Structured Reporting of Whole-Body Trauma CT Scans Using Checklists: Diagnostic Accuracy of Reporting Radiologists Depending on Their Level of Experience

Strukturierte Befundung von Polytrauma-CTs mit Checklisten: Diagnostische Genauigkeit in Abhängigkeit von der Erfahrung des Radiologen

Authors

Lena Marie Dendl^{1, 2}, Antonia Maria Pausch³, Patrick Hoffstetter⁴, Christian Dornia⁵, Josef Höllthaler⁶, Antonio Ernstberger⁷, Roland Becker⁸, Sebastian Kopf⁸, Stephan Schleder⁹, Andreas G. Schreyer¹

Affiliations

- Institute for Diagnostic and Interventional Radiology, Brandenburg Medical School Theodor Fontane, Brandenburg a. d. Havel, Germany
- 2 Department of Radiology, Johanniter-Hospital Treuenbrietzen, Germany
- 3 Department of Radiology, University Hospital Regensburg, Germany
- 4 Department of Radiology, Asklepios-Klinikum Bad Abbach, Germany
- 5 Radiology, MVZ für Diagnostische Radiologie Düsseldorf, Germany
- 6 Radiology, dia.log Altötting, Germany
- 7 Department of Trauma Surgery, Klinikum Osnabrück GmbH, Osnabruck, Germany
- 8 Department of Orthopaedics and Traumatology,
 Medizinische Hochschule Brandenburg Theodor Fontane,
 Brandenburg an der Havel, Germany
- 9 Department of Radiology, Barmherzige-Brüder-Klinikum Sankt Elisabeth Straubing GmbH, Straubing, Germany

Key words

structured reporting, CT, trauma

received 21.12.2020 accepted 08.06.2021 published online 04.08.2021

Bibliography

Fortschr Röntgenstr 2021; 193: 1451–1459 DOI 10.1055/a-1541-8265 ISSN 1438-9029 © 2021. Thieme. All rights reserved. Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Correspondence

Prof. Andreas G. Schreyer Institute for Diagnostic and Interventional Radiology, Brandenburg Medical School Theodor Fontane, Hochstraße 29, 14770 Brandenburg a. d. Havel, Germany Tel.: +49/33 81/41 26 00 Fax: +49/33 81/41 26 09 andreas.schreyer@mac.com

ABSTRACT

Purpose Evaluation of the diagnostic accuracy of a checkliststyle structured reporting template in the setting of wholebody multislice computed tomography in major trauma patients depending on the level of experience of the reporting radiologist.

Materials and Methods A total of 140 major trauma scans with the same protocol were included in this retrospective study. In a purely trial-intended reading, the trauma scans were analyzed by three radiologists with different levels of experience (resident, radiologist with 3 years of experience after board certification, and radiologist with 7 years of experience after board certification). The aim was to fill in the checklist 1 template within one minute to immediately diagnose management-altering findings. Checklist 2 was intended for the analysis of important trauma-related findings within 10 minutes. Reading times were documented. The final radiology report and the documented injuries in the patient's medical record were used as gold standard.

Results The evaluation of checklist 1 showed a range of falsenegative reports between 5.0% and 11.4% with the resident showing the highest accuracy. Checklist 2 showed overall high diagnostic inaccuracy (19.3–35.0%). The resident's diagnostic accuracy was statistically significantly higher compared to the radiologist with 3 years of experience after board certification (p = 0.0197) and with 7 years of experience after board certification (p = 0.0046). Shorter average reporting time resulted in higher diagnostic inaccuracy. Most of the missed diagnoses were fractures of the spine and ribs. **Conclusion** By using a structured reporting template in the setting of major trauma computed tomography, less experienced radiologists reach a higher diagnostic accuracy compared to experienced readers.

Key Points:

- In the setting of a pure trial reading, the diagnostic inaccuracy of template-based reporting of major trauma CT examinations is high.
- Fractures in general and especially of the vertebral bodies and ribs were the most commonly missed diagnoses.
- In a study setting, less experienced radiologists seem to reach a higher diagnostic accuracy when using a structured reporting approach.

Citation Format

 Dendl LM, Pausch AM, Hoffstetter P et al. Structured Reporting of Whole-Body Trauma CT Scans Using Checklists: Diagnostic Accuracy of Reporting Radiologists Depending on Their Level of Experience. Fortschr Röntgenstr 2021; 193: 1451–1459

ZUSAMMENFASSUNG

Ziel Evaluation einer strukturierten Befundung der Ganzkörper-Multislice-Computertomografie (MSCT) bei Patienten mit angenommener Polytraumatisierung anhand von Checklisten bezüglich ihrer diagnostischen Genauigkeit in Abhängigkeit der Berufserfahrung der auswertenden Radiologen. Material und Methoden In einem retrospektiven Studiendesign wurden 140 konsekutive Schockraum-CT in die Studie eingeschlossen. Diese wurden in einer Studienauswertung von 3 Radiologen (Weiterbildungsassistent im 1. Ausbildungsjahr und 2 Radiologen mit 3 und 7 Jahren Berufserfahrung als Fachärzte) mittels Checklisten ausgewertet. Checkliste 1 soll innerhalb von maximal einer Minute akut lebensbedrohliche Diagnosen und Checkliste 2 innerhalb von maximal 10 Minuten systematisch die wichtigsten Befunde erfassen. Auswertungszeiten wurden dokumentiert. Als Goldstandard dienten der schriftliche radiologische Befund und die Arztbriefe.

Ergebnisse Die Auswertung von Checkliste 1 ergab zwischen 5,0 und 11,4% falsch negative und 0,7 und 1,4% falsch positive Befunde. Dabei wies der am wenigsten erfahrene Auswerter die höchste diagnostische Genauigkeit auf. Bei Checkliste 2 fiel eine hohe diagnostische Ungenauigkeit mit einer Falsch-Negativ-Rate zwischen 19,3 und 35,0% auf. Die Befundgenauigkeit des Weiterbildungsassistenten war im Vergleich zu dem Facharzt seit 3 Jahren (p = 0,0197) und seit 7 Jahren (p = 0,0046) statistisch signifikant höher. Mit kürzerer Auswertungszeit stieg die Fehlerquote der Befunder deutlich an. Alle 3 Befunder übersahen am häufigsten Frakturen der Wirbelkörper und Rippen.

Schlussfolgerung Unerfahrene Befunder erzielen bei Einsatz einer Checkliste zur Erfassung der wichtigsten Diagnosen von Schockraum-Patienten eine signifikant bessere Befundgenauigkeit im Vergleich mit erfahreneren Auswertern.

Introduction

Based on data from the Trauma Registry of the German Society for Trauma Surgery from the year 2019, over 32 580 patients with major trauma or over 17 664 patients with an ISS (injury severity score) \geq 16 can be expected each year in Germany [1].

The currently valid S3 polytrauma guidelines were created to structure the complex procedures used in the care of major trauma patients. They also serve as a template for the creation of individual treatment algorithms in the individual hospitals [2, 3]. Clinical pathways are based on the ATLS concept (Advanced Trauma Life Support) [4]. The early clinical treatment phase is determined by interdisciplinary interaction between trauma surgery, anesthesiology, and radiology as well as further disciplines depending on the clinical picture of the polytrauma. The initial emergency room phase which focuses on stabilization of the cardiovascular system is typically followed by evaluation of the injury pattern via whole-body computed tomography [5].

Particularly in major trauma patients, performing whole-body CT at this early time can help to significantly lower the mortality rate [6]. Because it is necessary to immediately detect life-threatening injury patterns, reliable radiology reporting is essential. Reporting with the highest diagnostic reliability must also be guaranteed during night and weekend shifts, which are often primarily covered by residents [7–9]. In addition, the participation of different disciplines and colleagues in the care of major trauma

patients requires reliable communication of diagnoses [5, 10]. The goal of structured reporting is to achieve the necessary structure of findings and clear communication of data in text form to the referring physician or to physicians participating in the care of polytrauma patients [11]. In studies, structured reporting approaches have proven successful, particularly in complex clinical settings [11–13].

A two-phase checklist for major trauma reporting was created at our hospital (level I trauma center) on an interdisciplinary basis (trauma surgery, anesthesiology, neurosurgery, visceral surgery, and radiology) for structured reporting in the setting of wholebody CT examinations. The principle of two-phase reporting is also recommended in the currently valid trauma radiology recommendations of the Royal College of Radiologists (RCR, United Kingdom) [14]. The goal of this interdisciplinary checklist is to optimize information exchange, while improving the processes and accuracy of complex major trauma reporting, particularly in the case of professional novices.

To evaluate the clinical value of a checklist for physicians with varying levels of education as part of structured reporting, we examined the quality of the results of checklist reports with respect to major trauma CT examinations by three radiologists with varying levels of professional experience (first-year resident with 10 months of CT experience, radiologist with a total of 8 years of professional experience, and radiologist with a total of 12 years of professional experience).

The present retrospective, monocentric study was performed using anonymized CT datasets and was approved by the responsible ethics committee.

Patients

A full-text database search of the radiology information system (RIS, Nexus.medRIS, Version 8.42, Nexus, Villingen-Schwenningen, Germany) was performed to identify all patients who underwent multidetector CT (MDCT) with a major trauma protocol (► **Table 1**) due to suspected polytraumatization or a corresponding trauma mechanism within a period of six months. 140 patients were retrospectively identified.

CT examination

The indication for major trauma CT examination was determined in accordance with the recommendations of the S3 polytrauma guidelines in all included patients [2]. The CT examination was performed using a standardized protocol (**► Table 1**). In individual cases additional reconstruction (syngo Imaging, VB36a, Siemens, Erlangen, Germany) was performed and archived for special clinical issues. The retrospective trial reading was performed on two high-resolution monitors (Radioforce RX 220, EIZO Europe, Mönchengladbach, Germany).

Structured reporting

For structured reporting, two checklists were created on an interdisciplinary basis for fast (phase 1) and comprehensive (phase 2) evaluation of major trauma CT examinations. The treating disciplines (anesthesiology, trauma surgery, visceral surgery, neurosurgery, otorhinolaryngology, oral and maxillofacial surgery, and ophthalmology) were actively involved in the creation of these two checklists under the direction of radiology. This resulted in the creation of two major trauma CT checklists that meet the clinical requirements of a certified national level I trauma center. Reporting times of 1 minute for the phase 1 checklist and 10 minutes for the phase 2 checklist were defined on an interdisciplinary basis at our hospital as optimal for the established workflows of major trauma care. In the clinical routine, the phase 2 checklist is not a replacement for the final radiology report.

Phase 1 checklist (60 seconds)

The phase 1 checklist is used to record acute life-threatening trauma to the head, chest and abdomen, and pelvis as well as to check the position of an endotracheal tube (**Fig. 1**). Reporting by the responsible radiologist should be performed within the first 60 seconds at the CT scanner using the already acquired axial scans. The primary injuries should be identified so that the optimal diagnostic and therapeutic approach can be determined without delay. In the study, readers had access to transverse scans with a slice thickness of 5 mm in the soft tissue kernel and window. Readers were able to adjust the window settings, and medical-grade displays were used for the evaluation. CT scout images were not used to complete the phase 1 checklist.

▶ Table 1 Standardized polytrauma CT protocol (technique, acquisition, and reformatting).

technique	
CT scanner	2 × 128-slice spiral CT scanner (Somatom Definition Flash, Siemens, Erlangen, Germany)
preparation	positioning of the arms on the body
examination region	 native CCT contrast-enhanced scan from skull to pelvis (in the case of clinical suspicion of fractures of the extremities, particularly of the upper and lower leg, expansion of the scan region)
contrast medium	 120 ml Accupaque 350i. v. Delivery rate of 3 ml/s Delay 55 s
acquisition of con	trast-enhanced scans
scan mode	spiral CT
slice collimation	128 × 0.75 mm
pitch	0.6
automatic dose control	 automatic modulation tube voltage (Care kV; ref. 120 kV) tube current (Care Dose; ref. 30 mAs)
reconstructions	
native CCT	axial (based on the base of the skull) with a 5-mm slice thickness and soft tissue kernel (H20 s soft)
total scan	axial with a 5-mm slice thickness in the soft tissue kernel (kernel B31f medium soft)
CTA head and neck	axial with a 1-mm slice thickness
thorax	additionally, axial with 5-mm slice thickness with a hard kernel (B60f sharp)
secondary reconstructions	 cervical spine: multiplanar reconstruction (MPR) sagittal with a 2-mm slice thickness with a hard kernel (B60f sharp) thoracic spine and lumbar spine: MPR sagittal with a 3-mm slice thickness with a hard kernel (B60f sharp) CTA head and neck: Maximum intensity projection (MIPs) axial, sagittal, and coronal with a 10-mm slice thickness, 5-mm slice interval in the soft tissue kernel (D20f soft)

Phase 2 checklist (10 minutes)

The maximum time needed to complete the phase 2 checklist should be less than 10 minutes. At this time, important diagnoses should be made so that the further treatment plan can be defined. When creating this checklist, the focus was on clinically relevant pathologies in polytrauma patients that directly affect the further course of action (▶ Fig. 2). The checklist is not intended to replace the comprehensive report. In the clinical routine, a comprehensive report that is available in the radiology information system (RIS) is still generated.

The checklist is structured by organ region from cranial to caudal with the intention of achieving the greatest possible clarity for the reader and for those treating the patient. The side of the

Nein	CCT nativ			
	Intrazerebrale Blutung			
	Ödem			
	Einklemmung			
Nein	CT Thorax	links	rechts	
	Pneumothorax			
	Tubuslage nicht korrekt			
	Tubuslage <u>nicht</u> korrekt Blutung Thorax			
	Blutung Thorax			ere 🗆
Nein	Blutung Thorax CT Abdomen			ere
Nein	Blutung Thorax CT Abdomen Lazeration Leber		Milz Ni	
Nein	Blutung Thorax CT Abdomen Lazeration Leber Freie Flüssigkeit Oberbauch		Milz Ni Unterbauch	

▶ Fig. 1 Checklist phase 1, which should be performed within one minute using the initial transverse series of the major trauma CT scan.

injury was reported in the case of sites with a direct effect on further treatment, e. g., in the case of placement of a chest drain for a pneumothorax. Vertebral body fractures are classified based on severity ("vertebral body fracture", "involvement of the posterior edge", "constriction of the spinal canal", or "luxation fracture"), with the level of the fractured vertebral body being documented. The checklist also addresses the extremities and differentiates between the upper arm, lower arm, and wrist or hand and the upper leg, lower leg, and foot. However, it only records the presence of fractures and traumatic blood vessel ruptures. Written notes can be added for any organ region as needed. Possible secondary findings are intentionally not included in the checklists.

In the retrospective trial reading, the datasets and reconstructions in **Table 1** were available for phase 2 checklist reporting. These correspond to the datasets and reconstructions archived on a standard basis in the PACS at our institute.

	CCT nativ						rbelsäule							
Nein		Ja		Ja	Wirbel	körper-Frakt	ur		anten-Frakt			aleinengun	g	Verhakte Luxationsfraktur
	SAB	□ li. □ re.		20		□ Th1 □ Th2			Th1 Th2			Th1		
	Intracerebrale Blutung					Th3			Th3			Th3		03
	Epidurale Blutung	□ li. □ re.			□ C4	Th4	🗆 L4	□ C4	Th4	🗆 L4	🗆 C4	Th4	□ L4	□ C4
	Subdurale Blutung	□ li. □ re.			□ CS □ C6	Th5 Th6	015		Th5	015		Th5		
		□ li. □ re.			0 07	Th7		C7	Th7			Th7		00
	Akute Ischämie		er. 11		_	Th8		10	Th8			Th8		
0	Ödem	□ li. □ re.	Einklemmung			Th9			Th9			Th9		
-	Bulbusperforation	□ li. □ re.	Orbitafraktur			□ Th10			□ Th10			□ Th10		
	Mittelgesichtsfraktur	□ li. □ re.	Retroorbitales Hämi	atom 🗆 li. 🗆 re	-	□ Th11			□ Th11			□ Th11		
	Aneurysma	□ li. □ re.	0.2022.0			□ Th12			□ Th12			□ Th12		
	V.a. Carotisdissektion	□ li. □ re.	Notiz											
	V.a. Vertebralisdissekti					Erfass	t Extre	mitäter	1	Fr	akturen	Gefäß	abriss/	Dissektion
	Kalottenfraktur	□ li. □ re.				Ja				Ja		Ja		
	Intracerebrale Luft	□ li. □ re.					Ober	arm			li. 🗆 re.	□ li.	🗆 re.	
	Schädelbasisfraktur	□ li. □ re.			E.		Unte				li. 🗆 re.		🗆 re.	
	Felsenbeinfraktur	🗆 li. 🗆 re.			Erfasst im CT		Hand	igelenk			li. □ re.	□ li.	🗆 re.	
	Kondylenfraktur CO				tt in		Hand				li. 🗆 re.		🗆 re.	
	HWS-Fraktur				fas		Ober	schenke	2		li. 🗆 re.	□ li.	🗆 re.	
	CO-C1 Dissoziation				T		Unte	rschenk	el		li. □ re.	□ li.	🗆 re.	
	CT Thorax						Fuß				li. □ re.	□li.	🗆 re.	
Nein		Ja												
	Aortendissektion	asc. 🗆												
		desc.												
	Lungenembolie	□ li. □ re.	Notiz											
	Herzbeuteltamponade		HOUL											
	V.a. Zwerchfellriss	□ li. □ re.												
	Pneumothorax	□ li. □ re.												
	Lungenkontusion	□ li. □ re.												
		□ li. □ re.												
	Aspiration													
	Pneumomediastinum													
	Hāmatothorax / relativer Pleuraerguss	🗆 li. 🗆 re.												
	Rippen-Serien-Fraktur	🗆 li. 🗆 re.												
	Fehllage des Tubus													
	to mitgeteilt/korrigiert													
	Thoraxdrainage	□ li. □ re.												
ч □	∽ Fehllage Drainage													
	CT Abdomen													
Nein		Ja												
	Lazeration	Leber	Milz	□ Niere										
	Hämatom	intraperitoneal	retroperitoneal	L Niele										
		wo:	L retroperitoneal		-									
	Arterielle Blutung				-									
	Freie Flüssigkeit (Blut?)	Oberbauch	Unterbauch											
	Freie Luft abdominell	Oberbauch	Unterbauch		_									
	V.a. Ischämie	Darm	Niere											
	Beckenfraktur	🗆 vorne 🗆 hinten												
	- Beckenring	□li. □re.	Notiz											
	- Acetabulum	🗆 li. 🗆 re.												

Fig. 2 a Checklist phase 2, which should be performed within 10 minutes. b Checklist phase 2, which should be performed within 10 minutes.

Based on the phase 1 and phase 2 checklists (▶ Fig. 1, 2), the CT examinations of 140 patients with suspected polytraumatization were evaluated by three radiologists with different levels of education in a retrospective trial reading. The readers were blinded to the type of trauma, the trauma mechanism, and the results of the initial physical examination. The trial reading was performed over a period of 4 weeks on workstations. The radiologists were able to choose the time point to perform reporting.

Reader 1 was a first-year radiology resident with 10 months of CT experience. Reader 2 was a board-certified radiologist with 3 years of experience after board certification. Reader 3 was a board-certified radiologist with 7 years of experience after board certification and clinical specialization in musculoskeletal imaging.

All 3 readers were presented with the anonymized CT datasets for the 140 polytrauma patients along with the phase 1 and phase 2 checklists. The readers were asked to complete the two checklists within the specified time of 60 seconds and 10 minutes, respectively, under study conditions and to document the time needed in seconds. However, the process was not stopped if the defined reading times were exceeded.

The checklist results were evaluated after electronic transfer to a spreadsheet (Excel for Mac 2016, Microsoft; Redmont, WA, USA). Based on the data in the hospital information system (HIS; SAP-R/ 3 IS-H i.s.h.med), the sex and age of each patient, the trauma cause, and the ISS noted in the medical report were additionally recorded. To evaluate the correctness and accuracy of the trial reading, the final radiology report and the injuries documented in the transfer report from the emergency room and in other medical reports during hospitalization were used as the gold standard.

The phase 1 checklist was evaluated to determine whether findings were missed (false negative) or overdiagnosed (false positive) compared to the gold standard.

In the evaluation of the phase 2 checklist, the finding was categorized as (A) correct or (B) missed or incorrectly interpreted (> Table 2). In the case of discrepant findings (B), they were classified either as findings relevant for treatment and patient management (I) or findings not relevant for treatment (II). In constellation BI (missed diagnosis with high relevance for treatment and patient management), further classification was performed: (1) Diagnosis with high clinical relevance, (2) redundant finding (e.g., a vertebral body fracture was correctly diagnosed but an adjacent fracture was not described), and (3) missed diagnoses with minimal effect on further treatment not requiring any primary further intervention. For example, undiagnosed vertebral body fractures, a lacerated spleen, and malpositioning of an endotracheal tube were classified as BI1. Contralateral rib fractures in already diagnosed unilateral rib fractures was classified as BI2. A minor pneumothorax in the case of already diagnosed serial rib fracture was classified as BI3. In contrast, a missed minimal pulmonary contusion was classified as BII. The present study focused on treatment-relevant diagnoses and particularly on diagnoses with high clinical relevance.

Table 2 Classification of the evaluation of the accuracy of the checklist reports in comparison to the final radiology and medical reports (gold standard) for the phase 2 checklist.

А		correct diagnosis
В		missed diagnosis (false negative)
	T	treatment-relevant
	1	high therapeutic relevance
	2	redundant finding
	3	low therapeutic relevance
	II	not treatment-relevant

The spreadsheet software (Excel for Mac 2016, Microsoft; Redmont, WA, USA) and the GraphPad software (GraphPad Quick-Calcs (47)) were used for the statistical tests. The Chi-squared test was used for statistical evaluation. A p-value of <0.05 was considered statistically significant.

Results

Patients

All 140 identified patients were able to be included in the study. 40 of the 140 patients were female (28.6%) and 100 were male (71.4%). The median age of the patients was 38 years, and the average age was 43 years (range 2–93 years). The cause of trauma was a car accident in 40% of patients. Motorcycle accidents (5.7%), bicycle accidents (5.7%), and pedestrian accidents (5.7%) were documented as further types of traffic accident. A fall from a significant height (> 3 m) was the reason for the major trauma CT examination in 10% of cases and a fall from a minimal height (< 3 m) in 19.3% of cases. A suicide attempt was suspected in 3.6% of patients.

CT did not reveal any relevant trauma injury in 17 patients (12.1%). 62 patients (44.3%) required surgical intervention, and 12 patients (9%) died during hospitalization. The average ISS was 16.7 (range: 0–75).

Evaluation of the phase 1 checklist

On average, the radiology resident (reader 1) needed 51 seconds to complete the phase 1 checklist. 7 diagnoses were missed (false negative) in 7 of 140 patients (5.0% of the patients) and there was 1 misdiagnosis (false positive (cerebellar tonsillar herniation)) in 1 patient (0.7%) (\triangleright **Table 3**). The false-negative findings include diagnosis of a hematothorax/chest bleeding (n = 3), an intracerebral hemorrhage (n = 2), a bilateral pneumothorax (n = 1), and a contralateral pneumothorax (n = 1).

On average, reader 2 needed 35 seconds to complete the phase 1 checklist with 16 false-positive findings in 16 patients (11.4%) and 2 false-positive findings (pelvic fracture as well as suspicion of intracerebral hemorrhage) in 2 patients (1.4%). The missed diagnoses include the following in descending order: Pelvic fracture (n = 4), chest bleeding (n = 4), lacerated spleen (n = 2),

pneumothorax (n = 2), ICB (n = 2), free fluid in the upper abdomen (n = 1), and lacerated liver (n = 1).

The most experienced reader (reader 3) needed an average of 62 seconds to complete the checklist, missed 9 diagnoses in 9 patients (6.4%), and incorrectly diagnosed (false positive) 2 injuries in 2 patients (1.4%) (suspicion of intracerebral hemorrhage or chest bleeding). The diagnoses of pneumothorax (n = 3), chest bleeding (n = 2), pelvic fracture (n = 2), lacerated spleen (n = 1), and free fluid in the upper abdomen (n = 1) were classified as false-negative findings.

Evaluation of the phase 2 checklist

Reader 1 (resident) needed an average of 6 minutes and 19 seconds to complete the phase 2 checklist. In total, 34 findings in 27 patients were missed, corresponding to a false-negative rate of 19.3 % of the examinations (► **Table 4**). 18 cases (52.9 %) were classified as treatment-relevant (BI). Discrepancies in 10 cases were classified as BI1 (highly relevant for the further course of treatment). An occipital condyle fracture (► **Fig. 3**), a vertebral body fracture, and a serial rib fracture were missed (► **Table 5**).

Reader 2 needed an average time of 2 minutes and 45 seconds to complete the phase 2 checklist. The false-negative rate com-

pared to the gold standard was 32.1% with a total of 60 missed diagnoses in 45 patients (**> Table 4**). 55 findings (91.7%) were classified as BI, i. e., treatment-relevant, including 13 cases of missed vertebral body fractures and 5 undiagnosed cases of serial rib fracture.

Reader 3 required on average 5 minutes and 37 seconds to complete the phase 2 checklist. In total, 67 findings in 49 patients were missed (35.0% false-negative rate). 61 of these cases (91.0%) were classified as treatment-relevant (**> Table 3**). Vertebral body fractures in 11 cases and serial rib fracture in 9 cases were also the main missed diagnoses here (**> Fig. 4**).

Further examples of false-negative findings in category BI1 include lacerated liver and spleen, pulmonary embolism, intestinal ischemia, midfacial fracture, and malpositioning of an endotracheal tube. However, these diagnoses were only missed in n = 1 for all readers.

The difference in result quality was statistically compared using the Pearson's Chi-squared test. The difference in diagnostic accuracy for the resident compared to the radiologist with 3 years of experience after board certification (p = 0.0197) and compared to the radiologist with 7 years of experience after board certification (p = 0.0046) was statistically significant.

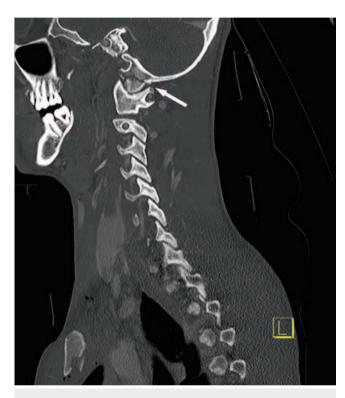
Table 3 Evaluation of the phase 1 checklist of the first-year resident and the two board-certified radiologists with 3 and 7 years of experience after board certification with documentation of average reporting time and false-positive and false-negative rates. Percentage of patients with false-positive and false-negative results in relation to the study population of 140 patients.

phase 1 (< 60 s) in n = 140 patients	reader 1 (resident)	reader 2 (radiologist with 3 years of experience after board certification)	reader 3 (radiologist with 7 years of experience after board certification)
average reading time (Range)	51 s (30–100 s)	35 s (21–70 s)	62 s (25–120 s)
missed diagnoses (false negative)	7 (5.0 %)	16 (11.4%)	9 (6.4 %)
misdiagnosis (false positive)	1 (0.7 %)	2 (1.4%)	2 (1.4 %)

► Table 4 Evaluation of the phase 2 checklist of the first-year resident and the two board-certified radiologists with 3 and 7 years of experience after board certification with documentation of the average reading time and analysis of missed diagnoses regarding their therapeutic consequences (as defined in the classification of ► Table 2).

phase 2 (< 10 min.) in n = 140 patients	reader 1 (resident)	reader 2 (radiologist with 3 years of experience after board certification)	reader 3 (radiologist with 7 years of experience after board certification)
average reading time (Range)	6 min 19 s (3 min 1 s to 15 min 10 s)	2 min 45 s (1 min 30 s to 5 min 41 s)	5 min 37 s (3 min 10 s to 10 min 20 s)
missed diagnoses (false negative)	34	60	67
BI1 (important missed diagnoses)	10	38	45
BI2 (redundant findings)	2	5	3
BI3 (low therapeutic relevance)	6	12	13
BII (not treatment-relevant)	16	5	6

The missed diagnoses in category BI2 were subdural hematoma, vertebral body fractures, serial rib fracture, and large pulmonary contusion, provided that these pathologies were already diagnosed on the opposite side or in the close proximity in the case of vertebral body fractures. In addition, a fracture of the lower arm was not diagnosed in n = 1. The most common diagnoses in category BI3 were pneumothorax (in the case of diagnosed serial rib fracture), pulmonary contusion (in the case of diagnosed serial rib fracture), serial rib fracture (in the case of



▶ Fig. 3 Sagittal reformation of the cervical spine displayed in bone window of a 22-year-old polytraumatized male patient who was involved in a car accident. The image shows the slightly dislocated fracture of the right occipital condyle (white arrow), which was missed in the trial reading and classified as highly treatment-relevant (B11).

diagnosed hematothorax), petrous bone fracture (in the case of diagnosed calvarial and basal skull fracture), and fractures of multiple vertebral processes. The most common BII finding was minor intracranial bleeding in the case of intracranial hemorrhage already diagnosed in close proximity.

Discussion

The goal of the present study was to evaluate the diagnostic accuracy of structured reporting based on two checklists for initial evaluation of major trauma CT scans and to check for possible differences depending on the level of education and experience of the reader.

Missed, i. e., false-negative, findings were most common in the case of the phase 1 checklist. These missed diagnoses were primarily pathologies of the thorax, such as pneumothorax and hematothorax. The frequency of undiagnosed trauma injuries (5 % for reader 1) and (11.4 % for reader 2) can be partly explained by the very short reading times. On average, reader 2 spent 35 seconds



▶ Fig. 4 Axial CT of the chest displayed in bone window of a 66-yearold male trauma patient. The patient experienced a fall from a height of 2 meters. Slightly dislocated fractures of the 3rd-6th left ribs (white arrows) were missed and classified as a treatment-relevant diagnosis of high clinical significance (B11).

► Table 5 Overview of the most commonly missed diagnoses in the phase 2 checklist with high therapeutic relevance (category BI1; see ► Table 2)
that were missed by at least one reader.

discrepant BI1 findings phase 2 checklist in n = 140	reader 1 (resident)	reader 2 (radiologist with 3 years of experience after board certification)	reader 3 (radiologist with 7 years of experience after board certification)
vertebral body fracture	2	13	11
serial rib fracture	2	5	9
calvarial fracture	1	3	3
basal skull fracture	-	3	3
femoral fracture	1	-	4
humerus fracture	-	1	2

completing the phase 1 checklist. A further possible reason for the results could be that only axial scans with a 5-mm slice thickness in the soft tissue kernel were available in phase 1 in order to simulate the conditions of clinical application, i. e., real-time reading at the CT scanner, to the greatest extent possible. In addition, the discrepant findings were not further categorized with respect to relevance in the study. Therefore, for example, an undiagnosed pneumothorax was classified as discrepant, i. e., a false-negative finding, regardless of size.

The high rate of missed diagnoses for all 3 readers seen in the evaluation of the phase 2 checklist was striking. Therefore, falsenegative findings in the checklist in relation to the gold standard were seen in 19.3 % to 35.0 % of the patients. It is important to mention that the relevance of findings was classified conservatively. Therefore, for example, serial rib fracture or pneumothorax is classified as BI1, i.e., missed trauma injury with high relevance for treatment and patient management. A further possible explanation for the high rate of false-negative findings is that the readers were blinded to the clinical data. With knowledge of the trauma mechanism and clinical presentation of the patient, the focus can be on the most probable region of injury, particularly during initial, fast reporting. Moreover, our trial reading was performed under time pressure. In the case of the maximum reading time of 10 minutes for the phase 2 checklist, diagnoses will likely also be missed in the clinical routine. Even in the case of its planned use in the clinical routine, the checklist is not a replacement for the final radiology report that would take significantly longer than 10 minutes particularly in major trauma patients. The RCR recommends a maximum of 60 minutes to write the final radiology report for CT examinations of polytrauma patients [14]. In their trial reading of trauma CT examinations, West et al. report a 12% error rate regarding relevant misdiagnoses [15]. The lower error rate compared to our study could be due to the fact that only abdominal trauma CT scans were included in the evaluation. In their study, Hillier et al. compare the reports produced by residents in the clinical routine from a total of 331 CT examinations with the gold standard report generated by the attending physician. This study also showed a high discrepancy rate of 21.5 %. However, there examinations were not limited here to trauma examinations, which may explain the lower error rate [16].

The fact that the evaluation of the phase 2 checklist in our study only showed false-negative findings could be explained by the fact that reading was performed under time pressure and as a pure trial reading. Therefore, in the case of the selected study design, the probability of perception errors in which case a finding is missed by the radiologist is significantly higher than interpretation errors in which case a detected pathology is interpreted incorrectly [17]. The abovementioned study by Hillier et al. also primarily showed false-negative findings (69.0 %) [16].

With regard to the phase 2 checklist, fractures primarily of the vertebral body and ribs were missed. Additional studies in the emergency setting also had the same result [17, 18]. The high number of missed fractures is probably also due to the fact that thin-slice secondary reconstructions or targeted views of possible fractures were not available for the trial reading.

Particularly in the case of the two experienced radiologists, most false-negative findings were in category BI1, i. e., they were

important treatment-relevant findings. The reason for this could be that the checklist intentionally only addressed initial trauma care and not secondary findings. In addition, it should be noted here that treatment-relevant findings of high clinical relevance were classified conservatively. Therefore, mildly dislocated serial rib fractures were already classified as BI1.

In total, the number of missed redundant findings "BI2", e.g., contralateral mild intracranial bleeding, was low for all 3 readers. This may indicate that the checklists may be able to reduce the "satisfaction of search" effect, i. e., the phenomenon that after one pathology is diagnosed further pathologies are often no longer diagnosed, in relation to such redundant findings. However, the high rate of missed findings in the present trial reading indicates that the "satisfaction of search" error is not reduced by the checklists [19–21].

Between the 3 readers, there were statistically significant differences regarding the reading times for the phase 2 checklists and the diagnostic accuracy. Therefore, with an average reading time of 2 minutes and 45 seconds, reader 2 (radiologist, 3 years of experience after board certification) had the lowest reading time but a high discrepancy rate of 32.1%, while the least experienced reader (first year resident) required the most time with an average reading time of 6 minutes and 19 seconds but had the lowest discrepancy rate compared to the gold standard (19.3%). Longer reading times therefore correlate with lower discrepancy rates. However, on average, all three readers remained under the time of 10 minutes defined for completing the phase 2 checklist in the present study setting. The connection between increased error rates and shorter reading times due to increasing time pressure shown in other studies can only be used as an explanation on a conditional basis [22].

It was striking that the least experienced reader, a first-year resident, had the lower error rate. These study results show both advantages and disadvantages of structured reporting. As Gunderman et al. stated in 2014, structured and thus guided reporting is useful primarily when learning to interpret images report radiologic examinations [23]. The changed workflow can greatly affect the diagnostic ability of experienced radiologists, who typically use their individual reporting patterns, particularly for reporting in the case of complex polytrauma CT examinations. The rigid format of structured reporting or, in our case, checklists runs the risk of being too different from proven reporting processes since there is increased focus on completing the checklist, thereby possibly making it more difficult to detect pathologies [24]. Therefore, there was also speculation that radiologists with many years of experience may be critical of standardized reporting approaches [13, 25]. By introducing checklist reporting of cervical spine CT examinations, Eaton et al. were able to achieve greater diagnostic accuracy, but the difference was not statistically significant [26].

Our study seems to support the theory that a structured reporting approach is advantageous particularly in the training phase since the most inexperienced reader had the highest accuracy. To prove that radiology profits most from structured reporting during training, a comparison with "conventional" reporting would have to be performed. A clear limitation of our study is that reading was performed as pure trial reading of 140 major trauma scans, resulting in the risk of lower diagnostic accuracy compared to reporting under real conditions. Therefore, it is conceivable, for example, that the resident wanted to perform well in the trial reading, but the experienced readers had less time for the study and were interrupted more often. It would be interesting in this context to compare the accuracy of conventional reporting and checklist-based reporting under identical conditions in the study setting. In addition, determination of the reasons for the sometimes very short reading times would be informative.

A prospective study of the checklists under real reporting conditions for major trauma CT scans would thus be desirable, particularly to avoid the effect of pure trial reading and the lack of integration in the clinical context.

CLINICAL RELEVANCE OF THE STUDY

- When used purely for study purposes and under time pressure, a two-phase checklist for initial trauma CT reporting results in high diagnostic inaccuracy, particularly in the case of experienced radiologists.
- Less experienced radiologists can achieve higher diagnostic accuracy compared to experienced readers when using checklists for the reporting of complex examinations.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

- (DGU) DGfU. Jahresbericht 2019; TraumaRegister DGU. http://www.trau maregister-dgu.de/fileadmin/user_upload/traumaregister-dgu.de/docs/ Downloads/Jahresbericht_2019.pdf; Stand: 20.12.2020
- [2] Unfallchirurgie DGfr. S3-Leitlinie Polytrauma/Schwerverletzten-Behandlung; AWMF Register-Nr. 012/019. https://www.awmf.org/uploads/ tx_szleitlinien/012-019I_S3_Polytrauma_Schwerverletzten-Behan dlung_2017-08.pdf; Stand 20.12.2020
- [3] Frink M, Lechler P, Debus F et al. Multiple Trauma and Emergency Room Management. Dtsch Arztebl Int 2017; 114: 497–503. doi:10.3238/arztebl.2017.0497
- [4] Collicott PE, Hughes I. Training in advanced trauma life support. JAMA 1980; 243: 1156–1159
- [5] Braunschweig R, Wawro W. Integrierte Bildgebung bei Polytrauma. Trauma und Berufskrankheit 2005; 7: S207–S213. doi:10.1007/s10039-004-0970-x
- [6] Wirth S, Hebebrand J, Basilico R et al. European Society of Emergency Radiology: guideline on radiological polytrauma imaging and service (short version). Insights Imaging 2020; 11: 135 doi:10.1186/s13244-020-00947-7
- [7] Stevens KJ, Griffiths KL, Rosenberg J et al. Discordance rates between preliminary and final radiology reports on cross-sectional imaging studies at a level 1 trauma center. Acad Radiol 2008; 15: 1217–1226. doi:10.1016/j.acra.2008.03.017

- [8] Ruchman RB, Jaeger J, Wiggins EF 3rd et al. Preliminary radiology resident interpretations versus final attending radiologist interpretations and the impact on patient care in a community hospital. Am J Roentgenol 2007; 189: 523–526. doi:10.2214/Am J Roentgenol.07.2307
- [9] Branstetter BFt, Morgan MB, Nesbit CE et al. Preliminary reports in the emergency department: is a subspecialist radiologist more accurate than a radiology resident? Acad Radiol 2007; 14: 201–206. doi:10.1016/ j.acra.2006.11.001
- [10] Bosmans JM, Weyler JJ, De Schepper AM et al. The radiology report as seen by radiologists and referring clinicians: results of the COVER and ROVER surveys. Radiology 2011; 259: 184–195
- [11] Larson DB, Towbin AJ, Pryor RM et al. Improving consistency in radiology reporting through the use of department-wide standardized structured reporting. Radiology 2013; 267: 240–250. doi:10.1148/radiol.12121502
- [12] Hales BM, Pronovost PJ. The checklist–a tool for error management and performance improvement. J Crit Care 2006; 21: 231–235. doi:10.1016/ j.jcrc.2006.06.002
- [13] Schwartz LH, Panicek DM, Berk AR et al. Improving communication of diagnostic radiology findings through structured reporting. Radiology 2011; 260: 174–181. doi:10.1148/radiol.11101913
- [14] Radiologist TRCo. Standards of practice and guidance for trauma radiology in severely injured patients, Second edition 2015. https://www.rcr. ac.uk/system/files/publication/field_publication_files/bfcr155_traumar adiol.pdf; Stand: 20.12.2020
- [15] West OC, Anderson J, Lee JS et al. Patterns of diagnostic error in trauma abdominal CT. Emerg Radiol 2002; 9: 195–200. doi:10.1007/s10140-002-0225-8
- [16] Hillier JC, Tattersall DJ, Gleeson FV. Trainee reporting of computed tomography examinations: do they make mistakes and does it matter? Clinical Radiology 2004; 59: 159–162. doi:10.1016/s0009-9260(03)00309-x
- [17] Guly HR. Diagnostic errors in an accident and emergency department. Emerg Med | 2001; 18: 263–269
- [18] Pinto A, Reginelli A, Pinto F et al. Errors in imaging patients in the emergency setting. Br J Radiol 2016; 89: 20150914 doi:10.1259/bjr.20150914
- [19] Berbaum KS, Franken EA Jr, Dorfman DD et al. Satisfaction of search in diagnostic radiology. Invest Radiol 1990; 25: 133–140
- [20] Fleck MS, Samei E, Mitroff SR. Generalized "satisfaction of search": adverse influences on dual-target search accuracy. J Exp Psychol Appl 2010; 16: 60–71. doi:10.1037/a0018629
- [21] Schartz KM, Madsen MT, Kim J et al. Trauma in CT: The Role of Severe Injury on Satisfaction of Search Revised. J Am Coll Radiol 2016; 13: 973– 978 e974. doi:10.1016/j.jacr.2016.04.014
- [22] Terreblanche OD, Andronikou S, Hlabangana LT et al. Should registrars be reporting after-hours CT scans? A calculation of error rate and the influencing factors in South Africa. Acta Radiol 2012; 53: 61–68. doi:10.1258/ar.2011.110103
- [23] Gunderman RB, McNeive LR. Is structured reporting the answer? Radiology 2014; 273: 7–9. doi:10.1148/radiol.14132795
- [24] Weiss DL, Langlotz CP. Structured reporting: patient care enhancement or productivity nightmare? Radiology 2008; 249: 739–747. doi:10.1148/radiol.2493080988
- [25] Langlotz CP. Structured radiology reporting: are we there yet? Radiology 2009; 253: 23–25. doi:10.1148/radiol.2531091088
- [26] Lin E, Powell DK, Kagetsu NJ. Efficacy of a checklist-style structured radiology reporting template in reducing resident misses on cervical spine computed tomography examinations. J Digit Imaging 2014; 27: 588–593. doi:10.1007/s10278-014-9703-2