

Aus der
Klinik und Poliklinik für Radiologie
Klinikum der Ludwig-Maximilians-Universität München



**Strukturierte Befundung von CT-Untersuchungen
des Thorax unter Verwendung der CO-RADS-Klassifikation –
Vergleich mit konventionellen Freitextbefunden**

Dissertation
zum Erwerb des Doktorgrades der Medizin
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

vorgelegt von
Magdalena Regehr

aus
Bremen

2024

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

Erster Gutachter:	Prof. Dr. Johannes Rübenthaler
Zweiter Gutachter:	Prof. Dr. Thomas Pfluger
Dritter Gutachter:	Prof. Dr. Gernot Schulte-Altendorneburg
Promovierte Mitbetreuerin:	Priv. Doz. Dr. Christine Schmid-Tannwald
Dekan:	Prof. Dr. med. Thomas Gudermann
Tag der mündlichen Prüfung:	25.06.2024

Inhaltsverzeichnis

Zusammenfassung	V
Abstract (English)	VI
Abbildungsverzeichnis	VII
Tabellenverzeichnis	VIII
Abkürzungsverzeichnis	IX
1 Einleitung und Problemstellung	1
1.1 Epidemiologie und Symptome der Coronavirus-Krankheit 2019	1
1.2 Diagnostik der Coronavirus-Krankheit 2019	2
1.2.1 Reverse-Transkriptase-Polymerase-Kettenreaktion (RT-PCR)	2
1.2.2 Computertomographie (CT)	3
1.3 Strukturierte Befundung	4
1.4 CO-RADS-Klassifikation	8
1.5 Fragestellung und Zielsetzung der Dissertation	15
2 Material und Methoden	16
2.1 Studienaufbau	16
2.1.1 Datenerhebung	16
2.1.2 Ein- und Ausschlusskriterien	16
2.1.3 Bildakquise	17
2.1.4 Erstellung eines Templates	17
2.1.5 Bildanalyse und Erstellung der Befundberichte	19
2.1.5.1 Strukturierte Befunde	20
2.1.5.2 Konventionelle Befunde	20
2.1.6 Bewertung der Befundberichte	20
2.1.6.1 Fragebogen	21
2.1.7 Ethikantrag	22
2.2 Statistische Auswertung	22
2.2.1 Bearbeitung der erfassten Daten	22
2.2.2 Statistische Auswertung der Daten	22
3 Ergebnisse	25
3.1 Patientenkohorte	25
3.1.1 Demographie	25
3.1.2 Grunderkrankungen/Vorerkrankungen	25
3.1.3 Symptome bei Vorstellung	26
3.1.4 Entzündungsparameter bei Aufnahme	26

3.2	Vergleich der Befundtypen	26
3.2.1	Formale Aspekte.....	27
3.2.1.1	Strukturiertheit.....	27
3.2.1.2	Befundlänge	27
3.2.1.3	Verständlichkeit.....	27
3.2.1.4	Fehlende Informationen im Befund	27
3.2.1.5	Fehlende Informationen in der Beurteilung	28
3.2.2	Klinische Relevanz	29
3.2.2.1	Beantwortung der klinischen Fragestellung	29
3.2.2.2	Entscheidung zum weiteren diagnostischen und/oder therapeutischen Vorgehen	29
3.2.3	Gesamtqualität.....	30
3.2.3.1	Vertrauen in die Information aus dem Befund.....	30
3.2.3.2	Gesamtqualität des Befunds	30
3.3	Korrelation zwischen CO-RADS-Stadien und RT-PCR- Testergebnissen.....	30
4	Diskussion.....	31
4.1	Zusammenfassung der Ergebnisse.....	31
4.2	Formale Aspekte	31
4.2.1	Struktur und Klarheit	31
4.2.2	Angemessene Befundlänge.....	32
4.2.3	Verständlichkeit	32
4.2.4	Vollständigkeit/fehlende Informationen	33
4.3	Klinische Relevanz und Entscheidung über das weitere diagnostische und/oder therapeutische Vorgehen	34
4.4	Gesamtqualität und Vertrauen in die Information	34
4.4.1	Korrelation zwischen radiologischem Befund und RT-PCR- Testergebnis	35
4.5	Limitationen.....	36
4.5.1	Schlussfolgerung	37
4.6	Ausblick.....	37
	Literaturverzeichnis	40
	Anhang A	48
	Danksagung.....	49
	Affidavit.....	50

Zusammenfassung

Hintergrund

Nachdem radiologische Befundberichte seit Jahrzehnten als Freitexte formuliert wurden, gab es in den letzten Jahren Bestrebungen, die strukturierte Befundung im radiologischen Alltag zu etablieren. Ihre Vorteile gegenüber konventionellen Prosa-berichten im Hinblick auf Klarheit, Qualität, Verständlichkeit und Vollständigkeit konnten bereits in diversen Studien belegt werden.

Ziel

Ziel dieser Arbeit ist es, zu untersuchen, ob im Kontext der Coronavirus-Krankheit-2019 (COVID-19) die strukturierte Befundung unter Anwendung der COVID-19 Reporting and Data System (CO-RADS) -Klassifikation dem konventionellen Prosa-bericht im Hinblick auf formale Aspekte, klinische Relevanz und Gesamtqualität überlegen ist.

Material & Methoden

Für eine Kohorte von 50 Patientinnen und Patienten wurden zusätzlich zu den bestehenden Freitextbefunden (CRs) strukturierte Befunde (SRs) erstellt und nach der CO-RADS-Klassifikation eingestuft. Daraufhin bewerteten zwei Pneumologinnen beide Befundungsmethoden hinsichtlich formaler Aspekte, klinischer Relevanz und Gesamtqualität auf Likert-Skalen. Zur statistischen Analyse wurden Wilcoxon- und McNemar-Tests verwendet.

Ergebnisse

Strukturierte Befunde erhielten signifikant höhere Bewertungen für sämtliche formalen Aspekte. Es ergab sich ein signifikanter Unterschied in der Strukturiertheit ($p < 0,001$), in der Angemessenheit der Befundlänge ($p < 0,001$), der Verständlichkeit ($p = 0,044$) sowie bezüglich fehlender Informationen ($p < 0,001$). Die klinische Relevanz betreffend ergab sich kein signifikanter Unterschied zwischen CRs und SRs ($p > 0,05$). Die Gesamtqualität wurde für SRs signifikant höher bewertet als für CRs ($p < 0,05$).

Zusammenfassung

Nach Bewertung der Pneumologinnen sind strukturierte Befunde dem konventionellen Freitextbefund hinsichtlich formaler Aspekte und Gesamtzufriedenheit überlegen. Bezüglich der klinischen Relevanz zeigt sich kein signifikanter Unterschied.

Abstract (English)

Background:

For decades, radiological reports were formulated as free texts. In recent years, however, there have been increasing efforts to establish structured reports in the everyday radiological practice. Its advantages over conventional prose reports (CRs), particularly with regard to clarity, quality, comprehensibility, and completeness, have already been demonstrated in various studies.

Objective:

The objective of this study is to investigate whether, in the context of the 2019 coronavirus disease (COVID-19) pandemic, structured reporting (SRs) using the COVID-19 Reporting and Data System (CO-RADS) classification is superior to conventional prose reporting in terms of formal aspects, clinical relevance, and overall quality.

Methods:

A cohort of 50 patients with existing free-text reports was examined. In the context of this study, structured reports were generated for these patients and they were subsequently classified according to the CO-RADS classification. Two pulmonologists then evaluated both reporting methods with respect to formal aspects, clinical relevance, and overall quality on Likert scales. Wilcoxon and McNemar tests were used for statistical analysis.

Results:

Structured reports received significantly higher ratings for all formal aspects. There was a significant difference in structuredness ($p < 0.001$), adequacy of report length ($p < 0.001$), comprehensibility ($p = 0.044$), and missing information ($p < 0.001$). Regarding clinical relevance, there was no significant difference between CRs and SRs, ($p > 0.05$). Overall quality was rated significantly higher for SRs than for CRs, ($p < 0.05$).

Summary:

According to the pulmonologists, structured reports are superior to conventional free-text reports in terms of formal aspects and overall quality. There was no significant difference regarding clinical relevance of the different report types.

Abbildungsverzeichnis

Abbildung 1:	Exemplarische CT einer CO-RADS-Kategorie 5.....	11
Abbildung 2:	Exemplarische CT einer CO-RADS-Kategorie 5.....	12
Abbildung 3:	Exemplarische CT einer CO-RADS-Kategorie 2.....	13
Abbildung 4:	Exemplarische CT einer CO-RADS-Kategorie 2.....	14
Abbildungen 5a) und 5b):	Exemplarischer Auszug des CO-RADS- Templates aus dem Smart Reporting Tool.	19

Tabellenverzeichnis

Tabelle 1:	Erläuterung der CO-RADS-Level 0-6.....	10
Tabelle 2:	Vor- und Grunderkrankungen der Patientenkohorte.	25
Tabelle 3:	Symptome des Patientenkollektivs in Anzahl und %.	26
Tabelle 4:	Vollständigkeit der Befundformen SRs versus CRs im Hinblick auf verschiedene Items mit Signifikanz der Unterschiede.....	28
Tabelle 5:	Vollständigkeit der Beurteilung von CRs und SRs in Gesamtzahl n und %.....	29

Abkürzungsverzeichnis

AI	Künstliche Intelligenz (engl. artificial intelligence)
Et al.	Lat. et alii, und andere
AGIT	Arbeitsgemeinschaft Informationstechnologie
ARDS	Acute Respiratory Distress Syndrome, akutes Lungenversagen
Bspw.	Beispielsweise
Bzw.	Beziehungsweise
Ca.	Circa
COVID-19	Coronavirus-Krankheit-2019
CO-RADS	COVID-19 Reporting and Data System
CR/CRs	Konventioneller Befund (engl. conventional report) / konventionelle Befunde (engl. conventional reports)
CT	Computertomographie
DD	Differentialdiagnose
DIN	Deutsche Industrienorm
DIVI	Deutsche Interdisziplinäre Vereinigung für Intensiv- und Notfallmedizin
DRG	Deutsche Röntgengesellschaft
Engl.	Englisch
HRCT	Hochauflösende Computertomographie (engl. high resolution computed tomography)
IS	Informationssystem
ILD	Interstitielle Lungenerkrankung (engl. interstitial lung disease)
IQA	Interquartilsabstand
KI	Konfidenzintervall
KIS	Krankenhausinformationssystem
LAMP	ist ein Akronym für den kombinierten Einsatz von Programmen: Betriebssystem Linux, Webserver Apache, Datenbank MySQL, Programmiersprache PHP
LMU	Ludwig-Maximilians-Universität
Mdn	Median

MRT	Magnetresonanztomographie
NZ	Nicht zutreffend
PACS	Picture Archiving and Communication System
PAT-ID	Patientinnen/-enidentifikationsnummer
RIS	Radiologieinformationssystem
PCR	Polymerase-Kettenreaktion
RT-PCR	Reverse-Transkriptase-Polymerase-Kettenreaktion (engl. reverse transcriptase polymerase chain reaction)
RKI	Robert Koch-Institut
RNA	Ribonukleinsäure
RSNA	Radiological Society of North America
SARS-CoV2	Severe Acute Respiratory Syndrome Coronavirus 2
SD	Standardabweichung (engl. standard deviation)
SR/SRs	Strukturierter Befund (engl. structured report) / strukturierte Befunde (engl. structured reports)
SW	Spannweite
Tbl.	Tabelle
u.a.	Unter anderem
WHO	Weltgesundheitsorganisation (engl. World Health Organization)
z.B.	Zum Beispiel
ZNA	Zentrale Notaufnahme

1 Einleitung und Problemstellung

1.1 Epidemiologie und Symptome der Coronavirus-Krankheit 2019

Das Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) wurde erstmalig im Dezember 2019 in Wuhan, China, offiziell dokumentiert. Dieses durch Tröpfcheninfektion verbreitete Virus verursacht die Coronavirus-Krankheit-2019 (COVID-19), die sich zumeist als Atemwegserkrankung mit Symptomen wie Schnupfen, Husten, Fieber und/oder pulmonaler Beteiligung in Form einer Lungenentzündung manifestiert. Zudem sind allgemeine Symptome wie Apathie, Appetit-, Geruchs-, Geschmacks- und Gewichtsverlust, Schmerzen in Form von Kopf- und Gliederschmerzen sowie Übelkeit bzw. Erbrechen mit dem Auftreten des Virus assoziiert (1).

COVID-19 äußert sich charakteristischerweise in variablen Verlaufsformen – grundsätzlich kann dabei zwischen milden, moderaten und schweren Verläufen unterschieden werden. Laut Schätzung des Robert Koch-Instituts (RKI) bekunden 55% bis 85% der Infizierten spürbare Beschwerden und zeigen typische Symptomkomplexe (2). Während sich milde Verläufe symptomlos oder in leichten Erkältungssymptomen manifestieren, können schwere Verläufe bis hin zu einem akuten Lungenversagen (Acute Respiratory Distress Syndrome; ARDS) mit letaler Folge führen (3).

Seit dem Ausbruch von COVID-19 in China Ende Dezember 2019 hat sich das hochkontagiöse Virus in einem sehr schnellen Tempo kontinentübergreifend in mehr als 190 Ländern verbreitet. Am 11. März 2020 wurde die weltweite Ausbreitung als COVID-19-Pandemie deklariert und galt lange Zeit als nicht unter Kontrolle. Zum jetzigen Zeitpunkt (4.10.2023) haben sich laut Weltgesundheitsorganisation (WHO) insgesamt ca. 770 Millionen Menschen infiziert und es sind ca. 6,9 Millionen Menschen mit nachgewiesener SARS-CoV-2-Infektion gestorben (4).

In Deutschland wurde der erste Patient mit dem neuartigen Coronavirus SARS-CoV-2 am 27. Januar 2020 laborchemisch detektiert und dokumentiert. Allein in Deutschland haben sich laut RKI ca. 32 Millionen Menschen infiziert und dabei sind 146.000 Menschen gestorben (1).

Das Intensivregister der deutschen Interdisziplinären Vereinigung für Intensiv- und Notfallmedizin (DIVI) meldete deutschlandweit einen Höchststand von über 5.500 intensivmedizinischen Behandlungsfällen Erwachsener mit COVID-19 (Stand November 2022) (5).

1.2 Diagnostik der Coronavirus-Krankheit 2019

1.2.1 Reverse-Transkriptase-Polymerase-Kettenreaktion (RT-PCR)

Die Diagnose COVID-19 wird in der Regel mittels Reverse-Transkriptase-Polymerase-Kettenreaktion (RT-PCR) gestellt. Für den diagnostischen Nachweis von SARS-CoV-2 mittels RT-PCR werden typischerweise nasopharyngeale Abstriche genutzt. Außerdem werden Proben aus den tieferen Atemwegen, wie bspw. einer bronchoalveolären Lavage (BAL), oder Sputum verwendet. Studien haben gezeigt, dass die zuletzt genannten Proben im Vergleich die größte Positivrate aufweisen (6). Die RT-PCR dient als quantitative Vervielfältigungsmethode von Nukleinsäuren in Echtzeit. Die RT-PCR beruht auf dem Prinzip der konventionellen Polymerase-Kettenreaktion (PCR) und ermöglicht eine Quantifizierung und einen Nachweis von Ribonukleinsäure (RNA) durch Einbau fluoreszierender Sonden (7). Nach dieser Methode werden bei der Quantifizierung Fluoreszenzen (Nukleinsäure-Templates) gemessen, die sich proportional zum Anstieg der PCR-Produkte vermehren. Derzeit findet der RT-PCR-Nachweis aufgrund seiner Vorteile als spezifischer und einfacher quantitativer Assay für den Nachweis aller Coronaviren inklusive SARS-CoV-2 bevorzugt Verwendung. Darüber hinaus ist die Echtzeit-RT-PCR schnell und sensitiv, welches für die Diagnosestellung einer frühen Infektion äußerst hilfreich ist (8-16).

Die quantitative RT-PCR hat sich als diagnostischer Goldstandard etabliert, liefert jedoch mitunter falsch negative Befunde und ist abhängig von Zeitpunkt und Durchführung des tiefen Oro- oder Nasopharyngealabstriches sowie von der Dauer der Symptome, der Viruslast und der Qualität der untersuchten Probe (17). Die Viruslast bei COVID-19 erreicht kurz vor oder zu Beginn der Symptome ihren Höhepunkt und nimmt dann ab (18). Die Übertragbarkeit liegt durchschnittlich zwischen 2 und 5 Tagen und die Inkubationszeit zwischen 4 und 12 Tagen (19). Die Sensitivität der RT-PCR korreliert mit der RNA-Menge in jeder Probe. Folglich ist sie für die Abklärung des Verdachts auf eine akute Infektion bei hoher

Viruslast kurz vor Beginn und während der Symptomatik am aussagekräftigsten, wohingegen die Sensitivität des Tests im sehr frühen oder späten Infektionsstadium deutlich geringer ist (20). Zudem vergehen bis zum Erhalt des Testergebnisses oft mehrere Stunden. Dies belastet die Notaufnahmen, in denen betroffene Patientinnen und Patienten untergebracht sind, bevor sie auf eine Normal- oder COVID-19-Station verlegt werden können (6, 21). Um eine schnellere Entscheidungsfindung zu ermöglichen, bietet sich in diesen Fällen als zusätzliche diagnostische Maßnahme die Computertomographie (CT) des Thorax an.

1.2.2 Computertomographie (CT)

Heutzutage gilt die CT als eines der wichtigsten radiologischen Verfahren zur nichtinvasiven Diagnostik. Mit der Entwicklung des ersten Computertomographen 1971 setzten der 1919 in England geborene Ingenieur Godfrey N. Hounsfield und der 1924 geborene, südafrikanisch-US-amerikanische Physiker Allan McLeod Cormack einen Meilenstein für die Radiologie. Für diese Arbeit wurden 1979 beide voneinander unabhängig tätigen Forscher mit dem Nobelpreis für Medizin prämiert (22).

Bei der CT handelt es sich um ein Röntgen-Schnittbildverfahren. Dies ermöglicht eine detaillierte, überlagerungsfreie Abbildung der untersuchten Körperregion. Die Untersuchung erfolgt, während der Patient auf einer Liege durch einen kurzen Tunnel, die sogenannte Gantry (Abtasteinheit), gefahren wird. In der Abtasteinheit rotiert eine Röntgenröhre um die Patientinnen und Patienten und fertigt multiple Projektionen aus unterschiedlichen Richtungen an. Mit Hilfe von Detektoren können hierbei Intensitätsunterschiede der Röntgenstrahlung differenziert und gemessen werden. Aus den unterschiedlichen Absorptionswerten ergeben sich digitale Schnittbilder (11,12). Bis zum heutigen Tag hat sich die CT in den vergangenen Jahren mit Einführung der Spiral-CT und der Erweiterung vom Einzel- zum Mehrzeiler stetig weiterentwickelt, wodurch die Methode zeitlich beschleunigt und qualitativ optimiert wurde (23).

Die CT ist außerdem eine Untersuchungsmodalität, die aufgrund ihrer breiten Verfügbarkeit, kurzen Untersuchungsdauer und hohen Sensitivität bei der Detektion atypischer Pneumonien einen hohen Stellenwert bei der Diagnostik der COVID-19-Pneumonie innehat (24). Dabei dient die CT nicht nur zur Detektion einer

COVID-19-Pneumonie, sondern sie erlaubt auch die Einschätzung von Ausmaß und Schweregrad der Lungenerkrankung sowie die Detektion möglicher Komplikationen, wie beispielsweise eines Pneumothorax (25). Basierend auf dem CT-Befund lassen sich somit Risikostratifizierung und Therapieentscheidung vornehmen (26), weshalb dieses bildgebende Verfahren eine Hauptrolle bei Primärdiagnostik und Follow-up von Patientinnen und Patienten mit COVID-19 spielt (27-31). Sowohl die Fleischner Society als auch die Deutsche Röntgengesellschaft (DRG) empfehlen bei Patientinnen und Patienten mit COVID-19-typischen Symptomen, deutlicher klinischer Verschlechterung und wegweisender klinischer Vorgeschichte eine hochauflösende Computertomographie (HRCT) des Thorax ohne Kontrastmittel (32).

Eine Vergleichsstudie zum Thema CT-Befunde von RT-PCR-negativen und RT-PCR-positiven Fällen bei COVID-19-Patientinnen und -Patienten schlussfolgerte, dass die CT eine wesentliche Rolle für die Diagnose, Isolierung und Behandlung in Fällen von COVID-19 trage (33).

1.3 Strukturierte Befundung

Radiologische Befundberichte werden traditionell als Freitext verfasst. Bereits Anfang des 20. Jahrhunderts wurde der radiologische Befund im Prosastil jedoch aufgrund seiner Heterogenität hinterfragt (34, 35). Entsprechend wurden erste Versuche unternommen, den Befund in Sprache und Aufbau zu standardisieren. Zunächst blieben diese Bestrebungen jedoch ohne Erfolg. Der radiologische Befund bestand trotz technischen Fortschritts durch Weiterentwicklung sowie durch erhebliche Technologisierung und trotz steigender Befundzahlen sowie Arbeitsverdichtung in der Radiologie unverändert im individuellen Prosastil fort (36). Seit den 80er Jahren wurden die aus der radiologischen Befundung im Prosastil resultierenden Schwierigkeiten in diversen Studien thematisiert und wissenschaftlich beleuchtet:

Konventionelle Freitextberichte (engl. conventional reports, CRs) bergen den Nachteil, dass sie in Form, Länge, Sprache und Inhalt individuell stark von ihrem jeweiligen Verfasser abhängen. So führt ein persönlicher Schreib- und Diktierstil mit unklar definiertem Vokabular ohne Längenvorgabe zu einer markanten Inhomogenität der konventionellen Berichte. Die Vergleichbarkeit zwischen den CRs

individueller Verfasser ist kaum gegeben (37, 38). Zudem bieten eine unterschiedliche Expertise der befundenden Ärzte in spezifischen Fachbereichen mit klinischem Hintergrundwissen und unterschiedliche Jahre an Berufserfahrung eine große Variationsbreite. Je nach Praxiserfahrung des berichtenden radiologischen Arztes beantworten die CRs die klinische Fragestellung möglicherweise formal und inhaltlich uneindeutig für die Kliniker. So kann eine Entscheidung über weitere Behandlungsindikationen basierend auf uneindeutigen oder unvollständigen CRs erschwert werden. Auch eine wissenschaftliche Auswertung ist mit Freitextbefunden im Prosastil deutlich erschwert (39).

Die Radiological Society of North America (RSNA) versuchte, den Nachteilen der Freitextbefunde im Jahr 2006 mit der Publikation eines radiologischen Vokabulariums namens RadLex zu begegnen (40). Das Wörterbuch definiert und erläutert medizinische Begriffe in radiologischen Befunden, um eine standardisierte Sprache zu etablieren (34, 35, 41, 42).

Die Fleischner Society lancierte erstmalig 1984 (und erneut 1996) ebenfalls ein Glossar und Empfehlungen für die Thoraxradiographie bzw. CT, welche zuletzt 2008 aktualisiert wurden (43). Eine Übertragung dieser Begriffe ins Deutsche wurde in enger Anlehnung an die Terminologie der Fleischner Society 1996 in einer richtungsweisenden Übersichtsarbeit publiziert (44).

Neben der Verwendung eines einheitlichen, definierten Vokabulars ist die strukturierte Gliederung ein Ansatz zur qualitativen Verbesserung des radiologischen Befundberichts. Dabei können nach Hackländer drei verschiedene Stufen der strukturierten Befundung unterschieden werden (45):

Die erste Stufe befasst sich mit der thematischen Strukturierung des Befundes in einzelne Abschnitte. Sie beinhaltet eine praxisorientierte, systematische Gliederung entsprechend der deutschen Röntgenverordnung in Unterpunkte, wie zum Beispiel bei einer CT des Thorax und Abdomens in „Thorax“, „Abdomen“ und „Skelett“ (35). In der zweiten Stufe wird die bestehende Gliederung um die Verwendung vorgefertigter Textbausteine ergänzt. Die dritte Stufe fordert zusätzlich die explizite Verwendung präformierter Textbausteine aus kontrollierten radiologischen Vokabularien wie bspw. dem RadLex oder Glossar der Fleischner Society.

Zur Umsetzung einer strukturierten Befundung bietet sich neben Checklisten die Verwendung einer Befundvorlage, eines sogenannten Templates, an (46, 47). Besagte Templates können beispielsweise auf den von der RSNA konzipierten Webseiten RadReport und RadElement oder auf der Internetseite der Firma Smart Reporting erstellt, entnommen und der Allgemeinheit zur Verfügung gestellt werden (25, 30, 48-50).

Eine Metaanalyse von 37 Studien kam zu dem Schluss, dass strukturierte Befundsysteme die Interpretation und Berichterstellung von bildgebenden Untersuchungen vereinfachen. Ferner bieten sie einen Rahmen für eine konsistente Generierung von Empfehlungen und verbessern die Qualität der Patientenversorgung (51).

Aktuelle Literatur im Fachbereich Onkologie belegt eine deutliche Verbesserung der Qualität der Berichte sowie in der Übermittlung von Ergebnissen durch strukturierte Befundung. Am Beispiel von Patientinnen und Patienten mit hepatozellulärem Karzinom (HCC) konnte gezeigt werden, dass die Verwendung einer auf zuvor definierten Kriterien basierenden Befundvorlage zu vollständigeren und präziseren Berichten führte als herkömmliche Freitextberichte (52). Diese Studienergebnisse decken sich mit Beobachtungen bei Rektum- und Bauchspeicheldrüsenkrebs (53, 54). Mehrere Studien konnten unabhängig voneinander zeigen, dass die Verwendung strukturierter Befundvorlagen zu einer signifikanten Verbesserung von Verständlichkeit und Vollständigkeit der radiologischen Berichte führte (53-55). Auch außerhalb der Onkologie wurde erwiesen, dass eine strukturierte Befundung zu einer relevanten Qualitätssteigerung führt. Gutachten für Magnetresonanztomographie (MRT)-Untersuchungen der Schulter, CT-Angiographie des Lungenkreislaufs und kardiovaskuläre Bildgebung profitieren von der Verwendung strukturierter Befundvorlagen (56-58). Fachdisziplinübergreifend enthielten die entsprechenden Berichte deutlich mehr relevante Informationen und Kliniker zogen die strukturierten Berichte dem Freitextformat vor.

Auch aus Sicht der europäischen und deutschen Fachgesellschaften bietet der Gebrauch einer strukturierten Befundung diverse Vorteile:

Formale/strukturelle Vorteile

Durch direkte Elemente wird eine Vollständigkeit des Befundes gewährleistet und implizite Angaben werden vermieden. Der formale und inhaltliche Aufbau ist unabhängig vom Verfasser. Für Weiterbildungsassistenten kann ein standardisierter Befundinhalt als übersichtlicher und strukturierter Leitfaden dienen. Eine korrekte und krankheitsspezifische Terminologie in der Vorlage kann ebenfalls hilfreich sein, da so Mehrdeutigkeiten verhindert werden. Zudem kann durch den Einsatz genormter radiologischer Lexika der Befundinhalt semantisch erschlossen werden. Strukturierte Befunde (engl. structured reports, SRs) erleichtern die Erstellung von Hyperlinks zwischen Befund und Bildmaterial (59, 60).

Inhaltliche und klinische Vorteile

Eine zuvor mit den Klinikern konsentierete Befundvorlage kann dediziert auf die klinische Fragestellung und Optionen der Weiterbehandlung eingehen. Auch können Krankheitsverläufe mit Hilfe einer Zeitachse im strukturierten Befund vereinfacht chronologisch dargestellt und verglichen werden. Durch die Verwendung sogenannter Actionable Reports können strukturierte Befunde mit einem automatisierten Vermerk versehen werden, der eine Nachkontrolle pathologischer Befunde vereinfacht. Bei einer multidisziplinären Verwendung der SRs ist auch eine Feedback-Funktion von Klinikern und Pathologen an Radiologen denkbar (50, 51).

Forschung und Ausblick

Bereits erhobene und in einer Datenbank gespeicherte, strukturierte Daten können in der sekundären Datennutzung wissenschaftlich ausgewertet und verglichen werden. Zukünftig ist auch eine potenzielle Verknüpfung der strukturierten Befunde mit klinischen Informationen, z. B. Laborparametern, denkbar. Dies könnte eine integrierte computergestützte Diagnosefindung unterstützen, um daraus klinische Konsequenzen leichter abzuleiten (48-50).

Der radiologische Befund dient in erster Linie der Kommunikation zwischen Radiologinnen und Radiologen und zuweisenden bzw. weiterbehandelnden Ärztinnen und Ärzten (61).

Auch aktuell bestehen weiterhin Bestrebungen, von der derzeit noch überwiegend praktizierten Freitextbefundung zu einer strukturierten Befundung überzugehen. Ziel ist es, die Kommunikation zwischen radiologischen und klinischen Ärztinnen und Ärzten sowie die Patientenversorgung zu verbessern, indem Qualität und Lesbarkeit der Ergebnisse gesteigert und vereinheitlicht werden (62).

1.4 CO-RADS-Klassifikation

Um behandelnden Ärztinnen und Ärzten bereits vor Erhalt des RT-PCR-Ergebnisses eine Entscheidungsgrundlage für weitere therapeutische Maßnahmen zur Verfügung zu stellen, entwickelte die COVID-19-Arbeitsgruppe für standardisierte Befundung der Niederländischen Radiologischen Gesellschaft (COVID-19 Standardized Reporting Working Group of the Dutch Radiological Society) Anfang 2020 ein standardisiertes Befundungsschema, mithilfe dessen sich die Wahrscheinlichkeit für eine COVID-19-Infektion anhand von CT-Untersuchungen des Thorax ermitteln lässt: das COVID-19 Reporting and Data System (CO-RADS) (63).

Die CO-RADS-Klassifikation beschreibt Veränderungen des Lungenparenchyms und kategorisiert die Wahrscheinlichkeit für das Vorliegen einer SARS-CoV-2-Infektion. Das Klassifikationsschema gliedert sich in insgesamt sieben Kategorien (von 0 bis 6). Die einzelnen Graduierungen bewerten den Verdacht auf Lungenbeteiligung bei COVID-19 auf einer Skala von 1 (sehr gering) bis 6 (nachgewiesene Infektion), wobei die Kategorie 0 aufgrund unzureichender Qualität bspw. aufgrund eines Artefaktes „nicht auswertbar“ bedeutet. Das Schema ist insbesondere für den Einsatz bei Patientinnen und Patienten mit mittelschweren bis schweren Symptomen von COVID-19 vorgesehen (63, 64).

Jedem CO-RADS-Level lassen sich entsprechende CT-morphologische Merkmale zuordnen:

- Das Level CO-RADS 1 liefert nur einen sehr geringen Verdacht auf eine COVID-19-Pneumonie. Es zeigen sich hier in der CT nicht-infektiöse Erkrankungen oder ein Normalbefund. Pathologien, die diese Einordnung rechtfertigen, beinhalten bspw. Emphyseme, Lungentumoren, Fibrose oder perifissurale

Lymphknoten. Interstitielle Septumverdickungen mit Pleuraerguss und Milchglastrübungen zählen ebenfalls zu dieser Kategorie, wenn sie als pulmonal-venöse Stauung mit Lungenödem interpretiert werden (65, 66).

- Das CO-RADS-Level 2 birgt einen geringen Verdacht auf pulmonale Beteiligung bei COVID-19. Im CT zeigen sich Hinweise für eine andere Lungeninfektion, wie z.B. Veränderungen im Rahmen einer Bronchiolitis, verdickte Bronchuswände, zentrilobuläre Knötchen, lobäre oder segmentale Konsolidierungen und Kavitationen.
- Nicht eindeutige Zeichen einer pulmonalen Beteiligung bei COVID-19 finden sich in der Kategorie CO-RADS 3. Sie beinhaltet CT-Veränderungen, die sowohl durch eine COVID-19-Pneumonie, als auch durch andere virale Pneumonien oder nicht-infektiöse Ursachen bedingt sein können. CT-Merkmale sind Milchglastrübungen mit perihilärem, homogen ausgedehntem oder fleckigem, aber nicht zentrilobulärem oder subpleuralem Verteilungsmuster. Zudem gehören Milchglastrübungen mit interlobulären Septumverdickungen, wenn diese nicht eindeutig stauungsbedingt sind, und ein isoliertes Konsolidierungsmuster, welches am ehesten als Zeichen für eine organisierende Pneumonie steht, in die Kategorie CO-RADS 3.
- Ein hoher COVID-19-Verdacht manifestiert sich in der Kategorie CO-RADS 4. Hierbei zeigen sich typische Anzeichen einer COVID-19-Pneumonie, die jedoch mit den Erscheinungsformen anderer viraler Pneumonien überlappen. Wegweisend sind multifokale Milchglastrübungen wie bei Level CO-RADS 5 (siehe weiter unten), die jedoch entweder keinen Kontakt zur viszeralen Pleura, eine unilaterale oder aber peribronchovaskuläre Lokalisation aufweisen. Zudem beinhaltet das Level CO-RADS 4 COVID-19-typische Veränderungen, wenn diese in schwer und diffus vorerkrankten Lungen auftreten.
- Die Kategorie CO-RADS 5 beschreibt einen sehr hohen Verdacht auf eine COVID-19-Infektion und erfasst CT-Befunde, die für die entsprechende Pneumonie pathognomonisch sind. Hierzu zählen Milchglastrübungen mit oder ohne Konsolidierungen. Charakteristisch zeigen sich ein multifokales, bilaterales Verteilungsmuster mit Kontakt zur Pleura, wobei subpleurale Ausparungen auftreten dürfen. Häufig finden sich verdickte Lungengefäße im Bereich der Milchglastrübungen. Als Ausdruck der Veränderung der pulmonalen Verdichtungen im zeitlichen Verlauf der COVID-19-Pneumonie fallen in die Kategorie CO-RADS 5 zudem intralobuläre Septumverdickungen sowie assoziierte Zeichen einer organisierenden Pneumonie.

- Die Kategorie CO-RADS 6 beschreibt die bereits oben genannten, für COVID-19 typischen pulmonalen Veränderungen in der CT, fordert jedoch zusätzlich einen Erregernachweis mittels Realtime-PCR.

Eine Zusammenfassung der CO-RADS-Klassifikation findet sich in der untenstehenden **Tabelle 1**.

Tabelle 1: Erläuterung der CO-RADS-Level 0-6.

CO-RADS-Kategorie	Wahrscheinlichkeit für pulmonale Beteiligung	Beschreibung
CO-RADS 0	nicht beurteilbar	Die Untersuchungsqualität ist für die Beantwortung der Fragestellung nicht ausreichend.
CO-RADS 1	sehr niedrig	Kein Nachweis infektypischer Lungenpathologien.
CO-RADS 2	niedrig	Die pulmonalen Veränderungen sind typisch für eine Infektion mit einem anderen Erreger als SARS-CoV-2.
CO-RADS 3	nicht eindeutig	Die pulmonalen Veränderungen sind vereinbar mit COVID-19, aber auch mit anderen Erkrankungen.
CO-RADS 4	hoch	Die pulmonalen Veränderungen sind typisch für COVID-19, aber auch andere virale Pneumonien.
CO-RADS 5	sehr hoch	Die pulmonalen Veränderungen sind COVID-19-typisch.
CO-RADS 6	sicher	Die SARS-CoV-2-Pneumonie wurde mittels PCR-Test nachgewiesen.

Folgende CT-Beispielbefunde der Lunge aus unserer Patientenkohorte illustrieren exemplarisch zwei unterschiedliche CO-RADS-Kategorien:

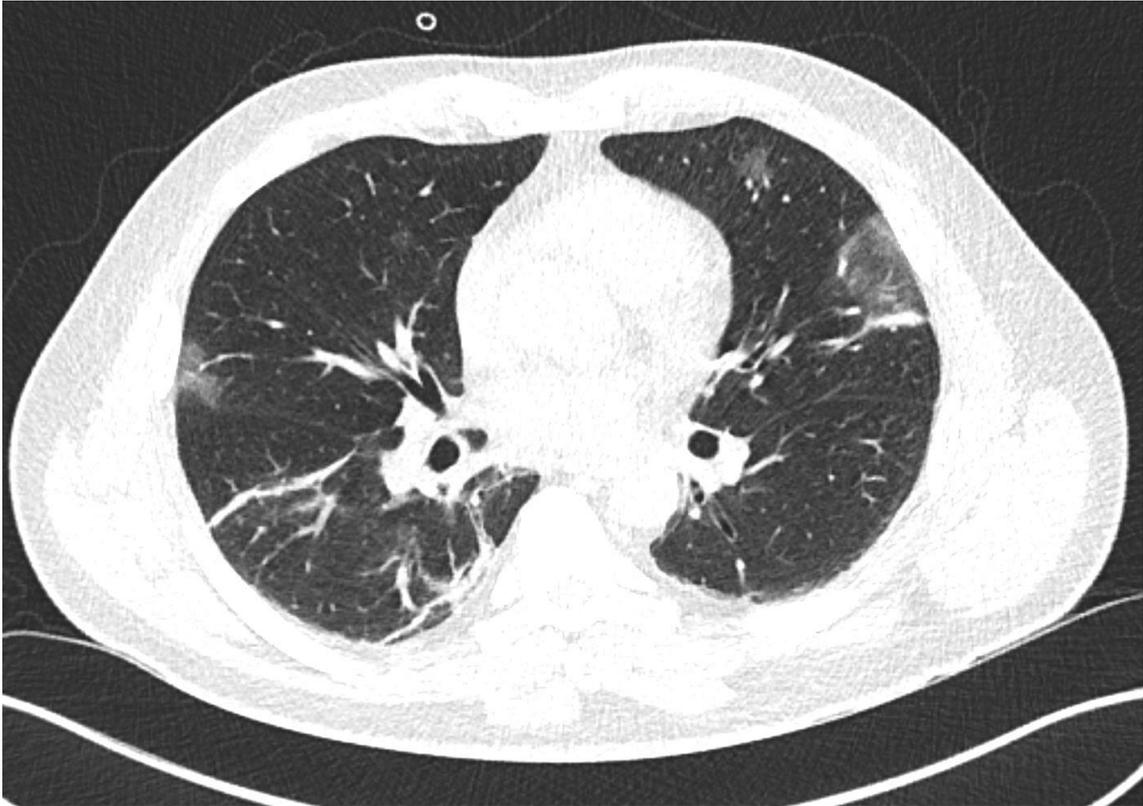


Abbildung 1: Exemplarische CT einer CO-RADS-Kategorie 5. Es zeigen sich bilaterale Milchglatrübungen, überwiegend mit Kontakt zur Pleura. Nebenbefundlich sind beidseits streifige Konsolidierungen abzugrenzen, die mit Plattenatelektasen vereinbar sind. Die für COVID-19 typischen pulmonalen Veränderungen untermauern eine sehr hohe Wahrscheinlichkeit für eine COVID-19-Infektion und entsprechen somit der CO-RADS-Kategorie 5. (Quelle: Picture Archiving and Communication System (PACS), Patientinnen/-enidentifikationsnummer (PAT-ID) 8).

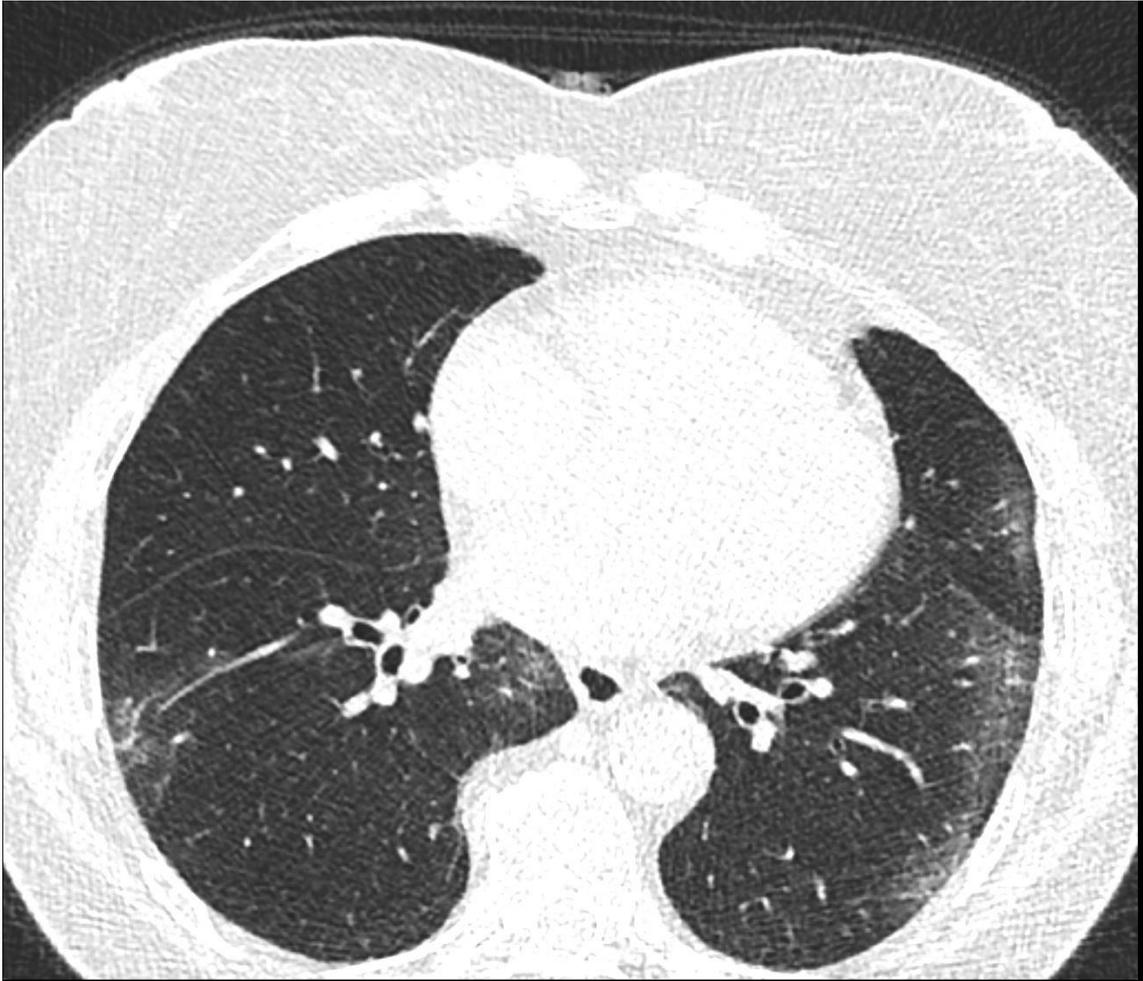


Abbildung 2: Exemplarische CT einer CO-RADS-Kategorie 5. Es zeigen sich bilaterale Milchglatrübungen mit Kontakt zur Pleura, links pulmonal (rechts im Bild) stellenweise mit subpleuraler Aussparung. Die pulmonalen Veränderungen weisen eine sehr hohe Wahrscheinlichkeit für eine COVID-19-Infektion im CO-RADS-Stadium 5 auf. (Quelle: PACS, PAT-ID 5).

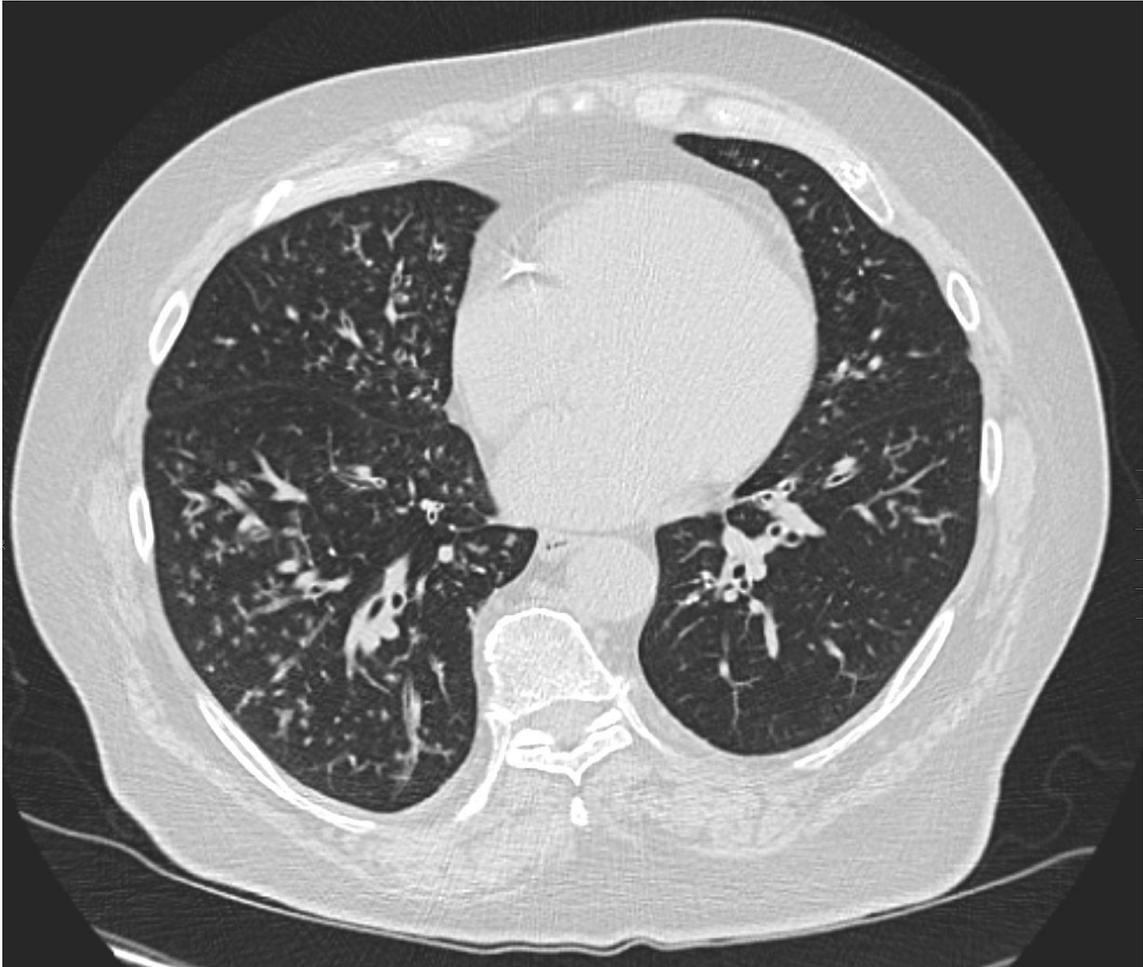


Abbildung 3: Exemplarische CT einer CO-RADS-Kategorie 2. Die CT-Aufnahme zeigt bilaterale, multifokale Bronchialwandverdickungen sowie zentrilobuäre Noduli im Sinne eines Tree-in-bud-Zeichens. Der Befund ist passend zu einer ausgedehnten Bronchiolitis und Bronchitis. Die Wahrscheinlichkeit für eine pulmonale Beteiligung im Rahmen einer COVID-19-Infektion ist niedrig: CO-RADS-Kategorie 2. Die pulmonalen Veränderungen sind typisch für eine Infektion mit einem anderen Erreger als SARS-CoV-2. (Quelle: PACS, PAT-ID 4).

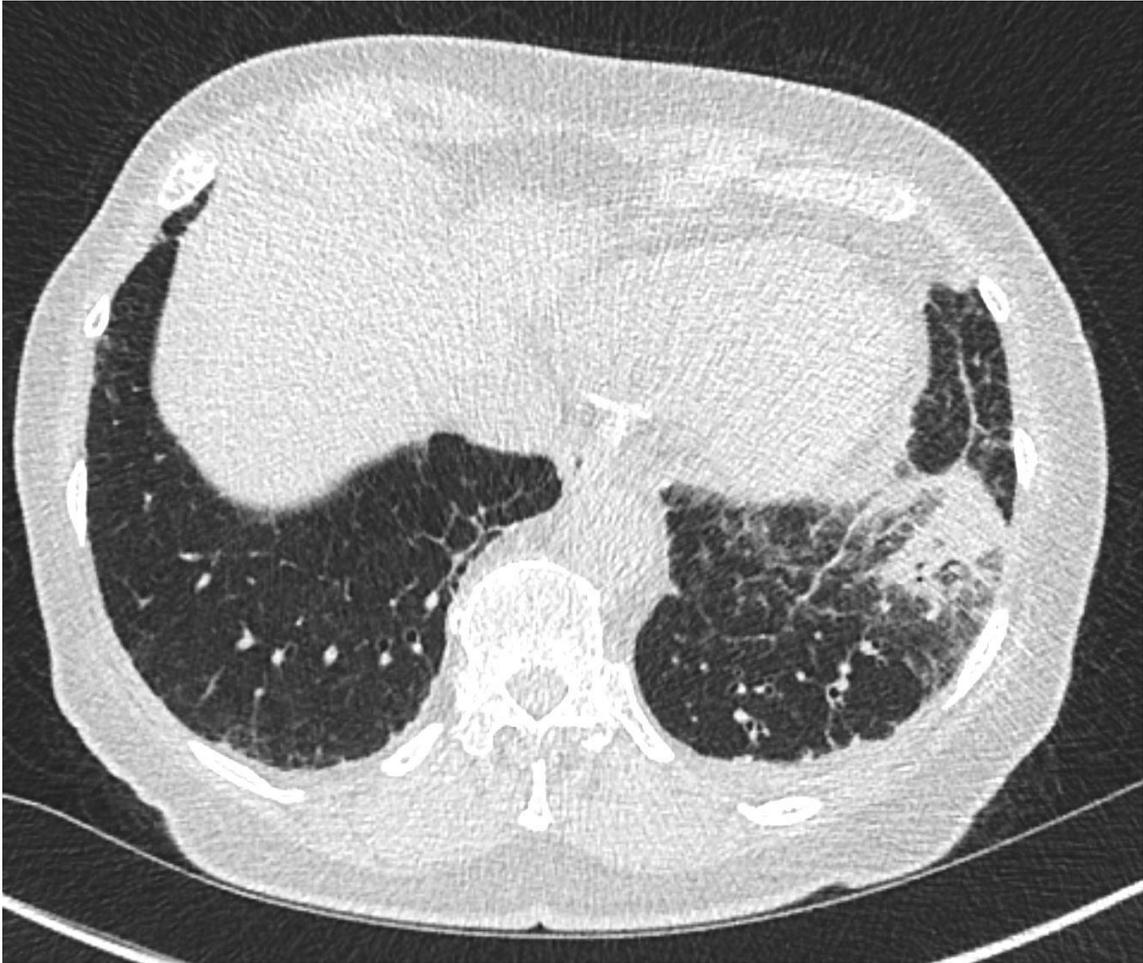


Abbildung 4: Exemplarische CT einer CO-RADS-Kategorie 2. Das CT-Bild zeigt eine unilaterale, unifokale, segmentale Konsolidierung im linken Unterlappen. Zudem finden sich in beiden Unterlappen verdickte Interlobulärsepten und geringe Bronchiektasen. Der Befund ist passend zu einer Bronchopneumonie im linken Unterlappen und zusätzlichen fibrotischen Veränderungen. Es besteht eine niedrige Wahrscheinlichkeit für eine pulmonale Beteiligung im Rahmen einer COVID-19-Infektion: CO-RADS-Stadium 2. (Quelle: PACS, PAT-ID 2)

Eine prognostische und diagnostische Vergleichsstudie zum Thema CT-Erstbefunde und CO-RADS-Stadium bei COVID-19-Patientinnen und Patienten hinsichtlich PCR-, Entzündungs- und Gerinnungsparametern von Yildirim et al. wies nach, dass die Verwendung einer Thorax-CT zur Durchführung des CO-RADS-Stagings hinsichtlich der Generierung diagnostischer und auch prognostischer Informationen hilfreich ist (67).

In einer statistischen Metaanalyse wurde systematisch die diagnostische Leistung der CO-RADS-Klassifikation bei mit COVID-19-Infizierten aus insgesamt 24 Studien mit 8.382 Patientinnen und Patienten evaluiert. Insgesamt zeigte CO-RADS dabei einen akkuraten Nachweis von COVID-19; die gepoolte Sensitivität und Spezifität und die Fläche unter der Kurve von CO-RADS ≥ 3 zum Nachweis von COVID-19 betrugen 0,89, 0,68 bzw. 0,87 (68). Bereits in der Originalarbeit zur Entwicklung und Evaluation der CO-RADS-Klassifikation konnte zudem eine beträchtliche Interobserver-Reliabilität nachgewiesen werden, insbesondere für die Kategorien 1 und 5 (63).

1.5 Fragestellung und Zielsetzung der Dissertation

Nachdem radiologische Befundberichte jahrzehntelang als Freitexte formuliert wurden, gab es in den letzten Jahren zunehmende Bestrebungen, die strukturierte Befundung im radiologischen Alltag zu etablieren. Ihre Vorteile gegenüber konventionellen Prosaberichten insbesondere im Hinblick auf Klarheit, Verständlichkeit, Vollständigkeit und Gesamtqualität konnten in diversen Studien belegt werden (37, 69). Ziel dieser Studie ist es, zu untersuchen, ob auch im Kontext der COVID-19-Pandemie die strukturierte Befundung unter Anwendung der CO-RADS-Klassifikation dem konventionellen Prosabericht im Hinblick auf Qualität, Verständlichkeit und klinischen Nutzen überlegen ist.

2 Material und Methoden

2.1 Studienaufbau

Im Rahmen einer monozentrischen, retrospektiven Kohortenanalyse wurden radiologische Befundberichte von CT-Thorax-Untersuchungen potenziell infizierter Patientinnen und Patienten, die sich während der ersten deutschen Coronawelle mit dem Verdacht auf COVID-19 in der Notaufnahme des Ludwig-Maximilians-Universität (LMU) Klinikums München vorstellten, verglichen. In diesem Zeitraum lagen hohe Inzidenzwerte und somit eine hohe Vortestwahrscheinlichkeit vor (70).

2.1.1 Datenerhebung

Über eine Volltextsuche, die die Stichworte „COVID-19“ und „CT Thorax“ beinhaltete und die auf den Zeitpunkt der ersten deutschen Coronawelle begrenzt war, wurden im Radiologieinformationssystem (RIS) des LMU Klinikums potentielle Studienpatientinnen und Studienpatienten ermittelt. Die ersten 50 Patientinnen und Patienten, die den Einschlusskriterien entsprachen, wurden in die Studie eingeschlossen. Klinische und radiologische Befunde der Probandinnen und Probanden wurden im klinischen Informationssystem (LAMP-IS, LMU Klinikum, München, Deutschland) bzw. RIS (NEXUS RIS, NEXUS Enterprise Imaging GmbH, Freiburg, Deutschland) erhoben, Bildmaterial wurde dem Picture Archiving and Communication System (PACS, Visage Client Version 7.1.18, Visage Imaging GmbH, Berlin, Deutschland) des LMU Klinikums entnommen. Die Datenerhebung erfolgte dabei stets in anonymisierter Form und, ebenso wie die Datenverarbeitung, in Übereinstimmung mit der Deklaration von Helsinki (Stand: 2013).

2.1.2 Ein- und Ausschlusskriterien

In unsere Studie eingeschlossen wurden 50 konsekutive erwachsene Patientinnen und Patienten, die ab dem 1. März 2020 in der Zentralen Notaufnahme des LMU Klinikums München behandelt wurden und aufgrund des Verdachts auf eine SARS-CoV-2-Infektion eine CT des Thorax erhielten. Ausgeschlossen wurden Patientinnen und Patienten, bei denen zum Zeitpunkt der CT-Untersuchung bereits ein RT-PCR-Testergebnis vorlag, sowie Patientinnen und Patienten, bei denen der

SARS-CoV-2-Infektionsstatus nicht im weiteren Verlauf mittels RT-PCR geklärt wurde.

2.1.3 Bildakquise

Die CT-Untersuchung erfolgte stets in Rückenlage mit über den Kopf elevierten Armen. Die Scans wurden nativ in tiefer Inspiration akquiriert, wenn lediglich die Frage nach einer COVID-19-Pneumonie beantwortet werden sollte, und mit Kontrastmittel in pulmonalarterieller Phase, wenn gleichzeitig der Verdacht auf eine Lungenarterienembolie bestand. Die Scanrange wurde jeweils individuell an die Patientenanatomie angepasst, um eine vollständige Abbildung sowohl der kranialen Lungenapices, als auch der dorsobasalen Recessus costodiaphragmatici sicherzustellen. Um eine optimale Bildauflösung zu erreichen, wurde auch axial die Anpassung des Sichtfelds (engl. field of view, FOV) an das Körpervolumen der einzelnen Patientin oder des einzelnen Patienten vorgenommen und die Bildakquise erfolgte mit einer Schichtdicke von ≤ 1 mm.

2.1.4 Erstellung eines Templates

Für die Erstellung strukturierter Befundberichte wurde ein Template (Englisch für „Vorlage“, „Schablone“) auf der Homepage Smart Reporting (Smart Reporting GmbH, München, Deutschland, <http://www.smart-radiology.de>) kreiert. Diese Plattform gewährt Zugriff auf bestehende Vorlagen und modifizierbare Module für die Erstellung strukturierter und leitlinienkonformer Befunde. Zudem besteht die Option der eigenen Templateanfertigung mit Hilfe eines „Template Creators“ (71). Das Template dient dazu, vorlagenbasierte strukturierte Befundberichte zu erstellen und diese geordnet wiederzugeben.

Das für diese Arbeit erstellte CO-RADS-Template gliedert sich in folgende vier Abschnitte: Prozedur, klinische Angaben, Befund und Beurteilung.

Zu jedem der Unterpunkte wurden einzelne Auswahlmöglichkeiten als Reiter zum Anklicken gebildet. Den einzelnen Auswahlmöglichkeiten wurden Textbausteine zugeordnet, die sich in Abhängigkeit von der durch den Befunder oder die Befunderin getroffenen Auswahl zur einem Befundtext zusammenfügen:

- Im Abschnitt Prozedur wird die jeweilige CT-Technik mit Blick auf Kontrastmittelgabe, Bildqualität und Voruntersuchungen beschrieben.
- Im Unterpunkt Klinische Angaben werden das Patientenalter und das Patientengeschlecht abgefragt.
- Im Befundteil werden gezielt Pathologien folgender anatomischer Strukturen abgefragt: Lunge, Bronchien, Pleura, Lymphknoten, Herz, große thorakale Gefäße, Oberbauch, Weichteilmantel/übrige Weichteile und Knochen.
- Die Beurteilung fordert eine konkrete Einschätzung des Befundes anhand der CO-RADS-Klassifikation 0-6. Nebenbefunde finden ebenfalls in der Beurteilung im Freitextformat Platz.

Das erstellte Template ermöglicht die strukturierte Befundung von CT-Aufnahmen des Thorax durch Radiologinnen und Radiologen, indem diese begleitend zur Bildbetrachtung systematisch zutreffende Felder anklicken. In der Vorschau und Endübersicht entsteht dabei ein vollständig ausformulierter, strukturierter Befund. Die Vorlage bietet zusätzlich die Option, fehlende Informationen manuell in ein gesondertes Freitextfeld einzugeben.

Die untenstehenden **Abbildungen** illustrieren eine strukturierte Berichtsvorlage, die mit Hilfe der Online-Software Smart Reporting erstellt wurde.

a)

The screenshot shows the 'CO-RADSv1MR' interface. The left sidebar lists anatomical regions: Lunge, Bronchien, Pleura, Lymphknoten, Herz, Große thorakale Gefäße, Oberbauch, Weichteilmantel/ übrige Weichteile, and Knochen. The main area contains a table of findings with dropdown menus for selection. The right sidebar shows the generated report (Beurteilung) for the lung findings, including a summary and detailed pathological descriptions.

b)

The screenshot shows the 'CO-RADSv1MR' interface in the 'Beurteilung' step. The left sidebar lists 'CO-RADS-Klassifikation' and 'Nebenbefunde'. The main area shows a dropdown menu for selecting the CO-RADS category, with 'CO-RADS 3' selected. The right sidebar shows the generated report (Beurteilung) for the selected category, including a summary and detailed pathological descriptions.

Abbildungen 5a) und 5b): Exemplarischer Auszug des CO-RADS-Templates aus dem Smart Reporting Tool.

Auf der **Abbildung 5a)** sind links und mittig Entscheidungsbäume mit mehreren anklickbaren Elementen und Unterelementen zu sehen. Aus diesen Eingaben werden auf der rechten Seite automatisch semantische Sätze im Resultatfeld generiert.

In einem weiteren Schritt kann die befundete Lunge entsprechend der CO-RADS-Klassifikation kategorisiert werden, wie in **Abbildung 5b)** ersichtlich. Die Beurteilung mit ausgewählter CO-RADS-Kategorie wird am Ende des generierten Berichts angezeigt, hier exemplarisch für einen Befund der CO-RADS-Kategorie 3.

2.1.5 Bildanalyse und Erstellung der Befundberichte

Die CT-Untersuchungen der in die Studie eingeschlossenen Patientinnen und Patienten wurden in Kenntnis der klinischen Angaben und Fragestellungen aus

den ursprünglichen Untersuchungsanforderungen, nicht jedoch der RT-PCR-Testergebnisse befundet. Die Bildbetrachtung erfolgte dabei in dem PACS. Die Lungen wurden im Lungenfenster (Level -600, Breite 1500), Mediastinum, Oberbauch und Weichteilmantel im Weichteilfenster (Level 70, Breite 370) und knöchernen Strukturen im Knochenfenster (Level 500, Breite 1500) betrachtet und beurteilt; neben axialen Schichtstapeln standen dabei auch multiplanare Rekonstruktionen zur Verfügung.

2.1.5.1 Strukturierte Befunde

Die Erstellung der SRs erfolgte pseudonymisiert durch eine auf Thoraxbildgebung spezialisierte Radiologin mit neun Jahren Berufserfahrung. Sie nutzte bei der Befundung das für diese Studie erstellte Template auf der Online-Plattform Smart Reporting. Die durch das Anklicken entsprechender Elemente im Entscheidungsbaum und gegebenenfalls Einfügung von Zusatzinformationen als Freitext generierten Befunde wurden aus dem Template in Microsoft Word (Microsoft® Word für Mac Version 16.75, Redmond, Washington, United States of America) exportiert.

2.1.5.2 Konventionelle Befunde

Als konventionelle Freitextberichte wurden die im RIS bereits vorhandenen Originalberichte verwendet, die im Rahmen der klinischen Routine während des Notaufnahmearaufenthalts der eingeschlossenen Patientinnen und Patienten als Freitext erstellt worden waren. Diese CRs wurden ursprünglich durch radiologische Assistenzärztinnen oder Assistenzärzte verfasst, anschließend durch radiologische Fach- oder Oberärztinnen oder -ärzte überprüft und gegebenenfalls korrigiert. Für die Studiaauswertung wurden die CRs pseudonymisiert aus dem RIS exportiert.

2.1.6 Bewertung der Befundberichte

Die Freitext- und strukturierten Befunde wurden durch zwei Pneumologinnen unabhängig voneinander in zufälliger Reihenfolge gelesen und anhand eines Fragebogens bewertet.

Zusätzlich zu den Befunden wurden beiden Bewerterinnen Informationen zur klinischen Symptomatik (kurze Notfallanamnese aus dem System, Verdachtsdiagnose bzw. Indikation sowie klinische Angaben) der Patientinnen und Patienten sowie die ursprünglich von der anfordernden Ärztin oder dem anfordernden Arzt formulierte Fragestellung zur Verfügung gestellt.

2.1.6.1 Fragebogen

Der Fragebogen diente der Evaluation der Freitextbefunde und strukturierten Befunde. Er beinhaltet insgesamt acht Fragen und gliedert sich in drei Abschnitte:

Der erste Abschnitt befasst sich in den Fragen 1-4 mit den formalen Aspekten. Neben der Strukturiertheit, der Angemessenheit der Befundlänge und einer verständlichen Formulierung wird dabei auf das Fehlen von Informationen eingegangen. Im zweiten Abschnitt wird mit Frage 5 und 6 die klinische Relevanz erhoben. Hierbei wird die Beantwortung der klinischen Fragestellung und die Entscheidung über das weitere diagnostische und therapeutische Vorgehen examiniert. Der letzte Abschnitt prüft in den Fragen 7 und 8 die Gesamtqualität des Befundes. Er fordert eine Einschätzung zum Vertrauen in die Informationen beider Befundarten und eine subjektive Bewertung der Gesamtqualität. Insgesamt werden sieben der acht Fragen auf einer Likert-Skala bewertet. Es handelt sich hierbei um eine ordinalskalierte Ratingskala mit fünf Merkmalsausprägungen von „stimme voll zu“ über „stimme eher zu“, „stimme teilweise zu“, „stimme eher nicht zu“ bis „stimme überhaupt nicht zu“. Lediglich Frage 4 klassifiziert die Ausprägungen der Variablen auf Nominalskalenniveau. Sie beinhaltet neun Ankreuzoptionen für fehlende Informationen (einfaches Ankreuzen) sowie ein freies Textfeld zum Ergänzen (72). Der detaillierte Fragebogen findet sich im Anhang.

Wie bereits erwähnt wurden für 50 Patientinnen und Patienten Daten aus dem RIS erhoben. Die entsprechenden Freitextbefunde lagen dort vor und die strukturierten Befunde wurden anhand des bereits erläuterten Templates erstellt. Mit Hilfe des Fragebogens galt es, diesen Datensatz von insgesamt 50 Patientinnen und Patienten pro Befundart zu evaluieren. Beide Bewerterinnen erhielten daher je 100 Fragebögen zu den Befundarten (50 SRs sowie 50 CRs) mit jeweils acht Fragen. So wurden insgesamt 200 Fragebögen, 100 Stück je Bewerterin, beantwortet, die im Folgenden genauer betrachtet werden.

2.1.7 Ethikantrag

Für die Durchführung des retrospektiven Forschungsprojektes auf Grundlage der durch die Radiologie erhobenen Daten liegt ein positives Votum der Ethikkommission der LMU München vor.

2.2 Statistische Auswertung

2.2.1 Bearbeitung der erfassten Daten

Die aus dem RIS und Krankenhausinformationssysteme (KIS) erhobenen Patientinnen- und Patientendaten sowie die Daten der Evaluationsbögen wurden pseudonymisiert in Excel-Tabellen (Microsoft® Excel für Mac Version 16.75.2, Redmond, Washington, United States of America) erfasst. Für die statistische Berechnung wurde der händisch erstellte Datensatz in ein dafür geeignetes Format überführt. Der generierte Datensatz wurde in der Tabelle nach Patienten/innenidentifikationsnummer (PAT-ID) und Bewerterin A bzw. Bewerterin B geordnet, in eine Statistiksoftware überführt und in einem .sav-Format gespeichert. Der Datensatz wurde mit Hilfe verschiedener statistischer Verfahren analysiert, auf die im Folgenden eingegangen wird.

2.2.2 Statistische Auswertung der Daten

Die statistische Auswertung erfolgte unter der Verwendung der SPSS (Statistical Package for the Social Science) Statistics® Software in der Version 29.0.0. (IBM Corp., Armonk, U.S.A).

Mit Methoden der deskriptiven Statistik wurden Häufigkeiten, Lage- und Streuungsmaße der Daten ermittelt. Entsprechend wurden Patientencharakteristika wie Geschlecht, Alter, Vorliegen einzelner Symptome oder koinzidenter Erkrankungen bzw. zuvor diagnostizierter Grunderkrankungen dargestellt und auch die Bewertungen der verschiedenen Befundtypen durch beide Pulmologinnen beschrieben. Für ordinalskalierte stetige Variablen (Frage 1-3 und 5-8) wurden neben den Mittelwerten die Mediane (Mdn), Standardabweichungen (SD), Spannweiten (SW) sowie die Interquartilsabstände (IQA) für SRs und CRs berechnet. Es wurden die Mediane der beiden Bewerterinnen für jeden Scan und jede der Aussagen 1-3 und 5-8 berechnet.

Frage 4 beinhaltet ein nominales Skalenniveau. Aufgrund multipler Antwortmöglichkeiten im Freitextfeld (4.10) der Frage 4 wurde diese zur schematischen Auswertung in Unterpunkte untergliedert. Die Ergebnisse der Fragen 4.1. bis 4.10 wurden als Häufigkeiten (n) und Prozentwerte (%) angegeben.

Anschließend wurden diese Daten auf statistisch signifikante Unterschiede geprüft: Für die Fragen mit ordinalskalierten und abhängigen Variablen wurde der Wilcoxon-Vorzeichen-Rang-Test gewählt. Der Test berechnet, ob sich die beiden Reportformate signifikant voneinander unterscheiden. Es handelt sich um einen nichtparametrischen Test (73). Für die nominalskalierte Frage 4 wurde der McNemar-Test ausgewählt. Der McNemar-Test ist ein nichtparametrischer Test bei verbundenen Stichproben. Er prüft ein dichotomes Merkmal (bspw. vorhanden versus fehlt) bei abhängigen Daten (74, 75).

In einer Sekundäranalyse des Datensatzes wurde der Zusammenhang zwischen der radiologischen Beurteilung im Hinblick auf das Vorliegen von COVID-19 und dem PCR-Testergebnis beleuchtet. Hierzu wurde den CRs zunächst eine der Kategorien 1 bis 3 mit folgender Bedeutung zugeordnet:

- Kategorie 1: „kein radiologischer Anhalt für eine SARS-CoV-2-Infektion“
- Kategorie 2: „radiologisch nicht ganz eindeutig, möglicherweise SARS-CoV-2-Infektion, jedoch auch andere Pathologie denkbar“
- Kategorie 3: „radiologischer Verdacht auf SARS-CoV-2-Infektion“.

Dabei erfolgte die Zuordnung der CRs zu einer der oben genannten Kategorien gemäß dem Wortlaut in der Beurteilung.

Daraufhin erfolgte eine dichotome Aufteilung der SRs und CRs in Befunde, gemäß derer der Patient oder die Patientin wahrscheinlich SARS-CoV-2-positiv war, und Befunde, gemäß derer keine Infektion vorlag. Dabei wurden sowohl für SRs, als auch für CRs zwei verschiedene Szenarien getestet, indem uneindeutige Befunde, also Kategorie 2 bei CRs und CO-RADS 3 bei SRs, jeweils einmal als „COVID-19-positiv“ und einmal als „COVID-19-negativ“ gewertet wurden. In einer ersten Analyse galten also SRs mit den Levels CO-RADS 1 und 2 sowie CRs der Kategorie 1 als „COVID-19-negativ“ und SRs mit den Levels CO-RADS 3, 4 und 5 sowie CRs der Kategorie 2 und 3 als „COVID-19-positiv“. In einer zweiten Analyse wurden SRs mit den Levels CO-RADS 1, 2 und 3 sowie CRs

der Kategorie 1 und 2 als „COVID-19-negativ“ und SRs mit den Levels CO-RADS 4 und 5 sowie CRs der Kategorie 3 als „COVID-19-positiv“ gewertet.

Anschließend wurden die Sensitivität und Spezifität der beiden radiologischen Befundarten für die Erkennung RT-PCR-positiver COVID-19-Patientinnen und -Patienten bestimmt.

3 Ergebnisse

3.1 Patientenkohorte

Fünzig konsekutive Patientinnen und Patienten, die sich ab dem 01. März 2020 mit Verdacht auf COVID-19 in der Notaufnahme des LMU Klinikums vorstellten und dort aus diesem Grund eine CT des Thorax erhielten, wurden in die Studie eingeschlossen.

3.1.1 Demographie

Die Geschlechterverteilung zeigte das Verhältnis von [n = 18] 36 % Frauen zu [n = 32] 64 % Männern. Das Durchschnittsalter lag bei 64,8 Jahren, wobei die Altersspanne bei 33 bis 80 Jahren lag.

3.1.2 Grunderkrankungen/Vorerkrankungen

Insgesamt handelte es sich um ein heterogenes Patientenkollektiv, in dem 90 % (n = 45) Individuen zuvor diagnostizierte Vor- bzw. Grunderkrankungen mitbrachten, welche in **Tabelle 2** näher ausgeführt werden.

Tabelle 2: Vor- und Grunderkrankungen der Patientenkohorte.

Vor/Grunderkrankung	Patientenanzahl n =	Patienten in %
Arterieller Hypertonus	22	44
Kardiale Vorerkrankungen	21	42
Onkologische Vorerkrankungen/ Tumor	14	28
Endokrinologie: Diabetes mellitus	9	18
Pulmonale Vorerkrankungen (Asthma bronchiale, COPD)	7	14
Nierentransplantation	3	6

3.1.3 Symptome bei Vorstellung

Die Patientinnen und Patienten litten bei ihrer Vorstellung in der Notaufnahme durchschnittlich seit 6,9 Tagen (Zeitspanne 0 bis 56 Tage) an Symptomen, die in **Tabelle 3** im Einzelnen aufgeführt werden.

Tabelle 3: Symptome des Patientenkollektivs in Anzahl und %.

Symptome	Patientenanzahl n =	Patienten in %
Husten	28	56
Fieber	28	56
Dyspnoe, Atembeschwerden	23	46
Allgemeinzustandsminderung, Abgeschlagenheit	21	42
Gastrointestinale Beschwerden	11	22
Kopf-, Muskel-, Gliederschmerzen	6	12
Thoraxschmerz, thorakaler Druck	4	8
Synkope, Schwindel	3	6

3.1.4 Entzündungsparameter bei Aufnahme

Bei Aufnahme erhielten die Patientinnen und Patienten eine Blutentnahme, bei der laborchemisch die Entzündungsparameter u.a. in Form von C-reaktivem Protein (CRP) und Leukozyten erfasst wurden. Hierbei lag der durchschnittliche CRP-Wert bei 6,74 mg/dL, Referenzbereich <0,5, Standardabweichung (SD) \pm 13,38/dL. Der Mittelwert der Leukozyten lag bei 7,89 G/L bei einer SD von \pm 34,65 G/L.

3.2 Vergleich der Befundtypen

Insgesamt wurden 200 Fragebögen (50 SRs und 50 CRs, jeweils durch zwei Pneumologinnen einmal bewertet) ausgefüllt und vollständig beantwortet. Die Rücklaufquote betrug 100%.

3.2.1 Formale Aspekte

Die Fragen 1-4 befassen sich mit den formalen Aspekten beider Befundformen.

3.2.1.1 Strukturiertheit

Bezüglich der Strukturiertheit (Frage 1) wurden die SRs mit einem Median von 1 (Interquartilsabstand (IQA) 0; Spannweite (SW) 1-2,5), die CRs mit einem Median von 2 (IQA 0,5; SW 1-3) bewertet. Es zeigte sich ein signifikanter Unterschied zu Gunsten der SRs, $p < 0,001$.

3.2.1.2 Befundlänge

Hinsichtlich der Befundlänge (Frage 2) zeigte sich ein signifikanter Unterschied zwischen beiden Befundarten, $p < 0,001$. Die SRs (Mdn 1; IQA 0,5; SW 1-2,5) wurden im Vergleich zu CRs (Mdn 1,5; IQA 0,5; SW 1-3) als angemessener bewertet.

3.2.1.3 Verständlichkeit

Die SRs (Mdn 1; IQA 0,5; SW 1-2) wurden im Vergleich zu CRs (Mdn = 1,25; IQA 0,5; SW 1-2,5) als signifikant verständlicher formuliert bewertet, $p = 0,044$.

3.2.1.4 Fehlende Informationen im Befund

Die Frage 4 befasst sich mit fehlenden Informationen im Befund und bietet dabei neun Ankreuzoptionen, von denen eine mit einer Freitextlücke versehen ist. Die Ergebnisse beider Befundtypen sowie die einzelnen Kategorien und das jeweilige Signifikanzniveau werden in der folgenden **Tabelle 4** aufgeführt.

Tabelle 4: Vollständigkeit der Befundformen SRs versus CRs im Hinblick auf verschiedene Items mit Signifikanz der Unterschiede.

Kategorie	SRs in %	CRs in %	p	Ergebnis signifikant ja/nein, Signifikanzniveau 0,05
4.1: Pneumonie	100	100	p = 1,000	nein
4.2: Rundherde/Raumforderung	100	60	p < 0,001	ja
4.3: Mögliche weitere pulmonale Pathologien (Ödem, Emphysem, Fibrose...)	100	67	p < 0,001	ja
4.4: Pleura	100	91	p < 0,004	ja
4.5: Bronchien	100	62	p < 0,001	ja
4.6: Lymphknoten	98	99	p = 1,000	nein
4.7: Herz	99	99	p = 1,000	nein
4.8: Große thorakale Gefäße	98	76	p < 0,001	ja
4.9: Andere	97	89	p = 0,008	ja

Die Bewertung von Frage 4 unterschied sich zwischen SRs und CRs in den Unterkategorien 4.2 Rundherde oder Raumforderung, 4.3 Mögliche weitere pulmonale Pathologien, 4.4 Pleura, 4.5 Bronchien, 4.8 Große thorakale Gefäße und 4.9. Andere signifikant.

Die Frage 4 wurde auch im Gesamtvergleich zwischen SRs und CRs als signifikant unterschiedlich bewertet: Beim strukturierten Befund (SR) wurden 99,1% als vollständig, beim konventionellen Befund (CR) 82,6% als vollständig beschrieben. SRs wurden somit gegenüber CRs als vollständiger bewertet, p < 0,001.

3.2.1.5 Fehlende Informationen in der Beurteilung

Um auf die Anmerkungen der Bewerterinnen einzugehen, wurde ein neuer Unterpunkt 4.10 angelegt. Eine der Bewerterinnen hätte sich die zusätzliche Erwähnung verschiedener Punkte aus dem Befund in der Beurteilung gewünscht, die andere Bewerterin nicht (p < 0,001). Eine Aufstellung der Informationen, die einer der Pneumologinnen zufolge im CR und SR auch in der Beurteilung hätten thematisiert werden sollen, findet sich in **Tabelle 5**.

Der Unterschied in der Einschätzung der Vollständigkeit der Beurteilung zwischen SRs und CRs ist signifikant, $p < 0,001$. Bei den SRs fehlen in den Unterpunkten der Frage 4.10 mehr Informationen als bei den CRs, die CRs werden in dieser Frage 4.10.1 bis 4.10.6 somit als vollständiger bewertet.

Tabelle 5: Vollständigkeit der Beurteilung von CRs und SRs in Gesamtzahl n und %.

Kategorie	n = % von SRs	n = % von CRs
4.10.1 Stauung, Erguss, Pleuraerguss	91	98
4.10.2: Kardiovaskulär: Gefäße, rechts atriale Dilatation	96	99
4.10.3: Informationen über Rundherde, Metastasen, Infiltrate	89	98
4.10.4: Knochenpathologien, Alter der Deckplatteneinsenkungen, Traumafolgen	96	99
4.10.5: Leberzirrhose und Aszites	99	100
4.10.6: Interstitielle Lungenerkrankung (ILD)-Exazerbation als Differentialdiagnose (DD), apikale Fibrose, Emphysem, Erkrankungen des Lungengerüsts	97	100

3.2.2 Klinische Relevanz

Die klinische Relevanz der jeweiligen Befundform wurde in den Aussagen 5 und 6 evaluiert.

3.2.2.1 Beantwortung der klinischen Fragestellung

Im Hinblick auf die Beantwortung der klinischen Fragestellung ergab sich statisch zwischen SRs (Mdn 1, IQA 0,5; SW 1-2,5) und CRs (Mdn 1; IQA 0,5, SW 1-3) kein signifikanter Unterschied ($p = 0,488$).

3.2.2.2 Entscheidung zum weiteren diagnostischen und/oder therapeutischen Vorgehen

Bezüglich der Entscheidung zum weiteren diagnostischen und/oder therapeutischen Vorgehen gab es keinen signifikanten Unterschied zwischen SRs (Mdn 1, IQA1; SW 1-3) und CRs (Mdn 1,5; IQA 1; SW 1-3), $p = 0,616$.

3.2.3 Gesamtqualität

Die Fragen 7 und 8 bewerten die Gesamtqualität des Befundes.

3.2.3.1 Vertrauen in die Information aus dem Befund

Die SRs (Mdn 1,5; IQA 0,5; SW 1-2) wurden im Vergleich zu CRs (Mdn 1,5; IQA 0,5; SW 1-3,5) bezogen auf das Vertrauen in die Information aus dem Befund signifikant besser bewertet, $p < 0,001$.

3.2.3.2 Gesamtqualität des Befunds

Die Gesamtqualität der SRs (Mdn 1,5; IQA 0,5; SW 1-2) wurde im Vergleich zu CRs (Mdn 2; IQA 0,5; SW 1,5-3) als signifikant überlegen bewertet, $p < 0,001$.

3.3 Korrelation zwischen CO-RADS-Stadien und RT-PCR-Testergebnissen

Erste Analyse

In der ersten Analyse, in der die CO-RADS-Levels 3, 4 und 5 als „COVID-19-positiv“ gewertet wurden, betrug die Sensitivität der SRs 100%. Die entsprechende Spezifität betrug 61%. Die Sensitivität für die CRs, bei denen sowohl Kategorie 2, als auch 3 als „COVID-19-positiv“ gewertet wurden, betrug 95%. Die entsprechende Spezifität belief sich auf 50%.

Zweite Analyse

In der zweiten Analyse, in der SRs mit den Levels CO-RADS 4 und 5 als „COVID-19-positiv“ gewertet wurden, betrug die Sensitivität 91%. Die entsprechende Spezifität lag bei 100%. Die Sensitivität für die CRs, bei denen lediglich die Kategorie 3 als „COVID-19-positiv“ gewertet wurden, betrug 95%. Die entsprechende Spezifität belief sich auf 96%.

4 Diskussion

4.1 Zusammenfassung der Ergebnisse

In der Gesamtschau wurde der SR gegenüber dem CR hinsichtlich der zu evaluierenden Merkmale besser bewertet. Bezüglich der formalen Aspekte ergab sich ein signifikanter Unterschied in der Strukturiertheit, in der Angemessenheit der Befundlänge, in der Verständlichkeit und bezüglich fehlender Informationen im Befund. Daraus ergibt sich, dass der strukturierte Befund (SR) dem konventionellen Befund (CR) hinsichtlich dieser evaluierten Merkmale überlegen ist. Die klinische Relevanz betreffend ergab sich kein signifikanter Unterschied zwischen CR und SR. In der Gesamtqualität wurde der SR besser als der CR bewertet.

4.2 Formale Aspekte

4.2.1 Struktur und Klarheit

Bezugnehmend auf die formalen Aspekte konnten bereits frühere Studien belegen, dass die strukturierte Befundung der konventionellen überlegen ist: So gelang es Naik et al. in einer Untersuchung, die Präferenz und insgesamt eine bessere Bewertung durch zuweisende Ärztinnen und Ärzte und Radiologinnen und Radiologen aufgrund eines attraktiveren und übersichtlicheren Layouts bzw. strukturierteren Formats standardisierter Befunde gegenüber dem Freitextbefund herauszustellen (46). In einer Studie von Schwartz et al. wurden außerdem im Rahmen von CT-Untersuchungen des Thorax, Abdomens und Beckens Vorteile der strukturierten Befundung im Sinne signifikant höherer Zufriedenheitswerte zuweisender Ärzte auf einer zehnstufigen Likert-Skala bzgl. Klarheit und Genauigkeit der Aussagen aufgezeigt (35, 76). Auch Sabel et al. berichteten von einer signifikant höheren Zufriedenheit der Klinikerinnen mit der Klarheit der SRs gegenüber den CRs für CT bei Lungenembolie und Angiographie der unteren Extremität (77). Eine weitere Studie zum Thema strukturierte Befundung bezogen auf die Becken-MRT zur Beurteilung des Operationsergebnisses nach ileoanaler Pouchanlage konnte ebenfalls mehr Klarheit der SRs belegen (78). Die hier angeführten Studienergebnisse stehen im Einklang mit unseren Erkenntnissen (58).

4.2.2 Angemessene Befundlänge

In unserer Studie wurde die Befundlänge bei den SRs als angemessener beurteilt als bei den CRs. Zu einem übereinstimmenden Ergebnis kamen Spiro et. al. in ihrer Studie zu CT-Untersuchungen von Lungentransplantierten (49). Eine Studie zum Thema Bauch-Becken-CT zeigte zwar, dass sich unter Verwendung einer strukturierten im Vergleich zur konventionellen Berichterstattung weniger Fehler im Bericht einschleichen, was möglicherweise klinisch signifikante unerwünschte Ergebnisse in der Patientenversorgung reduziert. Die Berichtslänge auf Grundlage der Wortanzahl unterschied sich in dieser Studie jedoch nicht signifikant zwischen den beiden Befundstilen (79). Dem gegenüber steht eine Untersuchung zur quantitativen Verbesserung der Hirntumor-MRT durch strukturierte Befundung: Hier konnte neben einer reduzierten Befundgesamtlänge auch eine geringere Verwendung von Füllwörtern wie „möglicherweise“ (3,8 % vs. 0,6 %, $p < 0,05$) und „wahrscheinlich“ (49,8 % vs. 28,5 %, $p < 0,01$) beim SR nachgewiesen werden (80).

4.2.3 Verständlichkeit

Studienübergreifend konnte in multiplen Untersuchungen eine bessere Verständlichkeit unter Einsatz des SR gegenüber dem CR verzeichnet werden (38, 49, 58, 81). Auch unsere Ergebnisse stimmen mit dieser Erkenntnis überein. Ein weiterer Aspekt, der in einer Studie zum CT-Staging des diffusen großzelligen B-Zell-Lymphoms in diesem Zusammenhang verdeutlicht wurde, ist die schnellere Informationsextraktion aus SRs im Vergleich zu CRs (mediane Bewertung auf der 10-Punkte-Likert-Skala = 9 vs. 6; IQA 7-10 vs. 4-8) (81).

Ein weiterer Vorteil, der sich aus dem templategestützten SR unter Gebrauch eines standardisierten, verständlichen Vokabulars ergibt, ist die damit einhergehende Möglichkeit einer eindeutigen, vollautomatisierten Übersetzung in eine andere Sprache (38). Insbesondere zu Zeiten der Globalisierung mit grenzübergreifender Medizin stellt die vereinfachte, präzise Übersetzungsmöglichkeit auf Grundlage strukturierter Befunde einen erheblichen Nutzen dar (36).

4.2.4 Vollständigkeit/fehlende Informationen

Naik et al. konnten in einer Studie nachweisen, dass die Zufriedenheit von Radiologinnen und Radiologen steigt, je detaillierter und vollständiger eine strukturierte Befundvorlage konzipiert ist (35, 46). Diese Einschätzung wird in einer Studie von Schwartz et al., in der die inhaltliche Zufriedenheit auf einer Likert-Skala mit 7,61 für herkömmliche Berichte und 8,33 für strukturierte Berichte signifikant unterschiedlich ($p < 0,0001$) bewertet wurde, bekräftigt (76). In einer Untersuchung templatebasierter Röntgen-Thoraxaufnahmen wurden die SRs, verglichen mit den konventionellen Befunden, ebenfalls als vollständiger und nützlicher bewertet (82). Die zuletzt aufgeführte Studie deckt sich mit unseren Studienresultaten, in denen die SRs als vollständiger als die CRs beurteilt wurden.

Es bestehen wegweisende Anhaltspunkte dafür, dass die strukturierte Befundung ein geeignetes Hilfsmittel für die Reduktion diagnostischer Fehler darstellt. Dies wurde auch in zahlreichen weiteren Studien belegt. So half beispielsweise die strukturierte Befundung in einer retrospektiven Überprüfung von 3.000 MRT-Untersuchungen der Lendenwirbelsäule, klinisch signifikante extraspinale Befunde bei 28,5 % der Patientinnen und Patienten zu identifizieren, die im ursprünglichen unstrukturierten Bericht nicht enthalten waren (37, 83). Eine Studie über nicht frakturbezogene Befunde der Halswirbelsäule in der CT wies ebenfalls eine verbesserte Diagnoserate durch vorlagenbasierte strukturierte Befundung in Form einer Checkliste auf (47). Eine Reduktion diagnostischer Fehler und somit eine verbesserte Gesamtqualität gelang es außerdem Roskopf et al. in ihrer Studie über die MRT des Knies zu belegen (84).

Ein neuer Aspekt in Bezug auf die Vollständigkeit ergab sich in unserer Studie mit Blick auf den Beurteilungsteil strukturierter Befundberichte. Im Freitextteil nominierte eine der beiden Bewerterinnen die Detailtiefe der Beurteilung im SR: Sie hätte sich die Erwähnung einzelner Aspekte, die nicht im Zusammenhang mit COVID-19 standen und die die Radiologin nur im Befundteil beschrieben hatte, auch in der Beurteilung gewünscht. Zwar teilte die zweite Pneumologin diese Ansicht nicht. Nichtsdestotrotz verdeutlicht die diskrepante Einschätzung zwischen den Bewerterinnen das Maß an Subjektivität, dem die Bewertung radiologischer Befundberichte unterworfen ist. Daraus lässt sich ableiten, dass für eine zufriedenstellende Kooperation zwischen behandelnden und radiologischen Ärztinnen

und Ärzten eine engmaschige, konsensbasierte Abstimmung der Befundinhalte unter Einbezug aktueller Leitlinien essentiell ist. Diese Einschätzung deckt sich mit dem Beschluss der DRG, die Erarbeitung strukturierter Befundvorlagen im Rahmen eines interdisziplinären Konsensprozesses zu vollziehen. So wurde im Jahr 2017 die operative Koordinierung der Initiative dem Arbeitskreis strukturierte Befundung innerhalb der Arbeitsgemeinschaft Informationstechnologie (AGIT) der DRG übertragen (39).

4.3 Klinische Relevanz und Entscheidung über das weitere diagnostische und/oder therapeutische Vorgehen

Eine Verbesserung der klinischen Relevanz durch Einführung der strukturierten Befundung wird in der aktuellen wissenschaftlichen Literatur kontrovers diskutiert. Während beispielsweise die Untersuchung von Kim et al. zum Thema Knochendichtemessung eine verbesserte Beantwortung der klinischen Fragestellung im SR ergab, konnten in der Studie von Spiro et al. zur weiteren Diagnostik und Therapie nach Lungentransplantation keine signifikanten Unterschiede zwischen beiden Befundarten festgestellt werden (49, 85). Diese Ergebnisse stehen in Einklang mit unseren Auswertungen, in denen sich die Beantwortung der klinischen Fragestellung und folglich das weitere diagnostische und/oder therapeutische Vorgehen in beiden Befundformen nicht signifikant voneinander unterscheiden.

4.4 Gesamtqualität und Vertrauen in die Information

In einer vorangegangenen Studie über die Verwendung des SR für CT-Thorax bei COVID-19-Patientinnen und Patienten gelang es Stanzione et al., im Vergleich zum CR eine höhere Zufriedenheit des überweisenden Arztes nachzuweisen. Insgesamt konnte die Generierung quantitativ und qualitativ hochwertiger Informationen im SR gegenüber dem CR herausgestellt werden (31). Eine Optimierung der Befundqualität durch SRs wurde zudem in einer Studie über MRT-Staging-Aufnahmen von Rektumkarzinomen nachgewiesen. Die Arbeit prognostiziert ein deutliches Verbesserungspotential für die Befundqualität im Bereich der Onkologie unter Einsatz von SRs (53). Diese früheren Forschungsergebnisse stehen in Einklang mit unserer Studie, in der die Gesamtqualität der SRs höher

als die Qualität der CRs bewertet wurde. Zahlreiche andere Studien bestätigten diese Ergebnisse (49, 56, 58, 81, 85-89).

In Hinblick auf das Vertrauen in die Informationen verglichen Brook et al. beide Berichtformen bei CT-Befunden für das Staging und die subjektive Beurteilung der Resektabilität beim Adenokarzinom des Pankreas. Die Autoren kamen dabei zu dem Ergebnis, dass die Chirurgeninnen und Chirurgen der Information bzgl. der Tumorsektabilität aus den strukturierten Berichten mehr Vertrauen schenken, als der aus den nicht strukturierten Befunden (37, 54). Diese Einschätzung deckt sich mit unseren Ergebnissen, in denen die SRs bezogen auf Vertrauen in die Information aus dem Befund signifikant besser bewertet wurden.

4.4.1 Korrelation zwischen radiologischem Befund und RT-PCR-Testergebnis

Erste Analyse

In der ersten Analyse wurden die CRs mit einem CO-RADS-Level von mindestens 3 und die CRs, die mindestens Kategorie 2 entsprachen, als „COVID-19-positiv“ bewertet. Dabei wiesen sowohl die SRs als auch die CRs eine hohe Sensitivität von 100% respektive 95% auf. Daraus ergibt, dass die meisten „COVID-19-positiven“ Patientinnen und Patienten also solche erkannt wurden. Die Spezifität in dieser ersten Analyse war allerdings sowohl bei SRs (mit 61%) wie auch bei CRs (mit 50%) vergleichsweise gering. Somit wurde eine relativ hohe Anzahl von „COVID-19-negativen“ Patientinnen und Patienten als „COVID-19-positiv“ diagnostiziert.

Zweite Analyse

In der zweiten Analyse wurden die CRs mit einem CO-RADS-Level von mindestens 4 und die CRs, die mindestens der Kategorie 3 entsprachen, als „COVID-19-positiv“ bewertet. In Korrelation zum RT-CR-Testergebnis wiesen damit sowohl die SRs als auch die CRs eine hohe Sensitivität mit 91% (SRs) respektive 95% (CRs) auf. Die Spezifität in der zweiten Analyse war allerdings sowohl bei

SRs (mit 100%) als auch bei CRs (mit 96%) deutlich höher als in der ersten Analyse. Es zeigten sich folglich deutlich weniger falsch positive Ergebnisse im Vergleich zum ersten Szenario.

Zusammenfassend lässt sich folgern, dass in der zweiten Analyse die Spezifität deutlich gestiegen ist, während die Sensitivität auf einem ähnlich hohen Niveau von über 90% blieb. Im Kontext einer Pandemie mit begrenzten klinischen Kapazitäten wäre eine Herangehensweise gemäß der zweiten Analyse, bei der unklare Fälle als „nicht positiv“ im Sinne der infektiologischen Fragestellung gewertet wurden, empfehlenswert. Man würde damit die große Mehrheit „Covid-19 positiver“ Patientinnen und Patienten erfassen und behandeln können. Gleichzeitig würden nur geringe Kapazitäten für falsch positiv diagnostizierte Patientinnen und Patienten aufgewendet.

Diese Erkenntnisse decken sich mit denen eines Reviews über 51 Studien, welches illustriert, dass insbesondere CO-RADS ≥ 3 mit 0,89 (95 % Konfidenzintervall (KI) 0,85–0,93) eine hohe Sensitivität als Screening-Instrument infizierter Personen aufweist und folglich die Ausbreitung von COVID-19 eindämmen könnte. Die Kategorie CO-RADS ≥ 4 hatte in der Studie eine hohe Spezifität von 0,84 (95 % KI 0,78–0,88) und bildete die eindeutige Diagnose von COVID-19 besser ab. Weiterhin wurde dargelegt, dass die Konsistenz von CO-RADS zwischen verschiedenen Zentren und Lesern überprüft und Gegenstand künftiger Studien sein sollte (68).

4.5 Limitationen

Die vorliegende Arbeit wurde professionell sowie sorgfältig geplant und durchgeführt, ist jedoch einigen Limitationen unterworfen. Aufgrund des retrospektiven, monozentrischen Studiendesigns und der relativ geringen Fallzahl sind die Ergebnisse dieser Arbeit unter Vorbehalt zu betrachten. Da, wie in den letzten Kapiteln aufgezeigt, bereits bestehende Studien in diesem Kontext zu ähnlichen Ergebnissen kommen, wie unsere Analysen, halten wir diese dennoch für verlässlich und sie sollten in weiteren Studien verifiziert werden.

Die Tatsache, dass alle konventionellen Berichte von unterschiedlich erfahrenen Radiologinnen und Radiologen erstellt worden sind, stellt ein potenzielles Bias

dar, das sich negativ auf die Vergleichbarkeit und Bewertung der entsprechenden Befunde auswirken kann. Jedoch wurden alle CRs unter oberärztlicher Leitung des LMU Klinikums professionell kontrolliert, was ein ausreichendes Mindestmaß an Befundqualität sicherstellt und zudem die tägliche Praxis der Klinikroutine widerspiegelt.

Diese Arbeit ermöglicht somit keinen abschließenden Vergleich zur Überlegenheit der strukturierten gegenüber der herkömmlichen Befundung bei Verdacht auf COVID-19, liefert jedoch wegweisende Aspekte und klare Aussagen.

4.5.1 Schlussfolgerung

Die hier vorliegende Studie liefert neue Erkenntnisse zur strukturierten CT-Befundung unter Anwendung der CO-RADS-Klassifikation gegenüber dem konventionellen Freitextbefund im Kontext der COVID-19-Pandemie. Explizit zu diesem Thema wurden bisher noch keine Studien durchgeführt. Unsere Observationen zeigen, dass durch strukturierte Befundung formale Aspekte wie die Strukturiertheit, die Angemessenheit der Befundlänge, die Vollständigkeit und die Verständlichkeit radiologischer Berichte optimiert werden können. Insgesamt kann damit eine höhere Gesamtqualität der Befundung erzielt werden. Diese Überlegenheit von SRs gegenüber CRs deckt sich mit multiplen vorherigen Publikationen unterschiedlicher bildgebender Verfahren wie CT, MRT und Röntgen bei einer Vielzahl von klinischen Fragestellungen (35, 55, 58, 87, 89, 90).

4.6 Ausblick

Wie der gesamte medizinische Sektor unterliegt auch die Radiologie dem derzeitigen Wandel durch zunehmende Informationsmengen und wachsende Technologisierung. Damit einher gehen eine höhere Arbeitsbelastung bei steigendem Befundumsatz, zunehmend herausfordernde Aufgaben und Fragestellungen für Radiologinnen und Radiologen sowie Klinikerinnen und Kliniker (36). Um dieser Mehrbelastung zu begegnen und um Arbeitsprozesse zu optimieren, existieren bereits dynamische Softwarelösungen, die auf künstlicher Intelligenz (engl. artificial intelligence, AI) beruhen.

Wenngleich sich viele Lösungen der AI noch im Evaluationsstadium befinden, so steht bereits fest, dass für den künftigen Einsatz die Fertigung repräsentativer, einheitlicher Datensätze unabdingbar ist (91) – somit ist ein standardisiertes Vokabular, wie es im SR zu finden ist, für diese neuen und wegweisenden Technologien eine Grundvoraussetzung.

Mit Hilfe von AI-Software ist es bspw. möglich, den exakten Prozentsatz der betroffenen Lungenareale im CT-Bild einer COVID-19-positiven Person genau zu bestimmen. Die Befundungssoftware der meisten Kliniken erlaubt es jedoch noch nicht, diesen Wert nahtlos in den Radiologiebefund zu integrieren. Dieser Sachverhalt illustriert neben einigen anderen Hindernissen die derzeitige Lage und weist auf eine erforderliche Prozessoptimierung hin, um die Anwendung von AI-Software benutzerfreundlich und zeiteffizient zu gestalten (92, 93). Ein naheliegender zukünftiger Schritt könnte die Implementierung der herkömmlichen sprachgesteuerten Software in die SR-Software sein, um sich Synergien durch gleichzeitiges Betrachten, Diktieren und Befüllen zunutze zu machen und so neben der Qualitätssteigerung auch eine Zeitersparnis beim Befunden zu erreichen. Der erfolgreiche Versuch, eine flexible Software in einem Krankenhaus zu etablieren, ging in einer Studie mit hohen Zufriedenheitswerten seitens der Radiologinnen und Radiologen einher (89, 94). In diesem Zusammenhang wäre es lohnenswert, Vorlagen mit neuen optimierten Softwarelösungen in zukünftiger Forschung zu untersuchen. Dabei könnte perspektivisch eine strukturierte Befundung zum Einsatz kommen und in Kombination mit AI zu erheblichen Effizienzsteigerungen führen.

Wenngleich AI noch nicht in erheblichem Maße im Klinikalltag etabliert ist, trägt ihr zukünftiger Einsatz zur Auswertung radiologischer Bildmaterialien wegweisendes Potential in Richtung einer erfolgsversprechenden, individuellen und modernen Medizin (95). Eine größere Verbreitung der strukturierten Befundung durch eine Weiterentwicklung und Feinjustierung der AI wird von Experten erwartet (96).

Derzeitige Hindernisse für diese Entwicklung bestehen insbesondere noch in der Regulatorik: Hierbei stößt man auf diverse Teillösungen, die eine Zusammenführung und Implementierung in funktionierende Strukturen erfordern. Abgesehen von grundlegenden Standards zur Umsetzung in IT-Systemen bedarf es der Validierung von Lernalgorithmen. Zudem hat die Schnittschnelle zwischen Radiologie und Klinik bei der Erstellung, Anpassung und Überarbeitung von Templates für eine leitliniengerechte Medizin große Bedeutung. Neben einer maßgebenden

Veränderungsbereitschaft der Ärztinnen und Ärzte bedarf es der Etablierung thematischer, technischer, ethischer und juristischer Standards sowie der Klärung regulatorischer Rahmenbedingungen wie bspw. Datenschutz sowie Vergütung der Leistungen durch die Krankenkassen in Kliniken und Praxen. Ziel sollte ein ganzheitlicher, lösungsorientierter Ansatz sein, der verschiedene Technologien verbindet und dabei keine Kompromisse bei Genauigkeit, Vollständigkeit, Arbeitsabläufen und Kosten-Nutzen-Verhältnis eingeht (97).

Insgesamt könnte eine künftige Verwendung der strukturierten Befundung eine vereinfachte Quantifizierung und Zugänglichkeit zu Patientendaten bei höheren qualitativen Standards gewährleisten (60). Die strukturierte Befundung stellt dabei ein geeignetes Werkzeug dar, um die Grundlage dafür zu schaffen, dass durch automatisierte Algorithmen zuverlässig Informationsextraktion und Qualitätssicherung aus radiologischen Befunden erfolgen können. Dadurch könnten die Patientenversorgung optimiert und die Forschung weiterentwickelt werden (56, 98). Die strukturierte Befundung in Kombination mit den beschriebenen Technologien birgt insofern das Potential und eine Chance, die aktuellen bzw. herkömmlichen Arbeitsabläufe in der Radiologie revolutionär zu verändern und zu optimieren.

Die Corona-Pandemie hat uns auf drastische Art und Weise vor Augen geführt, dass Gesundheitssysteme weltweit an ihre Kapazitätsgrenzen stoßen können.

Hierbei hat die CT-Thorax in der Diagnostik von COVID-Patientinnen und Patienten mangels ausreichender sowie verlässlicher Labortests eine Schlüsselrolle eingenommen. Da es nicht ausgeschlossen ist, dass uns zukünftig ähnliche Pandemien ereilen werden, gilt es, Methoden und Verfahren zu entwickeln, die die Effizienz in Gesundheitssystemen insbesondere in solchen Situationen steigern und auch die Erhebung und Analyse von Daten vereinfachen. Wie in dieser Studie unter Beweis gestellt, kann die strukturierte Befundung von CT-Untersuchungen des Thorax unter Verwendung der CO-RADS-Klassifikation dabei perspektivisch einen relevanten Beitrag leisten.

Literaturverzeichnis

1. Robert Koch-Institut [Internet]. Epidemiologisches Bulletin: SARS-CoV-2 in Deutschland und Ziele von Infektionsschutzmaßnahmen; 2020 [cited 2022 Aug 22]. Available from: https://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2020/Ausgaben/07_20.pdf?__blob=publicationFile
2. WHO. COVID-19 Dashboard [Internet]. Geneva: World Health Organization; 2020 [cited 2022 Aug 22]. Available online: <https://covid19.who.int/> (last cited:).
3. Deutsches Zentrum für Infektionsforschung e.V. [Internet]. Coronavirus; 2022 [2022 Aug 23]. Available from: <https://www.dzif.de/de/glossar/coronavirus>
4. WHO, WHO Coronavirus (COVID-19) Dashboard [Internet]. 2022 [cited 2022 AUG 22]. Available from: <https://covid19.who.int/?mapFilter=cases>
5. Robert Koch-Institut, DIVI e.V. [Internet]. Zeitreihen zu Fallzahlen und Intensivkapazitäten der Erwachsenen-Intensivstationen; 2022 [cited 2022 Nov 01]. Available from: <https://www.intensivregister.de/#/aktuelle-lage/zeitreihen>
6. Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, et al. Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *JAMA*. 2020;323(18):1843-4.
7. Rahbari R, Moradi N, Abdi M. rRT-PCR for SARS-CoV-2: Analytical considerations. *Clin Chim Acta*. 2021;516:1-7.
8. Shen M, Zhou Y, Ye J, Abdullah Al-Maskri AA, Kang Y, Zeng S, et al. Recent advances and perspectives of nucleic acid detection for coronavirus. *J Pharm Anal*. 2020;10(2):97-101.
9. Noh JY, Yoon SW, Kim DJ, Lee MS, Kim JH, Na W, et al. Simultaneous detection of severe acute respiratory syndrome, Middle East respiratory syndrome, and related bat coronaviruses by real-time reverse transcription PCR. *Arch Virol*. 2017;162(6):1617-23.
10. Wan Z, Zhang Y, He Z, Liu J, Lan K, Hu Y, et al. A Melting Curve-Based Multiplex RT-qPCR Assay for Simultaneous Detection of Four Human Coronaviruses. *Int J Mol Sci*. 2016;17(11):1880.
11. Lu X, Whitaker B, Sakthivel SK, Kamili S, Rose LE, Lowe L, et al. Real-time reverse transcription-PCR assay panel for Middle East respiratory syndrome coronavirus. *J Clin Microbiol*. 2014;52(1):67-75.
12. Corman VM, Eckerle I, Bleicker T, Zaki A, Landt O, Eschbach-Bludau M, et al. Detection of a novel human coronavirus by real-time reverse-transcription polymerase chain reaction. *Euro Surveill*. 2012;17(39): 1-6.

13. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill.* 2020;25(3):23-30.
14. TFS [Internet]. Thermo Fisher Scientific Scientific; 2023 [cited 2023 Jan 23] [RT-PCR use]. Available from: <https://www.thermofisher.com/de/de/home/life-science/pcr/real-time-pcr/real-time-pcr-learning-center/real-time-pcr-basics/real-time-vs-digital-vs-traditional-pcr.html>
15. Leontine, Marian, Nijhuis M, Karin, Ad, Rozenberg-Arska M, et al. Polymerase Chain Reaction Is More Sensitive than Viral Culture and Antigen Testing for the Detection of Respiratory Viruses in Adults with Hematological Cancer and Pneumonia. *Clinical Infectious Diseases.* 2002;34(2):177-83.
16. Rovida F, Percivalle E, Zavattoni M, Torsellini M, Sarasini A, Campanini G, et al. Monoclonal antibodies versus reverse transcription-PCR for detection of respiratory viruses in a patient population with respiratory tract infections admitted to hospital. *J Med Virol.* 2005;75(2):336-47.
17. STAKOB, Feldt T, Jensen B, Guggemos W, Kellner N, Kobbe R, Koch T et al. [Internet]. Hinweise zu Erkennung, Diagnostik und Therapie von Patienten mit COVID-19; 2022 [cited 2022 Apr 06]. Available from: https://www.rki.de/DE/Content/Kommissionen/Stakob/Stellungnahmen/Stellungnahme-Covid-19_Therapie_Diagnose.pdf?__blob=publicationFile
18. COVID-19-Patienten schon zweieinhalb Tage vor Symptombeginn infektiös. *Pneumo News.* 2020;12(3):27-28.
19. Petersen E, Koopmans M, Go U, Hamer DH, Petrosillo N, Castelli F, et al. Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. *Lancet Infect Dis.* 2020;20(9):e238-e44.
20. Yuce M, Filiztekin E, Ozkaya KG. COVID-19 diagnosis - A review of current methods. *Biosens Bioelectron.* 2021;172:112752.
21. Li Y, Yao L, Li J, Chen L, Song Y, Cai Z, et al. Stability issues of RT-PCR testing of SARS-CoV-2 for hospitalized patients clinically diagnosed with COVID-19. *J Med Virol.* 2020;92(7):903-8.
22. Radiologienetz [Internet]. Geschichte der Computertomographie; 2023 [cited 2023 Sep 9]. Available from: <https://www.radiologie.de/ge%c2%adschich%c2%adte-der-com%c2%adpu%c2%adter%c2%adto%c2%admo%c2%adgra%c2%adphie/>
23. Hsieh J, Flohr T. Computed tomography recent history and future perspectives. *J Med Imaging (Bellingham).* 2021;8(5):052109.

24. Rubin GD, Ryerson CJ, Haramati LB, Sverzellati N, Kanne JP, Raof S, et al. The Role of Chest Imaging in Patient Management during the COVID-19 Pandemic: A Multinational Consensus Statement from the Fleischner Society. *Radiology*. 2020;296(1):172-80.
25. Leonardi A, Scipione R, Alfieri G, Petrillo R, Dolciami M, Ciccarelli F, et al. Role of computed tomography in predicting critical disease in patients with covid-19 pneumonia: A retrospective study using a semiautomatic quantitative method. *Eur J Radiol*. 2020;130:109202.
26. Pfeifer N, Zaboli, A., Ciccariello, L., Bernhart, O., Troi, C., Fanni Canelles, M., Ammari, C., Fioretti, A., Turcato, G., 2022. Nomogramm zur Risikostratifizierung von COVID-19-Patienten mit interstitieller Pneumonie in der Notaufnahme. *Medizinische Klinik - Intensivmedizin und Notfallmedizin*. 2022;117:120-8.
27. Pontone G, Scafuri S, Mancini ME, Agalbato C, Guglielmo M, Baggiano A, et al. Role of computed tomography in COVID-19. *J Cardiovasc Comput Tomogr*. 2021;15(1):27-36.
28. The Canadian Association of Radiologists [Internet]. Canadian Society of Thoracic Radiology and the Canadian Association of Radiologists' Statement on COVID -19. 2022 [cited 2022 Sep 11]. Available from: <https://car.ca/news/canadian-society-of-thoracic-radiology-and-canadian-association-of-radiologists-statement-on-covid-19/>
29. ACR. Recommendations for the use of Chest Radiography and Computed Tomography (CT) for Suspected COVID-19 Infection. 2020 [cited 2022 Sep 13]. Available from: <https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection>
30. Long C, Xu H, Shen Q, Zhang X, Fan B, Wang C, et al. Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? *Eur J Radiol*. 2020;126:108961.
31. Stanzione A, Ponsiglione A, Cuocolo R, Rumolo M, Santarsiere M, Scotto R, et al. Chest CT in COVID-19 patients: Structured vs conventional reporting. *Eur J Radiol*. 2021;138:109621.
32. Valentin B, Steuwe A, Wienemann T, Andree M, Keitel V, Ljimini A, et al. Applicability of CO-RADS in an Anonymized Cohort Including Early and Advanced Stages of COVID-19 in Comparison to the Recommendations of the German Radiological Society and Radiological Society of North America. *Rofo*. 2022;194(8):862-72.
33. Ozkarafakili MA, Ozkurt H, Bardakci MI, Akilli IK, Yanc U, Altuntas Y, et al. Comparison of chest computed tomography findings of RT-PCR negative and RT-PCR positive cases in COVID-19 patients. *Clin Imaging*. 2022;82:7-12.

34. Gagliardi RA. The evolution of the X-ray report. *AJR Am J Roentgenol.* 1995;164(2):501-2.
35. Gassenmaier S. Strukturierte Befundung in der Radiologie – Einfluss auf Vollständigkeit und Qualität der Befunde [Dissertation]. München: LMU München; 2019. ediss:25227.
36. Sigl B, Herold C. Strukturierte Befundung in der Radiologie – Chance für die radiologische Jugend? *Radiologe.* 2021;61(5):487-9.
37. Ganeshan D, Duong PT, Probyn L, Lenchik L, McArthur TA, Retrouvey M, et al. Structured Reporting in Radiology. *Acad Radiol.* 2018;25(1):66-73.
38. Sobez LM, Kim SH, Angstwurm M, Stormann S, Pforringer D, Schmidutz F, et al. Creating high-quality radiology reports in foreign languages through multilingual structured reporting. *Eur Radiol.* 2019;29(11):6038-48.
39. Kotter E, Pinto Dos Santos D. [Structured reporting in radiology: German and European radiology societies' point of view]. *Radiologe.* 2021;61(11):979-85.
40. Langlotz CP. RadLex: a new method for indexing online educational materials. *Radiographics.* 2006;26(6):1595-7.
41. Durack JC. The value proposition of structured reporting in interventional radiology. *AJR Am J Roentgenol.* 2014;203(4):734-8.
42. Reiner BI, Knight N, Siegel EL. Radiology reporting, past, present, and future: the radiologist's perspective. *J Am Coll Radiol.* 2007;4(5):313-9.
43. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Muller NL, Remy J. Fleischner Society: glossary of terms for thoracic imaging. *Radiology.* 2008;246(3):697-722.
44. Wormanns D, Hamer OW. [Glossary of Terms for Thoracic Imaging – German Version of the Fleischner Society Recommendations]. *Rofo.* 2015;187(8):638-61.
45. Hackländer T. Strukturierte Befundung in der Radiologie. *Radiologe.* 2013;53(7):613-7.
46. Naik SS, Hanbidge A, Wilson SR. Radiology reports: examining radiologist and clinician preferences regarding style and content. *AJR Am J Roentgenol.* 2001;176(3):591-8.
47. 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(5):588-93.
48. RADElement RA. Common Data Elements(CDEs)for Radiology. 2023 [cited 2023 Jan 23]. Available from: <https://www.radelement.org/>

49. Spiro JE, Ceelen F, Kneidinger N, Sommer WH, Schinner R, Dinkel J, et al. Structured Reporting of Computed Tomography Examinations in Post-Lung Transplantation Patients. *J Comput Assist Tomogr.* 2021;45(6):959-63.
50. Radiological Society of North America [Internet]. Radiology Reporting Template License; 2022 [cited 2022 Dec 25]. Available from: <https://www.radreport.org/>.
51. Salehi S, Abedi A, Balakrishnan S, Gholamrezanezhad A. Coronavirus disease 2019 (COVID-19) imaging reporting and data system (COVID-RADS) and common lexicon: a proposal based on the imaging data of 37 studies. *Eur Radiol.* 2020;30(9):4930-42.
52. Flusberg M, Ganeles J, Ekinci T, Goldberg-Stein S, Paroder V, Kobi M, et al. Impact of a Structured Report Template on the Quality of CT and MRI Reports for Hepatocellular Carcinoma Diagnosis. *J Am Coll Radiol.* 2017;14(9):1206-11.
53. Sahni VA, Silveira PC, Sainani NI, Khorasani R. Impact of a Structured Report Template on the Quality of MRI Reports for Rectal Cancer Staging. *AJR Am J Roentgenol.* 2015;205(3):584-8.
54. Brook OR, Brook A, Vollmer CM, Kent TS, Sanchez N, Pedrosa I. Structured Reporting of Multiphasic CT for Pancreatic Cancer: Potential Effect on Staging and Surgical Planning. *Radiology.* 2015;274(2):464-72.
55. Norenberg D, Sommer WH, Thasler W, D'Haese J, Rentsch M, Kolben T, et al. Structured Reporting of Rectal Magnetic Resonance Imaging in Suspected Primary Rectal Cancer: Potential Benefits for Surgical Planning and Interdisciplinary Communication. *Invest Radiol.* 2017;52(4):232-9.
56. Gassenmaier S, Armbruster M, Haasters F, Helfen T, Henzler T, Alibek S, et al. Structured reporting of MRI of the shoulder – improvement of report quality? *Eur Radiol.* 2017;27(10):4110-9.
57. Khorasani R, Bates DW, Teeger S, Rothschild JM, Adams DF, Seltzer SE. Is terminology used effectively to convey diagnostic certainty in radiology reports? *Acad Radiol.* 2003;10(6):685-8.
58. Sabel BO, Plum JL, Kneidinger N, Leuschner G, Koletzko L, Raziorrouh B, et al. Structured reporting of CT examinations in acute pulmonary embolism. *J Cardiovasc Comput Tomogr.* 2017;11(3):188-95.
59. Nobel JM, Kok EM, Robben SGF. Redefining the structure of structured reporting in radiology. *Insights Imaging.* 2020;11(1):10.
60. European Society of R. ESR paper on structured reporting in radiology. *Insights Imaging.* 2018;9(1):1-7.
61. Reiner BI. Strategies for radiology reporting and communication : part 4: quality assurance and education. *J Digit Imaging.* 2014;27(1):1-6.

62. Kröger B, Bunck, Pinto dos Santos, Maintz, Persigehl, Düber et al. Strukturierte Befundung in der Radiologie [Internet]. 2022 [cited 2022 Feb 19]. Available from: <https://www.strukturierte-befundung.de/>
63. Prokop M, van Everdingen W, van Rees Vellinga T, Quarles van Ufford H, Stoger L, Beenen L, et al. CO-RADS: A Categorical CT Assessment Scheme for Patients Suspected of Having COVID-19-Definition and Evaluation. *Radiology*. 2020;296(2):E97-E104.
64. Radiology Bea. COVID-19 CO-RADS classification [Internet]. 2020 [cited 2022 Sep 10]. Available from: <https://radiologyassistant.nl/chest/covid-19/corads-classification>.
65. Penha D, Pinto EG, Matos F, Hochhegger B, Monaghan C, Taborda-Barata L, et al. CO-RADS: Coronavirus Classification Review. *J Clin Imaging Sci*. 2021;11:9.
66. Bueno J, Landeras L, Chung JH. Updated Fleischner Society Guidelines for Managing Incidental Pulmonary Nodules: Common Questions and Challenging Scenarios. *Radiographics*. 2018;38(5):1337-50.
67. Yildirim Ayaz E, Coskun ZU, Kaplan M, Bulut AS, Yesildal M, Ankarali H, et al. Comparison of Initial CT Findings and CO-RADS Stage in COVID-19 Patients with PCR, Inflammation and Coagulation Parameters in Diagnostic and Prognostic Perspectives. *J Belg Soc Radiol*. 2022;106(1):67.
68. Liu G, Chen Y, Runa A, Liu J. Diagnostic performance of CO-RADS for COVID-19: a systematic review and meta-analysis. *Eur Radiol*. 2022;32(7):4414-26.
69. Pinto Dos Santos D, Hempel JM, Mildenerberger P, Klockner R, Persigehl T. Structured Reporting in Clinical Routine. *Rofo*. 2019;191(1):33-9.
70. Schilling J, Tolksdorf K, Marquis A, Faber M, Pfoch T, Buda S, et al. [The different periods of COVID-19 in Germany: a descriptive analysis from January 2020 to February 2021]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2021;64(9):1093-106.
71. Smart Reporting [Internet]. 2022 [cited 2023 Mar 24]. Available from: <https://www.smart-reporting.com/en/>.
72. Völkl K, Korb C. Variablen und Skalenniveaus. Wiesbaden: Springer Fachmedien; 2018.
73. Wilcoxon F. Individual Comparisons by Ranking Methods. In: Kotz S, Johnson NL, editors. *Breakthroughs in Statistics: Methodology and Distribution*. New York: Springer; 1992. p. 196-202.
74. Pembury Smith MQR, Ruxton GD. Effective use of the McNemar test. *Behavioral Ecology and Sociobiology*. 2020;74:133-42.

75. Field A. *Discovering statistics using IBM SPSS statistics*. Los Angeles, London, New Delhi: Sage; 2013.
76. Schwartz LH, Panicek DM, Berk AR, Li Y, Hricak H. Improving communication of diagnostic radiology findings through structured reporting. *Radiology*. 2011;260(1):174-81.
77. Sabel BO, Plum JL, Czihal M, Lottspeich C, Schonleben F, Gabel G, et al. Structured Reporting of CT Angiography Runoff Examinations of the Lower Extremities. *Eur J Vasc Endovasc Surg*. 2018;55(5):679-87.
78. Ginocchio LA, Dane B, Smereka PN, Megibow AJ, Remzi FH, Esen E, et al. Structured versus non-structured reporting of pelvic MRI for ileal pouch evaluation: clarity and effectiveness. *Abdom Radiol (NY)*. 2023;48(9):2978-2985.
79. McFarland JA, Elkassema AMA, Casals L, Smith GD, Smith AD, Gunn AJ. Objective comparison of errors and report length between structured and freeform abdominopelvic computed tomography reports. *Abdom Radiol (NY)*. 2021;46(1):387-93.
80. Zhang JY, Weinberg BD, Hu R, Saindane A, Mullins M, Allen J, et al. Quantitative Improvement in Brain Tumor MRI Through Structured Reporting (BT-RADS). *Acad Radiol*. 2020;27(6):780-4.
81. Schoeppe F, Sommer WH, Norenberg D, Verbeek M, Bogner C, Westphalen CB, et al. Structured reporting adds clinical value in primary CT staging of diffuse large B-cell lymphoma. *Eur Radiol*. 2018;28(9):3702-9.
82. Marcovici PA, Taylor GA. Journal Club: Structured radiology reports are more complete and more effective than unstructured reports. *AJR Am J Roentgenol*. 2014;203(6):1265-71.
83. Quattrocchi CC, Giona A, Di Martino AC, Errante Y, Scarciolla L, Mallio CA, et al. Extra-spinal incidental findings at lumbar spine MRI in the general population: a large cohort study. *Insights Imaging*. 2013;4(3):301-8.
84. Rosskopf AB, Dietrich TJ, Hirschmann A, Buck FM, Sutter R, Pfirrmann CW. Quality management in musculoskeletal imaging: form, content, and diagnosis of knee MRI reports and effectiveness of three different quality improvement measures. *AJR Am J Roentgenol*. 2015;204(5):1069-74.
85. Stoleriu MG, Gerckens M, Obereisenbuchner F, Zaimova I, Hetrod J, Mavi SC, et al. Automated quantitative thin slice volumetric low dose CT analysis predicts disease severity in COVID-19 patients. *Clin Imaging*. 2021;79:96-101.
86. Armbruster M, Gassenmaier S, Haack M, Reiter M, Norenberg D, Henzler T, et al. Structured reporting in petrous bone MRI examinations: impact on report completeness and quality. *Int J Comput Assist Radiol Surg*. 2018;13(12):1971-80.

87. Ernst BP, Hodeib M, Strieth S, Kunzel J, Bischof F, Hackenberg B, et al. Structured reporting of head and neck ultrasound examinations. *BMC Med Imaging*. 2019;19(1):25.
88. Dickerson E, Davenport MS, Syed F, Stuve O, Cohen JA, Rinker JR, et al. Effect of Template Reporting of Brain MRIs for Multiple Sclerosis on Report Thoroughness and Neurologist-Rated Quality: Results of a Prospective Quality Improvement Project. *J Am Coll Radiol*. 2017;14(3):371-9.
89. Geyer T, Rubenthaler J, Marschner C, von Hake M, Fabritius MP, Froelich MF, et al. Structured Reporting Using CEUS LI-RADS for the Diagnosis of Hepatocellular Carcinoma (HCC) – Impact and Advantages on Report Integrity, Quality and Interdisciplinary Communication. *Cancers (Basel)*. 2021;13(3): 534.
90. Kim SH, Sobez LM, Spiro JE, Curta A, Ceelen F, Kampmann E, et al. Structured reporting has the potential to reduce reporting times of dual-energy x-ray absorptiometry exams. *BMC Musculoskelet Disord*. 2020;21(1):248.
91. Langs G, Attenberger U, Licandro R, Hofmanninger J, Perkonigg M, Zusag M, et al. [Machine learning in radiology: Terminology from individual timepoint to trajectory]. *Radiologe*. 2020;60(1):6-14.
92. Pinto dos Santos D, Fatehi M. Structured reporting in radiology. Cham, Switzerland: Springer; 2022.
93. Prayer F, Rohrich S, Pan J, Hofmanninger J, Langs G, Prosch H. [Artificial intelligence in lung imaging]. *Radiologe*. 2020;60(1):42-7.
94. Larson DB, Towbin AJ, Pryor RM, Donnelly LF. Improving Consistency in Radiology Reporting through the Use of Department-wide Standardized Structured Reporting. *Radiology*. 2013;267(1):240-50.
95. Murray JM, Kaissis G, Braren R, Kleesiek J. [A primer on radiomics]. *Radiologe*. 2020;60(1):32-41.
96. Mentzel HJ. [Artificial intelligence in image evaluation and diagnosis]. *Monatsschr Kinderheilkd*. 2021;169(8):694-704.
97. Bosmans JM, Peremans L, Menni M, De Schepper AM, Duyck PO, Parizel PM. Structured reporting: if, why, when, how-and at what expense? Results of a focus group meeting of radiology professionals from eight countries. *Insights Imaging*. 2012;3(3):295-302.
98. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018;18(8):500-10.

Anhang A

Strukturierte Befundung nach CO-RADS – Evaluationsbogen

Pat.-ID: CO-RADS

Befundart: konventioneller Befund
 strukturierter Befund

Formale Aspekte (1= stimme voll zu; 5 = stimme gar nicht zu)

- | | 1 | 2 | 3 | 4 | 5 |
|---------------------------------------------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. Der Befund ist klar strukturiert. | <input type="radio"/> |
| 2. Die Länge des Befundes angemessen. | <input type="radio"/> |
| 3. Der Befund ist gut verständlich formuliert. | <input type="radio"/> |
| 4. Zu folgenden Aspekten fehlen Informationen im Befund (einfaches Ankreuzen): | | | | | |
| <input type="radio"/> Pneumonie | | | | | |
| <input type="radio"/> Rundherde/ Raumforderung | | | | | |
| <input type="radio"/> Mögliche weitere pulmonale Pathologien (Ödem, Emphysem, Fibrose, ...) | | | | | |
| <input type="radio"/> Pleura | | | | | |
| <input type="radio"/> Bronchien | | | | | |
| <input type="radio"/> Lymphknoten | | | | | |
| <input type="radio"/> Herz | | | | | |
| <input type="radio"/> Große thorakale Gefäße | | | | | |
| <input type="radio"/> Andere: | | | | | |

Klinische Relevanz (1= stimme voll zu; 5 = stimme gar nicht zu)

- | | | | | | |
|-----------------------------------------------------------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 5. Der Befund beantwortet die klinische Fragestellung. | <input type="radio"/> |
| 6. Der Befund lässt eine Entscheidung über das weitere diagnostische und/oder therapeutische Vorgehen zu. | <input type="radio"/> |

Gesamtqualität (1= stimme voll zu; 5 = stimme gar nicht zu)

- | | | | | | |
|--------------------------------------------------------------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 7. Ich habe Vertrauen in die Informationen aus dem Befund. | <input type="radio"/> |
| 8. Wie bewerten Sie die Gesamtqualität des Befunds? (1= sehr gut; 5 = sehr schlecht) | <input type="radio"/> |

Danksagung

Mein herzlicher Dank gilt allen, die mich bei der Fertigstellung der vorliegenden Arbeit stets unterstützt haben.

Ich danke Herrn Prof. Dr. Jens Ricke und der Klinik und Poliklinik für Radiologie am Klinikum der Universität München für die Möglichkeit, auf dem Gebiet der Radiologie promovieren zu dürfen.

Herrn Professor Dr. med. Johannes Rübenthaler danke ich für die freundliche Bereitstellung des Themas sowie die Ermöglichung einer vertrauensvollen und professionellen Zusammenarbeit an der Klinik und Poliklinik für Radiologie.

Mein besonderer Dank für die Betreuung gilt Frau Dr. med. Judith Eva Spiro, die sie trotz ihrer Belastung in Klinik und Forschung tatkräftig übernommen hat. Sie stand sowohl bei fachlichen Fragen als auch bei allen anderen Unklarheiten jederzeit hilfsbereit und konstruktiv beratend zur Verfügung.

Für die Kooperation und Mithilfe der Bewertung der Befundformen anhand der Fragebögen danke ich PD Dr. med. Katrin Milger-Kneidinger und Frau Dr. med. Michaela Barnikel.

Abschließend möchte ich mich von ganzem Herzen bei meiner Familie und Freunden bedanken, die mich bei der Erstellung der Dissertation ermutigt und unterstützt haben.

Affidavit



Regehr, Magdalena

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

Strukturierte Befundung von CT-Untersuchungen des Thorax unter Verwendung der CO-RADS-Klassifikation – Vergleich mit konventionellen Freitextbefunden

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 20.01.2024
Ort, Datum

Magdalena Regehr
Unterschrift Doktorandin bzw. Doktorand