Multizentrische Analyse des Einflusses der strukturierten Befunderhebung auf die Interrater-Reliabilität von Sonographien des Kopf-Hals-Bereiches

Dissertation

zur Erlangung des Doktorgrades (Dr. med.)

der Medizinischen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität

Bonn

Carla Teresa Dörsching, geb. Eckrich

aus Hanau

2026

Angefertigt mit der Genehmigung

der Medizinischen Fakultät der Universität Bonn

1. Gutachter: PD Dr. Benjamin Philipp Ernst

2. Gutachter: Prof. Dr. Dariusch-Reza Hadizadeh Kharrazi

Tag der mündlichen Prüfung: 21.05.2026

Aus der Klinik und Poliklinik für Hals-Nasen-Ohrenheilkunde

Meinen Eltern Martina und Martin

Inhaltsverzeichnis

Abkürzungsverzeichnis		**6**
1.	**Einleitung**	**8**
1.1	Überblick	8
1.2	Fragestellung	26
2.	**Material und Methoden**	**28**
2.1	Studiendesign	28
2.2	Aufklärung und Einverständnis der Patienten	30
2.3	Patientenkollektiv	31
2.4	Untersuchungsablauf	32
2.5	Datenverarbeitung und statistische Auswertung	38
3.	**Ergebnisse**	**39**
3.1	Vollständigkeit strukturierter Befunde	39
3.2	Interrater-Reliabilität	40
3.3	Benutzerzufriedenheit	42
4.	**Diskussion**	**44**
4.1	Status Quo	45
4.2	Befundvollständigkeit	45
4.3	Interrater-Reliabilität und Interrater-Agreement	46
4.4	Benutzerzufriedenheit	49
4.5	Limitationen	50
4.6	Ausblick	51
5.	**Zusammenfassung**	**54**
6.	**Abbildungsverzeichnis**	**56**
7.	**Literaturverzeichnis**	**58**
8.	**Erklärung zum Eigenanteil**	**73**
9.	**Veröffentlichungen**	**74**
10.	**Danksagung**	**75**

Abkürzungsverzeichnis

A-Mode	*Amplitude-Mode-Scan*
ACR	*American College of Radiology*
B-Mode	*Brightness-Mode-Scan*
BDA	*Big-Data*-Analyse
BGB	Bürgerliches Gesetzbuch
BI-RADS	*Breast Imaging Reporting and Data System*
BMG	Bundesministerium für Gesundheit
BMV-Ä	Bundesmantelvertrag - Ärzte
CT	Computertomographie
CTPA	Computertomographische Pulmonalangiographie
DEGUM	Deutsche Gesellschaft für Ultraschall in der Medizin
DICOM SR	*Digital Imaging and Communications in Medicine Structured Reporting Standard*
DIN	Deutsches Institut der Normung
DSGVO	Datenschutz - Grundverordnung
EDV	Elektronische Datenverarbeitung
ePA	Elektronische Patientenakte
ESR	*European Society of Radiology*
FESS	*Functional Endoscopic Sinus Surgery*
FTR	*Free text reporting,* Freitextbefunderhebung
G-BA	Gemeinsamer Bundesausschuss
HNO	Hals-Nasen-Ohrenheilkunde
IRA	*Interrater Agreement*, Interrater-Übereinstimmung

IRR	*Interrater Reliability*, Interrater-Reliabilität
IT	Informationstechnologie
KBV	Kassenärztlichen Bundesvereinigung
KI	Künstliche Intelligenz
KIS	Krankenhausinformationssystem
LIMS	*Laboratory Information Management Systems*
LLM	*Large Language Models*
M-Mode	*Motion-Mode-Scan*
MBO-Ä	Musterberufsordnung - Ärzte
MKG	Mund-Kiefer-Gesichtschirurgie
MRT	Magnetresonanztomographie
MVZ	Medizinisches Versorgungszentrum
PACS	*Picture archiving and communication system*
QM	Qualitätsmanagement
QMS	Qualitätsmanagementsystem
RIS	Radiologieinformationssystem
SD	Standardabweichung
SGB	Sozialgesetzbuch
SOP	Strukturierte Operationsplanung
SR	*Structured Reporting,* Strukturierte Befunderhebung
StrlSchG	Strahlenschutzgesetz
VAS	Visuelle Analogskala
VFSS	Videofluoroskopische Schluckstudie

1. Einleitung

1.1 Überblick

1.1.1 Sonographie

Als Ultraschall bezeichnet man mechanische Schwingungen mit einer Frequenz oberhalb von 20 kHz, die für das menschliche Ohr nicht wahrnehmbar sind. Während Ultraschall seine Anwendung bereits in anderen Bereichen wie dem Militär Anfang des 20. Jahrhundert fand, etablierte er sich in der Medizin als bildgebendes Verfahren erst in den 1940er Jahren. Seither wird er in verschiedenen medizinischen Fachrichtungen eingesetzt. Dabei beruht die Möglichkeit der Visualisierung von Strukturen durch Ultraschallwellen auf dem sogenannten reziproken piezoelektrischen Effekt. Die Entdeckung dieses Effektes geht auf Pierre und Jacques Curie zurück (Shampo und Kyle, 1995).

Ultraschallwellen entstehen durch Schwingungen spezieller Kristalle, welche sich im Schallkopf des Ultraschallgerätes befinden. Diese werden durch Spannung in Bewegung versetzt (sog. piezoelektrischer Effekt). Die Sonde des Schallkopfes gibt diese Schwingungen in Form von Schallwellen ab (Transmission) und empfängt diese auch wieder (Rezeption). Die empfangenen Wellen werden in elektrische Impulse umgewandelt und anschließend weiterverarbeitet. Schallwellen breiten sich in unterschiedlichen Medien entsprechend ihrer Dichte unterschiedlich schnell aus. Treffen diese Impulse auf Grenzflächen kommt es je nach Veränderung des Schallwellenwiderstandes (Impedanz) zu einer unterschiedlich starken Ausbreitung bzw. Reflexion der Schallwellen (Dössel, 2016; Pohlmann, 1948). Die empfangenen Ultraschallwellen können mittels verschiedener Darstellungsmethoden sichtbar gemacht werden. Erstmalig stellte der Neurologe Theodore Karl Dussik in Zusammenarbeit mit seinem Bruder, einem Radiotechniker, Ultraschallwellen dar und setzte damit den Grundstein für den Einsatz des Ultraschalls als diagnostisches Mittel in der Medizin (Dussik et al., 1958).

Die heute genutzte Ultraschalltechnologie umfasst verschiedene Verfahren, die sich in ihrer Funktionsweise und Anwendung unterscheiden. Vom eindimensionalen A-Mode (*Amplitude-Mode*) und M-Mode (*Motion-Mode*) über den zweidimensionalen B-Mode

(Brightness-Mode) bis hin zur Doppler- und Duplexsonographie, die zusätzlich Flussinformationen abbilden (Booth et al., 1966; Dössel, 2016; Gehring und von Bibra, 1998; Reinert, 2000; Schäberle, 2016; Welkoborsky et al., 1997). Untersuchungsergebnisse stehen somit als unmittelbare Informationen zur Verfügung, die Interpretation kann unmittelbar erfolgen und eventuelle Therapieoptionen zeitnah eingeleitet werden.

Die Sonographie bringt eine Vielzahl grundlegender Vorteile mit sich und hat in den letzten Jahrzenten daher immer mehr an Bedeutung gewonnen. Das Verfahren zeichnet sich durch Flexibilität, Mobilität und Kosteneffizienz aus und ist bei korrekter Anwendung vollkommen unschädlich (Welkoborsky et al., 1997, 2009). Dank kontinuierlicher Verbesserungen in der Bildgebung bietet der Ultraschall heute eine Bildqualität mit hohem diagnostischem Wert. Ultraschall ermöglicht die suffiziente Visualisierung multipler Organe und die quantitative Erhebung zentraler Parameter wie Lage, Größe, Form, Binnenstruktur und sogar Hämodynamik (Dössel, 2016). Als Untersuchung ist sie leicht verfügbar und nahezu ubiquitär einsetzbar (Enders et al., 2011a).

Sonographie ist nicht-invasiv, frei von ionisierender Strahlung und weist eine hohe Akzeptanz unter den Patienten* auf, was die Verwendung auch bei Kindern, Schwangeren und Menschen mit Klaustrophobie ermöglicht (Schuler et al., 2010). Ferner ist von großem Vorteil, dass sich Ultraschalluntersuchungen beliebig oft wiederholen lassen, sprich eine longitudinale Befundkontrolle ohne großen Aufwand oder Belastung gewährleistet ist.

In der Hals-Nasen-Ohrenheilkunde (HNO) hat sich aufgrund der zu untersuchenden Halsweichteilstrukturen insbesondere der B-Mode-Ultraschall, ergänzt durch die farbcodierte Duplexsonographie, als bildgebendes Verfahren der ersten Wahl und diagnostischer Standard etabliert (Heyduck et al., 2018). Als wertvolles diagnostisches Werkzeug wird dieser daher oft im Rahmen der Tumornachsorge eingesetzt (Adibelli et al., 1998; Bialek et al., 2006; Künzel et al., 2017; Welkoborsky et al., 1997). Die zu untersuchenden Strukturen liegen meist oberflächlich und sind für den Untersucher gut zugänglich. Dadurch können Pathologien im Rahmen einer dynamischen Untersuchung

* In der vorliegenden Promotionsschrift wird aus Gründen der besseren Lesbarkeit ausschließlich die männliche Form verwendet. Sie bezieht sich explizit auf Personen aller Geschlechter (m/w/d).

in verschiedenen Ebenen exakt dargestellt und kategorisiert werden. Technische Weiterentwicklungen und die Einführung der farbcodierten Duplexsonographie, des dreidimensionalen Ultraschalles, der Elastrographie oder der Kontrastmittel-verstärkten Sonographie ermöglichen es noch exaktere Aussagen zu anatomischen Gegebenheiten zu treffen und daraus erweiterte diagnostische Interpretationsmaxima abzuleiten (Harding et al., 2011). Die Nutzung des Ultraschalls erfolgt ferner nicht nur als rein diagnostisches Mittel sondern auch oft im Rahmen von Interventionen, wie beispielsweise ultraschallgesteuerter Punktion von Lymphknoten, Gefäßen, Körperhöhlen oder Hohlorganen. Die direkte sonographische Unterstützung erleichtert die Untersuchung durch Visualisierung der Anatomie und reduziert so nachweislich die Rate an Fehlpunktionen (Mojoli et al., 2019).

Trotz dieser Vorzüge weist die Sonographie relevante Limitationen auf. Artefakte durch luft- oder knochendichte Strukturen beeinträchtigen die Bildqualität erheblich (Bock et al., 2019; Bönhof, 2016). Hinzu kommen die komplexen anatomischen Verhältnisse im Kopf-Hals-Bereich, welche die Untersuchung erschweren und fundiertes anatomisches Wissen einfordern. Die genaue Topographie und Bedeutung von Strukturen gilt es richtig zu identifizieren, zu interpretieren und zu benennen (Wallis und McCoubrie, 2011). Gerade unerfahrene Untersucher stellt dies vor Herausforderungen, weil grundlegende Fähigkeiten teilweise noch fehlen (Gmeinwieser et al., 1988; Meudt et al., 1986). Grundsätzlich gilt die Ultraschalldiagnostik als stark untersucherabhängiges Verfahren (Gandhi et al., 2006; Nasrallah, 2015; Wallis und McCoubrie, 2011). Die Qualität und Aussagekraft von Ultraschalluntersuchungen hängt maßgeblich vom Ausbildungsstand, der praktischen Erfahrung, der Weiterbildung sowie der routinemäßigen Anwendung durch den Untersucher ab. Methodische Unterschiede, variable Untersuchungspositionen und der unterschiedliche Einsatz technischer Ausstattung führen zusätzlich zu einer hohen Variabilität der Ergebnisse. Ferner kann eine unzureichende Berichtqualität die diagnostische Aussagekraft der Untersuchung zusätzlich negativ beeinflussen. So erschwert eine schlecht strukturierte oder unvollständige Befunddokumentation insbesondere bei Verlaufsuntersuchungen die Vergleichbarkeit über einen längeren Zeitraum. Eine unklare und uneinheitliche Befundformulierung kann zu Missverständnissen zwischen dem Untersucher und Befundempfänger führen, was fehlerhaften klinischen Entscheidungen mit sich bringen kann (Scott et al., 2005).

Da die Qualität der Ultraschalldiagnostik maßgeblich vom Untersucher abhängt, kommt dessen Ausbildung und Qualifikation eine zentrale Bedeutung zu. In Deutschland existiert jedoch weder im Rahmen der universitären Ausbildung noch während der Facharztweiterbildung ein einheitlicher Standard zur Strukturierung der Aus- und Weiterbildung. Um dieser hohen Untersucherabhängigkeit in der Ultraschaldiagnostik Rechnung zu tragen, wurde 1972 die Deutsche Gesellschaft für Ultraschall in der Medizin (DEGUM e.V.) gegründet. Sie hat sich zur Aufgabe gemacht, Wissenschaft und Forschung sowie die Aus-, Fort- und Weiterbildung im Bereich der medizinischen Ultraschalldiagnostik zu fördern und hat sich mittlerweile zu einer der wichtigsten Ultraschallgesellschaften weltweit entwickelt: „Zweck der Gesellschaft ist die Förderung von Wissenschaft und Forschung, der Volks- und Berufsbildung sowie des öffentlichen Gesundheitswesens und der öffentlichen Gesundheitspflege auf dem Gebiet des Ultraschalls, seiner Verfahren und Anwendungen in der Medizin“ heißt es in § 2 Absatz 1 der DEGUM Satzung (Deutsche Gesellschaft für Ultraschall in der Medizin, 2022). Ihr Ziel ist es die Ultraschalldiagnostik qualitativ und quantitativ stetig zu verbessern (GKV-Spitzenverband, 2025a).

Um der Untersucherabhängigkeit entgegenzuwirken, hat die DEGUM ein Zertifizierungssystem entwickelt, das sich in drei Stufen gliedert:

- Stufe I (Basisdiagnostik)
- Stufe II (spezialisierte Diagnostik)
- Stufe III/Kursleiter (Höchstqualifizierte, wissenschaftlich orientierte Kompetenz im entsprechenden Fachbereich)

Die Studie von Heese und Görg (2006) untersuchte erstmalig, ob die Untersuchungserfahrungen eines Untersuchers DEGUM Stufe III zu einer Verbesserung der Ergebnisqualität führt (Heese und Görg, 2006). Es konnte gezeigt werden, dass die diagnostische Genauigkeit mit steigendem Qualifikationsniveau signifikant zunimmt. 94,5 % der Untersucher mit DEGUM Stufe III stellten die richtige Diagnose im Vergleich zu Untersuchern mit Fortbildungsstufe II, deren Quote bei 60,9 % bzw. 38,8 % bei Stufe-I-Untersuchern lag (Cohen, 1960). Nachweislich verbessert sich somit die Qualität der Diagnostik durch langjährige Erfahrung und Fortbildung. Die DEGUM kontrolliert und rezertifiziert die Einhaltung dieser Qualitätsstandards sowohl für einzelne Untersucher als

auch für ganze Abteilungen und Kliniken in regelmäßigen Abständen. In der vorliegenden Studie wurden daher ausschließlich Kopf-Hals-Ultraschallzentren mit DEGUM Kursleitern (Stufe III) eingeschlossen, um ein einheitlich hohes Qualifikationsniveau zu gewährleisten.

1.1.2 Befunddokumentation

Neben der Qualität des Untersuchers stellt die Qualität der Befunddokumentation einen ebenso zentralen Faktor für die diagnostische Wertigkeit der Sonographie dar. Standardisierte Befundformate können dazu beitragen, individuelle Unterschiede in der Befundung zu reduzieren und damit die Vergleichbarkeit zwischen den untersuchenden Personen verbessern.

Ein Befund ist definiert als ein durch eine Untersuchung festgestellter Zustand eines Patienten. Befunde können sich im Inhalt stark voneinander unterscheiden und werden im medizinischen Bereich in Form von Text, Bild, Video oder Ton festgehalten. Um eine lückenlose medizinische Versorgung zu gewährleisten, ist es wichtig die individuellen Informationen über den Patienten selbst, aber auch über sich anschließende Interventionen genau zu erfassen, zu ordnen und detailliert festzuhalten. Schriftliche Befunde sind dabei die unverzichtbare Grundlage für diagnostische und therapeutische Entscheidungen. Darüber hinaus sind Ärzte rechtlich verpflichtet, Befunde zu erstellen und zu archivieren. Eine ordnungsgemäße Befunddokumentation ist ebenso Teil der medizinischen Leistung wie die Untersuchung selbst. In Bezug auf diese Thematik sind Begrifflichkeiten wie Qualitätsmanagement (*quality management*, QM) aus dem Gesundheitswesen nicht mehr wegzudenken. QM beschriebt ein umfassendes Konzept, das sich mit der Sicherstellung und Verbesserung der Qualität von Dienstleistungen und Prozessen innerhalb einer Organisation befasst. Ziel ist es durch verschiedene Maßnahmen, Methoden und Standards die Patientenzufriedenheit zu erhöhen, die Effizienz zu steigern und die Fehlerquote zu minimieren. Die hierfür geltenden Anforderungen sind in der DIN EN ISO 9001 festgelegt. Seit 2004 sind alle Vertragsärzte, Psychotherapeuten und Medizinischen Versorgungszentren verpflichtet QM einzuführen und weiterzuentwickeln. Entsprechende Richtlinien, welche durch den Gemeinsamen

Bundesausschuss (G-BA) kontrolliert werden sind rechtlich festgelegt worden (§ 135a Abs. 2 Nr. 2 SGB V).

1.1.2.1 Qualitätsanforderungen an die Dokumentation gemäß der KV-Vereinbarung

Zur Darstellung der Qualitätsanforderungen an die Dokumentation in der Ultraschalldiagnostik wird im Folgenden der § 1 der aktuellen Ultraschall-Vereinbarung zitiert: „Diese Vereinbarung ist eine Maßnahme zur Qualitätssicherung, mit der die Qualität bei der Erbringung von Leistungen der Ultraschalldiagnostik gesichert werden soll. Die Vereinbarung regelt die fachlichen und apparativen Voraussetzungen für die Ausführung und Abrechnung dieser Leistungen im Rahmen der vertragsärztlichen Versorgung.“ (GKV-Spitzenverband, 2025b, § 1). Bereits am 1. April 2009 wurde die Ultraschall-Vereinbarung nach § 135 Abs. 2 SGB V rechtsverbindlich und bildet einen großen Bestandteil der vertragsärztlichen Qualitätssicherung. Sie gliedert sich neben dem Haupttext in sieben Anlagen. Neben den Anforderungen an die fachliche Kompetenz und die zu erbringenden Untersuchungsleistungen legt die Ultraschall-Vereinbarung auch die genauen Voraussetzungen an die apparative Ausstattung fest. Durch eine Abnahmeprüfung des Herstellers kann ein Nachweis zur Einhaltung der geforderten Anforderungen gegenüber der Kassenärztlichen Vereinigung (KV) erbracht werden.

§ 10 der Ultraschall-Vereinbarung definiert die Inhalte der ärztlichen Dokumentation: Neben Patienten- und Untersucheridentität müssen eine nachvollziehbare Untersuchungsindikation, die (Verdachts-)Diagnose sowie das weitere Vorgehen schriftlich festgehalten werden. Besonderheiten der Dokumentationspflicht im Schwangerschaftsverlauf (Mutterschafts-Richtlinien) sowie im Rahmen der sonographischen Früherkennungsuntersuchung der Säuglingshüfte werden gesondert aufgeführt (GKV-Spitzenverband, 2025c, Abs.2 und 3). § 11 der Ultraschall-Vereinbarung regelt die Überprüfung der ärztlichen Dokumentation. Zur Kontrolle führt die zuständige KV in regelmäßigen Abständen Stichprobenprüfungen der Schrift- und Bilddokumentationen von Ultraschalluntersuchungen durch, welche von der Ultraschall-Kommission beurteilt werden. Die Ergebnisse werden den Ärzten innerhalb von 4 Wochen mitgeteilt und reichen von „keine Beanstandung: Es erfolgt eine Bestätigung, dass die geprüften Leistungen den Qualitätsanforderungen entsprechen“ bis „schwerwiegende

Beanstandung: Es erfolgt eine schriftliche Mitteilung über festgestellte Mängel und Aufforderung, diese Mängel zukünftig zu vermeiden in Verbindung mit der Auflage für ein Beratungsgespräch oder eine Maßnahme (z.B. Fortbildungskurs, Hospitation) […]“ (GKV-Spitzenverband, 2025d, § 11 Abs.4 a-d). Trotz dieser gesetzlichen Vorgaben zeigt die Praxis, dass die tatsächliche Befundqualität erheblich variiert. Dieses Problem kann durch standardisierte Befundformate verbessert werden (Ernst et al., 2019b; Wallis und McCoubrie, 2011).

1.1.2.2 Interrater-Reliabilität

Um Aussagen über den Grad der Messgenauigkeit zu machen, muss die Reliabilität (Zuverlässigkeit) bestimmt werden. Grundsätzlich wird zwischen der Übereinstimmung eines Beobachters über eine gewisse Zeit (Intrarater-Reliabilität) und der Übereinstimmung unterschiedlicher unabhängiger Beobachter (Interrater-Reliabilität) unterschieden. Es existieren unterschiedliche statistische Verfahren, um die Interrater-Reliabilität (IRR) zu bestimmen. Das am häufigsten verwendete statistische Maß zur Bewertung dieser Beurteiler- Übereinstimmung ist der Cohens-Kappa-Koeffizient κ (Fleiss und Plewis 1982). Bei mehr als zwei Beurteilern (Ratern) wird dieses als Fleiss‘ Kappa definiert (McHugh 2012; Steinijans et al. 1997). Die IRR wurde lange als prozentuale Übereinstimmung zwischen den Ratern mittels einer Matrix gemessen. Cohen stellte 1960 jedoch die Hypothese auf, dass es zudem ein Maß an zufälligen Übereinstimmungen gibt, welche in der Statistik mitberücksichtigt werden müssen. So wird für die genauere Interpretation der IRR eine Kombination der prozentualen Übereinstimmung und der Kappa-Werte empfohlen (Koch und Landis 1977). Die Interpretation dieser Werte erfolgt nach Landis und Koch und wird als Zahl zwischen 0 und 1 angegeben. Landis und Koch beschreiben Kappa-Werte im Bereich von $\kappa = 0$ bis 0,2 als „schwach“, im Bereich von $\kappa =$ 0,21 bis 0,4 als „ausreichend“, von $\kappa = 0{,}41$ bis 0,6 als „mittelgradig“, $\kappa = 0{,}61$ bis 0,8 als „substanziell“, sowie $\kappa = > 0{,}8$ als „nahezu perfekt“ (Coelho Neto et al., 2015; Kottner et al., 2011; Landis und Koch, 1977). Einen absolut perfekt reliablen Wert in einer Studie mit mehreren Teilnehmern und mehreren Ratern zu erreichen ist jedoch nahezu unmöglich. Dies würde bedeuten, bei wiederholter Testdurchführung auf exakt das gleiche Ergebnis zu kommen. Daher werden kleinere Fehleranteile akzeptiert und die Messung trotzdem

als reliable Messung bezeichnet. Je kleiner der Messfehler, desto höher ist also die Reliabilität. Liegen sehr große Fehleranteile vor, wird die Messung als nicht reliabel bezeichnet. Die Notwendigkeit der Berücksichtigung von absoluten und relativen Messfehlern in Reliabilitätsstudien wird in der Studie von Brauer et al. (2005) besonders hervorgehoben, da solche Analysen essenziell sind, um die Genauigkeit und Konsistenz von Messverfahren und Beurteilungen zu bewerten (Brauer et al., 2005).

Um ein umfassendes Bild der Beurteilerkonsistenz zu erhalten ist ergänzend zur IRR auch die Interrater-Übereinstimmung (*Interrater-Agreement*, IRA) als statistische Größe zu nennen. Die IRA ist enger definiert und misst die absolute Übereinstimmung zwischen unterschiedlichen Beurteilern als Prozentwert (LeBreton und Senter, 2008).

In der Ultraschalldiagnostik ist die Bewertung der IRR von besonderer Relevanz, da es sich um ein stark untersucherabhängiges Verfahren handelt. Während der dynamischen Untersuchung bedarf es eines guten anatomischen und räumlichen Vorstellungsvermögens des Untersuchers. Zwar existieren in den meisten medizinischen Bereichen definierte Standardschnitte, jedoch ist es praktisch unmöglich die exakte Schallkopfposition erneut zu treffen, um den identischen Bildausschnitt oder die Messgröße abzubilden. Diese Tatsache erschwert grundlegend die Vergleichbarkeit von Befunden. In einer Studie über die Schilddrüsenknoten-Volumenbestimmung wurde genau dies exemplarisch gezeigt. Es wurde nur eine mäßige Übereinstimmung zwischen den verschiedenen Untersuchern ermittelt (Hagmann et al., 2011). Mehrere Studien belegen, dass die IRR durch standardisierte Untersuchungsprotokolle sowie durch kontinuierliche sonographische Schulungen und zunehmende Erfahrung der Untersuchenden signifikant verbessert werden kann (Birrenbach et al., 2004). Dieser Zusammenhang gilt nicht nur für die Ultraschalldiagnostik, sondern wurde auch in anderen Bereichen der radiologischen Bildgebung nachgewiesen (European Society of Radiology (ESR), 2018; Sinitsyn et al., 2014; Sistrom und Honeyman-Buck, 2005). Ob die Verwendung strukturierter Befundvorlagen – über die reine Untersuchungsebene hinaus – auch auf der Ebene der Befunddokumentation zu einer messbaren Verbesserung der IRR führt, wurde bislang nicht systematisch untersucht. Diese Forschungslücke ist Ausgangspunkt der vorliegenden Arbeit.

1.1.2.3 Konventionelle Befunderhebung

Die konventionelle Befunderhebung wird in der Literatur auch als Freitextbefunderhebung (*Free Text Reporting*, FTR) bezeichnet. Je strukturierter und standardisierter Befunde dokumentiert werden, desto einfacher ist es, Vergleiche zu ziehen und Krankheitsverläufe zu beurteilen. Eine Patientenakte mit den wesentlichen Informationen, Behandlungsabläufen und Befundergebnissen zu führen, ist eine essenzielle Aufgabe und wird durch das Patientenrechtegesetz sowie die Berufsordnung und Bundesmantelverträge der Ärzte geregelt (§ 630f BGB, § 57 BMV-Ä, § 10 MBO-Ä). Zwar gab es in den letzten Jahren auch Fortschritte im Hinblick auf die Digitalisierung des Gesundheitswesens, diese sind jedoch im Vergleich zu anderen Gesellschaftsbereichen weder weit fortgeschritten noch umfassend standardisiert. Größtenteils werden weiterhin Freitextberichte in narrativer Sprache und von Hand geschrieben erstellt. Dies gilt insbesondere für radiologische Befunde sowie für Berichte über Ultraschalluntersuchungen (Forghani et al., 2015; Sabel et al., 2017a). Viele Fachbereiche verwenden seit Jahren abteilungsspezifische Standardformulare und Vorlagen zur Erstellung eines Freitextberichtes. Das Schreiben eines Freitextberichtes ermöglicht es durch Nutzung des individuellen Vokabulars und des eigenen Stils des Untersuchers einen Befund rasch und fallorientiert zu erstellen, welcher sich auf das Wesentliche zentriert. Viele Ärzte sind durch die jahrelange Dokumentation im Freitextformat auf diese Art der Befunderhebung geschult. Die Berichtsqualität ist auch hier jedoch abhängig vom Erfahrungsstand des Untersuchers. Gerade jungen Assistenzärzten fällt es zu Beginn schwer Befunde strukturiert und verständlich zu formulieren und die Befunde auf das Wesentliche zu beschränken (Ernst et al., 2019a; Sabel et al., 2017b; Schoeppe et al., 2018). Die unzureichende Lesbarkeit handschriftlicher Freitextberichte beeinträchtigt zudem häufig deren Vergleichbarkeit und Dokumentationsqualität. Dies kann die Zufriedenheit der überweisenden Ärzte mindern und führt nicht selten zu vermeidbaren Rückfragen (Szilveszter et al., 2017a). Das Fehlen einer klaren Struktur führt zu einer großen Variabilität des Berichtinhaltes. Uneinheitlich eingesetzte Terminologie sowie Dokumentationslücken können nachhaltig die Diagnose und die sich anschließende Therapie negativ beeinflussen (Dunnick und Langlotz, 2008; European Society of Radiology (ESR), 2011; Langlotz, 2006; Morgan et al., 2014; Schöppe et al., 2018; Tuncyurek et al., 2019). Um die Variabilität und Unklarheiten in

Freitextbefunden zu reduzieren, wurden standardisierte und strukturierte Befundvorlagen entwickelt, die in der klinischen Praxis zunehmend zur Anwendung kommen (Dunnick und Langlotz, 2008; Ernst et al., 2019a, 2019b; European Society of Radiology (ESR), 2018a; Morgan et al., 2014).

1.1.2.4 Strukturierte Befunderhebung

Die strukturierte Befunderhebung (*Structured Reporting*, SR) beschreibt die systematische Erstellung von Befundberichten unter Verwendung eines standardisierten Formats, einheitlicher Terminologie und festgelegter Inhalte. Weiss und Bolos (2009) beschrieben SR anhand von drei Ebenen (Weiss und Bolos, 2009):

1. Strukturiertes Format: Definition von Abschnitten und Unterabschnitten
2. Konsequenter Aufbau des Befundes: Festlegung und Reihenfolge der zu befundenden Strukturen
3. Konsequente Anwendung der vorgesehenen Terminologie (standardisierter sprachlicher Aufbau)

Diese drei Ebenen der SR wurden von Nobel et al. (2020) um zwei weitere Stufen ergänzt und damit zu einem fünfstufigen Modell erweitert (Nobel et al., 2020):

4. Strukturiertes Layout: Darstellung des Befundes in einer strikten, vordefinierten Reihenfolge, um eine Gleichförmigkeit der Befunde zur erreichen bzw. aufrechtzuerhalten
5. Strukturierter Inhalt: festgelegte Art und Weise, in der der medizinische Inhalt des Befundes IT-gestützt angeordnet und im Bericht angezeigt wird

Die SR bietet eine Reihe von Vorteilen und Möglichkeiten gegenüber der FTR und wird als Strategie zur Verbesserung der Befundung verschiedener diagnostischer Modalitäten befürwortet, da sie mit einer Qualitätsverbesserung von Untersuchungsberichten einher geht (Bauer et al., 2018; Hawkins et al., 2014; Weiss und Bolos, 2009). Im Gegensatz zu FTR die durch Variabilität in Sprache, Länge und Stil gekennzeichnet sind, verfolgt SR das Ziel, Inhalte durch standardisierte Formate, festgelegte Terminologie und definierte Berichtstrukturen zu vereinheitlichen (Ganeshan et al., 2018; Khorasani et al., 2003). In der Regel erfolgt sie IT-gestützt, wodurch Anwender bei der Erstellung des Befunds

effizient unterstützt werden (Wallis und McCoubrie, 2011). Ziel der Implementierung strukturierter Vorlagen ist es, den steigenden klinischen, ökonomischen und regulatorischen Anforderungen gerecht zu werden unter der Voraussetzung einer breiten Akzeptanz innerhalb der ärztlichen Fachkreise.

Trotz der großen Bedeutung und des hohen Stellenwertes im klinischen Alltag mangelt es in vielen medizinischen Fachabteilungen an einer ausreichenden Ausbildung in der Befunderhebung (Larson, 2018). Daher erhofft man sich durch die Etablierung solcher standardisierter Konzepte diesen Mangel langfristig zu beseitigen (Ernst et al., 2019a; European Society of Radiology (ESR), 2011; Sinitsyn et al., 2014; Sistrom und Honeyman-Buck, 2005). Viele Studien haben in den letzten Jahren die Überlegenheit der SR gegenüber der FTR gezeigt, da diese die allgemeine Befundqualität, Vollständigkeit, Genauigkeit und Klarheit von Berichten nachweislich verbessern (Ernst et al., 2019a, 2019b; Reiner, 2009).

Durch Strukturierung lassen sich Mehrdeutigkeiten und Missverständnisse in Berichten vermeiden, wodurch Klarheit und Genauigkeit verbessert wird. Auch sprachliche und grammatikalische Fehler treten seltener auf, was den Korrekturaufwand reduziert (Ernst et al., 2019b). Gerade für unerfahrene Untersucher und Nicht-Muttersprachler kann dies besonders vorteilhaft sein (Cho et al., 2017; Lacerda et al., 2014; Lee und Whitehead, 2017). Ein weiterer Vorteil besteht in der verbesserten interdisziplinären Kommunikation. Standardisierte Berichte mit einer festgelegten, einheitlichen Struktur sind für zuweisende Ärzte verschiedener Fachrichtungen leichter nachvollziehbar und fördern eine interdisziplinäre Kommunikation und kooperative Entscheidungsfindung (Dunnick und Langlotz, 2008; Johnson et al., 2009; Kahn et al., 2013; Prasad und Chen, 2018; Reiner, 2009; Schoeppe et al., 2018; Sluijter et al., 2016). Durch die Verwendung von Vorlagen, die definierte Datenelemente abfragen, wird der Untersucher systematisch durch den Untersuchungs- und Befunderhebungsprozess geführt. Dies kann dazu beitragen, auch relevante, aber nicht im Fokus stehende Strukturen zu beurteilen. Daher profitieren gerade unerfahrene Untersucher von dem Aufbau strukturierter Befunde. Demzufolge ist es möglich, vollständigere sowie differenziertere Befunde zu generieren (Timor-Tritsch und Goldstein, 2005). Zudem ermöglicht SR durch Homologisieren der diagnostischen Erhebung eine bessere Vergleichbarkeit von Krankheitsverläufen, unterstützt die

Therapieplanung und reduziert potenziell unnötige Folgeuntersuchungen, was sich positiv auf die Patientenzufriedenheit ausübt (Ernst et al., 2019a; Gassenmaier et al., 2017; Nörenberg et al., 2017; Park et al., 2019; Sabel et al., 2017b). Studien zeigen, dass strukturierte Befunde von Untersuchern und zuweisenden Ärzten gegenüber FTR bevorzugt werden, da sie vollständiger und übersichtlicher sind und besser mit klinischen Leitlinien übereinstimmen (Cho et al., 2017; Lacerda et al., 2014). Nicht zuletzt verbessert SR die Verfügbarkeit relevanter Informationen (z. B. Anamnese, Vorbefunde, Bilddaten etc.) und erleichtert durch die standardisierte Struktur den Vergleich von Untersuchungsergebnissen (Bosmans et al., 2012). Dank zunehmender Digitalisierung sind strukturierte Berichte zudem mit Such- und Kategorisierungsfunktionen kombinierbar und eignen sich für qualitativ hochwertige wissenschaftliche Analysen (Bosmans et al., 2012; European Society of Radiology (ESR), 2018; Reiner, 2009).

Trotz der genannten Vorteile wird SR nicht uneingeschränkt positiv bewertet. Ein häufig genannter Kritikpunkt betrifft die Rigidität standardisierter Berichtsformate (Johnson et al., 2009; Pinto Dos Santos und Baeßler, 2018; Sistrom und Honeyman-Buck, 2005). Vor allem bei komplexen diagnostischen Fragestellungen kann dies zum Problem werden, wenn nicht alle klinischen Besonderheiten adäquat abgebildet werden können. Mehrere Studien zeigen zudem, dass FTR eine vergleichbare Qualität aufweisen können (European Society of Radiology (ESR), 2018). Kritisch angemerkt werden auch sprachliche Einschränkungen sowie eine zum Teil unzureichende Terminologie in frühen SR-Systemen. Diese Schwächen lassen sich jedoch durch sorgfältige Planung, fachspezifische Anpassung und interdisziplinäre Zusammenarbeit beheben (Sluijter et al., 2016). Ein weiterer Nachteil ist der initiale Mehraufwand bei der Einführung der SR Software. Besonders erfahrene Untersucher empfinden den Wechsel von etablierten Freitextsystemen zu SR als umständlich, was sich in einer verlängerten Befundungszeit und einem damit verbundenen Lernprozess äußert (Johnson et al., 2009; Sluijter et al., 2016). Studien belegen jedoch, dass sich dieser Zeitverlust nach einer Einarbeitungsphase ausgleicht und langfristig sogar effizientere Arbeitsabläufe möglich sind (Sickles et al., 2013).

Die Etablierung strukturierter Befunde erfordert eine enge interdisziplinäre Abstimmung, was herausfordernd sein kann. Damit alle essenziellen Informationen im Bericht enthalten sind und eine flächendeckende Vergleichbarkeit gewährleistet werden kann, bedarf es klar definierter Gliederungen, einheitlicher Textbausteine und Vokabularien (Austin et al., 1996; Fardon und Milette, 2001). Eine erfolgreiche Umsetzung setzt idealerweise die Zusammenarbeit eines fachübergreifenden Expertengremiums voraus, um dieses Vorhaben in die Tat umzusetzen und SR auf institutioneller Ebene erfolgreich zu implementieren (Al-Hawary et al., 2014; Larson et al., 2013). Jedoch hat sich die Etablierung der SR nicht in allen Fachbereichen als einfach erwiesen. Gerade unter Radiologen zeigt sich eine gewisse Uneinigkeit hinsichtlich der Akzeptanz von SR (Bosmans et al., 2012; Faggioni et al., 2017; Powell und Silberzweig, 2015). Kritisiert werden insbesondere komplexe Vorlagen mit zahlreichen Menüs und Unterkategorien, die den Fokus vom zu befundenen Bild auf die Befundvorlage verschieben können und sich somit negativ auf die Befundqualität auswirken (Ganeshan et al., 2018; Heikkinen et al., 2000; Johnson et al., 2009; Krupinski und Lund, 1997). Das sogenannte „*eye dwell*"-Phänomen beschreibt die potenzielle Gefahr, dass radiologische Auffälligkeiten übersehen werden, weil es zu einer Unterbrechung des visuellen Suchvorgangs kommt und die Aufmerksamkeit zu sehr auf die Befundvorlage gelenkt wird (Babu und Brooks, 2015) Die oben hervorgehobenen Herausforderungen in der Befundung beziehen sich vor allem auf hochkomplexe Bildgebungsmodalitäten mit sehr großem Beurteilungsspektrum und Beschreibung feiner Details (z. B. Computertomographie (CT) und Magnetresonanztomographie (MRT)), weniger jedoch auf standardisierte Untersuchungen wie Röntgen oder Sonographie (Eghtedari et al., 2020; Ghoshhajra et al., 2013; Gunderman und McNeive, 2014; Karim et al., 2013; Margolies et al., 2016).

Die Entwicklung und Implementierung von SR ist zudem mit einem erheblichen technischen und organisatorischen Aufwand verbunden. Bestehende Workflows müssen angepasst, neue Softwarelösungen in das bestehende IT-Netzwerk integriert und Mitarbeitende geschult werden (Ernst et al., 2019a; Shore et al., 2012). Die Zusammenarbeit mit der EDV-Abteilung sowie die Einigung auf einheitliche Terminologie und Formatierung erfordern Zeit und wie bereits erwähnt eine fachübergreifende Abstimmung. Ohne die angemessene Einbindung von Zuweisern und verschiedenen Fachgesellschaften besteht das Risiko, dass strukturierte Befunde wesentliche klinische

Fragestellungen unzureichend berücksichtigen und dadurch an Verständlichkeit und Akzeptanz im klinischen Alltag verlieren (Gunn et al., 2013). Neben diesen inhaltlichen und organisatorischen Herausforderungen ist auch der finanzielle Aspekt nicht zu vernachlässigen. Die Neuanschaffung entsprechender Softwarelösungen, deren Integration in das bestehende Netzwerk sowie die damit verbundenen Schulungs- und Wartungsarbeiten sind stets mit einem erheblichen zusätzlichen Kostenaufwand verbunden. Diese Aspekte machen eine sorgfältige Abwägung der Effizienz und Wirtschaftlichkeit eines neuen Systems unerlässlich (Cramer et al., 2014; European Society of Radiology (ESR), 2023; Larson, 2018).

Mit dem Aufkommen leistungsfähiger moderner Technologien eröffnen sich neue Möglichkeiten in der medizinischen Textverarbeitung. Durch den Fortschritt im Bereich Künstlicher Intelligenz (KI) und die Entwicklung von Large Language Models (LLM) eröffnen sich im klinischen Alltag deutlich mehr Möglichkeiten zur automatisierten Strukturierung und Qualitätsverbesserung von Befunden (European Society of Radiology (ESR), 2023; Mallio et al., 2023a; Miao et al., 2024). Der Einsatz dieser Technologien bringt jedoch hohe Anforderungen an Datenschutz, Datensicherheit und medizinrechtliche Zulassungen mit sich und ersetzt nicht das ärztliche Urteil, weshalb auch hier eine sorgfältige Prüfung hinsichtlich Wirtschaftlichkeit und Sicherheit erforderlich ist (Bizzo et al., 2023a; European Society of Radiology (ESR), 2023; Mallio et al., 2023a; Miao et al., 2024).

Die folgenden Abschnitte geben einen Überblick über medizinische Fachbereiche, in denen SR bereits erfolgreich implementiert wurde und deren Erfahrungen als Grundlage für die Einführung in der HNO-Sonographie dienen können.

1.1.2.4.1.1 Mammographie - der Grundstein der strukturierten Befunderhebung

Das Breast Imaging Reporting and Data System (BI-RADS) wurde in den 1980er Jahren durch das American College of Radiology (ACR) entwickelt, um die Befundberichterstattung in der Mammographie zu vereinheitlichen und unklare

Formulierungen sowie uneinheitliche Empfehlungen zu beseitigen (D'Orsi und Kopans, 1997; Scott, 1989).

Kernbestandteil ist ein von einem interdisziplinären Expertengremium entwickeltes Lexikon standardisierter Beschreibungsbegriffe für Bildgebungsmerkmale (Sickles et al., 2013). BI-RADS hat sich von einem Kommunikations- und Datenverfolgungssystem zu einer umfassenden Berichtsplattform und einem zentralen Instrument der Qualitätssicherung für Mammographie, MRT und Sonographie entwickelt (Eghtedari et al., 2020). Es gilt heute als Referenzmodell für die erfolgreiche Implementierung strukturierter Befundung und zeigt, dass kontrollierte Terminologie und standardisierte Berichtsformate die Leitlinienadhärenz fördern sowie Vollständigkeit, Genauigkeit und Konsistenz der Befunderstattung verbessern (Baker et al., 1995; Bialek et al., 2006; Forghani et al., 2015; Künzel et al., 2017; Moshtaghi et al., 2017; Prasad und Chen, 2018).

1.1.2.4.1.2 Strukturierte Befunderhebung in der Sonographie

Vor allem klinische Untersuchungen, welche einem vorgegebenen Arbeitsablauf folgen, profitieren von einer strukturierten Vorgehensweise bei der Befunderhebung. In Diagnostik und Nachsorge hat sich die Sonographie in den letzten Jahren zum Goldstandard zur Beurteuilung der Halsweichteile in der HNO-Diagnostik etabliert und spielt eine große Rolle beispielsweise in der Abklärung von Speicheldrüsenerkrankungen, Lymphadenopathien und Malignomen (Schoeppe et al., 2018). Die Erstellung von detailgenauen und vergleichbaren Befundberichten ist elementar und bildet die Grundlage für Folgeuntersuchungen und klinische Entscheidungen (Lin et al., 2014; Takes et al., 2020; Tuncyurek et al., 2019; Wiegand und Dietz, 2020). Strukturierte Berichte von Kopf-Hals-Sonographien weisen im Vergleich zu konventionellen Freitextberichten eine verbesserte Berichtqualität, Vollständigkeit und Detailliertheit auf (Ekberg und Olsson, 1997; Ernst et al., 2019a; Gates et al., 2006; Ott und Pikna, 1993). Wichtige Strukturen wurden in den Untersuchungen durch Nutzung des standardisierten Formats und der einheitlichen Terminologie weniger übersehen und detaillierter beschrieben. Auch in anderen Bereichen wie beispielsweise der gynäkologischen Sonographie konnten eben diese Vorteile der SR nachgewiesen werden (Günter, 2013; Szilveszter et al., 2017b).

1.1.2.4.1.3 Strukturierte Befunderhebung in anderen bildgebenden Verfahren

SR hat sich auch in anderen bildgebenden Verfahren als qualitätssteigerndes Instrument bewährt. In der MRT-Diagnostik konnte gezeigt werden, dass strukturierte Berichte die Gesamtqualität verbessern, mit einer höheren Nutzerzufriedenheit und effizienteren Arbeitsabläufen einhergehen (Weiss und Langlotz, 2008). Besonders deutlich wird das bei der MRT-Stadieneinteilung des Rektumkarzinoms, bei der eine Vielzahl von Daten ermittelt und dokumentiert werden muss. Wichtige Behandlungsentscheidungen, wie die Entscheidung zur neo- bzw. adjuvanten Chemotherapie oder einer Operation, werden anhand dieses *Stagings* getroffen (Nörenberg et al., 2017). Es konnte gezeigt werden, dass strukturierte MRT-Berichte die Sicherheit in der klinischen Entscheidungsfindung erhöhen und die interdisziplinäre Zusammenarbeit, insbesondere mit den chirurgischen Kollegen, erleichtern (Gassenmaier et al., 2017; Schöppe et al., 2018). Gleiches gilt für den Bereich der muskuloskelettalen Bildgebung, insbesondere bei Schultererkrankungen, wo strukturierte, klar verständliche Berichte Folgekonsultationen reduzieren und die Operationsplanung gezielt unterstützen (Bosmans et al., 2011, 2012; Plumb et al., 2009; Sabel et al., 2017).

Auch für CT-Untersuchungen liegen überzeugende Belege vor. Bei präoperativen Nebenhöhlen-CTs vor Funktionell Endoskopischer Sinuschirurgie (FESS) wurde nachgewiesen, dass Radiologen und HNO-Ärzte unterschiedliche Schwerpunkte bei der Beurteilung setzten und insbesondere die korrekte Identifikation planungsrelevanter Strukturen eine häufige Fehlerquelle darstellt (Braun et al., 2012; Mistry et al., 2016; Reiner et al., 2007; Vaid et al., 2011). SR kann hier zu einer signifikant höheren Befundvollständigkeit im Vergleich zu konventionellen Berichten und konventioneller Operationsplanung beitragen (Ernst et al., 2020a). Studien zeigen zudem, dass die Kombination von SR und strukturierter Operationsplanung die Arbeitsweise von Radiologen und HNO-Ärzten verbessert und gleichzeitig potenzielle Risiken während endoskopischer Eingriffe reduziert (Ernst et al., 2019a, 2019b, 2020b; Sabel et al., 2017; Schöppe et al., 2018).

Für die CT-Pulmonalangiographie bei akuter Lungenembolie sowie für die CT-Angiographie der unteren Extremitäten konnte ebenfalls ein positiver Effekt der SR auf Vollständigkeit, Verständlichkeit und klinischen Nutzen der Berichte nachgewiesen

werden; überweisende Ärzte bewerteten strukturierte Berichte als klarer und nachvollziehbarer (Sabel et al., 2017, 2018). Der Vorteil eines standardisierten detaillierten, tabellarischen Berichtes, welcher die Informationsextraktion vereinfacht, wird vor allem von überweisenden Ärzten bevorzugt und unterstreicht den Mehrwert für Kommunikation und Entscheidungsfindung im klinischen Alltag. Diese Ergebnisse stimmen mit vorherigen Publikationsergebnissen überein (Larson et al., 2013; Schwartz et al., 2011).

Darüber hinaus wurden Vorteile der SR auch für spezielle Untersuchungstechniken belegt. Für die videofluoroskopische Schluckstudie (VFSS) zeigte sich im Vergleich zu narrativen Berichten eine höhere Vollständigkeit und eine größere Spezifität im Hinblick auf die jeweilige klinische Fragestellung, was mit einer gesteigerten Gesamtzufriedenheit der betreuenden HNO-Ärzte und Logopäden einherging (Martin-Harris et al., 2008; Martin-Harris und Jones, 2008; Schoeppe et al., 2018).

1.1.3 Befundqualität und Befundvollständigkeit

Ein medizinischer Befund dokumentiert das Ergebnis aller diagnostischen Maßnahmen und bildet die Grundlage für weitere therapeutische Entscheidungen. Die Dokumentation der Anamnese, Diagnostik und Therapie ist rechtlich verbindlich geregelt (§ 630f BGB, § 57 BMV-Ä, § 10 MBO-Ä). Die medizinische Dokumentation dient ferner der Entscheidungshilfe, der haftungsrechtlichen Absicherung und der externen Plausibilitätsprüfung, etwa durch den Medizinischen Dienst. Eine vollständige und verständlich formulierte Befunddokumentation ist daher nicht nur im Interesse des behandelnden Arztes, sondern vor allem im Interesse des Patienten und weiterbetreuender Kollegen (Sistrom et al., 2004).

Eine unzureichende Berichterstellung kann die Diagnosefindung, die korrekte Klassifikation von Befunden sowie die anschließenden Therapieentscheidungen beeinträchtigen (Enders et al., 2011b; Lima et al., 2019). Trotz dieser zentralen Bedeutung wird dem Erlernen strukturierter Befunderhebung in der ärztlichen Aus- und Weiterbildung ein untergeordneter Stellenwert zugeordnet (Beach et al., 2010; Coltrera, 2014; Forghani et al., 2015; Künzel et al., 2017; Moshtaghi et al., 2017). Meist erlernen die Weiterbildungsassistenten die Interpretation und Befundung erst während der Ausübung

ihrer ärztlichen Tätigkeit auf eigene Faust oder durch Beobachtung erfahrener Kollegen oder Oberärzte, selten jedoch durch strukturierte Anleitung oder standardisierte Vorgaben. Ferner fehlt es im klinischen Alltag oft an der notwendigen Zeit für beaufsichtigtes Arbeiten und persönliche Anleitung, was im Widerspruch zur Bedeutung des praktischen Erlernens und der Befunderhebung steht.

1.1.3.1 Implementierung strukturierten Befunderhebung in der Aus-und Weiterbildung

Die frühzeitige Implementierung von SR im Rahmen der ärztlichen Ausbildung kann diesem Problem effektiv entgegenwirken und Vollständigkeit, Qualität und Vergleichbarkeit von Befunden signifikant verbessern (Ernst et al., 2020c). Besonders unerfahrene Anwender profitieren dabei von der klar strukturierten Vorlage. Sie erleichtert nicht nur die Befunderstellung, sondern führt Schritt für Schritt durch die Untersuchung und orientiert sich dabei an medizinischen Leitlinien und klinischen Standards. Diese schablonenartige Struktur fördert das Verständnis und die Sicherheit bei der Durchführung und zeigt nachweislich einen besonders ausgeprägten Lerneffekt bei Medizinstudierenden und Berufsanfängern (Ernst et al., 2019b). Die Studie von Ernst et al. (2020) untersucht den optimalen Zeitpunkt für die Implementierung von SR während der ärztlichen Ausbildung und zeigten, dass der größte Mehrwert durch eine frühzeitige Einführung bereits im Medizinstudium erzielt wird (Ernst et al., 2020c). Sowohl Studierende als auch ärztliche Weiterbildungsassistenten unterschiedlicher Erfahrungsstufen bewerteten die Nutzung strukturierter Befunde als hilfreich und qualitätssteigernd. Neben einer erhöhten Befundqualität konnte eine signifikante Reduktion übersehener Pathologien sowie eine Verbesserung der diagnostischen Sicherheit und Zeiteffizienz festgestellt werden (Ernst et al., 2019a, 2019b, 2020b; Sabel et al., 2017). Bemerkenswert ist, dass selbst erfahrene Ärzte von der Einführung strukturierter Vorlagen profitierten. Der Effekt der verbesserten Berichtsqualität hielt nämlich bis zur Gruppe der Ärzte mit umfangreicher sonographischer Erfahrung an, obwohl diese in der Erstellung von Freitextberichten über Jahre geschult waren. Dies weist darauf hin, dass nicht das Erfahrungsniveau, sondern das Format des Befundes selbst entscheidend für die Berichtsqualität ist (Ernst et al., 2020c).

Ein kontinuierlicher Einsatz in der Facharztweiterbildung kann den Lerneffekt nachhaltig verstärken und langfristig zu einer höheren diagnostischen Qualität in der klinischen Praxis führen (Ernst et al., 2019b, 2020b).

Hinsichtlich der Befundungszeit belegen mehrere Studien, dass die Nutzung der SR im Vergleich zur FTR mit einer Verkürzung der Befundungszeit einhergeht (Gunderman und McNeive, 2014). Dieser Effekt ist besonders bei unerfahrenen Assistenzärzten sowie Berufseinsteigern und Medizinstudenten zu beobachten, welche vom schablonenartigen Aufbau der SR und einem strukturierterem Arbeitsablauf deutlich profitieren (Ganeshan et al., 2018; Powell und Silberzweig, 2015; Sabel et al., 2017b). Durch Vorgabe der inhaltlichen Struktur verringert sich der Aufwand für inhaltliche Gliederung, Terminologie und Formulierung, was zu einer schnelleren Fertigstellung der Befunde und weniger Rückfragen an andere Ärzte führt (Ganeshan et al., 2018; Powell und Silberzweig, 2015; Sabel et al., 2017b). Zwar kann die initiale Einführung von SR – insbesondere bei erfahrenen Ärzten mit langjähriger Erfahrung in der Freitextbefundung – mit einem höheren Zeitbedarf verbunden sein, dieser Effekt relativiert sich jedoch mit zunehmender Vertrautheit und Routine im Umgang mit dem neuen Befundformat (Matsumoto et al., 2015). Die zeitliche Effizienz ist ferner abhängig von der Komplexität des Befundinhalts da die Darstellung komplexer Pathologien einen zusätzlichen Aufwand, insbesondere durch die Integration von Freitextelementen in die standardisierte Vorlage mit sich bringt (Ernst et al., 2019a, 2019b, 2020b; Sabel et al., 2017). Dieser Mehraufwand führt jedoch zu einer höheren Vollständigkeit und damit zu einer besseren Gesamtqualität und Aussagekraft der Befundberichte.

1.2 Fragestellung

Die Interrater-Reliabilität (IRR) gilt als relevanter Qualitätsparameter in der sonographischen Diagnostik und ist aufgrund der ausgeprägten Untersucherabhängigkeit des Verfahrens ein häufig kritisierter Schwachpunkt. Bisherige Studien konnten zeigen, dass die strukturierte Befunderhebung (SR) gegenüber der konventionellen Freitextbefundung (FTR) die Genauigkeit, Vollständigkeit und Detailliertheit von Befunden in der Kopf-Hals-Sonographie verbessert und von Anwendern klar bevorzugt wird (Ernst et al., 2019a, 2019b, 2020b, 2020c). Ob SR darüber hinaus auch die Übereinstimmung

zwischen unabhängigen Untersuchern – also die IRR – verbessert, wurde bislang nicht systematisch untersucht.

Die vorliegende Dissertationsarbeit soll diese Forschungslücke schließen. Ziel war es, im Rahmen einer prospektiven multizentrischen Studie an Kopf-Hals-Ultraschallzentren mit DEGUM-Kursleitern den Einfluss der SR auf video- und bilddokumentierte Ultraschalluntersuchungen anhand folgender drei Zielgrößen zu untersuchen:

1. Befundvollständigkeit
2. Benutzerzufriedenheit
3. Interrater-Reliabilität (IRR)

2. Material und Methoden

2.1 Studiendesign

Im Rahmen der vorliegenden Arbeit wurde eine prospektive, multizentrische Studie durchgeführt, um den Einfluss der SR auf Befundvollständigkeit, Benutzerzufriedenheit und IRR in der Kopf-Hals-Sonographie zu untersuchen. Hierzu wurden n=4 komplette Ultraschalluntersuchungen von typischen Kopf- und Halspathologien durchgeführt. Diese wurden anonymisiert als detaillierte Video- bzw. Bilddateien aufgezeichnet. Die vier Fälle umfassten eine Nachsorge eines Parotiskarzinoms (Fall1, Abb. 1 (a)), eine Untersuchung einer unklaren Raumforderung in der Parotis (Fall 2, Abb. 1 (b)), eine Untersuchung bei Verdacht auf zervikale Lymphknotenmetastasen (Fall 3, Abb. 1 (c)) sowie eine weitere Untersuchung einer suspekten Parotisraumforderung (Fall 4, Abb. 1 (d)).

Von diesen Untersuchungen erstellten zwei unabhängige DEGUM-Kursleiter strukturierte Befunde mithilfe der online-basierten SR-Vorlage, die als Referenzberichte dienten. Anschließend wurden die Video- und Bilddateien sowie Anweisungen zur Nutzung der Befundvorlage an n=9 Universitätskliniken in Deutschland übermittelt, deren Ultraschallabteilungen von DEGUM-Kursleitern geleitet werden. Zu Übungszwecken erfolgte zudem eine zusätzliche Untersuchung an einem gesunden Probanden ohne pathologische Befunde. Diese diente den Teilnehmern dazu, sich mit der SR-Vorlage vertraut zu machen und sollte als erstes ausgefüllt werden.

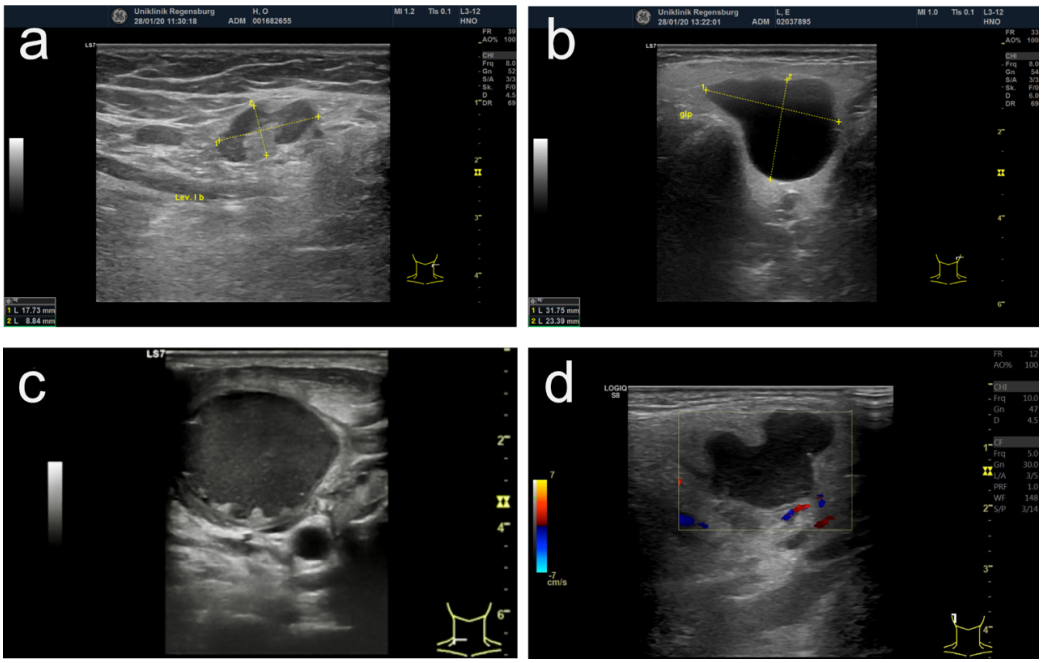


Abb. 1: Bilddateien der n=4 durchgeführten sonographischen Untersuchungen mit Darstellung typischer Kopf- und Halspathologien. (a) Fall 1: Nachuntersuchung bei Parotiskarzinom (b) Fall 2: Unklaren Raumforderung in der Parotis (c) Fall 3: Verdacht auf zervikale Lymphknotenmetastasen (d) Fall 4: Suspekte Parotisraumforderung

Bei den 9 teilnehmenden Kliniken handelte es sich um die

- Klinik und Poliklinik für Hals-Nasen-Ohren-Heilkunde, Universitätsklinikum Bonn
- Klinik und Poliklinik für Hals-Nasen-Ohren-Heilkunde, Universitätsklinikum Regensburg
- Klinik für Hals-, Nasen-, Ohrenheilkunde, Phoniatrie und Pädaudiologie, Universitätsklinik Aachen
- Klinik für Mund-, Kiefer- und Plastische Gesichtschirurgie, Universitätsklinikum Düsseldorf
- Hals-Nasen-Ohren-Klinik, Kopf- und Halschirurgie, Universitätsklinikum Erlangen
- Klinik für Hals-, Nasen-, Ohrenheilkunde, Universitätsklinikum Freiburg
- Klinik für Hals-, Nasen-, Ohrenheilkunde, Universitätsklinikum des Saarlandes
- Klinik für Hals-Nasen-Ohrenheilkunde, Kopf- und Halschirurgie, Universitätsmedizin Mannheim
- Klinik und Poliklinik für Hals-, Nasen-, Ohrenheilkunde, Klinikum rechts der Isar der Technischen Universität München

Jede der vier Untersuchungen wurde durch je einen Arzt pro Klinik mithilfe der SR-Vorlage befundet, woraus n=36 strukturierte Befunde resultierten. Alle 36 anonymisierten Berichte wurden anschließend durch zwei zertifizierte DEGUM-Kursleiter auf Vollständigkeit und IRR bewertet. Frühere Studien zeigen, dass SR der konventionellen Befundung mit Freitextberichten in Bezug auf Berichtqualität überlegen ist (Ernst et al., 2019a, 2019b, 2020b, 2020c). Da ausreichende Daten zur IRR der FTR vorliegen und diese für das Studienziel nur begrenzt relevant sind, wurde keine Kontrollgruppe mit FTR einbezogen.

Die Evaluation der Benutzerzufriedenheit der SR-Vorlage erfolgte zusätzlich anhand eines eigens entwickelten Fragebogens mit einer visuellen Analogskale (VAS) (Abb. 4) (Ernst et al., 2019a, 2019b, 2020b, 2020c). Alle Ergebnisse wurden schriftlich und digital dokumentiert und anschließend statistisch ausgewertet.

2.2 Aufklärung und Einverständnis der Patienten

Die n=4 Patienten, deren Ultraschalluntersuchungen für diese Studie anonymisiert per Video- bzw. Bild-Dokumentation festgehalten wurden, wurden von einem Studienarzt ausführlich über das Forschungsvorhaben informiert und hatten die Möglichkeit, offene Fragen zu klären. Neben dem wissenschaftlichen Ziel dient diese Studie unmittelbar dem Gesundheitsinteresse des Patienten sowie der Verbesserung der Befundqualität von Kopf-Hals-Sonographien. Die Option eines Teilnahmewiderrufs und ebenso die ärztliche Schweigepflicht wurden im Rahmen der Studienaufklärung erläutert. Bei den teilnehmenden Patienten ging neben der ausführlichen mündlichen Aufklärung ebenfalls eine schriftliche voraus. Ein unterschriebener Aufklärungsbogen lag von jedem Probanden vor. Diese Studie wurde als Teil eines mehrteiligen Studienvorhabens von der Ethikkommission des Fachbereichs Medizin der Johannes-Gutenberg-Universität Mainz unter dem Vorsitz von Frau Dr. Laufs geprüft, als ethisch unbedenklich deklariert und in der Folge genehmigt (Referenznummer: 2018-13255). Eine entsprechende Erweiterung des o.g. Ethikvotums für diese explizite Studie erfolgte formlos durch Prof. Dr. Wessler.

Alle Verfahren, die in Studien mit menschlichen Teilnehmern durchgeführt wurden, entsprachen den ethischen Standards des institutionellen und nationalen Forschungsausschusses sowie der Helsinki-Erklärung von 1964 und ihren späteren Änderungen oder vergleichbaren ethischen Standards.

2.3 Patientenkollektiv

Die an der Studie teilnehmenden Patienten (n=4) wurden im Rahmen geplanter Ultraschalluntersuchungen in der Hals-, Nasen-, Ohrenklinik und Poliklinik des Universitätsklinikums Mainz vorstellig. Die Rekrutierung der Patienten erfolgte ausschließlich im ambulanten Setting. Die Aufnahme in die Studie erfolgte über die Tumornachsorgesprechstunde, die allgemeinen HNO-Sprechstunden und die Speicheldrüsensprechstunde. Notwendigkeit der Vorstellung waren zum einen Routinekontrollen nach Kopf-Hals-Malignom oder aber die Abklärung einer zervikalen Lymphadenopathie bzw. Speicheldrüsenerkrankung. Es erforderte keine Änderungen der routinemäßigen Vorstellung in den Kliniken oder zusätzliche Kontrollen des Gesundheitszustandes des Patienten. Folgende Ein- und Ausschlusskriterien wurden festgelegt:

Einschlusskriterien:

- Alter > 18 Jahre
- Unterschriebene Einverständniserklärung
- Notwendigkeit einer Ultraschalluntersuchung der Halsweichteile im Rahmen
    - der Tumornachsorge nach Kopf-Hals-Malignom
    - der Abklärung einer zervikalen Lymphadenopathie
    - der Abklärung einer Speicheldrüsenerkrankung

Ausschlusskriterien:

- Alter < 18 Jahre
- Patient nicht einwilligungsfähig
- Bestehender Zweifel an der Einsichtsfähigkeit oder fehlende Einsichtsfähigkeit
- Keine klinische Notwendigkeit einer Sonographie der Halsweichteile
- Kein Zustand nach Kopf-Hals-Malignom
- Keine zervikale Lymphadenopathie
- Keine Speicheldrüsenerkrankung

2.4 Untersuchungsablauf

Die Befunde der Ultraschalluntersuchungen typischer Kopf- und Halspathologien wurden auf Bild- bzw. Videodateien festgehalten und an n=9 Kliniken (s.o.) zur Befundung mittels SR gesandt. Analysiert wurde jeweils die Vollständigkeit, sowie die IRR der Befunde. Die Benutzerzufriedenheit wurde mittels eines bestehenden Fragebogens erfasst und ausgewertet (Ernst et al., 2019a, 2019b, 2020b, 2020c).

2.4.1 Ultraschallgerät

Die Ultraschalluntersuchungen erfolgten im Rahmen klinischer Kontrollen in den jeweiligen Sprechstunden der HNO-Ambulanz des Universitätsklinikums Mainz. Alle Bilder wurden mit einem hochauflösenden Ultraschallgerät LOGIQ S8 (GE Healthcare, Solingen, Deutschland) aufgenommen. Als Ultraschallsonde wurde eine 9 MHz Linearsonde verwendet. Die Bilddokumentation beinhaltete die Darstellung der Pathologien in 2 Ebenen (transversal, sagittal) inklusive Maßangaben sowie eine duplexsonographische Darstellung. Es erfolgte eine Videodokumentation der Submandibular- und Parotislogen (rechts/links), des Mundbodens und der Gefäßnervenscheiden (rechts/links). Daraus konnte den befundenden Ärzten der gesamte Untersuchungsablauf zugänglich gemacht werden. Alle generierten Bilder und Videos wurden direkt nach der Anfertigung anonymisiert in einen *Cloud*-Speicher hochgeladen und zur Befundung freigegeben, wodurch eine dauerhafte digitale Archivierung ermöglicht wird.

2.4.2 Qualifikation des Untersuchers

Die Kopf-Hals-Sonographien wurden von zwei unabhängigen Experten (Facharzt für HNO mit DEGUM Stufe III/Kursleiter Kopf-Hals, aktuelles GCP-Zertifikat) durchgeführt und eine SR-Befundvorlage im Konsens erstellt.

Die teilnehmenden Ärzte waren acht Fachärzte der HNO sowie ein Facharzt der MKG. Diese neun an der Studie teilnehmenden Ärzte besitzen alle eine hohe Expertise in der Kopf-Hals-Sonographie und sind mindestens DEGUM Stufe II qualifiziert.

2.4.3 Verwendete Software

Die webbasierte Software der Firma Smart Reporting (Smart Reporting GmbH, Gewürzmühlstraße 11, 80538 München, Deutschland, http://www.smart-reporting.com) wurde verwendet, um eine spezifische Vorlage für die strukturierte Berichterstattung von Kopf-Hals-Ultraschalluntersuchungen zu entwerfen. Der Einsatz der Software in anderen medizinischen Fachbereichen (z.B. Videofluoroskopie, MRT, CT) wurde bereits erfolgreich untersucht und publiziert (Gassenmaier et al., 2017; Nörenberg et al., 2017; Sabel et al., 2017b; Schoeppe et al., 2018). In Zusammenarbeit mit Experten der Radiologie und HNO wurde diese Vorlage konzipiert. Der verwendete medizinische und linguistische Inhalt basiert auf den neuesten Empfehlungen der DEGUM zu Terminologie und Befundstrukturen. Sie enthalten die aktuellen Leitlinien und Klassifikationen. Die Untersucher werden durch anklickbare Entscheidungsbäume geführt und arbeiten sich systematisch durch die einzelnen Kategorien. Diese Kategorien setzten sich wiederum aus „Klinische Angaben“, „Voruntersuchungen“, „Prozedur“, „Befund“ und „Beurteilung“ zusammen. Die Anatomie sowie alle gängigen Kopf- und Halspathologien werden in Unterpunkten abgefragt (Abb. 2). Eine weitere Spezialisierung ist durch Erweiterung von anklickbaren Menüs möglich. Die Software generiert aus zuvor definierten Textphrasen in Echtzeit vollständige semantische Sätze, die übernommen und individuell angepasst werden können. Jede Befundung mit diesem Programm ist nach gleicher Struktur aufgebaut. Um jedoch ein höheres Maß an Flexibilität und Vollständigkeit zu ermöglichen, können zusätzliche freie Textelemente durch den Untersucher hinzugefügt werden. Weiterhin ermöglicht das Format den Zugriff auf eine implementierte Spracherkennung und -steuerung. Die Verknüpfung mit bereits bestehenden klinischen Netzwerksystemen sowie die Integration einer automatisierten Bildauswertung in die Software ermöglicht eine effiziente Befunderhebung.

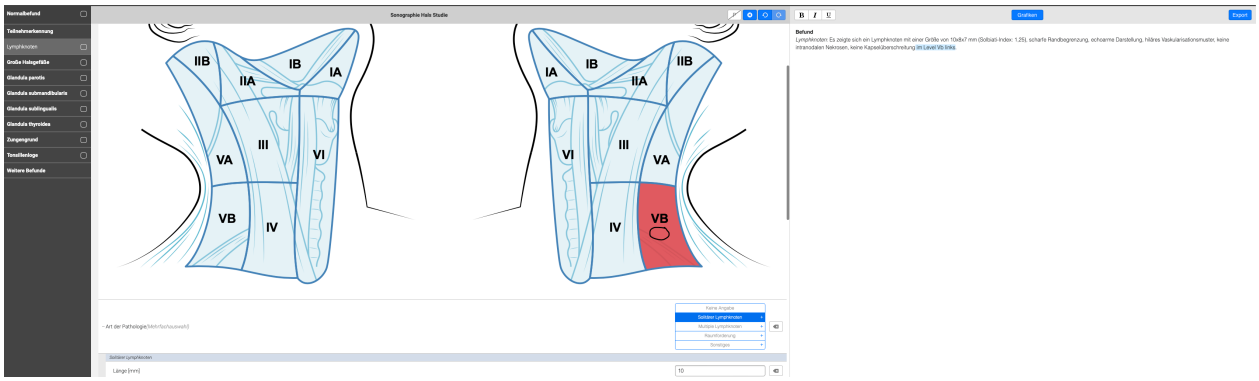


Abb. 2: Befundung eines solitären Lymphknotens mit 10x8x7 mm im Level Vb links mittels strukturierter Befunderhebung

2.4.4 Verwendete Formulare

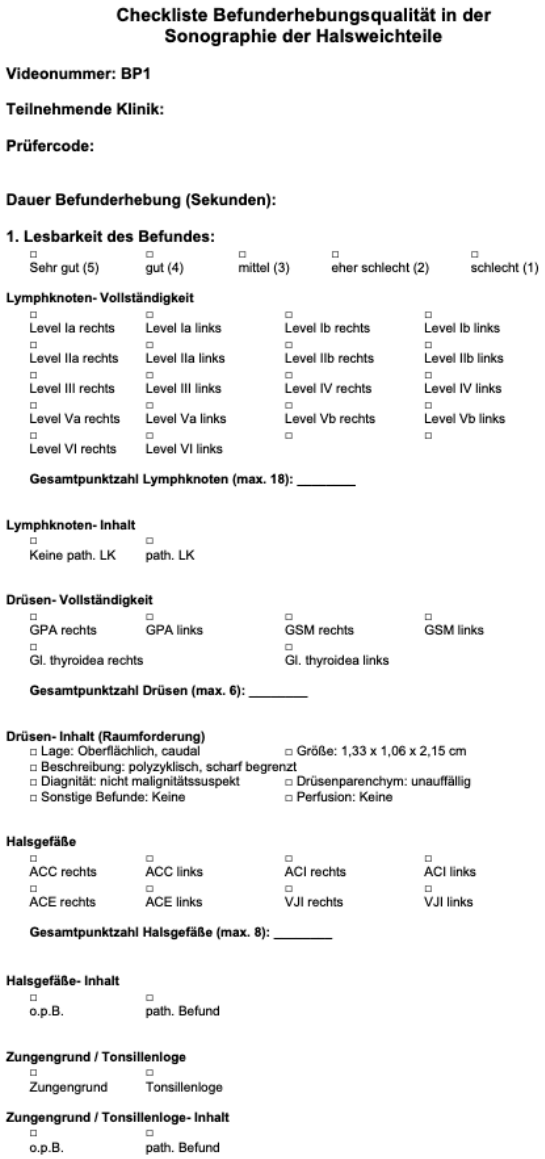
Zur Auswertung der Befunderhebungsqualität in der Sonographie der Halsweichteile kam ein eigens konzipierter Befundungsbogen zum Einsatz, welcher neben den studienrelevanten Feldern jedoch einem Standardfragebogen für Kopf-Hals-Sonographie entspricht (Abb. 3). Zu Beginn werden Formalitäten wie Name der teilnehmenden Klinik und der Prüfercode erfragt. Anschließend wird die Dauer der Befundung in Sekunden dokumentiert. Es folgt die Evaluation der Lesbarkeit des Berichtes. Hier können Punkte von 5 = „sehr gut“ bis 1 = „schlecht“ vergeben werden. Der sich anschließende Absatz befasst sich mit den obligat zu beschreibenden anatomischen Strukturen. Die Vollständigkeit von Lymphknoten, Drüsen und Halsgefäßen sowie die Erfassung des Zungengrundes und der Tonsillenloge wird durch eine Punktevergabe erfasst. Es wird 1 Punkt je Erwähnung einer Struktur vergeben.

Die Einteilung der zervikalen Lymphknoten erfolgt nach Robbins in sechs Regionen (Level I-VI) auf beiden Seiten (Robbins, 1999). Die Unterteilung der Drüsen beinhaltet die *Glandulae submandibulares*, die *Glandulae parotideae* und die *Glandulae thyroideae* beidseits. Die Halsgefäße sind unterteilt in die *Arteriae carotides communes*, *Arteriae carotides internae* und *externae* und die *Venae jugulares internae* auf beiden Seiten.

Unter „Inhalt“ wird nach möglichen Pathologien innerhalb eines Organsystems gefragt. Bei den Lymphknoten, den Halsgefäßen sowie dem Zungengrund/Tonsillenloge wird

lediglich zwischen einem „blanden“ und einem „pathologischen Befund“ unterschieden, während bei den Drüsen eine detailliertere Beschreibung erfolgt.

Die Punktevergabe auf diesem Befundungsbogen ermöglicht eine quantifizierbare Auswertung der Befundungsdauer, der Lesbarkeit der Berichte sowie der Vollständigkeit der dokumentierten anatomischen Strukturen.



Checkliste Befunderhebungsqualität in der Sonographie der Halsweichteile

Videonummer: BP1

Teilnehmende Klinik:

Prüfercode:

Dauer Befunderhebung (Sekunden):

1. Lesbarkeit des Befundes:

□ Sehr gut (5) □ gut (4) □ mittel (3) □ eher schlecht (2) □ schlecht (1)

Lymphknoten- Vollständigkeit

□ Level Ia rechts □ Level Ia links □ Level Ib rechts □ Level Ib links
□ Level IIa rechts □ Level IIa links □ Level IIb rechts □ Level IIb links
□ Level III rechts □ Level III links □ Level IV rechts □ Level IV links
□ Level Va rechts □ Level Va links □ Level Vb rechts □ Level Vb links
□ Level VI rechts □ Level VI links

Gesamtpunktzahl Lymphknoten (max. 18): _____

Lymphknoten- Inhalt

□ Keine path. LK □ path. LK

Drüsen- Vollständigkeit

□ GPA rechts □ GPA links □ GSM rechts □ GSM links
□ Gl. thyroidea rechts □ Gl. thyroidea links

Gesamtpunktzahl Drüsen (max. 6): _____

Drüsen- Inhalt (Raumforderung)

□ Lage: Oberflächlich, caudal □ Größe: 1,33 x 1,06 x 2,15 cm
□ Beschreibung: polyzyklisch, scharf begrenzt
□ Diagnität: nicht malignitätssuspekt □ Drüsenparenchym: unauffällig
□ Sonstige Befunde: Keine □ Perfusion: Keine

Halsgefäße

□ ACC rechts □ ACC links □ ACI rechts □ ACI links
□ ACE rechts □ ACE links □ VJI rechts □ VJI links

Gesamtpunktzahl Halsgefäße (max. 8): _____

Halsgefäße- Inhalt

□ o.p.B. □ path. Befund

Zungengrund / Tonsillenloge

□ Zungengrund □ Tonsillenloge

Zungengrund / Tonsillenloge- Inhalt

□ o.p.B. □ path. Befund

Abb. 3: Spezifischer Befundungsbogen für Sonographiebefunde der Halsweichteile im Rahmen der vorliegenden Promotionsarbeit

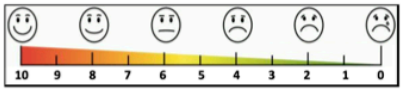
Zur Bewertung der Benutzerzufriedenheit mit der SR-Vorlage wurde ein bereits existierender Fragebogen mit 13 *Items* verwendet (Ernst et al., 2019a, 2029b, 2020b, 2020c) (Abb. 4). Anhand einer zehnteiligen VAS wurde eine subjektive Bewertung erfasst, wobei Abstufungen von 0 bis 10, entsprechend des Inhaltes der Frage, vorgegeben waren. Zu Beginn des Fragebogens wird erfragt, ob eine Computer-basierte Befunderhebung (Frage 1) beziehungsweise eine Computer-basierte strukturierte Befunderhebung (Frage 2) im klinischen Alltag generell Verwendung hat und ob sonographische Untersuchungen überhaut zu Diagnosezwecken erfolgen (Frage 3). Die nächsten Fragen befassen sich mit der Sinnhaftigkeit der SR (Frage 4) und ihrer Anwendung im Klinikalltag (Frage 5). Außerdem sollen die Studienteilnehmer einschätzen, ob SR zu einer Verbesserung der Befundung (Frage 6) und zu einer Zeitersparnis (Frage 7) führt. Falls Frage 7 mit „Nein“ beantwortet wird, soll der zeitliche Mehraufwand bewertet werden (Frage 8). Inwieweit eine SR-Vorlage Einsteigern beim Erlernen der Sonographie (Frage 9) und der Befundung (Frage 10) unterstützt, ist Inhalt der sich anschließenden *Items*. Die nächsten Abschnitte untersuchen die Meinungen der Studienteilnehmer zur Anwenderfreundlichkeit (Frage 11) und Übersichtlichkeit (Frage 12) des Programms. Abschließend steht dem Teilnehmer Platz für mögliche Verbesserungsvorschläge als Freitext zur Verfügung (Frage 13).

Evaluationsbogen Konventionelle Befunderhebung Kopf-Hals-Sonographie

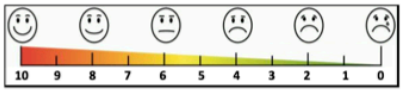
1.) Verwenden Sie im klinischen Alltag eine Computer-basierte Befunderhebung?
□ Ja □ Nein

2.) Verwenden Sie im klinischen Alltag eine Computer-basierte **STRUKTURIERTE** Befunderhebung?
□ Ja □ Nein

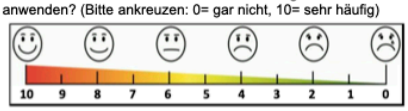
3.) Wie häufig verwenden Sie die Sonographie im klinischen Alltag? (Bitte ankreuzen: 0= gar nicht, 10= sehr häufig)



4.) Halten Sie eine strukturierte Befunderhebung für sinnvoll? (Bitte ankreuzen: 0= gar nicht sinnvoll, 10= sehr sinnvoll)



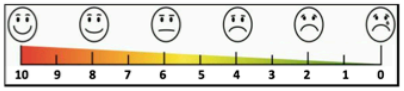
5.) Würden Sie eine strukturierte Befunderhebung im klinischen Alltag anwenden? (Bitte ankreuzen: 0= gar nicht, 10= sehr häufig)



E: Hr. Hodeib, P: PD Künzel, F: Dr. Ernst Datum: 01.02.2018 Seite 1 von 3

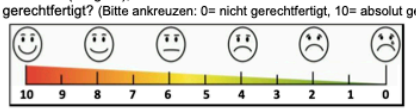
Evaluationsbogen Konventionelle Befunderhebung Kopf-Hals-Sonographie

6.) Glauben Sie, dass eine strukturierte Befunderhebung zur Verbesserung der Befundung führt? (Bitte ankreuzen: 0= nicht überzeugt, 10= sehr stark überzeugt)

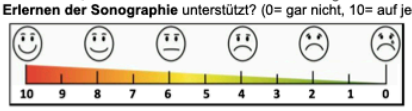


7.) Denken Sie, dass eine strukturierte Befunderhebung zu einer Zeitersparnis führt? (Bitte ankreuzen: 0= keine Zeitersparnis, 10= eindeutige Zeitersparnis)

8.) Falls nein (Frage 7), halten Sie den zeitlichen Mehraufwand für gerechtfertigt? (Bitte ankreuzen: 0= nicht gerechtfertigt, 10= absolut gerechtfertigt)



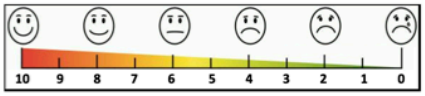
9.) Glauben Sie, dass eine strukturierte Befunderhebung den Einsteiger beim **Erlernen der Sonographie** unterstützt? (0= gar nicht, 10= auf jeden Fall)



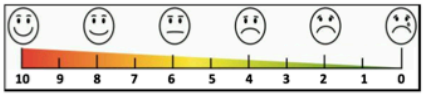
E: Hr. Hodeib, P: PD Künzel, F: Dr. Ernst Datum: 01.02.2018 Seite 2 von 3

Evaluationsbogen Konventionelle Befunderhebung Kopf-Hals-Sonographie

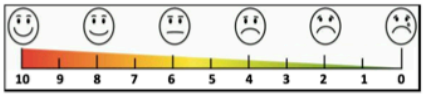
10.) Glauben Sie, dass eine strukturierte Befunderhebung den Einsteiger beim **Erlernen der Befundung** unterstützt? (0= gar nicht, 10= auf jeden Fall)



11.) Ist die vorliegende Version intuitiv und schnell erlernbar? (0= überhaupt nicht, 10= auf vollkommen intuitiv, praktisch sofort erlernbar)



12.) Ist die Menüführung übersichtlich? (0= gar nicht, 10= auf jeden Fall)



13.) Was würden Sie sich an Verbesserungen wünschen? (Freitext)

Abb. 4: Fragebogen zur Evaluation der Benutzerzufriedenheit bei der Anwendung einer strukturierten Befunderhebung in der Kopf-Hals-Sonographie mittels 10 Punkt VAS.

2.5 Datenverarbeitung und statistische Auswertung

Die gesammelten Daten dieser Studie wurden auf passwortgeschützten elektronischen Datenträgern gespeichert, darunter digitale Tabellen, die mit Microsoft® Excel (Microsoft Corporation, Redmond, Washington, USA) erstellt wurden. Der Zugriff war ausschließlich den Studienärzten vorbehalten. Die Auswertung erfolgte in anonymisierter Form mithilfe verschiedener statistischer Testverfahren.

Aus datenrechtlichen Gründen wurden die Patientennamen als dreistellige fortlaufende Codenummern generiert und mit den jeweiligen erhobenen Daten gespeichert. Ein entsprechendes Dokument, aus dem Codenummer und Patientendaten hervorgehen, stand lediglich dem Studienleiter zur Verfügung.

Die Anzahl der in dieser Studie erforderlichen Berichte wurde, basierend auf den von Sim und Wright beschriebenen Ansätzen, aus früheren Untersuchungen zur SR in der Kopf-Hals-Sonographie abgeleitet (Ernst et al., 2019a). Die Studie legte die Aussagekraft auf 80 % mit einem Signifikanzniveau von $\alpha = 0{,}05$ fest, mit dem Ziel, signifikante Unterschiede in der IRR darzustellen. Auf Grundlage früherer Daten mit einem Anteil von 50 % positiven Bewertungen sowie κ-Werten von 0,4 für die traditionelle Berichterstattung und 0,9 für SR wurde vor Studienbeginn eine erforderliche Fallzahl von 27 Bewertungen berechnet (Ernst et al., 2019a). Die IRR wurde mit Fleiss' κ gemessen und gemäß der von Landis und Koch vorgeschlagenen Klassifikation interpretiert (Landis und Koch, 1977). Die erhobenen Daten werden als Mittelwerte ± Standardabweichung (SD) angegeben. Zur Überprüfung der Normalverteilung wurde der Shapiro-Wilk-Test durchgeführt. Da einige Variablen nicht normalverteilt waren, kam für die Auswertung der Befundvollständigkeit und der Benutzerzufriedenheit der nichtparametrische Mann-Whitney-U-Test zum Einsatz. Für normalverteilte Größen, wie die Gesamtvollständigkeit und die IRA, wurde ein T-Test verwendet. Ein p-Wert von $<0{,}05$ wurde als statistisch signifikant angesehen. Zur Bewertung der IRR wurde Fleiss' κ berechnet. Alle statistischen Analysen wurden mit Prism 9.5.1 (GraphPad Software, Inc., San Diego, CA, USA) durchgeführt.

3. Ergebnisse

Die Rücksendequote der neun angeschriebenen Zentren lag im Rahmen der vorliegenden Studie bei 100 %. Die entsprechenden n=36 strukturierten Ultraschallbefunde wurden unter Supervision durch DEGUM-Kursleiter in Relation zur Befundvorlage ausgewertet. Evaluiert wurde die Vollständigkeit der Berichte, die IRR und die Benutzerzufriedenheit.

Nur fünf der neun teilnehmenden Abteilungen (55,6 %) bejahten die Verwendung eines digitalen Befunderhebungssystems im klinischen Alltag. Von diesen setzten nur drei Abteilungen (33,3 %) eine Form der SR ein (Abb. 5).

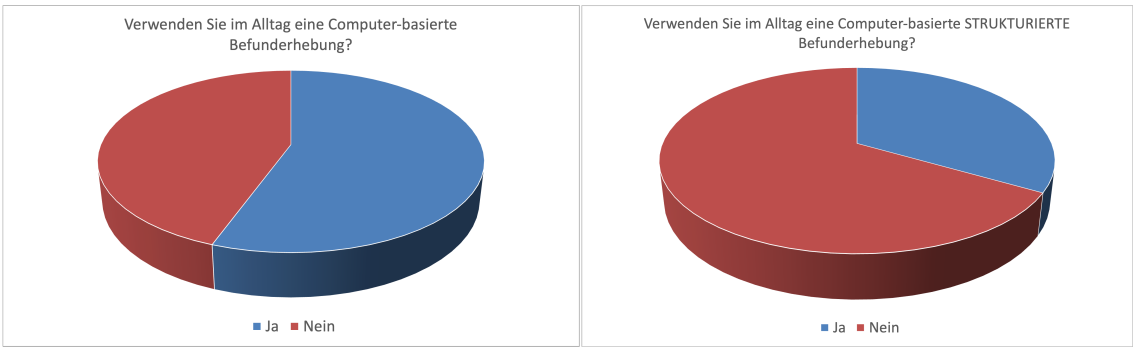


Abb. 5: Ergebnisdarstellung einer Ja/Nein- Umfrage zur Nutzung einer (a) Computer-basierten bzw. (b) Computer-basierten strukturierten Befunderhebung. Das blaue Segment repräsentiert den Anteil der Befragten die mit „Ja“ geantwortet haben, rot den Anteil der Befragten die mit „Nein“ geantwortet haben. Fünf der neun teilnehmenden Abteilungen gaben an eine Computer-basierte Befunderhebung im Allgemeinen zu benutzt ((a): 55,6 % „Ja“), jedoch verwendeten nur drei Abteilungen eine Computer-basierte strukturierte Befunderhebung ((b): 33,3 % „Ja“) (modifiziert nach Ernst et al.)

3.1 Vollständigkeit strukturierter Befunde

Die Analyse zeigte eine sehr gute mittlere Gesamtvollständigkeit der durch die Teilnehmer mittels SR-Vorlage erstellten Befunde (91,8 % ± 11,7 %). Betrachtet man die einzelnen Befundbereiche, so lag die Vollständigkeit bei der Beschreibung der Lymphknoten bei 87,3 % ± 65,1 %, der Blutgefäße bei 96,5 % ± 26,9 %, der großen Kopfspeicheldrüsen bei 97,7 % ± 9,6 % und der Pathologien bei 89,6 % ± 19,7 %.

In allen vier Fällen zeigte sich eine durchgehend sehr hohe Vollständigkeit der Berichte (Abb. 6). Im Einzelnen lag die Gesamtvollständigkeit bei 96,1 % ± 6,5 % für Fall 1 (Nachsorge nach Parotiskarzinom), 90,5 % ± 12,1 % für Fall 2 (Raumforderung in der Parotis), 87,9 % ± 12,9 % für Fall 3 (Zervikale Lymphknotenmetastase) und 92,8 % ± 12,6 % für Fall 4 (Raumforderung in der Parotis). Die Unterschiede zwischen den Fällen waren statistisch nicht signifikant.

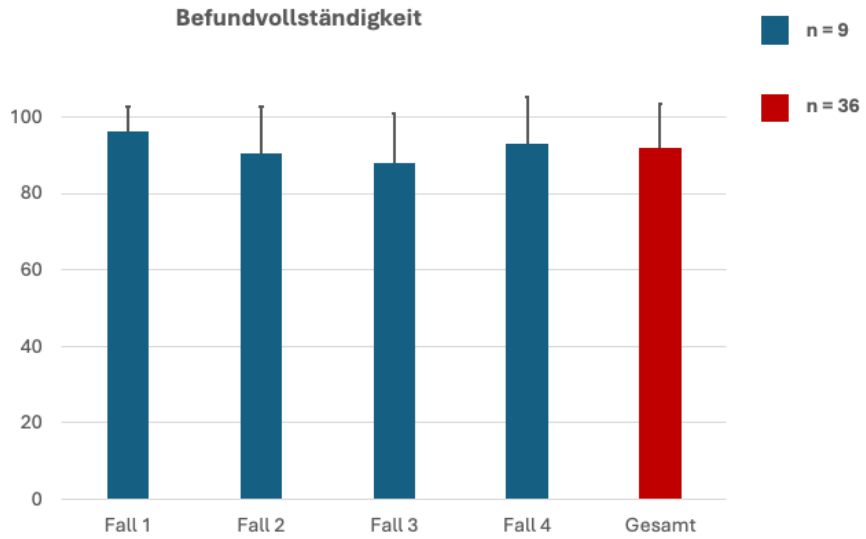


Abb. 6: Nachweis einer sehr hohen Befundvollständigkeit unter Verwendung einer strukturierten Befundvorlage (SR), bei allen 4 Fällen.

Durch die Befundung mittels SR als standardisiertes, softwaregesteuertes Programm gab es keinerlei Schwierigkeiten in Bezug auf die Lesbarkeit der Berichte (Lesbarkeit 100 %).

3.2 Interrater-Reliabilität

Der Einfluss der SR auf die IRR stand im Fokus der statistischen Auswertung. Zur Bestimmung der IRR wurde Fleiss' κ verwendet. Die IRR wurde für jede der vier durchgeführten Untersuchungen erhoben und zeigte Werte entsprechend der Klassifikation nach Landis und Koch von „substanziell“ (mind. κ=0,66) bis „nahezu perfekt“ (max. κ=0,92). Die Gesamtübereinstimmung ist insgesamt als „substanziell“ zu bewerten, mit einem κ von 0,72 und einem hohen prozentualen Übereinstimmungsgrad von 87,2 % ± 15,1 % (Bereich 50 – 100 %). Detailliert betrachtet zeigen sich für Fall 1 ein Fleiss' κ von

0,78 und eine IRA von 86,0 % ± 15,4 %, für Fall 2 ein Fleiss' κ von 0,92 und eine IRA von 96,9 % ± 5,4 %, für Fall 3 ein Fleiss' κ von 0,66 und eine IRA von 80,4 % ± 13,1 % sowie für Fall 4 ein Fleiss' κ von 0,74 und eine IRA von 85,7 % ± 18,2 % (Abb. 7, Abb. 8). Ein statistisch signifikanter Unterschied zwischen den Fragen in Bezug auf die IRR zeigte sich nicht. Ein statistisch signifikanter Unterschied der IRA wurde ausschließlich zwischen den Fällen 2 und 3 festgestellt ($p<0,05$, siehe Abb. 8).

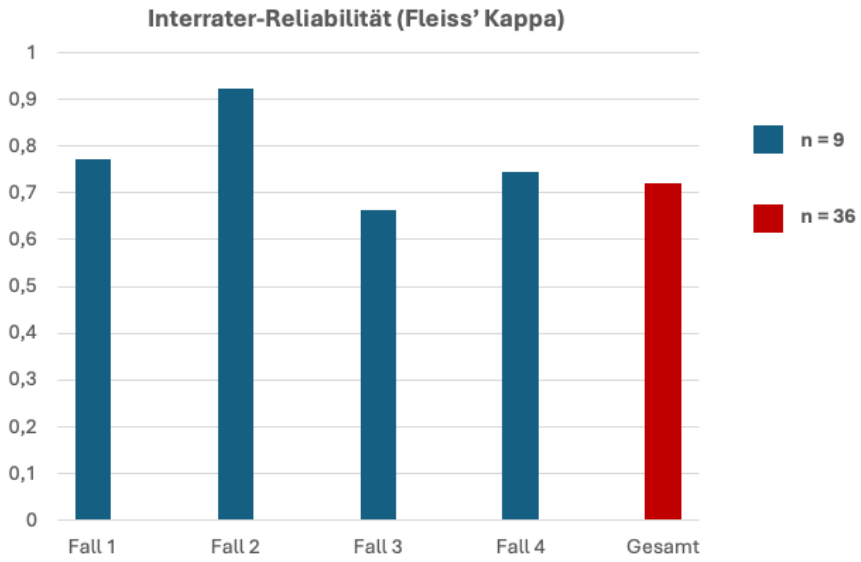


Abb. 7: Bestimmung der Interrater-Reliabilität (IRR) mittels Fleiss' Kappa: Die Verwendung einer strukturierter Befunderhebung (SR) führte in allen vier Fällen zu einer substanziellen bis nahezu perfekten IRR

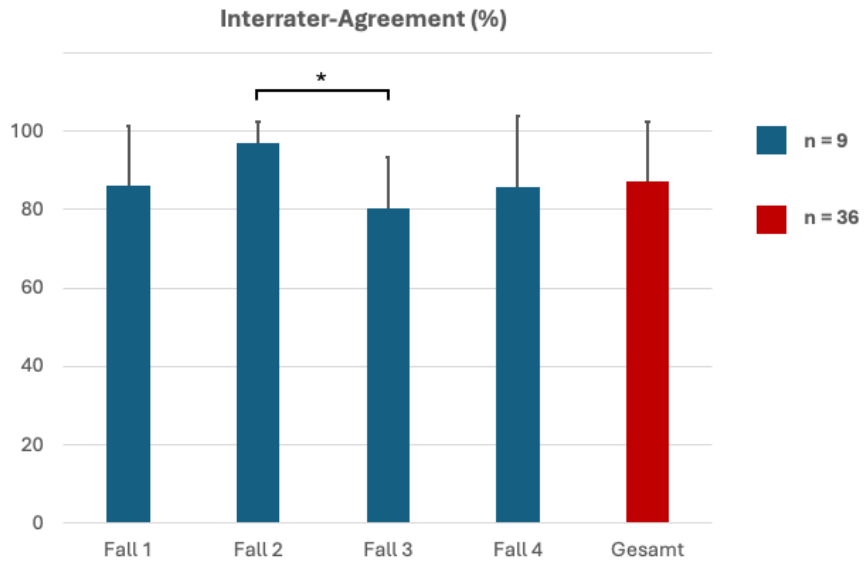


Abb. 8: Ergebnisse der Interrater-Übereinstimmung (IRA) mit einem hohen mittleren Übereinstimmungsgrad von 87,2 % ± 15,1 %. Abgesehen von den Fällen 2 und 3 zeigten die analysierten Fälle keine signifikanten Unterschiede in der Höhe der IRA. * $p<0,05$

3.3 Benutzerzufriedenheit

Die Auswertung der VAS-basierten Evaluationsfragebögen ergab insgesamt eine sehr hohe Zufriedenheit mit der SR-Vorlage (VAS 8,6 ± 1,8) (Abb. 9).

Auch die detaillierten Einzelbewertungen (Frage 3-12, Abb. 9) des Evaluationsbogens dienen der Erhebung von Einschätzungen zur SR. Sonographische Untersuchungen werden im klinischen Alltag „äußerst häufig“ eingesetzt (Frage 3; VAS 8,9 ± 2,5). Auffällig war, dass der Großteil der teilnehmenden Zentren bisher keine SR-Vorlage zu Diagnosezwecken benutzte, aber die Etablierung dieser als „sehr sinnvoll“ einschätzt (Frage 4; VAS 9,8 ± 0,6). SR würde laut den Befragten „sehr häufig“ im klinischen Alltag angewendet werden (Frage 5; VAS 9,9 ± 0,3). Zudem waren die bewertenden Personen „sehr stark überzeugt“ (Frage 6; VAS 9,6 ± 0,8), dass SR eine deutliche Steigerung der Befundqualität zur Folge hat.

Auch zeitliche Aspekte wurden differenziert bewertet. Die Befragten gingen von einer „eher wahrscheinlichen Zeitersparnis“ durch SR aus (Frage 7; VAS 7,2 ± 2). Der mit der Anwendung verbundene zeitliche Mehraufwand wurde von den Befragten als „sehr wahrscheinlich gerechtfertigt“ erachtet (Frage 8; VAS 9 ± 0,9). Weiterhin waren die

Befragten der Meinung, dass Berufseinsteiger bei der Durchführung einer sonographischen Untersuchung eine „deutliche Unterstützung“ durch Verwendung der SR erfahren (Frage 9; VAS 8,2 ± 2,6). Ferner waren die Befragten der Meinung, dass SR Berufseinsteiger auch bei der sich anschließenden Befunderstellung „sehr stark unterstützen“ kann (Frage 10; VAS 9 ± 1,3). Ergänzend wurde das Programm als „ziemlich intuitiv und schnell erlernbar“ (Frage 11; VAS 7,7 ± 0,9) sowie als „ziemlich übersichtlich in der Menüführung“ (Frage 12; VAS 7,3 ± 2,3) bewertet.

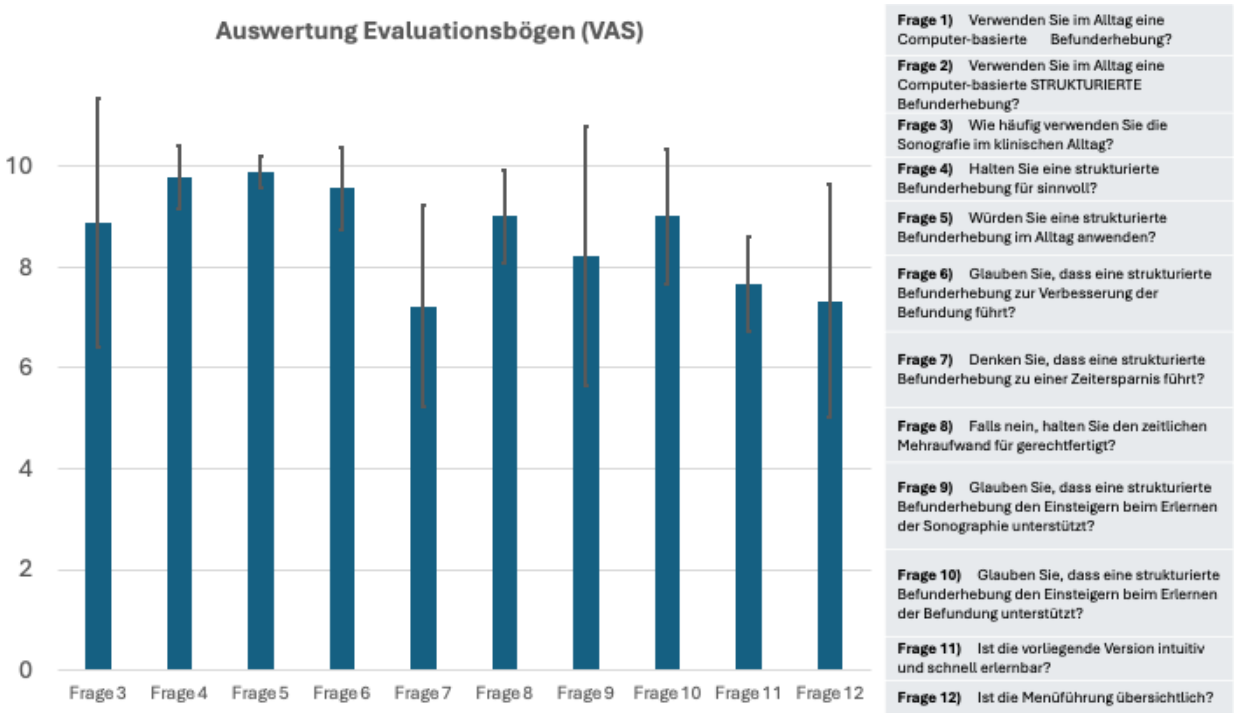


Abb. 9: Evaluation der Benutzerzufriedenheit mittels visueller Analogskala (VAS, 0 = trifft voll zu bis hin zu 10 = trifft gar nicht zu). Die in blauen Balken dargestellten Ergebnisse des Fragebogens zeigen eine durchweg positive Rückmeldung zur Einführung und Nutzung der strukturierten Befunderhebung. Der durchschnittliche Mittelwert aller Fragen beträgt VAS 8,6, mit einer Standardabweichung von 1,8

4. Diskussion

In der vorliegenden prospektiven multizentrischen Studie wurde erstmals der Einfluss der SR auf die IRR, die Befundvollständigkeit und die Benutzerzufriedenheit in der Kopf-Hals-Sonographie untersucht. Der Wert der Ultraschalluntersuchung als bildgebendes Verfahren wurde bereits umfassend untersucht (Heller und Dörges, 2006; Künzel et al., 2017; Mojoli et al., 2019; Welkoborsky et al., 1997). Dank des technologischen Fortschritts hat die Sonographie in den letzten Jahren an Spezifität und Anwendungsmöglichkeiten gewonnen, was einen zunehmenden diagnostischen Stellenwert der Untersuchung bedingt und es als Standardverfahren in der Primärdiagnostik, bei Kontrolluntersuchungen und in der Operationsplanung etabliert (Beach et al., 2010; Moshtaghi et al., 2017; Welkoborsky, 2009). Ein zentraler Kritikpunkt an der Sonographie als diagnostische Methode ist ihre im Vergleich zu anderen bildgebenden Verfahren, wie MRT oder CT, deutlich stärkere Untersucherabhängigkeit. Selbst die Etablierung sonographischer Standardschnitte führt nicht immer zu identischen Untersuchungsergebnissen (Müller et al., 2001; Schüle et al., 1999). Ferner können unterschiedliche Dokumentationsmodi eine longitudinale Beurteilung von Strukturen oder Krankheitsverläufen massiv erschweren.

Die Anwendung von SR bietet eine Möglichkeit, die Reproduzierbarkeit zu erhöhen. Wie auch schon in Vorarbeiten umfassend dargelegt, verbessert SR die Zeiteffizienz, Befundvollständigkeit und Lesbarkeit bildgebender Verfahren bei gleichzeitigem Lerneffekt signifikant (Ernst et al., 2019a, 2019b, 2020b, 2020c).

Welchen Einfluss SR auf die IRR hat, wurde in Studien bislang jedoch nicht systematisch untersucht. In der vorliegenden Studie werden nun erstmals Daten über die Auswirkung strukturierter Berichte auf die IRR bei Sonographien des Kopf-Hals-Bereiches vorgelegt.

Es zeigte sich, dass die Einführung von SR zu einer substanziellen bis nahezu perfekten IRR führt und mit einer hohen Befundvollständigkeit sowie einer durchgehend hohen Benutzerzufriedenheit einhergeht.

4.1 Status Quo

Die Auswertung des im Rahmen dieser Studie angewendeten Evaluationsfragebogens zeigte, dass sonographische Diagnoseverfahren in allen teilnehmenden Kliniken regelmäßig zur Anwendung kommen (Abb. 5). Dies unterstreicht den Stellenwert der Kopf-Hals-Sonographie als diagnostisches Mittel im klinischen Alltag. Allerdings zeigte sich auch, dass in einem Großteil der befragten Zentren bisher weder eine Computer-basierte noch eine strukturierte Computer-basierte Befunderhebung zu Diagnosezwecken genutzt wird. Dies überrascht insbesondere vor dem Hintergrund, dass es sich bei den teilnehmenden Zentren um ausgewiesene Ultraschallzentren mit starker Einbindung in die DEGUM handelt. Bemerkenswert ist daher, dass SR trotz dieser geringen Vorerfahrung von allen Teilnehmern als klinisch sinnvoll und alltagstauglich bewertet wurde – ein Befund, der im nächsten Abschnitt detailliert dargestellt wird.

4.2 Befundvollständigkeit

Die Befundvollständigkeit der mittels SR-Vorlage erstellten Berichte war in allen Kategorien sehr hoch. Konkret zeigte sich eine Gesamtvollständigkeit von 91,8 % ± 11,7 %. Die Vollständigkeit bei der Beschreibung von Blutgefäßen (96,5 %) und Drüsen (97,7 %) lag dabei deutlich höher als bei der Beschreibung von Lymphknoten (87,3 %) (Abb. 6). Dieser Unterschied lässt sich plausibel erklären: Die Beschreibung von Lymphknoten erfordert eine systematische Beurteilung mehrerer zervikaler Level beidseitig und ist damit komplexer und fehleranfälliger als die Beschreibung anatomisch klar abgrenzbarer Strukturen wie Drüsen oder Gefäße. Trotzdem ist auch dieser Wert als hoch einzustufen.

Ferner zeigten sich für die Befundung mittels SR als standardisiertes softwaregesteuertes Computerprogramm keinerlei Schwierigkeiten in Bezug auf die Lesbarkeit der Berichte. Dies mag zunächst trivial erscheinen, doch die Verwendung handgeschriebener Befundungsbögen mit und ohne nachträgliche Digitalisierung durch Einscannen ist in vielen deutschen Kliniken weiterhin gängige Praxis (Szilveszter et al., 2017b). Während handgeschriebene Bögen oft schwer lesbar sind, gewährleisten computergestützte SR-Berichte eine vollständige Lesbarkeit und bieten damit einen klaren Qualitätsvorteil im klinischen Alltag.

Diese Ergebnisse stehen im Einklang mit Vorarbeiten, welche zeigen, dass SR das Potential hat die Qualität von Ultraschallberichten im Hinblick auf Vollständigkeit und Reproduzierbarkeit deutlich zu verbessern (Ernst et al., 2019a, 2019b, 2020b; Timor-Tritsch und Goldstein, 2005). Die vorliegende Studie bestätigt und erweitert diese Befunde durch den multizentrischen Charakter: Die hohe Vollständigkeit wurde nicht in einem einzelnen Zentrum, sondern einheitlich über neun verschiedene Universitätskliniken mit unterschiedlichen Arbeitsroutinen hinweg erzielt.

4.3 Interrater-Reliabilität und Interrater-Agreement

Die Bestimmung der IRR mittels Fleiss' κ durch neun Untersucher ergab für die vier ausgewerteten Fälle „substanzielle“ bis „nahezu perfekte“ Übereinstimmungswerte (κ = 0,66–0,92) (Landis und Koch, 1977). Die Gesamt-IRR betrug κ = 0,72 bei einem mittleren prozentualen Übereinstimmungsgrad von 87,2 % ± 15,1 %, was nach Landis und Koch als „substanziell“ einzustufen ist. Damit liegt die Übereinstimmung der Beurteilungen auf einem hohen Niveau – ein deutliches Zeichen dafür, dass die standardisierte Befundung mithilfe von SR mit einer hohen IRR einhergeht. Dies ist ein klinisch bedeutsames Ergebnis, da die Sonographie allgemein als stark untersucherabhängiges Verfahren gilt und eine solche Übereinstimmung zwischen neun Untersuchern aus neun verschiedenen Zentren nicht selbstverständlich ist. Der bereits aufgeführte Kritikpunkt einer im Vergleich zu anderen bildgebenden Verfahren deutlich intensivierten Personenabhängigkeit der Befundung kann somit durch die Einführung von SR reduziert werden.

Ein signifikanter Unterschied wurde ausschließlich zwischen Fall 2 (Raumforderung Parotis) und Fall 3 (zervikale Lymphknotenmetastase) festgestellt ($p < 0,05$). Während Fall 2 nahezu perfekte Übereinstimmungswerte aufwies (κ = 0,92 „nahezu perfekt“; IRA = 96,9 %), lag die Übereinstimmung in Fall 3 deutlich niedriger (κ = 0,66 „substanziell“; IRA = 80,4 %). Dieses Ergebnis deutet darauf hin, dass neben dem standardisierten Befundformat auch die Komplexität und Eindeutigkeit des untersuchten Falles einen relevanten Einfluss auf die IRR hat. Klar abgrenzbare Befunde führten zu einer höheren Übereinstimmung zwischen den Untersuchern als komplexere oder schwerer zu interpretierende Veränderungen. Dieser Zusammenhang wurde auch in anderen Studien

beschrieben und sollte bei der Interpretation von IRR-Werten stets berücksichtigt werden (Cartwright et al., 2013; Youk et al., 2016).

Neben der Komplexität des Befundes spielt auch der Ausbildungsstand und die Erfahrung der Untersucher eine entscheidende Rolle für eine hohe IRR (Cartwright et al., 2013; Fleury et al., 2018). Je präziser die Untersuchung und je stärker der Fokus des Untersuchers auf Details, desto höher die nachweisbare IRR. Die Verwendung einer standardisierten Vorgehensweise kann sich hier als hilfreich erweisen, da sie nachweislich die Reproduzierbarkeit und Genauigkeit verbessert, was nicht nur für die Ultraschalldiagnostik, sondern auch für andere bildgebende Verfahren nachweisbar war (Birrenbach et al., 2004; Chiu et al., 2019).

Gerade vor dem Hintergrund der vergleichsweise untergeordneten Rolle der sonographischen und befunddokumentarischen Ausbildung sowie der hohen Untersucherabhängigkeit ist es bemerkenswert, dass die eingeschlossenen Untersucher aus neun verschiedenen Kliniken mit unterschiedlichen Arbeitsansätzen unter Verwendung der SR zu sehr ähnlichen Ergebnissen bezüglich der analysierten Strukturen und Pathologien gelangten. Bei der Kopf-Hals-Sonographie handelt es sich um ein hochdynamisches bildgebendes Verfahren. Anders als bei statischen Bildgebungsverfahren liefert die Sonographie eine Vielzahl von Bildern in Echtzeit, die sich durch Bewegungsdynamik, Lageveränderungen und physiologische Flussphänomene auszeichnen. Diese Dynamik bietet Vorteile für die Befundung, da sie eine objektivere Beurteilung beweglicher Strukturen wie des Herzens oder der Gelenke ermöglicht und Flüssigkeitsströme in Gefäßen mittels Dopplersonographie besser darstellbar sind. Zudem ist eine dynamische Ultraschalluntersuchung zeitsparender, objektiver und weniger abhängig vom Untersucher als statische Untersuchungen (O'Dell et al., 2015; Youk et al., 2016). In der Studie von Youk et al. (2016) zeigte sich eine höhere Übereinstimmung zwischen den Beurteilern bei dynamischer Bildgebung, insbesondere bei kleineren Läsionen, verglichen mit statischen Aufnahmen (Youk et al., 2016). Um eine möglichst realitätsnahe und standardisierte Bewertung zu ermöglichen, erhielten die Untersucher dieser Studie daher neben den statischen Bildern der Pathologien auch die zugehörigen Videosequenzen der einzelnen Untersuchungsschritte. Die Bereitstellung dieser zusätzlichen Informationen dürfte zu der hohen IRR beigetragen haben, da sie den

dynamischen Charakter der Sonographie besser abbildet und Interpretationsspielräume reduziert (Cartwright et al., 2013; Larson, 2021; O'Dell et al., 2015; Youk et al., 2016). Die Vorteile im Vergleich zur Befundung statischer Bilder könnten zudem darauf beruhen, dass wichtige Informationen zu Lagebeziehungen, Komprimierbarkeit, Elastizität oder Durchblutungsverhalten zusätzlich besser erfasst werden können (Cartwright et al., 2013; Youk et al., 2016).

Der strukturierte, checklistenartige Aufbau von SR erleichtert nicht nur die Anwendung im klinischen Alltag, sondern trägt wesentlich zur Qualitätssicherung bei (Ellis und Srigley, 2016; Srigley et al., 2013). Durch die systematische Erfassung wird sichergestellt, dass keine relevanten Informationen übersehen werden. Dies verbessert die Nachvollziehbarkeit und erhöht die Vergleichbarkeit zwischen den Untersuchungen.

SR ist vor allem dort unerlässlich, wo Beurteilung im zeitlichen Verlauf eine entscheidende Rolle spielt. Hier ist die Tumornachsorge ein prädestiniertes Beispiel. Neben der beschriebenen Steigerung der IRR ist eine klare Befundstruktur Voraussetzung für die verlässliche Beurteilung wiederholter Untersuchungen, da die Entscheidungsfindung zu weiterführender Diagnostik und Therapie meistens auf Basis der Befunddynamik und selten anhand von Einzelbefunden festgelegt wird. Nur eine im longitudinalen Verlauf vergleichbare Befunddokumentation ermöglicht eine adäquate Beurteilung dieser Dynamik. Die Aussagekraft klinischer Untersuchungen hängt dabei wesentlich von der präzisen Erfassung und Bewertung der erhobenen Befunde ab, beispielsweise bei der Beurteilung von Größenveränderungen über einen bestimmten Zeitraum. Dies sei hier kurz am Beispiel der zervikalen Lymphknotensonographie erläutert: Sonographische Verlaufskontrollen haben in der Nachsorge maligner Tumore im Kopf-Hals-Bereich einen zentralen Stellenwert (Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft und AWMF), 2021). Je nach zeitlichem Abstand zur Erkrankung und dem Abschluss der Therapie wird die Nachsorge in festgelegten zeitlichen Abständen in gleicher Weise wiederholt und unter anderem eine sonographische Untersuchung der Halsweichteile durchgeführt. Gerade bei der Beurteilung der Halslymphkoten ist Volumendynamik ein entscheidendes Kriterium zur Identifikation und Bewertung möglicher pathologischer Veränderungen. So kann ein initial als suspekt eingeschätzter Halslymphknoten (z.B. im

akuten Infekt) bei deutlicher Größenregredienz im zeitlichen Verlauf als unauffällig eingestuft und damit eine invasive Abklärung vermieden werden.

Für eine valide Verlaufskontrolle sind neben einer hohen IRR der Befundungsmodalität auch die diagnostischen Fähigkeiten der Untersuchenden, die Qualität der Dokumentation sowie die Reproduzierbarkeit und Interpretierbarkeit der Befunde entscheidend. Auch die Standardisierung wird durch das einheitliche SR-Format gefördert. Beides trägt entscheidend zu einer höheren IRR bei, da es individuelle Bewertungsspielräume reduziert.

Die Sonographie des Kopf-Hals-Bereiches stellt eine große Herausforderung dar, da sich eine komplexe Anatomie auf engem Raum befindet. Gerade in diesem Kontext bietet SR durch standardisierte Entscheidungspfade und klare Menüführung eine wesentliche Hilfestellung von welcher vor allem unerfahrenere Untersucher profitieren, da sie Schritt für Schritt durch die Untersuchung geleitet werden. Das Risiko, relevante Nebenbefunde zu übersehen oder zu vergessen, wird dadurch reduziert, was sich positiv auf die Befundqualität und damit auf die IRR auswirkt (Ernst et al., 2019a, 2019b, 2020b). Auch die Ergebnisse der vorliegenden Arbeit verdeutlichen, dass in allen untersuchten Kategorien eine hohe mittlere Gesamtvollständigkeit der SR-Befunde erreicht wurde.

4.4 Benutzerzufriedenheit

Die Auswertung des VAS-basierten Evaluationsfragebogens ergab eine sehr hohe Gesamtzufriedenheit mit der SR-Vorlage (VAS 8,6 ± 1,8). Gleichzeitig bewerteten alle Befragten die Einführung einer SR als sinnvoll, alltagstauglich, übersichtlich, intuitiv bedienbar und besonders hilfreich für Neueinsteiger (Abb 9). Dieses positive Feedback unterstreicht das Potenzial der SR, die klinische Dokumentation zu verbessern und standardisierter zu gestalten.

Besonders hervorzuheben ist, dass SR trotz der zum Zeitpunkt der Studie geringen Verbreitung in den teilnehmenden Zentren – nur drei von neun Abteilungen (33,3 %) setzten eine Form der SR ein – von allen Teilnehmern als klinisch sinnvoll (VAS 9,8 ± 0,6) und alltagstauglich (VAS 9,9 ± 0,3) bewertet wurde. Dies ist ein relevanter Befund, da er

zeigt, dass die Akzeptanz von SR nicht von bestehender Vorerfahrung abhängt, sondern auch bei erstmaliger Anwendung hoch ist.

Hinsichtlich der Zeiteffizienz bewertete die Mehrheit der Befragten SR als zeitersparend (VAS 7,2 ± 2,0) und ein anfänglicher Mehraufwand wurde als durch die verbesserte Befundqualität gerechtfertigt angesehen (VAS 9,0 ± 0,9). Dieses Ergebnis ist konsistent mit der Literatur, die zeigt, dass ein initialer Zeitaufwand bei der Einführung von SR nach einer Einarbeitungsphase entfällt und langfristig sogar Zeitersparnisse möglich sind (Nörenberg et al., 2017; Sluijter et al., 2016).

Darüber hinaus sahen die Teilnehmer in SR ein wertvolles Instrument für die Aus- und Weiterbildung: Sowohl die Unterstützung bei der Durchführung sonographischer Untersuchungen (VAS 8,2 ± 2,6) als auch bei der anschließenden Befunderstellung (VAS 9,0 ± 1,3) wurden hoch bewertet. Dies bestätigt frühere Erkenntnisse, dass SR insbesondere für unerfahrene Untersucher einen ausgeprägten Lerneffekt entfaltet und als strukturiertes Lehr- und Lernwerkzeug im klinischen Alltag eingesetzt werden kann (Ernst et al., 2019b, 2020c). Angesichts der bekannten Defizite in der sonographischen Ausbildung an deutschen Universitäten, sowie im Rahmen der Facharztweiterbildung gewinnt dieser Aspekt zusätzliche Relevanz (Benè und Bergus, 2014; Wallis und McCoubrie, 2011).

4.5 Limitationen

Die vorliegende Studie unterliegt einigen methodischen Einschränkungen, die bei der Interpretation der Ergebnisse berücksichtigt werden müssen.

Erstens wurde keine Kontrollgruppe mit FTR einbezogen. Da ausreichende Vergleichsdaten zur IRR der FTR aus der Literatur vorliegen und ein direkter Vergleich nicht primäres Studienziel war, erscheint diese Entscheidung vertretbar (Ernst et al., 2019b; Ernst et al., 2019a). Dennoch schränkt das Fehlen einer Kontrollgruppe die Aussagekraft hinsichtlich des Ausmaßes der durch SR erzielten IRR-Verbesserung ein.

Zweitens ist die Stichprobengröße mit n=4 Fällen und n=9 Untersuchern, resultierend in n=36 Befunden, begrenzt. Die Fallzahlkalkulation erfolgte auf Grundlage früherer Studien und war für die Beantwortung der primären Fragestellung ausreichend; eine

Generalisierung auf breitere Patientenpopulationen oder andere Pathologien des Kopf-Hals-Bereiches sollte jedoch mit Zurückhaltung erfolgen.

Drittens handelt es sich bei den eingeschlossenen Untersuchern ausschließlich um DEGUM-Kursleiter bzw. mindestens DEGUM Stufe II qualifizierte Fachärzte an Universitätszentren. Die Ergebnisse sind daher möglicherweise nicht direkt auf weniger erfahrene Untersucher oder nicht-universitäre Einrichtungen übertragbar. Es ist anzunehmen, dass SR bei weniger erfahrenen Untersuchern zu einem noch stärkeren Effekt auf die IRR und die Befundvollständigkeit führt, was in zukünftigen Studien untersucht werden sollte.

Viertens wurden alle Untersuchungen an einem einzigen Gerät durchgeführt (LOGIQ S8, GE Healthcare). Ob die Ergebnisse auf andere Geräteklassen oder Sondenfrequenzen übertragbar sind, bleibt offen.

4.6 Ausblick

Die vorliegenden Ergebnisse unterstreichen den Bedarf an einer klinikübergreifenden Implementierung von SR, zumal Freitextbefunde in vielen Kliniken weiterhin der Standard, sind, obwohl sie als weniger alltagstauglich gelten und eine größere Varianz in Struktur, Inhalt und Terminologie aufweisen, was die Vergleichbarkeit der Befunde zwischen den Untersuchern erschwert. Der Nutzen strukturierter Befunde konnte bereits in verschiedenen Fachbereichen nachgewiesen werden auch im Hinblick auf die IRR (Gassenmaier et al., 2017; Nörenberg et al., 2017; Park et al., 2019; Sabel et al., 2017a; Schoeppe et al., 2018).

Die Ergebnisse dieser Studie legen nahe, dass eine klinikübergreifende Implementierung von SR in der HNO-Sonographie die Befundqualität und die Vergleichbarkeit zwischen Untersuchern nachhaltig verbessern kann. Perspektivisch bietet die Maschinenlesbarkeit strukturierter Befunde Anknüpfungspunkte, beispielsweise für die Integration in die elektronische Patientenakte (ePA). Diese ermöglicht durch die digitale Archivierung medizinischer Informationen verlässliche Rückschlüsse auf Patienteninformationen, Diagnosen, Therapieverläufe und Behandlungsergebnisse und kann damit zu einer verbesserten Patientenversorgung beitragen (Dochow, 2019; Lux und Breil, 2017). Die

ePA befindet sich seit 2021 in der Weiterentwicklung. Nach einem erfolgreichen Abschluss einer Testphase startete sie bundesweit am 29. April 2025. Seit dem 1. Oktober 2025 wird die ePA im Rahmen eines Opt-out-Modells für gesetzlich Versicherte eingeführt und soll einen wichtigen Beitrag zu einer digital vernetzten Gesundheitsversorgung leisten. (Gesundheit, 2025).

Auf Basis der nachgewiesenen hohen IRR und Befundvollständigkeit erscheint SR auch für telemedizinische Anwendungsbereiche besonders geeignet: Die durch SR vermittelte Standardisierung und hohe Beurteilerübereinstimmung kann dazu beitragen, auch bei einem durch Telemedizin erweiterten Kreis betreuender Ärzte Vergleichbarkeit und Homogenität der Befunde zu erhalten (Hoffmann und van den Berg, 2013). Telemedizin umfasst verschiedene ärztliche Versorgungskonzepte, welchen gemein ist, dass medizinische Leistungen verschiedener medizinischer Fachdisziplinen über eine räumliche und/oder zeitliche Distanz erfolgen. Gerade in strukturschwachen Regionen mit begrenztem Zugang zu HNO-Fachärzten könnte SR so zu einer verbesserten flächendeckenden Versorgung beitragen. Ein weiterer Vorteil von SR liegt in der Möglichkeit zur automatisierten Textwiedergabe, wodurch grammatikalische und Syntaxfehler unter Verwendung der korrekten Terminologie automatisch minimiert, und Folgefehler vermeiden werden können (Nörenberg et al., 2017). Durch diese Eigenschaften könnte SR wesentlich dazu beitragen eine unzureichende Befundqualität aufgrund von mangelhaften Grammatik- und Sprachkenntnissen zu verhindern. Zwar haben einige Fachgesellschaften bereits Empfehlungen ausgesprochen, jedoch existieren bislang noch keine flächendeckend festgelegten nationalen und internationalen Standards in Bezug auf Terminologie und die abgefragten Befundstrukturen in der HNO-Sonographie – die Entwicklung solcher Standards wäre ein wichtiger nächster Schritt, um das Potenzial der SR im telemedizinischen Kontext voll auszuschöpfen.

Abschließend kann die zunehmende Verfügbarkeit moderner KI-gestützter Sprach- und Übersetzungsmodelle dabei helfen, den internationalen Austausch strukturierter Befunde weiter zu erleichtern. Studien zeigen, dass computerbasierte Übersetzungssysteme bereits seit Jahren in der Lage sind, hochwertige Befunde in verschiedenen Sprachen ohne relevante Qualitätsunterschiede zu generieren (Bosmans et al., 2012; Ranschaert und Binkhuysen, 2013; Ross et al., 2010; Sobez et al., 2019). Mit dem Aufkommen

moderner KI-basierter Sprachmodelle (LLMs) hat sich die Qualität automatischer Befundübersetzungen in der Medizin deutlich verbessert. Diese Technologien ermöglichen nicht nur eine simultane und präzise Übersetzung strukturierter Befunde, sondern unterstützen auch eine automatische Anpassung an die richtige medizinische Terminologie (Lyu et al., 2023; Mallio et al., 2023b). Solch eine Funktion macht eine klinische Zusammenarbeit und einen wissenschaftlichen Austausch auf nationaler und internationaler Ebene möglich, was in Zeiten steigender Migration und Globalisierung mehr und mehr an Bedeutung gewinnt (Sobez et al., 2019). Mehrsprachige SR-Berichte helfen dabei, Kommunikationsbarrieren zwischen Kollegen und Patienten zu überwinden, sodass Nachfragen reduziert und Missverständnisse vermieden werden können.

Diese Bereiche bleiben künftiger Forschung vorbehalten. Insbesondere multizentrische Studien mit größeren Fallzahlen, heterogeneren Untersucherpopulationen und einem direkten Vergleich mit FTR-Kontrollgruppen wären wünschenswert, um die hier erzielten Ergebnisse zu validieren und den Nutzen von SR in der HNO-Sonographie weiter zu festigen.

5. Zusammenfassung

Die Interrater-Reliabilität stellt in der sonographischen Diagnostik einen bekannten Schwachpunkt dar, da Ultraschalluntersuchungen erheblichen subjektiven Einflüssen des Untersuchers unterliegen. Ob die strukturierte Befunderhebung die Übereinstimmung zwischen unabhängigen Untersuchern verbessern kann, war bislang nicht systematisch untersucht worden. Ziel der vorliegenden prospektiven multizentrischen Studie war es, den Einfluss der Strukturierten Befunderhebung auf Befundvollständigkeit, Benutzerzufriedenheit und Interrater-Reliabilität in der Kopf-Hals-Sonographie zu untersuchen. Hierzu wurden n=4 video- und bilddokumentierte Ultraschalluntersuchungen typischer Kopf-Hals-Pathologien an n=9 deutsche Universitätskliniken mit DEGUM-Kursleitern übermittelt und dort mittels einer webbasierten SR-Vorlage befundet, woraus n=36 strukturierte Befunde resultierten.

Die Analyse zeigte eine sehr hohe mittlere Gesamtvollständigkeit der SR-Befunde von 91,8 % ± 11,7 %, mit durchgehend hohen Werten in allen vier Fällen und Befundkategorien. Die Lesbarkeit betrug 100 %. Hinsichtlich der IRR ergaben sich Fleiss'-κ-Werte zwischen 0,66 und 0,92, entsprechend einer „substanziellen“ bis „nahezu perfekten“ Übereinstimmung nach Landis und Koch. Der Gesamt-κ-Wert betrug 0,72 bei einem mittleren prozentualen Übereinstimmungsgrad von 87,2 % ± 15,1 %. Ein signifikanter Unterschied zeigte sich ausschließlich zwischen dem klar abgrenzbaren Fall 2 (Parotisraumforderung; κ = 0,92) und dem komplexeren Fall 3 (zervikale Lymphknotenmetastase; κ = 0,66), was auf den Einfluss der Befundkomplexität auf die Beurteilerübereinstimmung hinweist.

Die Benutzerzufriedenheit war mit einem mittleren VAS-Wert von 8,6 ± 1,8 durchgehend hoch. Der Einsatz sonographischer Untersuchungen im klinischen Alltag wurde als „äußerst relevant“ eingeschätzt (VAS 8,9 ± 2,5), die Etablierung von SR als „sehr sinnvoll“ (VAS 9,8 ± 0,6) und „alltagstauglich“ (VAS 9,9 ± 0,3) bewertet. Die Befragten gaben an, dass SR eine deutliche Steigerung der Befundqualität bewirke (VAS 9,6 ± 0,8). Eine Zeitersparnis durch SR wurde als „wahrscheinlich“ eingeschätzt (VAS 7,2 ± 2,0) und ein damit verbundener zeitlicher Mehraufwand als „gerechtfertigt“ erachtet (VAS 9,0 ± 0,9). Auch Lernaspekte wurden positiv bewertet: SR wurde als hilfreiche Unterstützung für Berufsanfänger sowohl bei der Untersuchungsdurchführung (VAS 8,2 ± 2,6) als auch bei

der Befunderstellung (VAS 9,0 ± 1,3) eingeschätzt. Das Programm wurde als „intuitiv erlernbar“ (VAS 7,7 ± 0,9) und „übersichtlich in der Menüführung“ (VAS 7,3 ± 2,3) bewertet.

Die Ergebnisse belegen, dass SR ein wirksames Instrument zur Standardisierung der Befunddokumentation in der HNO-Sonographie darstellt und mit einer hohen IRR, Befundvollständigkeit und Benutzerzufriedenheit einhergeht. Eine klinikübergreifende Implementierung erscheint auf Basis dieser Daten sinnvoll. Perspektivisch bietet SR darüber hinaus Anknüpfungspunkte für die Integration in die elektronische Patientenakte, telemedizinische Anwendungen, datenbasierte Forschung sowie den internationalen Befundaustausch – Bereiche, die weiterer systematischer Untersuchung bedürfen.

6. Abbildungsverzeichnis

Abb. 1: Bilddateien der n=4 durchgeführten sonographischen Untersuchungen mit Darstellung typischer Kopf- und Halspathologien. (a) Fall 1: Nachuntersuchung bei Parotiskarzinom (b) Fall 2: Unklaren Raumforderung in der Parotis (c) Fall 3: Verdacht auf zervikale Lymphknotenmetastasen (d) Fall 4: Suspekte Parotisraumforderung 29

Abb. 2: Befundung eines solitären Lymphknotens mit 10x8x7 mm im Level Vb links mittels strukturierter Befunderhebung 34

Abb. 3: Spezifischer Befundungsbogen für Sonographiebefunde der Halsweichteile im Rahmen der vorliegenden Promotionsarbeit..... 35

Abb. 4: Fragebogen zur Evaluation der Benutzerzufriedenheit bei der Anwendung einer strukturierten Befunderhebung in der Kopf-Hals-Sonographie mittels 10 Punkt VAS. 37

Abb. 5: Ergebnisdarstellung einer Ja/Nein- Umfrage zur Nutzung einer (a) Computer-basierten bzw. (b) Computer-basierten strukturierten Befunderhebung. Das blaue Segment repräsentiert den Anteil der Befragten die mit „Ja“ geantwortet haben, rot den Anteil der Befragten die mit „Nein“ geantwortet haben. Fünf der neun teilnehmenden Abteilungen gaben an eine Computer-basierte Befunderhebung im Allgemeinen zu benutzt ((a): 55,6 % „Ja“), jedoch verwendeten nur drei Abteilungen eine Computer-basierte strukturierte Befunderhebung ((b): 33,3 % „Ja“) (modifiziert nach Ernst et al.)..... 39

Abb. 6: Nachweis einer sehr hohen Befundvollständigkeit unter Verwendung einer strukturierten Befundvorlage (SR), bei allen 4 Fällen. 40

Abb. 7: Bestimmung der Interrater-Reliabilität (IRR) mittels Fleiss' Kappa: Die Verwendung einer strukturierter Befunderhebung (SR) führte in allen vier Fällen zu einer substanziellen bis nahezu perfekten IRR 41

Abb. 8: Ergebnisse der Interrater-Übereinstimmung (IRA) mit einem hohen mittleren Übereinstimmungsgrad von 87,2 % ± 15,1 %. Abgesehen von den Fällen 2 und 3 zeigten die analysierten Fälle keine signifikanten Unterschiede in der Höhe der IRA. * $p<0,05$ 42

Abb. 9: Evaluation der Benutzerzufriedenheit mittels visueller Analogskala (VAS, 0 = trifft voll zu bis hin zu 10 = trifft gar nicht zu). Die in blauen Balken dargestellten Ergebnisse des Fragebogens zeigen eine durchweg positive Rückmeldung zur Einführung und Nutzung der strukturierten Befunderhebung. Der durchschnittliche Mittelwert aller Fragen beträgt VAS 8,6, mit einer Standardabweichung von 1,8..... 43

7. Literaturverzeichnis

Adibelli ZH, Ünal G, Gül E, Uslu F, Koçak Ü, Abali Y. Differentiation of benign and malignant cervical lymph nodes: Value of B-mode and color Doppler sonography. Eur. J. Radiol. 1998; 28: 230–234

Al-Hawary MM, Francis IR, Chari ST, Fishman EK, Hough DM, Lu DS, Macari M, Megibow AJ, Miller FH, Mortele KJ, Merchant NB, Minter RM, Tamm EP, Sahani D V., Simeone DM. Pancreatic ductal adenocarcinoma radiology reporting template: consensus statement of the society of abdominal radiology and the american pancreatic association. Radiology. 2014; 146: 291–304

Austin JHM, Müller NL, Friedman PJ, Hansell DM, Naidich DP, Remy-Jardin M, Webb WR, Zerhouni EA. Glossary of terms for CT of the lungs: Recommendations of the Nomenclature Committee of the Fleischner Society. Radiology. 1996; 200: 327–331

Babu AS, Brooks ML. The Malpractice Liability of Radiology Reports: Minimizing the Risk. Radiographics. 2015; 35: 547–554

Bauer A, Lind K, Van Noort H, Myers M, Borgstede J. Ultrasound and Dual-Energy X-Ray Absorptiometry Report Transcription Error Rates and Strategies for Reduction. J Am Coll Radiol. 2018; 15: 1784–1790

Beach KW, Bergelin RO, Leotta DF, Primozich JF, Sevareid PM, Stutzman ET, Zierler RE. Standardized ultrasound evaluation of carotid stenosis for clinical trials: University of Washington Ultrasound Reading Center. Cardiovasc. Ultrasound. 2010; 8: 39

Benè KL, Bergus G. When learners become teachers: A review of peer teaching in medical student education. Fam. Med. 2014; 46: 783–787

Bialek EJ, Jakubowski W, Zajkowski P, Szopinski KT, Osmolski A. US of the Major Salivary Glands: Anatomy and Spatial Relationships, Pathologic Conditions, and Pitfalls. RadioGraphics [Internet]. 2006; 26: 745–763Available from: http://pubs.rsna.org/doi/10.1148/rg.263055024

Birrenbach S, Miller S, Stern W, Xydeas T, Pietsch-Breitfeld B, Belka C, Fersis N, Claussen CD, Müller-Schimpfle M. Klinischer stellenwert bildgebender verfahren im

ersten jahr nach brusterhaltender therapie des mammakarzinoms. RoFo Fortschritte auf dem Gebiet der Rontgenstrahlen und der Bildgebenden Verfahren. 2004; 176: 1423–1430

Bönhof JA. Ultrasound Artifacts – Part 1 Artefakte bei der Sonografie – Teil 1. Ultraschall in Med. 2016; 37: 140–153

Booth RAD, Goddard BA, Paton A. Measurement of fat thickness in man: a comparison of ultrasound, Harpenden calipers and electrical conductivity. Br J Nutr. 1966; 20: 719–725

Bosmans JML, 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 [Internet]. 2012; 3: 295–302Available from: http://link.springer.com/10.1007/s13244-012-0148-1

Brauer VFH, Eder P, Miehle K, Wiesner TD, Hasenclever H, Paschke R. Interobserver variation for ultrasound determination of thyroid nodule volumes. Thyroid. 2005; 15: 1169–1175

Cartwright MS, Demar S, Griffin LP, Balakrishnan N, Harris JM, Walker FO. Validity and reliability of nerve and muscle ultrasound. Muscle Nerve. 2013; 47: 515–521

Chiu LC, Leonardi M, Lu C, Mein B, Nadim B, Reid S, Ludlow J, Casikar I, Condous G. Predicting Pouch of Douglas Obliteration Using Ultrasound and Laparoscopic Video Sets: An Interobserver and Diagnostic Accuracy Study. J. Ultrasound Med. 2019; 38: 3155–3161

Cho SH, Cho YS, Choi IY, Ha H Il, Huh J, Hur BY, et al. Essential Items for Structured Reporting of Rectal Cancer MRI: 2016 Consensus Recommendation from the Korean Society of Abdominal Radiology. Korean J. Radiol. [Internet]. 2017; 18: 132–151Available from: https://www.kjronline.org/DOIx.php?id=10.3348/kjr.2017.18.1.132

Coelho Neto MA, Roncato P, Nastri CO, Martins WP. True Reproducibility of UltraSound Techniques (TRUST): Systematic review of reliability studies in obstetrics and gynecology. Ultrasound in Obstetrics and Gynecology. 2015: 14–20

Cohen J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement. Educ. Psychol. Meas. 1960; 20: 37–46

Coltrera MD. Clinician-Performed Thyroid Ultrasound. Otolaryngol. Clin. North Am. 2014; 47: 491–507

Cramer JA, Eisenmenger LB, Pierson NS, Dhatt HS, Heilbrun ME. Structured and Templated reporting: An overview. Appl. Radiol. 2014; 43

Deutsche Gesellschaft für Ultraschall in der Medizin. Satzung der Deutschen Gesellschaft für Ultraschall in der Medizin e.V. [Internet]. 2022:Available from: https://www.degum.de/fileadmin/dokumente/2022-DEGUM-Satzung_v29.06.2022.pdf

Dochow C. Telemedizin und Datenschutz. Medizinrecht [Internet]. 2019; 37: 636–648Available from: http://link.springer.com/10.1007/s00350-019-5295-7

D'Orsi CJ, Kopans DB. Mammography interpretation: the BI-RADS method. Am. Fam. Physician. 1997: 1548–1550

Dössel O. Bildgebende Verfahren in der Medizin. Bildgebende Verfahren in der Medizin. Springer Berlin Heidelberg, 2016:

Dunnick NR, Langlotz CP. The Radiology Report of the Future: A Summary of the 2007 Intersociety Conference. J Am Coll Radiol. 2008: 626–629

Dussik KT, Fritch DJ, Kyriazidou M, Sear RS. Measurements of articular tissues with ultrasound. Am. J. Phys. Med. Rehabil. 1958; 37: 160–165

Eghtedari M, Chong A, Rakow-Penner R, Ojeda-Fournier H. Current Status and Future of BI-RADS in Multimodality Breast Imaging, From the AJR Special Series on Radiology Reporting and Data Systems. AJR. Am. J. Roentgenol. 2020; 216: 860–873

Ekberg O, Olsson R. Dynamic Radiology of Swallowing Disorders. Endoscopy [Internet]. 1997; 29: 439–446Available from: http://www.thieme-connect.de/DOI/DOI?10.1055/s-2007-1004248

Ellis DW, Srigley J. Does standardised structured reporting contribute to quality in diagnostic pathology? The importance of evidence-based datasets. Virchows Archiv. 2016: 51–59

Enders J, Zimmermann E, Rief M, Martus P, Klingebiel R, Asbach P, Klessen C, Diederichs G, Wagner M, Teichgraber U, Bengner T, Hamm B, Dewey M. Reduction of

claustrophobia with short-bore versus open magnetic resonance imaging: a randomized controlled trial. PLoS One. 2011a; 6: e23494

Enders J, Zimmermann E, Rief M, Martus P, Klingebiel R, Asbach P, Klessen C, Diederichs G, Wagner M, Teichgräber U, Bengner T, Hamm B, Dewey M. Reduction of claustrophobia with short-bore versus open magnetic resonance imaging: A randomized controlled trial. PLoS One. 2011b: e23494

Ernst BP, Hodeib M, Strieth S, Künzel J, Bischof F, Hackenberg B, Huppertz T, Weber V, Bahr K, Eckrich J, Hagemann J, Engelbarts M, Froelich MF, Solbach P, Linke R, Matthias C, Sommer WH, Becker S. Structured reporting of head and neck ultrasound examinations. BMC Med. Imaging. 2019a; 19: 25

Ernst BP, Katzer F, Künzel J, Hodeib M, Strieth S, Eckrich J, Tattermusch A, Froelich MF, Matthias C, Sommer WH, Becker S. Impact of structured reporting on developing head and neck ultrasound skills. BMC Med. Educ. 2019b; 19: 102

Ernst BP, Reissig MR, Strieth S, Eckrich J, Hagemann JH, Döge J, Matthias C, Gouveris H, Rübenthaler J, Weiss R, Sommer WH, Nörenberg D, Huber T, Gonser P, Becker S, Froelich MF. The role of structured reporting and structured operation planning in functional endoscopic sinus surgery. PLoS One. 2020a; 15: e0242804

Ernst BP, Strieth S, Katzer F, Hodeib M, Eckrich J, Bahr K, Rader T, Künzel J, Froelich MF, Matthias C, Sommer WH, Becker S. The use of structured reporting of head and neck ultrasound ensures time-efficiency and report quality during residency. European Archives of Oto-Rhino-Laryngology [Internet]. 2020b; 277: 269–276Available from: http://link.springer.com/10.1007/s00405-019-05679-z

Ernst BP, Strieth S, Künzel J, Hodeib M, Katzer F, Eckrich J, Bahr K, Matthias C, Sommer WH, Froelich MF, Becker S. Evaluation of optimal education level to implement structured reporting into ultrasound training. Med. Ultrason. 2020c; 22: 281–287

European Society of Radiology (ESR). Good practice for radiological reporting. Guidelines from the European Society of Radiology (ESR). Insights Imaging [Internet]. 2011; 2: 93–96Available from: https://insightsimaging.springeropen.com/articles/10.1007/s13244-011-0066-7

European Society of Radiology (ESR). ESR paper on structured reporting in radiology. Insights Imaging [Internet]. 2018; 9: 1–7Available from: https://insightsimaging.springeropen.com/articles/10.1007/s13244-017-0588-8

European Society of Radiology (ESR). ESR paper on structured reporting in radiology—update 2023. Insights Imaging. 2023; 14: 199

Faggioni L, Coppola F, Ferrari R, Neri E, Regge D. Usage of structured reporting in radiological practice: results from an Italian online survey. Eur. Radiol. 2017; 27: 1934–1943

Fardon DF, Milette PC. Nomenclature and classification of lumbar disc pathology. Recommendations of the Combined task Forces of the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. Spine (Phila. Pa. 1976). 2001; 26: E93–E113

Fleury EFC, Gianini AC, Marcomini K, Oliveira V. The feasibility of classifying breast masses using a computer-assisted diagnosis (CAD) system based on ultrasound elastography and BI-RADS lexicon. Technol. Cancer Res. Treat. 2018; 17: 1–9

Forghani R, Yu E, Levental M, Som PM, Curtin HD. Imaging evaluation of lymphadenopathy and patterns of lymph node spread in head and neck cancer. Expert Rev. Anticancer Ther. 2015: 207–224

Gandhi TK, Kachalia A, Thomas EJ, Puopolo AL, Yoon C, Brennan TA, Studdert DM. Missed and delayed diagnoses in the ambulatory setting: A study of closed malpractice claims. Ann. Intern. Med. 2006; 145: 488–496

Ganeshan D, Duong PAT, Probyn L, Lenchik L, McArthur TA, Retrouvey M, Ghobadi EH, Desouches SL, Pastel D, Francis IR. Structured Reporting in Radiology. Acad. Radiol. 2018: 66–73

Gassenmaier S, Armbruster M, Haasters F, Helfen T, Henzler T, Alibek S, Pförringer D, Sommer WH, Sommer NN. Structured reporting of MRI of the shoulder – improvement of report quality? Eur. Radiol. [Internet]. 2017; 27: 4110–4119Available from: http://link.springer.com/10.1007/s00330-017-4778-z

Gates J, Hartnell GG, Gramigna GD. Videofluoroscopy and Swallowing Studies for Neurologic Disease: A Primer. RadioGraphics [Internet]. 2006; 26: e22–e22Available from: http://pubs.rsna.org/doi/10.1148/rg.e22

Gesundheit B für. Elektronische Patientenakte (ePA) für alle [Internet]. 2025:Available from: https://www.bundesgesundheitsministerium.de/themen/digitalisierung/elektronische-patientenakte/epa-fuer-alle.html

Ghoshhajra BB, Lee AM, Ferencik M, Elmariah S, Margey RJP, Onuma O, Panagia M, Abbara S, Hoffmann U. Interpreting the interpretations: The use of structured reporting improves referring clinicians' comprehension of coronary CT angiography reports. J Am Coll Radiol. 2013; : 432–438

GKV-Spitzenverband. Vereinbarung von Qualitätssicherungsmaßnahmen nach § 135 Abs. 2 SGB V zur Ultraschalldiagnostik (Ultraschall-Vereinbarung) [Internet]. Berlin, 2025aAvailable from: https://www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/aerztliche_versorgung/qualitaetssicherung/qualitaetssicherung_1/qs_sonstige_vereinbarungen/2025-04-01_Ultraschall.pdf

GKV-Spitzenverband. Vereinbarung von Qualitätssicherungsmaßnahmen nach § 135 Abs. 2 SGB V zur Ultraschalldiagnostik (Ultraschall-Vereinbarung), § 1 [Internet]. 2025bAvailable from: https://www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/aerztliche_versorgung/qualitaetssicherung/qualitaetssicherung_1/qs_sonstige_vereinbarungen/2025-04-01_Ultraschall.pdf

GKV-Spitzenverband. Vereinbarung von Qualitätssicherungsmaßnahmen nach § 135 Abs. 2 SGB V zur Ultraschalldiagnostik (Ultraschall-Vereinbarung), § 10 Abs. 2 und 3. 2025c

GKV-Spitzenverband. Vereinbarung von Qualitätssicherungsmaßnahmen nach § 135 Abs. 2 SGB V zur Ultraschalldiagnostik (Ultraschall-Vereinbarung), § 11 Abs. 4 a–d. 2025d

Gmeinwieser J, Golder W, Lehner K. [Quality assurance in sonography]. Rontgenpraxis [Internet]. 1988; 41: 279–282Available from: http://www.ncbi.nlm.nih.gov/pubmed/3051440

Gunderman RB, McNeive LR. Is structured reporting the answer? Radiology. 2014: 7–9

Gunn AJ, Alabre CI, Bennett SE, Kautzky M, Krakower T, Palamara K, Choy G. Structured feedback from referring physicians: A novel approach to quality improvement in radiology reporting. AJR Am J Roentgenol. 2013; 201: 853–857

Günter R. Frank Wenzel (Hrsg.), Handbuch des Fachanwalts Medizinrecht. Medizinrecht. 2013;

Hagmann CF, Halbherr M, Koller B, Wintermark P, Huisman T, Bucher HU. Interobserver variability in assessment of cranial ultrasound in very preterm infants. J. Neuroradiol. 2011; 38: 291–297

Harding U, Goeters C, Schmidt C. Ultraschall in der Anästhesie und Intensivmedizin - Ultraschall in der Intensivmedizin. Anasthesiologie Intensivmedizin Notfallmedizin Schmerztherapie. 2011:

Hawkins CM, Hall S, Zhang B, Towbin AJ. Creation and Implementation of Department-Wide Structured Reports: An Analysis of the Impact on Error Rate in Radiology Reports. J Digit Imaging. 2014; 27: 581–587

Heese F, Görg C. The Value of Highest Quality Ultrasound as a Reference for Ultrasound Diagnosis. UiM/EJU. 2006; 27: 220–224

Heikkinen K, Löyttyniemi M, Kormano M. Structure and content of 400 CT reports in four teaching hospitals using a new, windows-based software tool. Acta radiol. 2000; 41: 102–105

Heller G, Dörges V. Stellungnahme zur “Präklinischen Sonographie.” Notfall und Rettungsmedizin. 2006; 8

Heyduck A, Jecker P, Bozzato A. A german-wide inquiry about the significance of ultrasound in the head and neck area. Laryngorhinootologie. 2018; 97: 392–397

Hoffmann W, van den Berg N. Morbiditätsentwicklung im demographischen Wandel. Verfügbarkeit medizinischer Leistungen in der regionalen Versorgung. Chirurg. 2013; 84: 286–290

Johnson AJ, Chen MYM, Swan JS, Applegate KE, Littenberg B. Cohort Study of Structured Reporting Compared with Conventional Dictation. Radiology [Internet]. 2009; 253: 74–80Available from: http://pubs.rsna.org/doi/10.1148/radiol.2531090138

Kahn CE, Heilbrun ME, Applegate KE. From guidelines to practice: How reporting templates promote the use of radiology practice guidelines. J Am Coll Radiol. 2013; 10: 268–273

Karim S, Fegeler C, Boeckler D, H Schwartz L, Kauczor H-U, von Tengg-Kobligk H. Development, Implementation, and Evaluation of a Structured Reporting Web Tool for Abdominal Aortic Aneurysms. JMIR Res. Protoc. 2013; 2: e30

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: 685–688

Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, Roberts C, Shoukri M, Streiner DL. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. Int. J. Nurs. Stud. 2011; 48: 661–671

Krupinski EA, Lund PJ. Differences in Time to Interpretation for Evaluation of Bone Radiographs with Monitor and Film Viewing. Acad. Radiol. 1997; 4: 177–182

Künzel J, Bozzato A, Strieth S. Sonographie in der Nachsorge bei Kopf- und Halskarzinomen. HNO [Internet]. 2017; 65: 939–952Available from: http://link.springer.com/10.1007/s00106-017-0411-3

Lacerda TC, von Wangenheim CG, von Wangenheim A, Giuliano I. Does the use of structured reporting improve usability? A comparative evaluation of the usability of two approaches for findings reporting in a large-scale telecardiology context. J. Biomed. Inform. 2014; 52: 222–230

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics [Internet]. 1977; 33: 159–174Available from: http://www.ncbi.nlm.nih.gov/pubmed/843571

Langlotz CP. RadLex: a new method for indexing online educational materials. Radiographics [Internet]. 2006; 26: 1595–1597Available from: http://www.ncbi.nlm.nih.gov/pubmed/17102038

Larson CM. Editorial Commentary: Ultrasound-Guided Injection With Dynamic Assessment Could Improve Outcomes in Atypical and Revision Hip Arthroscopy Patients: But Today, Few Are Trained to Hear at Such High Frequencies. Arthroscopy. 2021: 136–138

Larson DB. Strategies for implementing a standardized structured radiology reporting program. Radiographics. 2018; 38: 1705–1716

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: 240–250

LeBreton JM, Senter JL. Answers to 20 questions about interrater reliability and interrater agreement. Organ. Res. Methods. 2008; 11

Lee B, Whitehead MT. Radiology Reports: What YOU Think You’re Saying and What THEY Think You’re Saying. Curr. Probl. Diagn. Radiol. [Internet]. 2017; 46: 186–195Available from: http://www.ncbi.nlm.nih.gov/pubmed/28069356

Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft DK, AWMF). S3-Leitlinie Diagnostik und Therapie des Mundhöhlenkarzinoms. Langversion 3.0. [Internet]. Berlin, 2021Available from: www.awmf.org/uploads/tx_szleitlinien/007-100OLl_S3-Diagnostik-Therapie-Mundhoehlenkarzinom_2021-03.pdf

Lima PH, Fan B, Bérubé J, Cerny M, Olivié D, Giard JM, Beauchemin C, Tang A. Cost-utility analysis of imaging for surveillance and diagnosis of hepatocellular carcinoma. AJR Am J Roentgenol. 2019; 213: 17–25

Lux T, Breil B. Digitalisierung im Gesundheitswesen: bessere Versorgungsqualität trotz Kosteneinsparungen. Wirtschaftsdienst. 2017; 97: 687–703

Lyu Q, Tan J, Zapadka ME, Ponnatapura J, Niu C, Myers KJ, Wang G, Whitlow CT. Translating radiology reports into plain language using ChatGPT and GPT-4 with prompt learning: results, limitations, and potential. Vis. Comput. Ind. Biomed. Art. 2023; 6: 9

Mallio CA, Sertorio AC, Bernetti C, Beomonte Zobel B. Large language models for structured reporting in radiology: performance of GPT-4, ChatGPT-3.5, Perplexity and Bing. Radiologia Medica. 2023a; 128: 808–812

Mallio CA, Sertorio AC, Bernetti C, Beomonte Zobel B. Large language models for structured reporting in radiology: performance of GPT-4, ChatGPT-3.5, Perplexity and Bing. Radiologia Medica. 2023b; 128: 808–812

Martin-Harris B, Jones B. The Videofluorographic Swallowing Study. Phys. Med. Rehabil. Clin. N. Am. [Internet]. 2008; 19: 769–785Available from: https://linkinghub.elsevier.com/retrieve/pii/S1047965108000594

Matsumoto M, Koike S, Kashima S, Awai K, Ojima T. Geographic distribution of radiologists and utilization of teleradiology in Japan: A longitudinal analysis based on national census data. PLoS One. 2015; 10: e0139723

McHugh ML. Interrater reliability: the kappa statistic. Biochem Med [Internet]. 2012; 22: 276–282Available from: http://www.biochemia-medica.com/en/journal/22/3/10.11613/BM.2012.031

Meudt R, Buser Ch, Bösch A. Untersuchungskonzept für die sonographische Organdiagnostik - Ein Beitrag zur Qualitätssicherung. Ultraschall in der Medizin. 1986; 7: 54–58

Miao J, Thongprayoon C, Suppadungsuk S, Garcia Valencia OA, Cheungpasitporn W. Integrating Retrieval-Augmented Generation with Large Language Models in Nephrology: Advancing Practical Applications. Medicina (Lithuania). 2024: 445

Mistry SG, Strachan DR, Loney EL. Improving paranasal sinus computed tomography reporting prior to functional endoscopic sinus surgery - An ENT-UK panel perspective. J Laryngol Otol. 2016: 962–966

Mojoli F, Bouhemad B, Mongodi S, Lichtenstein D. Lung ultrasound for critically ill patients. Am. J. Respir. Crit. Care Med. 2019: 701–714

Morgan TA, Helibrun ME, Kahn CE. Reporting initiative of the radiological society of north america: Progress and new directions. Radiology. 2014: 642–645

Moshtaghi O, Haidar YM, Mahmoodi A, Tjoa T, Armstrong WB. The Role of In-Office Ultrasound in the Diagnosis of Neck Masses. Otolaryngol. Head Neck Surg. 2017; 157: 58–61

Müller W, Lercher K, DeVaney TTJ, Giner F, Graf R. Technical errors in the application of Graf's hip sonography method. Ultraschall in der Medizin. 2001; 22: 48–54

Nasrallah HA. Consequences of misdiagnosis: inaccurate treatment and poor patient outcomes in bipolar disorder. J Clin Psychiatry. 2015: e1328

Nobel JM, Kok EM, Robben SGF. Redefining the structure of structured reporting in radiology. Insights Imaging. 2020; 11: 10

Nörenberg D, Sommer WH, Thasler W, D'Haese J, Rentsch M, Kolben T, Schreyer A, Rist C, Reiser M, Armbruster M. Structured Reporting of Rectal Magnetic Resonance Imaging in Suspected Primary Rectal Cancer. Invest Radiol [Internet]. 2017; 52: 232–239Available from: http://journals.lww.com/00004424-201704000-00006

O'Dell MC, Cassady C, Logsdon G, Varich L. Cinegraphic versus Combined Static and Cinegraphic Imaging for Initial Cranial Ultrasound Screening in Premature Infants. Pediatr. Radiol. 2015; 45: 1706–1711

Ott DJ, Pikna LA. Clinical and videofluoroscopic evaluation of swallowing disorders. AJR Am J Roentgenol [Internet]. 1993; 161: 507–513Available from: http://www.ajronline.org/doi/10.2214/ajr.161.3.8352094

Park S Bin, Kim M-J, Ko Y, Sim JY, Kim HJ, Lee KH. Structured Reporting versus Free-Text Reporting for Appendiceal Computed Tomography in Adolescents and Young Adults: Preference Survey of 594 Referring Physicians, Surgeons, and Radiologists from 20 Hospitals. Korean J. Radiol. [Internet]. 2019; 20: 246Available from: https://www.kjronline.org/DOIx.php?id=10.3348/kjr.2018.0109

Pinto Dos Santos D, Baeßler B. Big data, artificial intelligence, and structured reporting. Eur. Radiol. Exp. [Internet]. 2018; 2: 42Available from: http://www.ncbi.nlm.nih.gov/pubmed/30515717

Plewis I, Fleiss JL. Statistical Methods for Rates and Proportions. J. R. Stat. Soc. Ser. A (Gen.). 1982;

Pohlmann R. Ultraschall in der Medizin. Grenzgeb. Med. 1948; 1: 148–150

Powell DK, Silberzweig JE. State of Structured Reporting in Radiology, a Survey. Acad. Radiol. 2015; 22: 226–233

Prasad R, Chen B. Imaging Evaluation of the Head and Neck Oncology Patient. Cancer Treat. Res. [Internet]. 2018: 59–86Available from: http://link.springer.com/10.1007/978-3-319-65421-8_5

Ranschaert ER, Binkhuysen FHB. European Teleradiology now and in the future: Results of an online survey. Insights Imaging. 2013; 4: 93–102

Reiner BI. The challenges, opportunities, and imperative of structured reporting in medical imaging. J Digit Imaging [Internet]. 2009; 22: 562–568Available from: http://www.ncbi.nlm.nih.gov/pubmed/19816742

Reiner BI, Knight N, Siegel EL. Radiology Reporting, Past, Present, and Future: The Radiologist's Perspective. J Am Coll Radiol. 2007; 4: 313–319

Reinert S. Ultraschall in der Mund-, Kiefer- und GesichtschirurgieUltrasonography in oromaxillofacial surgery. Mund-, Kiefer- und Gesichtschirurgie. 2000; 1: 375–383

Robbins KT. Integrating radiological criteria into the classification of cervical lymph node disease. Archives of Otolaryngology - Head and Neck Surgery. 1999: 385–387

Ross P, Sepper R, Pohjonen H. Cross-border teleradiology-Experience from two international teleradiology projects. Eur. J. Radiol. 2010; 3: 20–25

Sabel B, Plum J, Kneidinger N, Leuschner G, Koletzko L, Raziorrouh B, Schinner R, Schoeppe F, Thierfelder K, Sommer W, Meinel F. Structured reporting of CT examinations in acute pulmonary embolism. J. Cardiovasc. Comput. Tomogr. 2017a; 11: 188–195

Sabel BO, Plum JL, Kneidinger N, Leuschner G, Koletzko L, Raziorrouh B, Schinner R, Kunz WG, Schoeppe F, Thierfelder KM, Sommer WH, Meinel FG. Structured reporting of CT examinations in acute pulmonary embolism. J. Cardiovasc. Comput. Tomogr. [Internet]. 2017b; 11: 188–195Available from: http://www.ncbi.nlm.nih.gov/pubmed/28259629

Schäberle W. Ultraschall in der Gefäßdiagnostik. Ultraschall in der Gefäßdiagnostik. 2016:

Schoeppe F, Sommer WH, Haack M, Havel M, Rheinwald M, Wechtenbruch J, Fischer MR, Meinel FG, Sabel BO, Sommer NN. Structured reports of videofluoroscopic swallowing studies have the potential to improve overall report quality compared to free text reports. Eur. Radiol. [Internet]. 2018; 28: 308–315Available from: http://link.springer.com/10.1007/s00330-017-4971-0

Schöppe F, Sommer WH, Schmidutz F, Pförringer D, Armbruster M, Paprottka KJ, Plum JLV, Sabel BO, Meinel FG, Sommer NN. Structured reporting of x-rays for atraumatic shoulder pain: Advantages over free text? BMC Med. Imaging. 2018; 18: 20

Schüle B, Wissel H, Neumann W, Merk H. Verlaufskontrollen von Huftbefunden im sonographischen Neugeborenenscreening. Ultraschall in der Medizin. 1999; 20: 161–164

Schuler A, Reuss J, Delorme S, Hagendorff A, Giesel F. Costs of clinical ultrasound examinations - an economical cost calculation and analysis. Ultraschall in der Medizin. 2010; 31: 379–386

Schwartz LH, Panicek DM, Berk AR, Li Y, Hricak H. Improving communication of diagnostic radiology findings through structured reporting. Radiology [Internet]. 2011; 260: 174–181Available from: http://www.ncbi.nlm.nih.gov/pubmed/21518775

Scott TE, Koziol L, Marshall L, Falconi C. Diploma and Degree. J Diagn Med Sonogr [Internet]. 2005; 21: 62–69Available from: http://journals.sagepub.com/doi/10.1177/8756479304272680

Shampo MA, Kyle RA. Karl Theodore Dussik—Pioneer in Ultrasound. Mayo Clin. Proc. [Internet]. 1995; 70: 1136Available from: https://linkinghub.elsevier.com/retrieve/pii/S002561961163437X

Shore MW, Rubin DL, Kahn CE. Integration of imaging signs into RadLex. J. Digit. Imaging. 2012; 25: 50–55

Sickles EA, Appleton CM, Burnside ES, Gavenonis SC. Breast imaging reporting and data system (BI-RADS) Atlas- Ultrasound 5th edn. American College of Radiology BI-RADS-Atlas. 2013: 1921

Sinitsyn VE, Komarova MA, Mershina EA. [Radiology report: past, present and future]. Vestn. Rentgenol. Radiol. [Internet]. 2014; : 35–40Available from: http://www.ncbi.nlm.nih.gov/pubmed/25782296

Sistrom C, Lanier L, Mancuso A. Reporting Instruction for Radiology Residents. Acad. Radiol. 2004; 11: 76–84

Sistrom CL, Honeyman-Buck J. Free text versus structured format: Information transfer efficiency of radiology reports. AJR. Am. J. Roentgenol. 2005; 185: 804–812

Sluijter CE, van Lonkhuijzen LRCW, van Slooten H-J, Nagtegaal ID, Overbeek LIH. The effects of implementing synoptic pathology reporting in cancer diagnosis: a systematic review. Virchows Arch. [Internet]. 2016; 468: 639–649Available from: http://www.ncbi.nlm.nih.gov/pubmed/27097810

Sobez LM, Kim SH, Angstwurm M, Störmann S, Pförringer D, Schmidutz F, Prezzi D, Kelly-Morland C, Sommer WH, Sabel B, Nörenberg D, Berndt M, Galiè F. Creating high-quality radiology reports in foreign languages through multilingual structured reporting. Eur. Radiol. 2019; 29: 6038–6048

Srigley J, Lankshear S, Brierley J, McGowan T, Divaris D, Yurcan M, Rossi R, Yardley T, King MJ, Ross J, Irish J, McLeod R, Sawka C. Closing the quality loop: Facilitating improvement in oncology practice through timely access to clinical performance indicators. JCO Oncol. Pract. 2013; 9: e255-61

Steinijans VW, Diletti E, Bömches B, Greis C, Solleder P. Interobserver agreement: Cohen's kappa coefficient does not necessarily reflect the percentage of patients with congruent classifications. Int. J. Clin. Pharmacol. Ther. 1997; 35: 93–95

Szilveszter B, Kolossváry M, Karády J, Jermendy ÁL, Károlyi M, Panajotu A, Bagyura Z, Vecsey-Nagy M, Cury RC, Leipsic JA, Merkely B, Maurovich-Horvat P. Structured reporting platform improves CAD-RADS assessment. J. Cardiovasc. Comput. Tomogr. 2017a; 11: 449–454

Szilveszter B, Kolossváry M, Karády J, Jermendy ÁL, Károlyi M, Panajotu A, Bagyura Z, Vecsey-Nagy M, Cury RC, Leipsic JA, Merkely B, Maurovich-Horvat P. Structured reporting platform improves CAD-RADS assessment. J. Cardiovasc. Comput. Tomogr. 2017b; 11: 449–454

Takes RP, Halmos GB, Ridge JA, Bossi P, Merkx MAW, Rinaldo A, Sanabria A, Smeele LE, Mäkitie AA, Ferlito A. Value and Quality of Care in Head and Neck Oncology. Curr. Oncol. Rep. [Internet]. 2020; 22: 92Available from: http://www.ncbi.nlm.nih.gov/pubmed/32651680

Timor-Tritsch IE, Goldstein SR. The complexity of a “complex mass” and the simplicity of a “simple cyst.” J. Ultrasound Med. 2005: 255–258

Tuncyurek O, Garces-Descovich A, Jaramillo-Cardoso A, Durán EE, Cataldo TE, Poylin VY, Gómez SF, Cabrera AM, Hegazi T, Beker K, Mortele KJ. Structured versus narrative reporting of pelvic MRI in perianal fistulizing disease: impact on clarity, completeness, and surgical planning. Abdominal Radiology. 2019; 44: 811–820

Wallis A, McCoubrie P. The radiology report — Are we getting the message across? Clin. Radiol. [Internet]. 2011; 66: 1015–1022Available from: https://linkinghub.elsevier.com/retrieve/pii/S0009926011002388

Weiss DL, Bolos PR. Reporting and dictation. Practical Imaging Informatics: Foundations and Applications for PACS Professionals. 2009: 147–162

Weiss DL, Langlotz CP. Structured reporting: Patient care enhancement or productivity nightmare? Radiology. 2008; 249: 739–747

Welkoborsky HJ. Ultrasound usage in the head and neck surgeon’s office. Curr. Opin. Otolaryngol. Head Neck Surg. 2009: 116–121

Welkoborsky H-J, Maurer J, Klimek L, Mann W. Neueste Entwicklungen in der Ultraschalldiagnostik - Teil 1: Neue technische Entwicklungen; - Ultraschalluntersuchung der Weichteilstrukturen des Halses. Laryngo-Rhino-Otologie. 1997; 76: 651–652

Wiegand S, Dietz A. Qualität in der Kopf-Hals-Onkologie. Laryngo-Rhino-Otologie [Internet]. 2020; 99: S60–S106Available from: http://www.thieme-connect.de/DOI/DOI?10.1055/a-1012-9396

Youk JH, Jung I, Yoon JH, Kim SH, Kim YM, Lee EH, Jeong SH, Kim MJ. Comparison of Inter-Observer Variability and Diagnostic Performance of the Fifth Edition of BI-RADS for Breast Ultrasound of Static versus Video Images. Ultrasound Med. Biol. 2016; 42: 2083–2088

8. Erklärung zum Eigenanteil

Die vorliegende Arbeit wurde an der Klinik und Poliklinik für Hals-, Nasen- und Ohrenheilkunde des Universitätsklinikums Bonn unter Betreuung von PD Dr. med. Benjamin Philipp Ernst durchgeführt.

Die Konzeption der Studie erfolgte durch PD Dr. med. Benjamin Philipp Ernst. Nach einer Einarbeitung in die Methodik der Kopf-Hals-Sonographie sowie das an der Klinik vorhandenen Ultraschallgerät (GE LOGIQ S8) habe ich die Datenerhebung eigenständig durchgeführt. Diese umfasste zunächst die Anfertigung und Dokumentation standardisierter Untersuchungen an einer gesunden Kontrollperson sowie anschließend die Durchführung und Dokumentation von Kopf-Hals-Sonographien bei Patienten, bei denen die Untersuchung ohnehin im Rahmen der klinischen Routine indiziert war.

Das erhobene Bild- und Videomaterial wurde von mir anonymisiert, aufbereitet und zu standardisierten Sequenzen zusammengestellt. Basierend auf in unserer Arbeitsgruppe etablierten Konzepten habe ich spezifische Auswertungschecklisten für die erhobenen Fälle erstellt und nach Rücksprache mit meinem Betreuer angewendet. Deren Durchführung sowie die Auswertung der erhobenen Daten, einschließlich der Analyse der Vollständigkeit der Befunde und der Erhebung der Interrater-Reliabilität, habe ich eigenständig übernommen.

Darüber hinaus habe ich einen in der Arbeitsgruppe bereits etablierten Fragebogen mit visueller Analogskala (VAS) an die aktuelle Studie angepasst und die eingegangenen Daten selbständig ausgewertet. Die Interpretation der Ergebnisse sowie die schriftliche Ausarbeitung der vorliegenden Dissertation erfolgten in enger Abstimmung mit meinem Betreuer PD Dr. med. Benjamin Philipp Ernst, jedoch selbständig von mir.

Ich versichere, die Dissertationsschrift selbständig verfasst zu haben und keine weiteren als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

9. Veröffentlichungen

Teile der Inhalte der vorliegenden Dissertationsschrift wurden bereits in der folgenden Publikation veröffentlicht:

Ernst BP, Dörsching C, Bozzato A, Gabrielpillai J, Becker S, Froelich MF, Kramer B, Sproll C, Schapher M, Goncalves M, Mansour N, Hofauer B, Sommer WH, von Scotti F, Weimer JM, Künzel J. Structured reporting of head and neck sonography achieves substantial interrater reliability. Ultrasound Int Open 2023; 9: E26–E32.

https://doi.org/10.1055/a-2173-3966

10. Danksagung

Wenn ich nun zurückblicke, hätte ich zwischenzeitlich fast selbst nicht mehr daran geglaubt, dass diese Arbeit einmal zu Ende geführt wird. Umso mehr freue ich mich, dass es nun so weit ist – und es ist vor allem an der Zeit, Danke zu sagen.

Mein besonderer Dank gilt meinem Doktorvater PD Dr. med. Benjamin Ernst. Danke für deine Zeit, deine Geduld und deine konstruktive Kritik. Danke auch dafür, dass du mir die Möglichkeit gegeben hast, dieses Projekt zu bearbeiten, mich stets mit Rat und Tat unterstützt und mir mit deiner positiven Art geholfen hast, den roten Faden nicht zu verlieren. Deine ehrliche Meinung, dein Vertrauen und deine große fachliche Kompetenz haben entscheidend dazu beigetragen, dass diese Arbeit gelingen konnte. Ohne dich wäre dieses Projekt nicht möglich gewesen.

Ein herzlicher Dank gilt auch meinen Freunden, die mich in dieser Zeit begleitet, mich immer wieder motiviert und mir durch ihr offenes Ohr Rückhalt gegeben haben. Ihr habt so sicherlich wesentlich zum Gelingen dieser Arbeit beigetragen.

Meinen Schwiegereltern Angelika und Rainer Dörsching danke ich für ihre große Hilfsbereitschaft. Ihr haltet mir und uns immer den Rücken frei und springt ein, wann immer es nötig ist.

Ganz besonders möchte ich mich bei meinen Eltern Martina und Martin Eckrich bedanken. Ihr seid immer an meiner Seite, habt mich motiviert und stets an meinen beruflichen Weg geglaubt. Ohne euch wäre all das nicht möglich gewesen.

Danke auch an meinen Bruder Jonas. Du bist und bleibst ein großes Vorbild – nicht nur für mich, sondern auch für deine Nichte und deinen Neffen. Du bist ein toller Arzt, ein Ratgeber und hast viele Hürden meiner Aus- und Weiterbildungszeit mit mir gemeistert.

An letzter, doch eigentlich an erster Stelle möchte ich meinem Ehemann Timo Dörsching danken. Danke, dass es dich gibt, dass du mir zuhörst, an mich glaubst und all das mit mir meisterst – auch wenn diese Reise durch Corona, zwei Kinder, Hochzeit und Hausumbau nicht einfacher wurde. Danke für deine Motivation, deine Ehrlichkeit, deine Aufmunterungen und deine fachliche Unterstützung.

Frieda und Emil – ihr seid meine größte Motivation und mein größter Stolz. Diese Arbeit ist auch für euch.