgery (24 patients), unnecessary investigations (1 patient) and unnecessary surgery (8 patients).

Results of statistical analyses

The full results of all analyses are included in Appendix C.

Predictors of provisional agreement with auditor (pooled non-surgical and non-surgical data)

Table 5 shows risks of major discrepancy for onsite Consultants, radiology registrars and offsite reporters separately in the surgical and non-surgical groups. Overall risks of major discrepancy were 5.6% in the surgical group and 2.8% in the non-surgical group. In each group major discrepancy risks were highest in offsite reporters and lowest in onsite Consultants, although these between group differences only achieved statistical significance in the surgical group (p=0.0003). There was no evidence that the major discrepancy risk ratios differed between the two groups (p= 0.36) suggesting results could be pooled. In the combined analysis major discrepancy risks were 44% higher (95% CI 5% lower to 118% higher) in registrars than onsite Consultants and 181% higher (95% CI 75% to

351% higher) in offsite reports than registrars (p=0.0001, joint test of differences). Restricting to major discrepancies where the patient came to harm numbers were reduced but the pattern of results was similar (sections 1.1.5 to 1.1.7 in Appendix C); for the pooled analysis the joint test of differences among the three groups was borderline statistically significant (p=0.061) with risks statistically significantly higher for the offsite group compared to the onsite Consultants (p=0.018). A similar pattern of discrepancy risk ratios was seen when all discrepancies, not just major discrepancies, were considered (sections 1.1.4 in Appendix C).

There was little evidence of differences in risks of discrepancy according to whether or not registrars discussed their interpretations with a Consultant (section 1.2 Appendix C). Among Consultants, discrepancy risks were lower in those with a GI interest or a GI sub-specialty than in those without such specialisation (section 1.3 Appendix C). Combining the two specialist groups, risk of a major discrepancy was 28% lower (95% CI 57% lower to 21% higher) and risk of discrepancy was 32% lower (95% CI 5% to 51%), with this latter difference achieving statistical significance (p=0.022).

Looking at institutional comparisons there was no evidence of differences in discrepancy risks between district general hospitals and teaching hospitals (section 1.4 Appendix C). There was evidence that major discrepancy risk ratios were higher in hospitals where on call registrar reporting was available (risks increased by 76% (95% CI 9% to 184%, p=0.021) in the pooled analysis). However, this difference was much reduced in magnitude and became non-statistically significant when adjusted for registrar/onsite Consultant/offsiter imbalances between institutions (section 1.5 Appendix C).

There was also evidence that major discrepancy risk ratios were higher in hospitals where on- call CT reporting by onsite Consultants was available (section 1.6 Appendix C). In both the non-surgical and surgical groups major discrepancy risks were lowest (2.3% in the non-surgical group, 3.8% in the surgical group) when on-call CT reporting by an onsite Consultant was fully available. When this was partially or not available risks were higher (3.4% and 3.6% respectively in the non-surgical group,

8.0% and 8.2% in the surgical group) although these between group differences only achieved statistical significance in the surgical group (p=0.0093). There was no evidence that the major discrepancy risk ratios differed between the two groups (p=0.56) suggesting results could be pooled. In the combined analysis major discrepancy risks were 85% (95% CI 20% to 188%) higher when on-call CT reporting was partially available and 90% (95% CI 6% to 239%) higher when this was not available compared to when it was fully available. These differences were somewhat reduced in magnitude when adjusted for registrar/onsite Consultant/offsiter imbalances between institutions with the overall test of adjusted differences between groups being only borderline statistically significant (p=0.066).

There was also evidence that discrepancy risks were higher when on call CT reporting was carried out by off- rather than onsite radiologists (risks increased by 61% (95% CI 6% to 145%, p=0.025) in pooled analysis). However, this difference was again reduced in magnitude and became non-statistically significant when adjusted for registrar/onsite Consultant/offsiter imbalances between institutions (section 1.7 Appendix C). There was no evidence that the availability of a speciality GI radiologist onsite, or that routine onsite Consultant review of outsourced CT on-call reports was associated with risks of discrepancy (sections 1.8 and 1.9 Appendix C).

Table 6 shows where discrepancy occurred between addendum, provisional and auditor reports in the subset of the data where an addendum report was available. There are five eventualities: all reports can agree, all can disagree, or any pair can agree whilst disagreeing with the third. The net benefit of the addendum can be assessed by comparing the number of occasions when the auditor agrees with the addendum but not the provisional with the number of occasions where the auditor agrees with provisional but not the addendum. In the non-surgical group there is net benefit from switching to an addendum report in terms of major discrepancies (19 resolved, 3 introduced) and in terms of all discrepancies (75 resolved, 26 introduced). Using conditional logistic regression both differences are statistically significant (p=0.006 major discrepancy, p<0.0001 all discrepancies).

In the surgical group there is again a strong net benefit in switching to an addendum, both in terms of major discrepancies (45 resolved, 2 introduced) and all discrepancies (72 resolved, 13 introduced). Using conditional logistic regression both differences are statistically significant ($p<0001$).

Availability of results pre-operatively

A written or validated RIS provisional report was available pre-operatively in 98.3% of patients (Table 1). A written or validated addendum report was only available pre-operatively in 64.3% of patients. In 45 patients with a major discrepancy at provisional report level the discrepancy was corrected at addendum. In 14/45 of these cases the addendum was not available pre-operatively; hence there were 14 cases of potentially avoidable major discrepancy (only 1/14 patients came to harm).

"Normal" CT and Laparotomy Findings

Twenty-two patients had a "normal" laparotomy. Of these 10 patients also had a "normal" provisional CT report with pathology reported in 12 patients (including cases of ischaemia, Crohn's disease, appendicitis, colitis). Twenty-three patients had a "normal" CT report and still proceeded to laparotomy. Of these patients 10 also had a normal laparotomy with pathology found in the remaining 13 (including 3 cases of appendicitis, 3 cases of ischaemic bowel, 1 abscess and 2 small bowel obstructions, "no response" in 4 patients).

Overall compliance with audit standards

These are documented in Table 1.

Overall registrars met the audit standard for correlation of provisional report with laparotomy (standard >80%, achieved 83.7%) but onsite Consultants narrowly missed their standard (standard >90%, achieved 87.2%). Offsite radiologists missed their target by a larger margin (standard >90%, achieved 78.9%).

Sensitivity/Specificity Data for the common pathologies

For results of these calculations please see appendix D.

~~The ten most common provisional report CT diagnoses were selected from the non-surgical and surgical groups. Sensitivity/specificity calculations were then undertaken using the auditor final diagnosis as reference standard (for definitions see earlier and also appendix C). Non-surgical results are found in table 7, surgical results in table 8. In addition, the ten most common provisional report CT diagnosis sensitivity/specificity calculations were then repeated, but using the laparotomy diagnosis as the reference standard (see table 9). CT was most sensitive in the diagnosis of appendicitis using both the auditor and laparotomy as reference standard (96.4%, 95.6% respectively). There was a considerable drop off however noted in relation to the diagnosis of ischaemic bowel when using the auditor as reference standard (89.5%) as opposed to laparotomy (72.5%).~~

Discussion

Discrepancy in radiological reporting is a complex issue and the causes of discrepancy are numerous, well recognised and often inter-related [5, 6]. Radiologist specific causes include faulty reasoning, lack of knowledge (particularly when working outside an individual's area of specialty expertise), failure of perception or poor communication of findings. System related factors are also important and a number of causes are recognised – staff shortages (with over reliance on locum radiologists), combined with excess workload, inexperience of staff and insufficient or inaccurate clinical and/or previous radiological information[5].

The investigation of discrepancy rates and related causes in radiology has been the subject of numerous publications with an emphasis on radiology registrar reporting, trauma and cranial CT[21-25]. There is variation in published rates for discrepancy in CT abdominal reporting and again these papers predominantly assess registrar reporting and there are differences in sample sizes and also

definitions of discrepancy. Allowing for this discrepancy rates for CT abdomen reporting range widely from < 0.1% to 18%[8, 12, 14, 17, 26-31]. A meta-analysis looking at discrepancy rates in adult CT (all types and including elective and emergency) demonstrated overall no significant differences in rates of discrepancy between a registrar and more senior radiologists, with a pooled discrepancy rate for abdominopelvic CT of 2.6%[32]. A recent study looking at abdominal CT in surgical patients found a 14% rate (146/1071 reports) of clinically important management changes following double/expert reading of initial CT reports[33]. There is a relative paucity of published literature pertaining to discrepancy in outsourced, offsite radiology, a large series published in 2005 looking at a radiological group practice, reported a discrepancy rate of 2.1% for CT of the abdomen/pelvis[20].

The demand for access to radiology services continues to increase year on year in the UK. Due to its high diagnostic accuracy and increased availability CT has experienced a rapid expansion in its roles both in and out of hours; a growth of 141% in CT scans was reported in the USA over a 10 year period[34]. Unfortunately the increased diagnostic imaging workload has not been matched by an increase in reporting radiologists. This is a situation which is particularly acute in the UK, but is also recognised worldwide. It is challenging to maintain a 24 hour service, 7 days a week and to ensure that emergency imaging, in particular CT, is reported in a timely and accurate manner. These service challenges have led to the development of other reporting models – registrars often provide the first tier of reporting, however increasingly hospitals have been looking at offsite/outsourced radiology reporting solutions, particularly during antisocial hours and weekends. Outsourcing is now widely used in the UK, but it is a worldwide phenomenon with remote reporting hubs in India, Australasia, Europe and the USA[35, 36].

Our study incorporated 4931 patients from 108 United Kingdom radiology departments. It explored in detail factors that might be related to increased major discrepancy at the level of the provisional (initial) radiology report on review by a CT auditor. When compared to an onsite Consultant there was a statistically significant increased risk of major discrepancy and major discrepancy with harm in

an offsite/outsourced CT report, this finding was consistent in both surgical and pooled data. Major discrepancy was also found to be more likely in the surgical group; registrars had a major discrepancy rate intermediate between onsite Consultants and offsite reporting radiologists. These findings are also reflected in failure of compliance with the major discrepancy audit standards. Offsite reporters narrowly missed the non-surgical major discrepancy standard (standard <5%, achieved 5.2%) but also missed the surgical major discrepancy (standard <5%, achieved 12.7%) and pooled (standard <5%, achieved 8.7%) standards. Both registrars and onsite Consultants were able to meet the recommended provisional report standards for major discrepancy. The overall major discrepancy rate (patient came to harm) standard was also missed in the surgical group (standard <1%, achieved 1.5%).

These results do raise important questions and there are no immediate or straightforward solutions. It is clear in the UK at least that the national shortage of Consultant radiologists is going to persist with no short or medium term answers to the workforce shortfall. Hence the offsite and outsourced solution is not only attractive but has become a necessity in many hospitals. The issue of concern is maintaining quality in the outsourced arena. Many of the reported factors associated with increased risk of discrepancy are particularly relevant to a remote, offsite reporter[5]. Factors particularly affecting offsite reporters include: excess workload, fatigue, exposure to a wide range of studies for reporting not reflecting their specialty training and experience, lack of clinical contact and clinical information, lack of access to previous imaging and problems with communication. There is no doubt that the presence of local clinical networks, were radiologists work closely with surgical teams, can enhance the quality of CT reporting. Lack of access to these established networks is a significant disadvantage to radiologists reporting remotely [37]. Close in-person collaboration between the reporting radiologist and the surgical team is associated with significant and also frequent changes in patient management, even when the radiological report is correct and contains the necessary diagnostic information [38]. There is another important potential side-effect of increased utilisation of outsourced reporting, namely reduced exposure of radiology trainees to on-call

experience. When managed appropriately involvement in out-of-hours, emergency radiology is an invaluable part of radiology training. Reducing this exposure, outsourcing is one important cause of this, is likely to have a significant and deleterious effect on training the radiologists of the future [39].

The audit also evaluated the addendum/supplementary report and availability of reports in surgical patients pre-operatively. This was partly in response to the recently published UK National Emergency Laparotomy Audit [40], which highlighted deficiencies in Consultant radiologist reported abdominopelvic CT prior to surgery (53%). This laparotomy audit did acknowledge that 24 hour contemporaneous reporting was available at all hospitals in the audit offering laparotomy, though the grade of reporting radiologist was not specified.

In our study a written/validated RIS provisional report was available pre-laparotomy in 98.3% of patients (standard 100%), but only 64.3% of addendum reports were available pre-operatively (standard 100%). Of note the majority of departments in the audit offered secondary review of registrar provisional reports with the issuing of an addendum, usually by the rostered CT Consultant radiologist the next morning. The majority (22/38) of departments utilising offsite CT reporters do not routinely review offsite on call CT reports.

The study did establish the value of the onsite Consultant addendum report with statistically significant benefits of switching to an addendum in terms of reducing both major and all discrepancies within provisional reports and in both non-surgical and more markedly the surgical groups. The timing of issue of addendum reports is also relevant (see earlier) for them to have appropriate clinical impact. A benefit in terms of reduction in major discrepancy in provisional reporting was also noted in the audit when evaluating the availability of a specialist GI radiologist. Clearly numbers of specialist GI radiologists are relatively limited but there may be a role for both addendum double reading and peer review by GI radiologists of abdominopelvic CT on-call where resources allow.

~~The final component of the study for discussion relates to the sensitivity/specificity of CT in the more common pathologies in the non-surgical and surgical groups. It is beyond the scope of this report to cover all pathologies in these areas but of note is the reduction in sensitivity of CT in the diagnosis of ischaemic bowel in the surgical group when using the auditor as reference standard (89.5%) when compared to laparotomy (72.5%); specificity was the same in both groups (99.5%). The specificity compares well with published data[39], with sensitivity reduced. The reasons for this are unclear but may reflect difficulties encountered when diagnosing early stages of intestinal ischaemia on CT and later correlated with laparotomy findings. Overall registrars met the audit standard for correlation of provisional report with laparotomy (standard >80%, achieved 83.7%) but onsite Consultants narrowly missed their standard (standard >90%, achieved 87.2%). Offsite radiologists missed their target by a larger margin (standard >90%, achieved 78.9%).~~

This study does have limitations. It was performed retrospectively and as such findings do rely on availability and accuracy of relevant documentation. The results reflect practice from 2013. Data was incomplete in some sections and also the prevalence of discussion of cases by registrars with onsite Consultants may not be fully reflected in the reports, possibly enhancing the accuracy of registrar reporting. The response rate overall of 58% bears favourable comparison other similar published studies in the literature[7, 41]. There was no evidence of significant response bias- only small percentage differences were evident between proportions of departments from teaching and non-teaching hospitals who did and did not respond, similar findings were found when looking at geographical response rates.

Conclusion

This study provides data on factors influencing discrepancy rate in the provisional (initial) radiological report in a large cohort of patients undergoing emergency abdominal CT. The lowest rate of discrepancy was found when reporting was undertaken by onsite Consultant radiologists. Statistically significant increases in the rates of major discrepancy and in patients coming to harm

were found when reporting was undertaken by a radiologist at a site remote from the image acquisition. Patients undergoing surgery were at a greater risk of major discrepancy and harm than non-surgical patients. These findings give cause for concern and should provide impetus for further consideration of optimal models of service provision for the reporting of emergency abdominal CT. It is clear that both the seniority and location of the reporter can have a significant effect on the accuracy of emergency CT reporting and hence patient outcomes. Radiological departments should also ensure that a robust and timely system of onsite Consultant addendum reporting is in place as a safety net for registrar and offsite reporters.

References

1. Brewer BJ, Golden GT, Hitch DC, Rudolph L, Wangensteen S. Abdominal pain: An analysis of 1000 consecutive cases in a University Hospital setting. Am J Surg; 1976; 131; 219-223

2. Sala E, Watson CJ, Beardsmoore C. A randomised controlled trial of routine early abdominal CT in patients presenting with non-specific acute abdominal pain. Clin Radiol; 2007; 62; 961-969
3. Tsushima Y, Yamada S, Aoki J. Effect of contrast enhanced CT on the diagnosis and management of the acute abdomen in adults. Clin Radiol; 2002; 57; 507-513
4. Rosen MP, Sands D, Longmaid H. Impact of abdominal CT on the management of patients presenting in the emergency department with acute abdominal pain. Am J Roentgenol; 2000; 174; 1391-1396
5. Brady A, Olaoide R, McCarthy P, McDermott R. Discrepancy and error in radiology: Concepts, causes and consequences; Ulster Med J. 2012; 81 (1); 3-9
6. Goddard P, Leslie A, Jones A, Wakely Cand Kabala J. Error in Radiology. British J Radiol; 2001; 74; 949-951
7. Howlett D, Drinkwater K, Lawrence D, Barter S, Nicholson T. Findings of a UK national audit evaluating image guided and image assisted liver biopsy. Part 1. Procedural aspects, diagnostic adequacy and accuracy. Radiology 2012; 265 (3); 821-822
8. Weir-McCall J, Shaw A, Knight A, Howlett D C. The use of pre-operative CT in the assessment of the acute abdomen. Ann R Coll Surg Eng; 2012; 94 (2); 102-107
9. Briggs R H, Rowbotham E, Johnstone AL, Chalmers AG. Provisional reporting of Polytrauma CT by on call radiology registrars, is it safe? Clin Radiol; 2010; 65:616-622
10. Agostin C, Durieux M, Milot L. Value of double reading of whole body CT in poly trauma Patients. J Radiol; 2008; 89: 325-330
11. Hillier JC, Tattersall DJ, Gleeson F V. Trainee reporting of CT examinations, do they make mistakes and does it matter? Clin Radiol; 2004; 59: 159-162
12. Tieng N, Grinberg F, Li SF. Discrepancies in interpretation of ED Body CT scans by radiology residents. Am J of Emerg Med; 2007: 25 (2): 45-48
13. Chung JH, Strigel RM, Chew AR, Albrecht E, Gunn ML. Overnight resident interpretation of

549 torso CT at a level 1 trauma centre: An analysis and review of the literature. Acad Radiol;
550 2009; 16 (9): 1155 – 1160.
551 14. Kang MJ, Sim MS, Shin TG et al. Evaluating the accuracy of emergency medicine resident
552 interpretations of abdominal CTs in patients with non-traumatic abdominal pain. J Korean
553 Med Sci; 2012; 27 (10): 1255 – 1260.
554 15. Carney E, Kempf J, DeCarvalho V, Yudd A, Nosher J. Preliminary interpretations of after-
555 hours CT and sonography by radiology residents versus final interpretations by body imaging
556 radiologists at a level 1 trauma centre. Am J Roentgenol; 2003; 181; 367 – 373.
557 16. Cooper VF, Goodhartz LA, Nemcek AA, Ryu RK. Radiology resident interpretations of on -
558 call imaging studies: the incidence of major discrepancies. Acad Radiol; 2008; 15: 1198 –
559 1204.
560 17. Walls J, Hunter N, Brasher PM, Ho SG. The DePICTORS study: discrepancies in preliminary
561 interpretation of CT scans between on-call residents and staff. Emerg Radiol; 2009; 16: 303-
562 308.
563 18. Ruutiainen AT, Scanlon MH, Itri JN. Identifying benchmarks for discrepancy rates in
564 preliminary interpretations provided by radiology trainees at an academic institution. J
565 AmColl Radiol; 2011; 8 (9): 644-648.
566 19. Ruchman RB, Jaeger J, Wiggins EF et al. Preliminary radiology resident interpretations versus
567 final attending radiologist interpretations and the impact on patient care in a community
568 hospital. AmJRoentgenol; 2007; 189 (3): 523 – 526.
569 20. Wong WS, Roubal I, Jackson DB, Paik WN, Wong VK. Outsourced teleradiology imaging
570 Services: on analysis of discordant interpretation in 124, 870 cases. J Am Coll Radiol; 2005; 2
571 (9): 804-809.
572 21. Yaniv G, Mozes O, Greenberg G, Bakon M, Hoffman C. Common sites and aetiologies of
573 resident's misinterpretation of head CT scans in the emergency department of a level 1
574 trauma centre.IMAJ; 2013; 15 (5); 221-225

22. Miyakoshi A, Nguyen Q, Cohen W, Talner L, Anzai I. Accuracy of preliminary interpretations of neurologic CT examinations by on-call radiology residents and assessment of patient outcomes at a level 1 trauma centre.J Am Coll Radiol; 2009; 6(12); 864-870
23. Strub W, Leach J, Tomsick T, Vagal A. Overnight preliminary head CT interpretations provided by residents: Locations of misidentified intracranial haemorrhage. Am J Neuroradiol; 2007; 28 (9); 1679-1682
24. Roszler M, McCarroll K, Rashid T, Donovan K, Kling G. Resident interpretation of Emergency CT scans. Invest Radiol; 1991; 26 (4); 374-376
25. Wysoki M, Nassar C, Koenigsberg R, Novelline R, Faro S, Faerber E. Head Trauma: CT scan interpretation by radiology residents versus staff radiologists. Radiology 1998; 208 (1); 125-128
26. De Witt J, Griffith-Richards S, Pitcher R. The accuracy of after hours registrar CT reporting in a South African tertiary teaching hospital. SA Journal of Radiology. 2014; 18 (1); 3; doi: 10.4102/sajr.v18i1.591
27. Ruma J, Klein K, Chong S et al. Cross sectional examination interpretation discrepancies between on-call diagnostic radiology residents and subspecialty faculty radiologists: Analysis by imaging modality and subspecialty. J Am Coll Radiol; 2011; 8 (6); 409 – 414
28. Strub W, Vagal A, Tomsick T, Moulton J; Overnight resident preliminary interpretations of CT examinations; Should the process continue? Emerg Radiol; 2006; 13 (1) 19-23
29. Lowe L, Draud K, Hernanz-Schulman M et al. Non-enhanced limited CT in children suspected of having appendicitis: Prospective comparison of attending and resident interpretations. Radiology; 2001; 221 (3); 755-759
30. Friedman S, Merman E and Chopra A. Clinical impact of diagnostic imaging discrepancy by radiology trainees in an urban teaching hospital emergency department.Int J Emerg Med; 2013; 6: 24-29

31. Terreblanche O, Androikou S, Hlabangana L, Brown T and Boshoff P. Should registrars be reporting after hours CT scans? A calculation of error rate and the influencing factors in South Africa; Acta Radiol; 2012; 53 (1); 61-68
32. Wu MZ, Mcinnes MD, Macdonald DB et al. CT in adults: Systematic review and meta-analysis of interpretation and discrepancy rates. Radiology; 2014; 270; 717-735
33. Lauritzen PM, Andersen JG, Stokke MV. Radiologist initiated double reading of abdominal CT: Retrospective analysis of the clinical importance of changes to radiology reports. BMJ Qual Saf; 2016; 0: 1-9; doi.10.1136/bmjqs-2015-004536
34. Levin D, Rao V, Parker L. Ownership or leasing of CT scanners by non-radiologist physicians: A rapidly growing trend that causes concern about self-referral. Journal of the Am J Coll Radiol; 2008; 5; 1206-1209
35. Burute N, Jankharia B; Teleradiology: The Indian perspective. Indian J Radiol; 2008; 19 (1); 16-18
36. Reinus W. American radiology and outsourcing. Radiology; 2007; 242; 654-657
37. Quint L, Quint D, Myles J. Frequency and spectrum of errors in final radiology reports generated with automatic speech recognition technology. J Am Coll Radiol; 2009; 6 (4); 282-283
38. Dickerson EC, Alam HB, Brown RKJ, Stojanovska J. In-Person Communication Between Radiologists and Acute Care Surgeons Leads to Significant Alterations in Surgical Decision Making. J Am Coll Radiol; 2016: 13(8); 943-949.
39. Little D, McCoubrie P. Learning Whilst On-Call: a Vital Part of Radiology Training? Clin Radiol 2016; 71; 921-924.
40. NELA project team. First patient report of the National Emergency Laparotomy Audit. Royal College of Anaesthetists; London; 2015

~~40. Jang K, Min K, Kim M et al. Diagnostic performance of CT in the detection of intestinal ischaemia associated with small bowel ischaemia associated with small bowel obstruction using maximal attenuation of region interest. AmJ Roentgenol; 2010; 184 (4); 857-863~~

41. Barter S, Jones K, Drinkwater K, Uberoi R, Tawn J. The UK nephrostomy audit: Can a voluntary registry produce robust performance data. Clin Radiol; 2008; 63 (8); 888-894