Manuscripts, Articles, Book Chapters and Other Papers

2-21-2024

Degree of Uncertainty in Reporting Imaging Findings for Necrotizing Enterocolitis: A Secondary Analysis from a Pilot Randomized Diagnostic Trial.

Alain Cuna
Children's Mercy Hospital

Disa Rathore

Kira Bourret

Erin K. Opfer
Children's Mercy Hospital

Sherwin S. Chan
Children's Mercy Hospital

Let us know how access to this publication benefits you

Recommended Citation

Article

Degree of Uncertainty in Reporting Imaging Findings for Necrotizing Enterocolitis: A Secondary Analysis from a Pilot Randomized Diagnostic Trial

Alain Cuna [1,2,*] **, Disa Rathore** [3] **, Kira Bourret** [3] **, Erin Opfer** [2,4] **and Sherwin Chan** [2,4]

1   Division of Neonatology, Children's Mercy Kansas City, Kansas City, MO 64108, USA
2   School of Medicine, University of Missouri-Kansas City, Kansas City, MO 64108, USA
3   School of Medicine, Kansas City University, Kansas City, MO 64106, USA
4   Department of Radiology, Children's Mercy Kansas City, Kansas City, MO 64108, USA
*   Correspondence: accuna@cmh.edu

Abstract: Diagnosis of necrotizing enterocolitis (NEC) relies heavily on imaging, but uncertainty in the language used in imaging reports can result in ambiguity, miscommunication, and potential diagnostic errors. To determine the degree of uncertainty in reporting imaging findings for NEC, we conducted a secondary analysis of the data from a previously completed pilot diagnostic randomized controlled trial (2019–2020). The study population comprised sixteen preterm infants with suspected NEC randomized to abdominal radiographs (AXRs) or AXR + bowel ultrasound (BUS). The level of uncertainty was determined using a four-point Likert scale. Overall, we reviewed radiology reports of 113 AXR and 24 BUS from sixteen preterm infants with NEC concern. The BUS reports showed less uncertainty for reporting pneumatosis, portal venous gas, and free air compared to AXR reports (*pneumatosis*: 1 [1–1.75] vs. 3 [2–3], $p < 0.0001$; *portal venous gas*: 1 [1–1] vs. 1 [1–1], $p = 0.02$; *free air*: 1 [1–1] vs. 2 [1–3], $p < 0.0001$). In conclusion, we found that BUS reports have a lower degree of uncertainty in reporting imaging findings of NEC compared to AXR reports. Whether the lower degree of uncertainty of BUS reports positively impacts clinical decision making in infants with possible NEC remains unknown.

Keywords: necrotizing enterocolitis; radiography; ultrasound; radiology report; language; diagnostic certainty; uncertainty

1. Introduction

Necrotizing enterocolitis (NEC) is an acute and life-threatening disease characterized by uncontrolled inflammation of the premature intestinal tract [1]. Despite extensive research, NEC remains a leading cause of morbidity and mortality in preterm infants. About 10% of preterm infants with birth weight less than 1500 g are expected to develop NEC, while mortality from NEC can be as high as 20% to 45%. Infants who survive NEC are also at high risk for serious morbidities including prolonged hospital stay, short bowel syndrome, poor nutrition and growth, and poor neurodevelopmental outcomes [2].

Because clinical and laboratory features are non-specific, the diagnosis of NEC is challenging and is typically dependent on identifying pathognomonic signs on imaging [3–5]. The presence of pneumatosis intestinalis or portal venous gas is generally sufficient to make a diagnosis in infants with clinical suspicion of NEC, while the presence of free air typically indicates NEC complicated by intestinal perforation that requires surgical intervention. Traditionally, abdominal radiographs (AXRs) have been used as the standard imaging test to identify the pathognomonic imaging findings of NEC. But, in recent years, bowel ultrasound (BUS) has steadily emerged as a helpful adjunct to AXR to aid in NEC diagnosis [6–8]. As a non-invasive imaging modality, BUS is free from radiation and well tolerated by preterm infants [9]. Like AXR, BUS can identify pneumatosis, portal venous

gas, and free air. In addition, BUS allows for real-time assessment of the intestinal wall, vascular perfusion, peristalsis, and abdominal fluid [10].

Unlike diagnostic tests that have discrete empiric values (such as complete blood counts or chemistry panel), imaging tests require interpretation by the radiologist, the findings of which are then conveyed to the ordering clinician via a written imaging report. The interpretation of imaging tests is not binary [11]. Often, there exists a degree of uncertainty regarding whether imaging findings are normal or abnormal [12]. As such, the language used in imaging reports needs to accurately convey the radiologist's uncertainty in identifying the presence or absence of important imaging findings [13]. Many factors can influence the degree of uncertainty conveyed in imaging reports, including technical limitations, insufficient clinical data, lack of established standards, education and training variability, and fear of malpractice [14]. Using language with a high degree of uncertainty can be a common source of miscommunication and misunderstanding that can lead to diagnostic errors, delayed clinical decision making, and adverse outcomes [15–18].

The degree to which uncertain language is used in reporting imaging findings for NEC can have important implications for clinical care. For example, highly uncertain phrases such as "cannot rule out pneumatosis" could influence the neonatologist's plan for treatment even when overall clinical suspicion for NEC is low. Conversely, decreasing uncertainty helps provide clarity, improve communication, and convey the diagnostic confidence of the radiologist regarding the presence or absence of key imaging findings. To the best of our knowledge, the degree to which uncertainty complicates diagnostic imaging reports for NEC has not been characterized.

In this study, we sought to quantify the degree of uncertainty in imaging reports derived from a pilot diagnostic randomized clinical trial (RCT) of infants with suspected NEC who were evaluated with either AXR alone or AXR + add-on BUS [19]. We hypothesized that BUS reports have lower uncertainty than AXR reports and that uncertainty in AXR reports could be lowered by adding information from BUS. We also evaluated whether uncertainty in imaging reports is affected by the experience of the interpreting radiologist.

2. Materials and Methods

2.1. Study Design

This was a secondary analysis using data from a previously completed pilot diagnostic RCT [19]. The study population comprised preterm infants $\leq$ 32 weeks' gestation at birth with suspected NEC who were randomized to either standard imaging with AXR alone or experimental imaging with AXR + add-on BUS. Infants with abdominal wall defects that prohibit 4-quadrant evaluation using BUS were excluded from the study. We evaluated uncertainty by analyzing de-identified imaging reports of AXR and BUS studies acquired during this pilot RCT. The institutional review board at our local institution approved the secondary analysis of this pilot RCT with waiver of informed consent.

2.2. Study Setting

The original pilot RCT was conducted in a level IV neonatal intensive care unit of a tertiary free-standing children's hospital with 24/7 coverage by neonatologists, pediatric radiologists, and pediatric surgeons. Infants randomized to the AXR group were evaluated with a portable AXR as per standard of care and consisted of anteroposterior view, with additional cross-table or left lateral decubitus view per neonatologist discretion. Infants randomized to the AXR + BUS group had a BUS performed within 6 h of the standard of care AXR. The BUS protocol consisted of standard grayscale, color Doppler, and spectral Doppler images of the abdomen supplemented with cine acquisitions in both transverse and sagittal planes. All imaging tests were performed by radiology technologists and interpreted by pediatric radiologists who were board-certified in pediatric radiology. At the time of the pilot RCT, the use of standard reporting templates (Table 1) for reporting NEC findings was encouraged but not mandated. These templates could be edited by the

radiologist to reflect their interpretation regarding the presence or absence of a finding as well as their degree of certainty.

Table 1. Reporting templates for NEC evaluation. The AXR template reports on four findings, while the BUS template reports on ten findings.

Imaging Modality	Template
AXR	There are no findings to suggest bowel obstruction, free intraperitoneal gas, or pneumatosis.
	There is no portal venous gas.
BUS	There is no bowel wall thickening (>2.7 mm).
	There is no bowel wall thinning (<1.0 mm).
	There is no bowel wall hyperechogenicity.
	There is no pneumatosis intestinalis.
	There is no portal venous gas.
	There is no pneumoperitoneum.
	There are no focal fluid collections with complex echoes.
	There is no free fluid.
	There is normal bowel wall perfusion.
	Peristalsis is present.

AXR, abdominal radiograph; BUS, bowel ultrasound; NEC, necrotizing enterocolitis.

2.3. Determination of Uncertainty

Three investigators (AC, DR, and KB) independently reviewed de-identified imaging reports and assigned uncertainty scores using a 4-point Likert scale adopted from the study by Reiner [20]. In this scale, a score of 1 denotes the lowest level of uncertainty; a score of 2 denotes minimal level of uncertainty; a score of 3 denotes intermediate level of uncertainty; and a score of 4 denotes the highest level of uncertainty (Table 2). We limited our evaluation to pneumatosis, portal venous gas, and free air because these findings are most specific for NEC and because these findings are assessed in both AXR and BUS. Disagreements in scoring were discussed as a group and resolved by consensus.

2.4. Outcomes and Variables of Interest

Our primary outcome was uncertainty scores for pneumatosis, portal venous gas, and free air. Secondary outcomes included proportion of imaging reports with all three imaging findings present and proportion of imaging reports that used standardized reporting templates. We also assessed whether other factors such as the addition of BUS and differences in radiology personnel (involvement of trainees, years of experience and subspecialty of the interpreting radiologist) impacted uncertainty scores.

2.5. Statistical Analysis

Data are presented as mean and standard deviation, median and interquartile range, or numbers and percentages. Normality of data was determined using the Shapiro–Wilk test. Differences in primary and secondary outcomes between the imaging modality (AXR vs. BUS) and randomization arm (AXR vs. AXR + BUS) were evaluated using the Mann–Whitney U test and chi-square test, as appropriate. To determine the effect of add-on BUS on quality of imaging reporting, uncertainty scoring of infants with AXR reports completed within 6–8 h before and after add-on BUS were compared using a paired Wilcoxon signed-rank test. To determine the effect of trainee involvement as well as years of experience and subspecialty of the interpreting radiologist on the uncertainty of medical reporting, uncertainty scores were first dichotomized to low uncertainty (scores of 1 or 2) or high uncertainty (scores of 3 or 4) followed by binomial logistic regression. All analyses

were performed in SPSS software (IBM SPSS Statistics for Windows, Version 24.0, Armonk, NY, USA), and statistical significance was set at $p < 0.05$.

Table 2. Scoring system for determining degree of uncertainty.

Score	Degree of Uncertainty	Example
1	Lowest level of uncertainty. No equivocation.	No pneumatosis.
		No portal venous gas.
		No pneumoperitoneum.
2	Low level of uncertainty. Minimal equivocation.	No definite pneumatosis.
		No supine evidence of pneumoperitoneum.
		No obvious free air.
3	Intermediate level of uncertainty. Intermediate degree of equivocation.	No findings to suggest pneumatosis, portal venous gas, or free air.
		Small area of possible linear lucency along the bowel wall.
		Suggestion of interim mural air within the bowel wall.
4	Highest level of uncertainty. High degree of equivocation.	Cannot exclude pneumatosis.
		No obvious pneumatosis although evaluation is limited.
		Mild mottled lucencies which may represent pneumatosis or stool.

3. Results

Overall, sixteen preterm infants with concern for NEC were randomized as part of the original pilot RCT. The mean gestational age and mean birth weight for the entire cohort was 27.2 ± 2.2 weeks and 1020 ± 373 g, respectively. The baseline characteristics were comparable between the groups (Table 3). The eight infants randomized to the AXR arm had 49 AXRs, while the eight infants randomized to the AXR + BUS arm had 64 AXRs and 24 BUSs. Thus, overall, we reviewed 113 AXR reports and 24 BUS reports. We had six instances of disagreement in uncertainty scoring—four with AXR reports and two with BUS reports.

The uncertainty scores for each main NEC finding described in the BUS and AXR reports are displayed in Figure 1. Overall, BUS had lower uncertainty scores for reporting each of the three main NEC findings than AXR (*pneumatosis*: 1 [1–1.75] vs. 3 [2–3], $p < 0.0001$; *portal venous gas*: 1 [1–1] vs. 1 [1–1], $p = 0.02$; *free air*: 1 [1–1] vs. 2 [1–3], $p < 0.0001$).
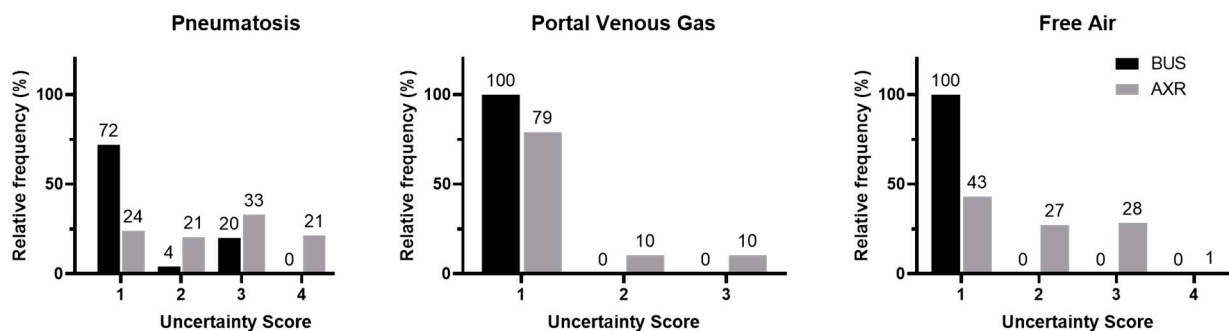
Figure 1. Frequency distribution of uncertainty scores for reporting pneumatosis, portal venous gas, and free air comparing BUS versus AXR. In all three imaging findings of NEC, BUS (black bars) was observed to have lower uncertainty scores than AXR (gray bars).

Table 3. Characteristics of the study participants from the original pilot trial.

Baseline Characteristics	AXR Group (*n* = 8)	AXR Plus BUS Group (*n* = 8)
Gestational age, weeks	26.9 ± 2.5	27.4 ± 2.1
Birth weight, grams	1056 ± 399	1022 ± 381
Male sex, no. (%)	5 (63)	2 (25)
White race, no. (%)	4 (50)	6 (75)
Small for gestational age, no. (%)	0 (0)	1 (13)
Maternal age, years [a]	25 ± 7	31 ± 7
Caesarian delivery, no. (%)	3 (38)	5 (63)
Apgar score < 5 at 1 min, no. (%) [b]	4 (50)	7 (88)
Apgar score < 5 at 5 min, no. (%) [b]	3 (38)	2 (25)
Antenatal corticosteroids, no. (%)	8 (100)	7 (88)
Surfactant, no. (%)	8 (100)	6 (75)
No. of NEC concern episodes	1 (1–3)	1 (1–3)
No. of imaging studies		
AXR	49	64
BUS	0	24

Data presented as number (percentage), mean ± standard deviation, or median (interquartile range). [a] One mother with unknown age; [b] two infants with unknown Apgar scores. AXR, abdominal radiograph; BUS, bowel ultrasound; NEC, necrotizing enterocolitis.

The BUS reports described all three main NEC imaging findings (pneumatosis, portal venous gas, and free air) in 96% of the reports (Table 4). The only exception was the omission of reporting about free air on one BUS report. In contrast, the AXR reports described all three main NEC findings in only 52% of the reports. The NEC finding most frequently absent in the AXR reports was portal venous gas (Table 4). The BUS reports also used a standardized reporting template more frequently than the AXR reports (Table 4).

Table 4. Comparison of reporting metrics between BUS and AXR.

	BUS *n* = 24	AXR *n* = 113	*p*
Complete report, *n* (%)	23 (96)	59 (52)	0.0001
Pneumatosis	24 (100)	111 (98)	1.0
Portal venous gas	24 (100)	67 (59)	0.0001
Free air	23 (96)	95 (84)	0.19
Use of standardized template, *n* (%)	21 (88)	18 (16)	0.0001

Data presented as number (percent); *p* value represents comparison with chi-square test. AXR, abdominal radiograph; BUS, bowel ultrasound.

To determine whether the addition of BUS impacted AXR reporting, we first compared the AXR reports from infants in the AXR arm to infants in the AXR + BUS arm. We found that the completeness of describing the three main NEC findings was similar between the two arms (Table 5). The use of a standardized template for NEC impression was also similar (Table 5).

Table 5. Comparison of reporting metrics of AXR reports from infants with and without add-on BUS.

	AXR + BUS Arm *n* = 64	AXR Arm *n* = 49	*p*
Complete report, *n* (%)	35 (55)	24 (49)	0.55
Pneumatosis	63 (98)	48 (98)	1.0
Portal venous gas	42 (66)	25 (51)	0.12
Free Air	54 (84)	41 (84)	0.92
Use of standardized template, *n* (%)	9 (14)	9 (18)	0.54

Data presented as number (percent). *p* value represents comparison with chi-square test. AXR, abdominal radiograph; BUS, bowel ultrasound.

We then evaluated AXRs that had a paired BUS (within 6 h of AXR) in the AXR + BUS arm. The change in uncertainty score from the AXR reports for pneumatosis, portal venous gas, and free air pre- and post-BUS is displayed in Figure 2. A paired Wilcoxon signed-rank test revealed no significant change in uncertainty scoring from AXRs conducted before and after the BUS study (Figure 2).
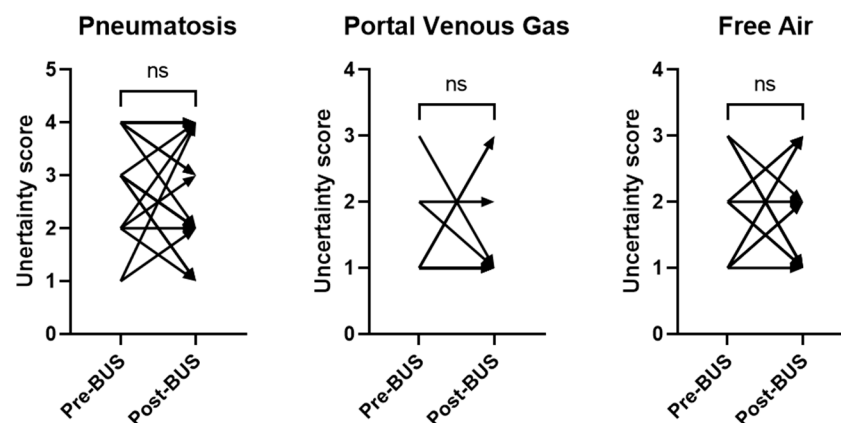
Figure 2. Change in uncertainty scores in AXR reports before and after add-on BUS. For all three imaging findings, uncertainty scores of AXR reports did not change significantly between baseline (pre-BUS) and after paired BUS examination (post-BUS). ns = not significant.

Lastly, we evaluated whether differences in experience, specialization, and trainee involvement played a role in the lower uncertainty scoring of BUS compared to AXR. Overall, 22 pediatric radiologists interpreted all imaging tests in the study. Of these, 11 reported on only AXR, 1 reported on only BUS, and 10 reported on both AXR and BUS. Their experience ranged from 5 to 24 years, and their specialization included body, cardiac, brain, interventional radiology, fetal, and musculoskeletal. Eight reports had trainee involvement. Using binomial logistic regression, we found that only BUS was associated with increased likelihood of lower-uncertainty reports (OR 5.64, 95% CI 1.88–16.9). Years of experience, specialization, and trainee involvement did not influence uncertainty scores.

4. Discussion

In this study, we sought to determine the level of uncertainty in radiology reports of preterm infants who underwent diagnostic imaging for suspected NEC. We found that the BUS reports had significantly less uncertainty in reporting pneumatosis, portal venous gas, and free air compared to the AXR reports. The rate of completeness in reporting imaging findings of NEC and the rate of using standardized reporting templates were also significantly higher in the BUS reports than in the AXR reports. The addition of paired BUS did not decrease the high level of uncertainty in the AXR reports. Years of experience, radiology subspecialty, or trainee involvement also did not affect the level of uncertainty.

The lower level of uncertainty in BUS reports may be explained by the superior technical images of BUS over AXR in NEC evaluation. Because ultrasound waves cannot pass through air, pneumatosis intestinalis and portal venous gas are easily detected as bright, echogenic foci on BUS. Several studies have found that BUS can identify portal venous gas and pneumatosis intestinalis earlier than AXR [21–23]. BUS is also capable of the real-time assessment of bowel loops in cross-section, allowing an opportunity to evaluate bowel wall thickness, peristalsis, and perfusion (with Doppler interrogation) [24]. In addition, BUS assessment occurs over ~30 min, whereas AXR is a one-time assessment [9,25]. Therefore, BUS is better able to capture dynamic and intermittent processes like portal venous gas. A constellation of thinning of the bowel wall, decreased or absent peristalsis, and decreased or absent perfusion can be indicative of impending bowel perforation [10,26]. The advantage of BUS in assessing abdominal structures in more detail compared to AXR is likely a major driver of the lower level of uncertainty we found in BUS reports.

In a similar fashion, the high level of uncertainty in AXR reports can be explained by the inherent challenges in identifying findings of NEC on plain film radiography with consistency. Sharma et al. [27] used computer-aided simulation to assess radiology trainees and found that only 28% were able to correctly identify pneumatosis intestinalis on standardized test AXRs. Di Napoli et al. [18] found poor reliability for NEC diagnosis and individual NEC imaging findings even among three expert radiologists who independently reviewed 297 AXRs of infants with and without NEC. Other studies have demonstrated similar poor agreement in identifying pathognomonic signs of NEC on AXR among radiologists, neonatologists, and pediatric surgeons [15–17]. These studies suggest that the inherent limitations of radiography play a major role in the higher level of uncertainty found in AXR reports for NEC evaluation.

A secondary objective of our study was to investigate whether the addition of BUS would help decrease the level of uncertainty in AXR reports. Contrary to our hypothesis, we found that the uncertainty in the AXR reports remained unchanged despite the addition of BUS. One possible explanation for our results is that the AXRs were interpreted independently and without consideration of the preceding BUS results. Although we cannot exclude this possibility, we believe this is unlikely because it is standard practice to review prior images for accurate interpretation of any follow-up imaging for NEC. Instead, a more likely explanation is that the inherent limitations of AXR preclude radiologists from using more certain language for reporting NEC findings. Other investigators have used standardized reporting tools to help overcome the challenges and limitations of AXR in NEC imaging. One such tool is the Duke Abdominal Assessment Scale (DAAS), a standardized 10-point numerical scale for reporting abnormal radiographic findings in NEC [28]. Initial studies demonstrated substantial improvement in inter-observer and intra-observer agreement among radiologists following DAAS implementation [29]. However, other studies found that a fair degree of disagreement persisted even with DAAS implementation [30]. Taken together, these findings suggest that strategies such as standardized reporting tools and adjunct BUS may not be sufficient to overcome the inherent limitations of AXR that result in the high level of uncertainty of reporting NEC findings.

We also investigated whether the experience of the interpreting radiologist impacted the level of uncertainty in imaging reports. In a single-center study, Callen et al. [31] used natural language processing to characterize the use of uncertainty terms of over 640,000 radiology reports by 171 interpreting radiologists. While substantial variability in the use of uncertainty terms was observed, the degree of variability could not be explained by differences in years of experience of radiologists. Similar results were observed by Crombe et al. [32] in a multicenter study of over 30,000 computed topography scans and magnetic resonance imaging interpreted by 165 radiologists. Despite having a smaller sample size, our study mirrors the findings of these studies in demonstrating no correlation between the use of uncertain language on imaging reports and the experience of the interpreting radiologist.

One way to potentially decrease the uncertainty in imaging reports is the addition of diagnostic certainty scales that convey the radiologist's confidence in their interpretation [33]. Several studies have reported how the voluntary adoption of diagnostic certainty scales by radiologists resulted in a modest to moderate increase in its use, suggesting the feasibility of this practice [34–36]. More importantly, studies have found that the radiologist's diagnostic certainty score correlated strongly with the subsequent confirmation of disease. In one study, Godwin et al. [37] employed diagnostic certainty scoring for the presence or absence of appendicitis based on CT scan findings and found that diagnostic certainty scoring correlated strongly with post-operative pathology. In another study, Wibmer et al. [38] demonstrated how the radiologist's diagnostic certainty score for the routine staging MRI of prostate cancer correlated with subsequent histopathology for extracapsular extension. Taken together, these studies indicate that the adoption of diagnostic certainty scales in imaging reports is feasible and can potentially increase the accuracy of diagnosis.

Furnishing more clinical information to the ordering physician can also decrease uncertainty by improving communication and offering valuable context to radiologists. In a retrospective study, Maizlin and Somers [39] investigated the impact of additional information from radiology technician notes compared to the original information from the imaging order requisition. Of the 250 radiographs and ultrasound tests reviewed, 47% of imaging order requisitions were either incomplete or absent. The added information from technician notes was deemed by radiologists to provide important value in more than two-thirds (173/250 = 69%) of the reviewed cases. In another study, Lacson et al. [40] compared the clinical information on imaging order requisition with the clinical information on the provider notes. Of the 315 magnetic resonance imaging and computed tomography scans reviewed, information on order requisitions was incomplete in 81% and discordant in 42% of reviewed cases. Incomplete and discordant clinical information was deemed by reviewing radiologists to negatively impact interpretation in 43% (135/315) of cases. Several other studies [41,42], including a systematic review [43], showed similar findings that the improved availability of clinical information is beneficial for optimal interpretation.

Whether improving uncertainty in reporting NEC imaging findings has important implications in clinical care remains unknown. The results from the original pilot RCT showed no difference in clinical outcomes between infants randomized to AXR versus AXR + BUS imaging [19]. However, this pilot study was designed for feasibility and not powered to detect differences in clinical outcomes. Studies of ultrasound for pediatric appendicitis suggested that uncertain interpretation was associated with poorer outcomes and higher medical costs compared to situations when more certain interpretation was provided [44,45]. Given the similarities between appendicitis and NEC, it is plausible that increasing certainty in reporting for NEC could translate to improved patient outcomes, particularly reductions in antibiotic use, bowel rest, and intravenous nutrition.

Our study has several strengths. To the best of our knowledge, this is the first study that measured the degree of uncertainty in the language used in reports of diagnostic imaging tests for NEC. Our results could be useful as reference data for future research or quality improvement studies regarding the uncertainty in imaging reports for NEC. Our results also provide supporting evidence that higher diagnostic certainty could be another advantage of BUS over AXR for evaluating NEC. The secondary use of data from a randomized diagnostic trial increases the validity of our results, as randomization eliminates the possibility that the lower uncertainty scores in the BUS reports could have been due to selection bias. Another strength of the study is our analysis of other factors that could influence uncertainty scores, such as the addition of paired BUS and the experience of the interpreting radiologist.

Our study is limited by the small sample size and the single-center study design of the original randomized diagnostic trial, which experienced challenges in recruitment due to the COVID-19 pandemic [19]. Despite this limitation, our study evaluated 137 imaging reports, which provided sufficient power to detect a significant difference in our primary outcome. Another limitation is that our study was conducted in a tertiary children's hospital with pediatric radiologists who were adept at interpreting AXR and BUS studies for NEC. Whether similar uncertainty scores are seen in institutions that lack pediatric radiologists is unknown. A third limitation is that we cannot exclude the possibility that the wide variation in uncertainty could be related to the differential use of structured reporting templates between BUS and AXR. While reporting templates are available for both AXR and BUS, radiologists may have relied more on standardized reporting templates for BUS because it is a relatively newer modality compared with AXR. Lastly, our study could not determine whether less uncertainty in reporting from BUS leads to improved patient outcomes. Future studies are needed to properly evaluate the relationships among communication of diagnostic uncertainty, management, and improved patient outcomes.

5. Conclusions

In conclusion, we found that BUS reports convey imaging findings of NEC with less uncertainty than AXR reports. These findings suggest that radiologists are more confident in identifying the presence or absence of key NEC findings based on BUS than on AXR. Future studies are needed to determine whether the lower degree of uncertainty with BUS translates to improved care of infants with suspected NEC.

Author Contributions: Conceptualization, A.C.; methodology, A.C. and S.C.; validation, A.C., D.R. and K.B.; formal analysis, A.C.; resources, S.C. and E.O.; data curation, D.R. and K.B.; writing—original draft preparation, A.C.; writing—review and editing, D.R., K.B., E.O. and S.C.; supervision, A.C.; funding acquisition, E.O. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to the retrospective nature of this study using de-identified data.

Informed Consent Statement: Patient consent was waived due to the retrospective nature of this study using de-identified data and minimal risk to subjects.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Neu, J.; Walker, W.A. Necrotizing enterocolitis. *N. Engl. J. Med.* **2011**, *364*, 255–264. [CrossRef]
2. Alganabi, M.; Lee, C.; Bindi, E.; Li, B.; Pierro, A. Recent advances in understanding necrotizing enterocolitis. *F1000Research* **2019**, *8*, 107. [CrossRef] [PubMed]
3. Zani, A.; Eaton, S.; Puri, P.; Rintala, R.; Lukac, M.; Bagolan, P.; Kuebler, J.F.; Hoellwarth, M.E.; Wijnen, R.; Tovar, J.; et al. EUPSA Network, International survey on the management of necrotizing enterocolitis. *Eur. J. Pediatr. Surg.* **2015**, *25*, 27–33. [CrossRef] [PubMed]
4. Ahle, M.; Ringertz, H.G.; Rubesova, E. The role of imaging in the management of necrotising enterocolitis: A multispecialist survey and a review of the literature. *Eur. Radiol.* **2018**, *28*, 3621–3631. [CrossRef] [PubMed]
5. Valpacos, M.; Arni, D.; Keir, A.; Aspirot, A.; Wilde, J.C.H.; Beasley, S.; De Luca, D.; Pfister, R.E.; Karam, O. Diagnosis and Management of Necrotizing Enterocolitis: An International Survey of Neonatologists and Pediatric Surgeons. *Neonatology* **2018**, *113*, 170–176. [CrossRef] [PubMed]
6. van Druten, J.; Khashu, M.; Chan, S.S.; Sharif, S.; Abdalla, H. Abdominal ultrasound should become part of standard care for early diagnosis and management of necrotising enterocolitis: A narrative review. *Arch. Dis. Child. Fetal Neonatal Ed.* **2019**, *104*, F551–F559. [CrossRef] [PubMed]
7. Chan, B.; Gordon, S.; Yang, M.; Weekes, J.; Dance, L. Abdominal Ultrasound Assists the Diagnosis and Management of Necrotizing Enterocolitis. *Adv. Neonatal Care* **2021**, *21*, 365–370. [CrossRef]
8. Kim, J.H. Role of Abdominal US in Diagnosis of NEC. *Clin. Perinatol.* **2019**, *46*, 119–127. [CrossRef]
9. Alexander, K.M.; Chan, S.S.; Opfer, E.; Cuna, A.; Fraser, J.D.; Sharif, S.; Khashu, M. Implementation of bowel ultrasound practice for the diagnosis and management of necrotising enterocolitis. *Arch. Dis. Child. Fetal Neonatal Ed.* **2021**, *106*, 96–103. [CrossRef]
10. Epelman, M.; Daneman, A.; Navarro, O.M.; Morag, I.; Moore, A.M.; Kim, J.H.; Faingold, R.; Taylor, G.; Gerstle, J.T. Necrotizing enterocolitis: Review of state-of-the-art imaging findings with pathologic correlation. *Radiographics* **2007**, *27*, 285–305. [CrossRef]
11. Brady, A.; Laoide, R.Ó.; McCarthy, P.; McDermott, R. Discrepancy and error in radiology: Concepts, causes and consequences. *Ulster Med. J.* **2012**, *81*, 3–9. [PubMed]
12. Wattamwar, K.; Garg, T.; Wheeler, C.A.; Burns, J. Diagnostic Uncertainty in Radiology: A Perspective for Trainees and Training Programs. *Radiographics* **2022**, *42*, E193–E196. [CrossRef]
13. Bruno, M.A.; Petscavage-Thomas, J.; Abujudeh, H.H. Communicating Uncertainty in the Radiology Report. *AJR Am. J. Roentgenol.* **2017**, *209*, 1006–1008. [CrossRef] [PubMed]
14. Reiner, B. Uncovering and improving upon the inherent deficiencies of radiology reporting through data mining. *J. Digit. Imaging* **2010**, *23*, 109–118. [CrossRef]

15. Mata, A.G.; Rosengart, R.M. Interobserver variability in the radiographic diagnosis of necrotizing enterocolitis. *Pediatrics* **1980**, *66*, 68–71. [CrossRef]

16. Rehan, V.K.; Seshia, M.M.; Johnston, B.; Reed, M.; Wilmot, D.; Cook, V. Observer variability in interpretation of abdominal radiographs of infants with suspected necrotizing enterocolitis. *Clin. Pediatr.* **1999**, *38*, 637–643. [CrossRef]

17. El-Kady, S.; Petel, D.; Baird, R. Inter-rater agreement in the evaluation of abdominal radiographs for necrotizing enterocolitis. *J. Pediatr. Surg.* **2014**, *49*, 733–735. [CrossRef]

18. Di Napoli, A.; Di Lallo, D.; Perucci, C.A.; Schifano, P.; Orzalesi, M.; Franco, F.; De Carolis, M.P. Inter-observer reliability of radiological signs of necrotising enterocolitis in a population of high-risk newborns. *Paediatr. Perinat. Epidemiol.* **2004**, *18*, 80–87. [CrossRef] [PubMed]

19. Cuna, A.; Chan, S.; Jones, J.; Sien, M.; Robinson, A.; Rao, K.; Opfer, E. Feasibility and acceptability of a diagnostic randomized clinical trial of bowel ultrasound in infants with suspected necrotizing enterocolitis. *Eur. J. Pediatr.* **2022**, *181*, 3211–3215. [CrossRef]

20. Reiner, B.I. Quantifying Analysis of Uncertainty in Medical Reporting: Creation of User and Context-Specific Uncertainty Profiles. *J. Digit. Imaging* **2018**, *31*, 379–382. [CrossRef]

21. Merritt, C.R.; Goldsmith, J.P.; Sharp, M.J. Sonographic detection of portal venous gas in infants with necrotizing enterocolitis. *AJR Am. J. Roentgenol.* **1984**, *143*, 1059–1062. [CrossRef]

22. Lindley, S.; Mollitt, D.L.; Seibert, J.J.; Golladay, E.S. Portal vein ultrasonography in the early diagnosis of necrotizing enterocolitis. *J. Pediatr. Surg.* **1986**, *21*, 530–532. [CrossRef]

23. Robberecht, E.A.; Afschrift, M.; De Bel, C.E.; Van Haesebrouck, P.J.; Van Bever, H.P.; De Wit, M.; Leroy, J.G. Sonographic demonstration of portal venous gas in necrotizing enterocolitis. *Eur. J. Pediatr.* **1988**, *147*, 192–194. [CrossRef] [PubMed]

24. Faingold, R. Technical aspects of abdominal ultrasound and color Doppler assessment of bowel viability in necrotizing enterocolitis. *Pediatr. Radiol.* **2018**, *48*, 617–619. [CrossRef] [PubMed]

25. Darge, K.; Anupindi, S.; Keener, H.; Rompel, O. Ultrasound of the bowel in children: How we do it. *Pediatr. Radiol.* **2010**, *40*, 528–536. [CrossRef] [PubMed]

26. Faingold, R.; Daneman, A.; Tomlinson, G.; Babyn, P.S.; Manson, D.E.; Mohanta, A.; Moore, A.M.; Hellmann, J.; Smith, C.; Gerstle, T.; et al. Necrotizing enterocolitis: Assessment of bowel viability with color doppler US. *Radiology* **2005**, *235*, 587–594. [CrossRef] [PubMed]

27. Sharma, P.G.; Rajderkar, D.A.; Sistrom, C.L.; Slater, R.M.; Mancuso, A.A. Bubbles in the belly: How well do radiology trainees recognize pneumatosis in pediatric patients on plain film? *Br. J. Radiol.* **2022**, *95*, 20211101. [CrossRef] [PubMed]

28. Hollingsworth, C.L.; Rice, H.E. The Duke Abdominal Assessment Scale: Initial experience. *Expert. Rev. Gastroenterol. Hepatol.* **2010**, *4*, 569–574. [CrossRef] [PubMed]

29. Coursey, C.A.; Hollingsworth, C.L.; Gaca, A.M.; Maxfield, C.; Delong, D.; Bisset, G. Radiologists' agreement when using a 10-point scale to report abdominal radiographic findings of necrotizing enterocolitis in neonates and infants. *AJR Am. J. Roentgenol.* **2008**, *191*, 190–197. [CrossRef] [PubMed]

30. Markiet, K.; Szymanska-Dubowik, A.; Janczewska, I.; Domazalska-Popadiuk, I.; Zawadzka-Kepczynska, A.; Bianek-Bodzak, A. Agreement and reproducibility of radiological signs in NEC using The Duke Abdominal Assessment Scale (DAAS). *Pediatr. Surg. Int.* **2017**, *33*, 335–340. [CrossRef]

31. Callen, A.L.; Dupont, S.M.; Price, A.; Laguna, B.; McCoy, D.; Do, B.; Talbott, J.; Kohli, M.; Narvid, J. Between Always and Never: Evaluating Uncertainty in Radiology Reports Using Natural Language Processing. *J. Digit. Imaging* **2020**, *33*, 1194–1201. [CrossRef]

32. Crombé, A.; Seux, M.; Bratan, F.; Bergerot, J.-F.; Banaste, N.; Thomson, V.; Lecomte, J.-C.; Gorincour, G. What Influences the Way Radiologists Express Themselves in Their Reports? A Quantitative Assessment Using Natural Language Processing. *J. Digit. Imaging* **2022**, *35*, 993–1007. [CrossRef]

33. Panicek, D.M.; Hricak, H. How Sure Are You, Doctor? A Standardized Lexicon to Describe the Radiologist's Level of Certainty. *AJR Am. J. Roentgenol.* **2016**, *207*, 2–3. [CrossRef]

34. Glazer, D.I.; Budiawan, E.; Burk, K.S.; Shinagare, A.B.; Lacson, R.; Boland, G.W.; Khorasani, R. Adoption of a diagnostic certainty scale in abdominal imaging: 2-year experience at an academic institution. *Abdom. Radiol.* **2022**, *47*, 1187–1195. [CrossRef]

35. Shinagare, A.B.; Alper, D.P.; Hashemi, S.R.; Chai, J.L.; Hammer, M.M.; Boland, G.W.; Khorasani, R. Early Adoption of a Certainty Scale to Improve Diagnostic Certainty Communication. *J. Am. Coll. Radiol.* **2020**, *17*, 1276–1284. [CrossRef]

36. Das, J.P.; Panicek, D.M. Added Value of a Diagnostic Certainty Lexicon to the Radiology Report. *Radiographics* **2021**, *41*, E64–E65. [CrossRef]

37. Godwin, B.D.; Drake, F.T.; Simianu, V.V.; Shriki, J.E.; Hippe, D.S.; Dighe, M.; Bastawrous, S.; Cuevas, C.; Flum, D.; Bhargava, P. A novel reporting system to improve accuracy in appendicitis imaging. *AJR Am. J. Roentgenol.* **2015**, *204*, 1212–1219. [CrossRef]

38. Wibmer, A.; Vargas, H.A.; Sosa, R.; Zheng, J.; Moskowitz, C.; Hricak, H. Value of a standardized lexicon for reporting levels of diagnostic certainty in prostate MRI. *AJR Am. J. Roentgenol.* **2014**, *203*, W651–W657. [CrossRef] [PubMed]

39. Maizlin, N.N.; Somers, S. The Role of Clinical History Collected by Diagnostic Imaging Staff in Interpreting of Imaging Examinations. *J. Med. Imaging Radiat. Sci.* **2019**, *50*, 31–35. [CrossRef] [PubMed]

40. Lacson, R.; Laroya, R.; Wang, A.; Kapoor, N.; Glazer, D.I.; Shinagare, A.; Ip, I.K.; Malhotra, S.; Hentel, K.; Khorasani, R. Integrity of clinical information in computerized order requisitions for diagnostic imaging. *J. Am. Med. Inform. Assoc.* **2018**, *25*, 1651–1656. [CrossRef] [PubMed]

41. Babcook, C.J.; Norman, G.R.; Coblentz, C.L. Effect of clinical history on the interpretation of chest radiographs in childhood bronchiolitis. *Investig. Radiol.* **1993**, *28*, 214–217. [CrossRef]
42. Berbaum, K.S.; Franken, E.A.; Dorfman, D.D.; Lueben, K.R. Influence of clinical history on perception of abnormalities in pediatric radiographs. *Acad. Radiol.* **1994**, *1*, 217–223. [CrossRef]
43. Castillo, C.; Steffens, T.; Sim, L.; Caffery, L. The effect of clinical information on radiology reporting: A systematic review. *J. Med. Radiat. Sci.* **2021**, *68*, 60–74. [CrossRef] [PubMed]
44. Binkovitz, L.A.; Unsdorfer, K.M.L.; Thapa, P.; Kolbe, A.B.; Hull, N.C.; Zingula, S.N.; Thomas, K.B.; Homme, J.L. Pediatric appendiceal ultrasound: Accuracy, determinacy and clinical outcomes. *Pediatr. Radiol.* **2015**, *45*, 1934–1944. [CrossRef] [PubMed]
45. Unsdorfer, K.M.L.; An, J.Y.; Binkovitz, L.A. Pediatric appendiceal ultrasound: Maintaining accuracy, increasing determinacy and improving clinical outcomes following the introduction of a standardized reporting template. *Pediatr. Radiol.* **2021**, *51*, 265–272. [CrossRef] [PubMed]