

# Strukturierte Befundung zur effizienten Erhebung von epidemiologischen Daten und Analyse der innerklinischen Prävalenz bei Lungenarterienembolien

## Structured reporting for efficient epidemiological and in-hospital prevalence analysis of pulmonary embolisms

Autorinnen/Autoren

Tobias Jorg<sup>1</sup>, Moritz C Halfmann<sup>1</sup>, Dirk Graafen<sup>1</sup>, Lukas Hobohm<sup>2</sup>, Christoph Düber<sup>1</sup>, Peter Mildenerger<sup>1</sup>, Lukas Müller<sup>1</sup>

### Institute

- 1 Department of Diagnostic and Interventional Radiology, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany
- 2 Center for Cardiology, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany

### Schlüsselwörter

Epidemiologie, CT-Angiographie, Embolie/Thrombose

### Keywords

epidemiology, CT-angiography, embolism/thrombosis

eingereicht 20.9.2023

akzeptiert nach Revision 26.3.2024

Artikel online veröffentlicht 2024

### Bibliografie

Fortschr Röntgenstr

DOI 10.1055/a-2301-3349

ISSN 1438-9029

© 2024, Thieme. All rights reserved.

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

### Korrespondenzadresse

Dr. Tobias Jorg

Department of Diagnostic and Interventional Radiology, University Medical Center of the Johannes Gutenberg University Mainz, Langenbeckstr. 1, 55131 Mainz, Germany  
tobias.jorg@unimedizin-mainz.de

 English version at: <https://doi.org/10.1055/a-2301-3349>.

### ZUSAMMENFASSUNG

**Ziel** Die strukturierte Befundung (Structured Reporting; SR) bietet neben Vorteilen hinsichtlich der Befundqualität als IT-basierte Methode auch die Möglichkeit umfangreiche, hochstrukturierte Datenmengen zu erheben und diese sekundär auszuwerten (Data-Mining). Im Rahmen dieser Studie wurde ein Data-Mining-Algorithmus genutzt, um epidemiologische Daten und innerklinische Prävalenzstatistiken der Lun-

genarterienembolie (LAE) durch Auswertung strukturierter CT-Befunde zu erheben.

**Methoden** Alle strukturierten Befunde zu LAE-CTs der letzten 5 Jahre (n = 2790) wurden aus der SR-Datenbank extrahiert und ausgewertet. Die Prävalenz einer LAE wurde für die Gesamtkohorte und in Abhängigkeit der Zuweisungsart und in Abhängigkeit des klinischen Zuweisers berechnet. Verteilungen der Lokalisationen der LAEs (zentral, lobär, segmental, subsegmental sowie linksseitig, rechtsseitig, beidseitig) wurden analysiert sowie das Auftreten einer Rechtsherzbelastung mit der Lokalisation korreliert.

**Ergebnisse** Die Prävalenz einer LAE in der Gesamtkohorte betrug 24% (n = 678). Das mediane LAE-Patientenalter betrug 71 Jahre (IQR 58 – 80) und die Geschlechterverteilung 1,2/1 (M/F). Ambulante Patienten zeigten eine geringere Prävalenz (23%) als Patienten von Normalstationen (27%) oder Intensivpatienten (30%). Chirurgisch zugewiesene Patienten zeigten eine höhere Prävalenz als internistisch zugewiesene (34% vs. 22%). Zentrale und beidseitige LAEs korrelierten mit einer signifikant höheren Rechtsherzbelastung im Vergleich zu peripheren und einseitigen Embolien.

**Schlussfolgerung** Data-Mining von strukturierten Befunden ist eine einfache Methode zur Erhebung von Prävalenzstatistiken, epidemiologischen Daten und Verteilung von Krankheitscharakteristika, wie am Beispiel der LAE gezeigt. Die so generierten Daten können in vielerlei Hinsicht hilfreich sein, wie beispielsweise zur internen klinischen Qualitätssicherung oder für wissenschaftliche Auswertungen. Um davon zu profitieren, ist die konsequente Nutzung von SR erforderlich und wird deshalb empfohlen.

### Kernaussagen

- SR-basiertes Data-Mining ermöglicht einfache epidemiologische Auswertungen bei LAEs.
- Die Prävalenz der LAE ist bei ambulanten und stationären Patienten unterschiedlich.
- Zentrale und beidseitige LAEs haben ein erhöhtes Risiko für eine Rechtsherzbelastung.

**Zitierweise**

- Jorg T, Halfmann MC, Graafen D et al. Structured reporting for efficient epidemiological and in-hospital prevalence analysis of pulmonary embolisms. *Fortschr Röntgenstr* 2024; DOI 10.1055/a-2301-3349

**ABSTRACT**

**Purpose** Structured reporting (SR) not only offers advantages regarding report quality but, as an IT-based method, also the opportunity to aggregate and analyze large, highly structured datasets (data mining). In this study, a data mining algorithm was used to calculate epidemiological data and in-hospital prevalence statistics of pulmonary embolism (PE) by analyzing structured CT reports.

**Methods** All structured reports for PE CT scans from the last 5 years (n = 2790) were extracted from the SR database and analyzed. The prevalence of PE was calculated for the entire cohort and stratified by referral type and clinical referrer. Distributions of the manifestation of PEs (central, lobar, segmental, subsegmental, as well as left-sided, right-sided, bilateral) were calculated, and the occurrence of right heart strain was correlated with the manifestation.

**Results** The prevalence of PE in the entire cohort was 24% (n = 678). The median age of PE patients was 71 years

(IQR 58–80), and the sex distribution was 1.2/1 (M/F). Outpatients showed a lower prevalence of 23% compared to patients from regular wards (27%) and intensive care units (30%). Surgically referred patients had a higher prevalence than patients from internal medicine (34% vs. 22%). Patients with central and bilateral PEs had a significantly higher occurrence of right heart strain compared to patients with peripheral and unilateral embolisms.

**Conclusion** Data mining of structured reports is a simple method for obtaining prevalence statistics, epidemiological data, and the distribution of disease characteristics, as demonstrated by the PE use case. The generated data can be helpful for multiple purposes, such as for internal clinical quality assurance and scientific analyses. To benefit from this, consistent use of SR is required and is therefore recommended.

**Key Points**

- SR-based data mining allows simple epidemiologic analyses for PE.
- The prevalence of PE differs between outpatients and inpatients.
- Central and bilateral PEs have an increased risk of right heart strain.

**Hintergrund**

Die Lungenarterienembolie (LAE) ist nach dem akuten Myokardinfarkt und dem Schlaganfall die dritthäufigste kardiovaskuläre Erkrankung [1]. Aufgrund der Gefahr eines Rechtsherzversagens durch Obstruktion der Lungenstrombahn geht sie unbehandelt mit einer hohen Letalität von ca. 30% einher. Da das Risiko an einer LAE zu versterben innerhalb der ersten Stunden nach Symptombeginn am höchsten ist, ist eine frühzeitige Diagnose von größter Bedeutung [2]. Die CT-Angiografie ist der international anerkannte Referenzstandard zur Diagnosestellung einer LAE [3, 4]. Die CT ist nicht nur in der Erkennung von Thromben in der Lungenstrombahn sehr sensitiv, sondern ermöglicht auch eine einfache Quantifizierung einer möglicherweise vorliegenden Rechtsherzbelastung, als Surrogatparameter für ein drohendes Rechtsherzversagen [4, 5]. Aufgrund der geschilderten Dringlichkeit ist eine schnelle und korrekte Befundung dieser CTs durch die Radiologie essenziell.

Ein Großteil der radiologischen Befunde, inklusive der für CT-Angiografien bei Verdacht auf LAE (LAE-CTs), werden noch immer in einer Freitextform verfasst. Freitextbefunde haben eine hohe Variabilität hinsichtlich Inhalt und Struktur. Zudem fehlen in Freitextbefunden nicht selten für den klinischen Zuweiser relevante Informationen [6, 7]. Neben der Freitextbefundung ist die strukturierte Befundung (engl. *Structured reporting*; SR) in den letzten Jahren in den Fokus gerückt. Vorteile von SR gegenüber Freitextbefundung konnten in dutzenden Studien belegt werden. Strukturierte Befunde sind unter anderem besser lesbar, besser vergleichbar, vollständiger und detaillierter als Freitextbefunde [6,

7, 8]. Dies konnte auch explizit für LAE-CTs belegt werden [9]. Sowohl klinische Zuweiser als auch Radiologen bevorzugten strukturierte Befunde gegenüber Freitextbefunden [10].

In der klinischen Routine konnte sich SR trotz der mannigfaltigen Vorteile dennoch bisher nicht durchsetzen. Die Nutzung ist in den meisten radiologischen Abteilungen gering [10]. Dieses Phänomen hängt sehr wahrscheinlich mit Einschränkungen zusammen, die im Rahmen einer SR-Implementation entstehen können. Die Einführung von SR-Lösungen in den klinischen *Reporting workflow* kann mitunter technisch schwierig und mit großem Aufwand verbunden sein. Zudem ist sie stark vom jeweiligen Anbieter des *Picture Archiving and Communication System (PACS)* und *Radiology Information System (RIS)* abhängig [11]. Weiter ist die Integration der Spracherkennung in die SR bisher unzureichend. Anders als Freitextbefunde, müssen strukturierte Befunde größtenteils händisch mit Maus und Tastatur befüllt werden, was einerseits unbequem sowie zeitaufwändig ist und andererseits von den eigentlichen Bilddaten ablenken kann [12].

Neben den Vorteilen bezüglich der Befundqualität bietet die strukturierte Befundung als IT-basierte Methode auch die Möglichkeit, große hoch strukturierte Datenmengen in Datenbanken automatisiert anzureichern und sekundär auszuwerten. Dieses (begrifflich aus dem Englischen kommende) *Data-Mining* von strukturierten Befunden kann zur Beantwortung vielfältiger Fragestellungen genutzt werden.

So können beispielsweise Daten zur Epidemiologie der Erkrankung einschließlich der Setting-spezifischen Prävalenz basierend auf den Befunddaten erhoben werden. Erste Anstrengungen in diesem Feld wurden in unserer Abteilung 2017 in einer Machbar-

keitsstudie unternommen. Diese umfasste knapp über 500 Patienten, die bei Verdacht auf LAE ein CT erhielten. Da die strukturierte Befundung von LAE-CTs zu diesem Zeitpunkt noch nicht in der klinischen Routine implementiert war, mussten alle zugehörigen Freitextbefunde retrospektiv händisch strukturiert werden [13]. Epidemiologische Daten, die primär aus strukturierten Befunden, die in der klinischen Routine erhoben wurden, generiert wurden, gibt es in der Literatur kaum [14]. Darüber hinaus eignet sich *Data-Mining* von strukturierten Befunden auch für wissenschaftliche Fragestellungen. So konnte basierend auf strukturierten Befunddaten ein neues Bewertungssystem zur Abschätzung der Thrombuslast bei LAE-Patienten entwickelt werden [15]. Auch für das Training und die Validierung von Künstlicher Intelligenz (KI) sind strukturierte Befunddaten ideal geeignet [16].

Aufbauend auf der beschriebenen Machbarkeitsstudie war das Ziel dieser Arbeit, alle in der klinischen Routine erhobenen strukturierten Befunde zu LAE-CTs der letzten 5 Jahre aus der Datenbank unserer Abteilung mit Hilfe eines *Data-Mining*-Algorithmus auszuwerten. Zur internen Qualitätssicherung sollten neben epidemiologischen Daten und Krankheitscharakteristika auch Erkenntnisse zu möglichen innerklinischen Unterschieden in der Prävalenz einer LAE gewonnen werden.

## Methoden

In unserer Klinik wurde 2016 eine web-basierte, *Integrating the Health Enterprise Management of Radiology Report Templates* (IHE MRRT)-kompatible Plattform zur SR entwickelt und für die Nutzung in der klinischen Routine implementiert [17]. Seitdem konnten zahlreiche Befundvorlagen (Templates) für verschiedene Untersuchungstypen in den Modalitäten Ultraschall, CT und MRT entwickelt und zur Nutzung gebracht werden [10]. Das Template zur Befundung der CT-Angiografie zum Lungenembolieausschluss wurde im 3. Quartal 2018 klinisch implementiert. ► **Abb. 1** zeigt die aktuelle Version der Befundmaske. Die Templates werden während des Befundvorgangs typischerweise mit Maus und Tastatur befüllt. Die Bestimmung der Größenverhältnisse von rechtem und linkem Ventrikel (RV/LV-Index) zur Quantifizierung einer möglichen Rechtsherzbelastung erfolgte durch Vermessung der jeweils maximalen Ventrikelbreite in axialer Schnittführung und wird durch das Template nach Eintragung beider Werte automatisch berechnet. Das Vorliegen einer Rechtsherzbelastung wurde im Einklang mit den Richtlinien der European Society of Cardiology als ein RV/LV-Index  $> 1$  definiert [18].

Alle in der klinischen Routine mit Hilfe von Templates erhobenen und freigegeben Befunde werden in der Befundungsplattform automatisch gespeichert und sind somit sekundären Auswertungen gut zugänglich. Von der Befundungsplattform können Befunde systematisch mit Hilfe der Software RapidMiner Studio (RapidMiner, Cambridge, USA) über einen zuvor definierten Zeitraum ausgelesen werden. Dabei können einerseits gezielt einzelne Bestandteile der Befunde (wie z. B. das Vorliegen einer Rechtsherzbelastung) oder aber auch der gesamte Inhalt der Befunde ausgelesen werden. Die Ergebnisse werden als .csv-Datei ausgegeben und können anschließend mit Hilfe einer Statistik-

software ausgewertet werden. ► **Abb. 2** zeigt eine grafische Darstellung des hier beschriebenen Arbeitsablaufs.

Für die in dieser Arbeit untersuchten Fragestellungen wurden die folgenden Inhalte aus dem LAE-Template ausgelesen: Patientenalter und -geschlecht, innerklinischer Zuweiser, Zuweisungsart (ambulant, Normalstation, Intensivstation), Vorliegen einer Lungenarterienembolie, Ausprägung einer möglichen Lungenarterienembolie (Hauptstamm, Lobärrarterie, Segmentarterie, subsegmentale Arterie), Seitangabe der Embolie (rechts, links, beidseitig), Vorliegen einer Rechtsherzbelastung (RV/LV-Index  $> 1$ ) sowie Datum und Uhrzeit der Untersuchung. Alle seit Implementation des Templates erstellten Befunde von 08/2018 bis 07/2023 wurden eingeschlossen.

Die statistischen Auswertungen wurden mit Hilfe der Software R (The R Foundation für Statistical Computing, Wien, Österreich) durchgeführt. Binäre und kategoriale Daten wurden als absolute Zahlen und Prozentsätze ausgedrückt. Kontinuierliche Daten wurden als Mittelwert und Standardabweichung (SD) angegeben. Nicht normalverteilte Daten wurden als Median und Interquartilsabstand (IQR) angegeben. Auf statistische Signifikanz wurde mittels des Kruskal-Wallis-Tests geprüft. Hierfür wurde das Signifikanzniveau auf  $\alpha = 0,05$  festgelegt.

## Ergebnisse

Insgesamt wurden 2790 strukturierte Befunde von Patienten, bei denen ein LAE-CT durchgeführt wurde, eingeschlossen.

Eine Lungenarterienembolie bestand bei 24 % ( $n = 678$ ) des Gesamtkollektivs. Das mediane Patientenalter betrug 71 Jahre (IQR: 58 – 80), das Verhältnis von Männern zu Frauen betrug 1,2/1 (► **Abb. 3**).

Es zeigten sich zum Teil deutliche Unterschiede in der Prävalenz einer LAE in Abhängigkeit der Zuweisungsart. Von den ambulant zugewiesenen Patienten, die insgesamt 69 % ( $n = 1913$ ) der Gesamtkohorte ausmachten, hatten nur 23 % ( $n = 441$ ) einen positiven Befund. Von den insgesamt 678 von Normalstationen zugewiesenen Patienten hatten 27 % ( $n = 182$ ) eine LAE. Im relativ kleinen Patientenkollektiv, dass von der(?) Intensivstation zugewiesen wurde ( $n = 199$ ), zeigte sich eine höhere Rate an positiven Befunden von 30 % ( $n = 59$ ) (► **Abb. 4A**). Auch zeigten sich Unterschiede in der Prävalenz einer Lungenarterienembolie in Abhängigkeit der zuweisenden klinischen Disziplin. Insgesamt 75 % ( $n = 2104$ ) der Patienten wurden von internistisch geführten Abteilungen zugewiesen. Von diesen hatten 22 % ( $n = 470$ ) eine LAE. Von den 453 chirurgisch zugewiesenen Patienten hatten 34 % ( $n = 155$ ) eine LAE. Die übrigen hier aggregiert als „Sonstige“ dargestellten Abteilungen wiesen 233 Patienten zu, von denen 24 % ( $n = 57$ ) eine LAE hatten (► **Abb. 4B**).

Von den 678 Patienten mit LAE hatten 32 % ( $n = 215$ ) eine zentrale Embolie, 25 % ( $n = 172$ ) eine lobäre Embolie, 33 % ( $n = 228$ ) eine segmentale Embolie und 10 % ( $n = 69$ ) eine subsegmentale Embolie. 65 % aller LAEs zeigten eine beidseitige Ausprägung ( $n = 444$ ), 10 % ( $n = 71$ ) eine einseitig linksseitige und 35 % ( $n = 167$ ) eine einseitig rechtsseitige Ausprägung. Eine Rechtsherzbelastung lag bei insgesamt 43 % ( $n = 292$ ) der Patienten mit LAE vor.

**Befund**

Voruntersuchung

Aufklärung

---

**Lungenarterien**

Kontrastmittelaussparung

<p><b>rechts</b></p> <input type="checkbox"/> subsegmental <input type="text" value="Segment"/> <input checked="" type="checkbox"/> segmental <input type="text" value="Segment"/> <input checked="" type="checkbox"/> lobär <input type="text" value="Segment"/> <input type="checkbox"/> Hauptstamm <input type="text" value="Segment"/>	<p><b>links</b></p> <input type="checkbox"/> subsegmental <input type="text" value="Segment"/> <input checked="" type="checkbox"/> segmental <input type="text" value="Segment"/> <input checked="" type="checkbox"/> lobär <input type="text" value="Segment"/> <input type="checkbox"/> Hauptstamm <input type="text" value="Segment"/>
---	--

Ghanima Score

Zeichen älterer Embolien  Dilatierte Bronchialarterien  wandständige Thromben  Webs

Durchmesser Tr. pulmonalis

---

**Herz**

Herzgröße

**Rechtsherzbelastung**

RV-Durchmesser  mm

LV-Durchmesser  mm

Ratio

Interventrikuläres Septum

Perikarderguss

Koronarsklerose

---

**Pleura**

Pleuraerguss

Pneumothorax

---

**Lungenparenchym**

Infiltrate

Belüftungsstörungen

Pulmonale Herdbefunde

Pulmonalvenöse Stauungszeichen

► **Abb. 1** Strukturiertes Befundungstemplate CT-Angiografie Lungenarterienembolie.

Eine weiterführende Auswertung zeigte, dass zentrale Embolien mit höheren RV/LV-Indizes als lobäre, segmentale oder subsegmentale Embolien assoziiert sind (► **Abb. 5A**). Zudem trat eine Rechtsherzbelastung mit 73 % bei zentralen Embolien deut-

lich häufiger auf als bei lobären (38 %), segmentalen (26 %) und subsegmentalen Embolien (13 %) (► **Abb. 5B**).

Analog dazu waren auch beidseitige Embolien mit höheren RV/LV-Indizes assoziiert verglichen mit einseitigen Embolien (► **Abb. 5C**). Bei 54 % aller beidseitigen Embolien lag eine Rechts-



► **Abb. 2** Grafische Darstellung des Arbeitsablaufs für die sekundäre Auswertung strukturierter Befunde (Data-Mining).

herzbelastung vor, während dies bei einseitigen Embolien nur in 27 % linksseitig und 19 % rechtsseitig der Fall war (► **Abb. 5D**).

## Diskussion

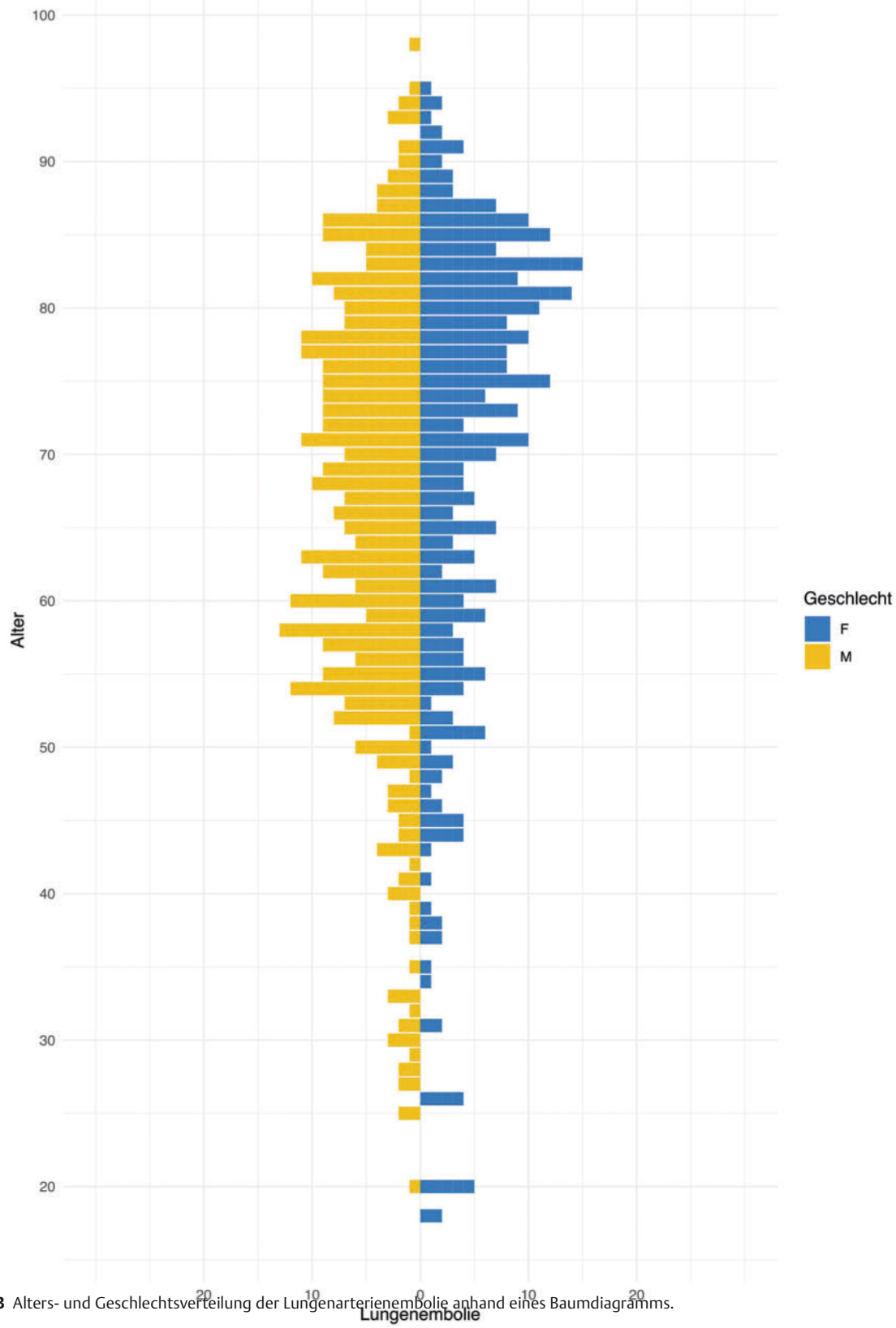
Die erhobenen Daten zeigen, dass mit Hilfe der konsequenten Anwendung von SR ohne großen Aufwand Prävalenzstatistiken und epidemiologische Daten erhoben werden können, wie hier am Beispiel der Lungenarterienembolie demonstriert. Die Daten bieten einen wertvollen Überblick über die Patientenpopulation und die Verteilung von Krankheitsmerkmalen und Krankheitsausprägung. Zudem bieten sie ein wichtiges Feedback für klinische Zuweiser und Radiologen.

Die in der untersuchten Kohorte ermittelte Alters- und Geschlechtsverteilung der LAE (medianes Patientenalter 71, Verhältnis von Männern zu Frauen 1,2 zu 1) zeigt Übereinstimmung mit mehreren anderen epidemiologischen Studien zur LAE [19, 20, 21]. Die ermittelte Prävalenz einer LAE von 24 % in der Gesamtkohorte deckt sich ebenfalls mit der 2017 durchgeführten Machbarkeitsstudie in der Freitextbefunde einer kleineren Kohorte ( $n = 500$ ) retrospektiv strukturiert wurden [13]. Eine amerikanische Arbeit, in der Freitextbefunde von über 500 LAE-CTs manuell ausgewertet wurden, beschreibt die Prävalenz einer LAE mit 9 % hingegen deutlich niedriger, schließt aber nur Patienten ein, die von einer Notaufnahme zugewiesen wurden [22]. Hierbei ist zudem zu beachten, dass in den USA aufgrund anderer Rechtsbedingungen, die Indikation zur CT mitunter niederschwelliger getroffen wird.

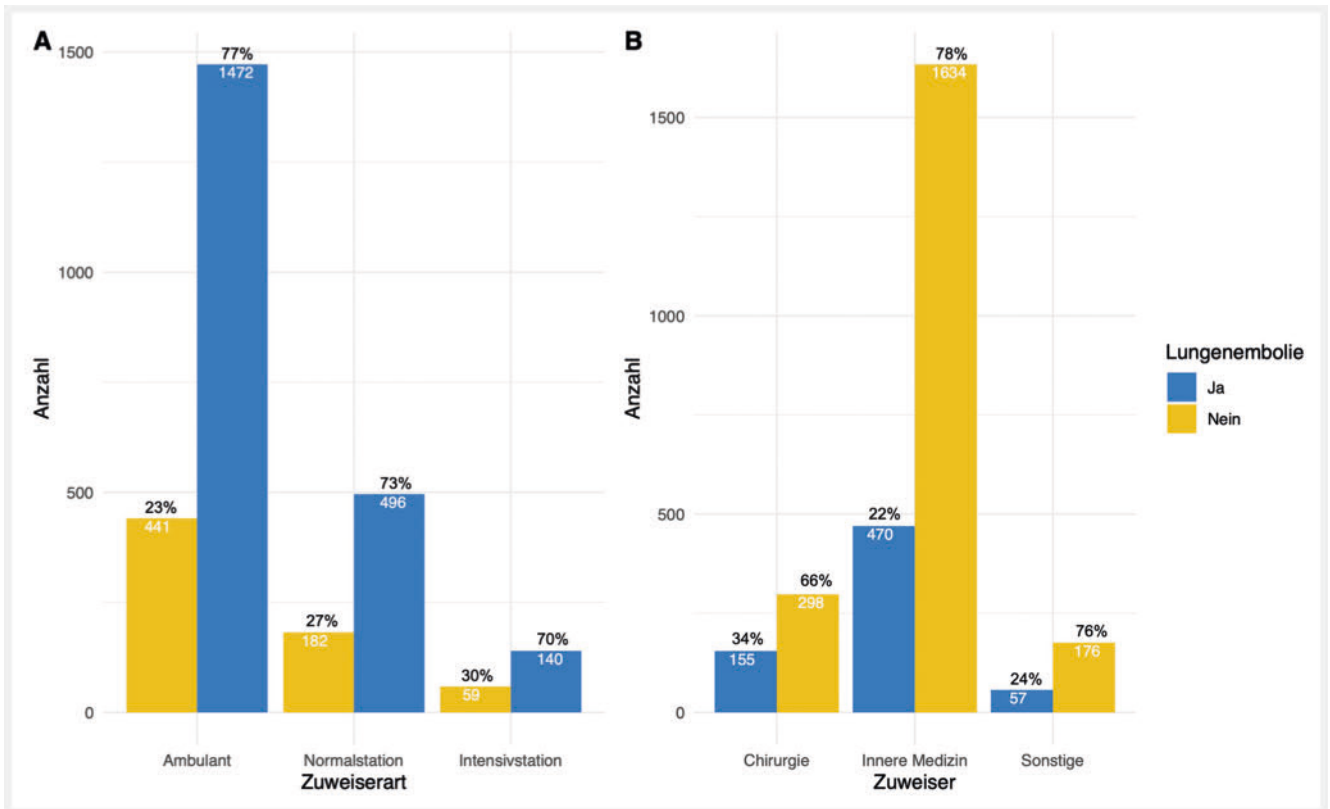
Auch in dieser Arbeit ließen sich Unterschiede in der Wahrscheinlichkeit für einen positiven Befund in Abhängigkeit der Zu-

weisungsart nachweisen. Die dabei niedrigere Rate von 23 % bei ambulanten Patienten im Gegensatz zu 27 % bei Patienten von Normalstation und 30 % bei Intensivpatienten lässt sich in erster Linie dadurch erklären, dass im stationären Setting bei einer neu aufgetretenen typischen Symptomatik etwa nach einer Operation oder längerer Immobilisation die Diagnose einer LAE naheliegender ist als im ambulanten Setting. In letzterem dient das LAE-CT nicht selten bei uneindeutiger Symptomatik und positivem D-Dimer im Labor eher zum Ausschluss einer LAE als zur Diagnosebestätigung. Diese Überlegungen werden auch durch die niedrigere Rate von positiven Befunden bei internistisch zugewiesenen Patienten (22 %) im Gegensatz zu chirurgisch zugewiesenen Patienten (34 %) gestützt. Eine signifikant zunehmende Rechtsherzbelastung bei zentralen Embolien verglichen mit peripher gelegenen Embolien konnte Hahiharan et al. für eine kleinere Kohorte von Patienten, bei der Bild- und Befunddaten aufwändig händisch ausgewertet wurden, belegen [23]. Dieser Zusammenhang konnte mit Hilfe des hier durchgeführten SR-basierten *Data-Mining*-Ansatzes an einer mehr als doppelt so großen Kohorte belegt werden.

Letztlich liegen die ermittelten Ergebnisse dieser Arbeit im Rahmen des erwartbaren. Die eigentliche Innovation liegt in der Methodik selbst, mit der diese generiert wurden. Diese ermöglicht es, Daten aus der Bildgebung schnell und ohne den Einsatz größerer personeller Ressourcen zu erfassen, was hingegen beim Einsatz von Freitextbefundung nur mühevoll unter Durchsicht einzelner Befunde möglich wäre. Auf der anderen Seite muss jedoch berücksichtigt werden, dass die primäre Erstellung der strukturierten Befunde in der klinischen Routine aufwändiger sein kann als konventionelle Freitextbefundung. Weiter besteht eine vielfäl-



► **Abb. 3** Alters- und Geschlechtsverteilung der Lungenarterienembolie anhand eines Baumdiagramms.



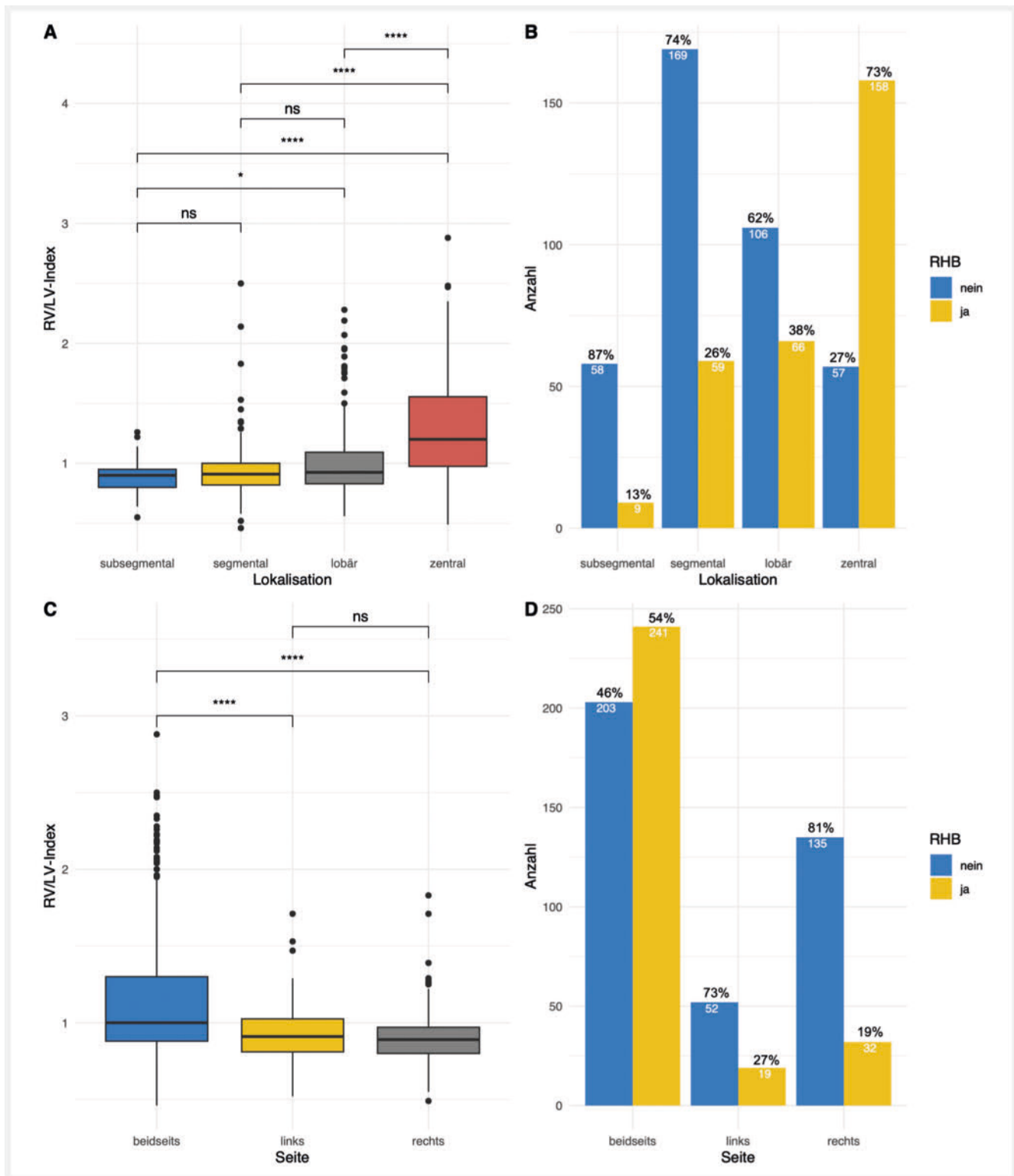
► **Abb. 4** Auftreten einer LAE in Abhängigkeit der Zuweisungsart (ambulant, Normalstation, Intensivstation (A)) und in Abhängigkeit der zuweisenden klinischen Abteilung (B).

tige Anwendbarkeit für die Daten. Sowohl für klinische Zuweiser als auch für den Radiologen ist es essenziell, Wahrscheinlichkeiten für das Vorliegen eines positiven Befunds bei radiologischen Untersuchungen zu kennen. Insbesondere bei Untersuchungen, bei denen Röntgenstrahlung zum Einsatz kommt, wie der CT, wären bei niedrigen Wahrscheinlichkeiten im Sinne des Strahlenschutzes Qualitätssicherungsmaßnahmen erforderlich. Weiter eignen sich bildgebungs-basierte strukturierte Daten bestens, um Registerdatenbanken, die es in verschiedenen europäischen Ländern für LAE-Patienten bereits gibt, anzureichern [24]. Diese beinhalten für gewöhnlich vor allem klinische Daten und könnten zusätzlich durch Bildparameter noch optimiert werden. Dies könnte auch die Entwicklung von KI-Modellen, beispielsweise zur Vorhersage von Rezidivrisiken, ermöglichen. Darüber hinaus könnten die Daten in Zukunft auch in im Rahmen der Medizininformatikinitiative an den Universitätsklinika eingerichteten Datenintegrationszentren aufgenommen werden, was eine standortübergreifende Nutzung für Forschungszwecke ermöglichen würde.

Neben des hier gezeigten *Data-Mining*-Ansatzes von strukturierten Befunden können auch Freitextbefunde retrospektiv durch die Nutzung von Natural Language Processing (NLP) zur sekundären Datennutzung ausgewertet werden [25, 26, 27]. NLP ist als KI-verwandte Technologie dazu in der Lage, automatisch Freitext zu analysieren, relevanten Inhalt zu extrahieren und zu strukturieren [27]. Die Methode ist jedoch einerseits dadurch eingeschränkt, dass die dafür genutzten Freitextbefunde nicht immer alle benöti-

gen Informationen enthalten [6, 7]. Andererseits haben sich NLP-Algorithmen in den letzten Jahren zwar deutlich verbessert, eine vollständig richtige Erkennung des Inhalts aus Freitextbefunden ist jedoch weiter hin nicht garantiert [26]. Darüber hinaus sind *Large Language Models*, wie Chat GPT-4, dazu in der Lage, ganze Freitextbefunde retrospektiv in einen strukturierten Befund umzuwandeln [28]. Letztlich ist der hier praktizierte SR-basierte Ansatz jedoch geeigneter, da vollständige Datensätze ohne die Notwendigkeit eines zusätzlichen Transformationsschrittes primär in einer hoch strukturierten Form generiert und direkt ausgewertet werden können.

Die vorliegende Studie hat Einschränkungen. Um eine hohe Aussagekraft von epidemiologischen Daten und Prävalenzstatistiken, die mittels des SR-basierten *Data-Mining*-Ansatzes erhoben wurden, zu gewährleisten, ist eine hohe klinische Nutzungsrate der SR essenziell. Für die meisten strukturierten Befundungstemplates konnten in unserer Klinik hohe Nutzungsraten nachgewiesen werden (z. B. Polytrauma CT 97 %, Prostata MRT 92 % oder Urolithiasis CT 91 % im Jahre 2022). Wohingegen die Nutzungsrate des LAE-Templates mit 58 % im Jahre 2022 relativ geringer ausfiel, seit der Implementierung des Templates im Jahre 2018 (hier 18 %) aber stetig anstieg [10]. Letztlich ist die SR nicht für alle Untersuchungstypen gleich gut geeignet und ihre Nutzung an unserer Klinik nicht verpflichtend. Im Falle des LAE-CTs bestehen in bis zu 33 % der Untersuchungen alternative Diagnosen [22]. Für die genaue Beschreibung alternativer Diagnosen und



► **Abb. 5** RV/LV-Index in Abhängigkeit der höchsten Lokalisation einer Embolie im pulmonalarteriellen Gefäßsystem (A) und Auftreten einer Rechtssherzbelastung in Abhängigkeit der höchsten Lokalisation einer Embolie im pulmonalarteriellen Gefäßsystem (B). RV/LV-Index in Abhängigkeit der Seite des Auftretens einer LAE (C) und Auftreten einer Rechtssherzbelastung in Abhängigkeit der Seite des Auftretens einer LAE (D). RHB = Rechtssherzbelastung.



bei sehr komplexen Fällen könnten Radiologen das Befundungstemplate ungeeignet finden und von einer Nutzung absehen. Gemäß dieser Annahme ist davon auszugehen, dass die tatsächliche Wahrscheinlichkeit eines positiven Befunds im LAE-CT niedriger liegt als im ausgewerteten Patientenkollektiv. Unabhängig der Anwendbarkeit auf einzelne Untersuchungstypen besteht weiterhin Potenzial, die SR zu optimieren und die klinische Anwendung somit weiter zu fördern. Dies könnte durch eine verbesserte Integration der Spracherkennung in strukturierte Befundungstemplates mit Hilfe von NLP, was Gegenstand aktueller Forschungsarbeiten ist, gelingen [29].

Darüber hinaus betrachtet diese Bildgebungs-basierte Studie die Erkrankung nur zu einem definierten Zeitpunkt. Aussagen zum Krankheitsverlauf können daher nicht getroffen werden. Letztlich wurden auch Patienten, die zur Diagnosestellung beziehungsweise zum Diagnoseausschluss eine andere Untersuchung als ein CT erhalten haben, wie beispielsweise ein MRT oder eine Szintigrafie, nicht eingeschlossen. Im spezifischen Falle der spezifischen Fragestellung LAE ist dieser Anteil an Patienten jedoch als relativ gering einzuordnen [30].

## Schlussfolgerung

SR ermöglicht es durch *Data-Mining* von Befunden ohne großen Einsatz von Zeit oder Ressourcen epidemiologische Daten und Prävalenzstatistiken zu erheben. Für das Fallbeispiel der Lungenarterienembolie konnten so zum Beispiel Unterschiede in der innerklinischen Prävalenz in Abhängigkeit des klinischen Zuweisers und der Zuweisungsart gezeigt werden. Die generierten Daten bieten vielfältige Anwendungsbereiche, wie beispielsweise zur internen Qualitätssicherung, wissenschaftlichen Auswertungen sowie zur Anreicherung von Registerdatenbanken. Um von diesen Vorteilen zu profitieren, ist die Implementation und konsequente Nutzung von SR unerlässlich und wird auch deshalb empfohlen.

## Klinische Relevanz

- Die konsequente klinische Anwendung von SR ermöglicht es durch *Data-Mining* ohne den Einsatz großer Ressourcen epidemiologische Daten und innerklinische Prävalenzstatistiken zu erheben
- Für das Fallbeispiel der Lungenarterienembolie können so Unterschiede in der innerklinischen Prävalenz in Abhängigkeit des Zuweisers und der Zuweisungsart, sowie Wahrscheinlichkeiten für das Vorliegen einer Rechtsherzbelastung in Abhängigkeit der Ausprägung der Embolie berechnet werden.
- Die Kenntnis von Wahrscheinlichkeiten für das Vorliegen eines positiven Befunds bei radiologischen Untersuchungen bietet sowohl für Radiologen als auch für klinische Zuweiser ein wichtiges Feedback.
- Die generierten Daten bieten darüber hinaus vielfältige Anwendungsbereiche, wie zum Beispiel zur inneren Qualitätssicherung, Strahlenschutz, wissenschaftlichen Auswertung oder zur Anreicherung von Registerdatenbanken.

## Abkürzungen

CT	Computertomografie
LAE	Lungenarterienembolie
SR	<i>Structured Reporting</i> ; strukturierte Befundung
IHE MRRT	Integrating the Health Enterprise Management of Radiology Report Templates
KI	Künstliche Intelligenz
NLP	<i>Natural Language Processing</i>
RV/LV-Index	Rechtsventrikulärer / Linksventrikulärer Index

## Interessenkonflikt

Die Autorinnen/Autoren geben an, dass kein Interessenkonflikt besteht.

## Literatur

- [1] Licha CRM, McCurdy CM, Maldonado SM et al. Current management of acute pulmonary embolism. *Ann. Thorac. Cardiovasc. Surg* 2020; 26: 65–71
- [2] Bělohávek J, Dytrych V, Linhart A. Pulmonary embolism, part I: Epidemiology, risk factors and risk stratification, pathophysiology, clinical presentation, diagnosis and nonthrombotic pulmonary embolism. *Exp. Clin. Cardiol* 2013; 18 (2): 129–138
- [3] Duffett L, Castellucci LA, Forgie MA. Pulmonary embolism: Update on management and controversies. *BMJ* 2020; 370: m2177
- [4] Sin D, McLennan G, Rengier F et al. Acute pulmonary embolism multimodality imaging prior to endovascular therapy. *Int. J. Cardiovasc. Imaging* 2021; 37: 343–358
- [5] Moore AJE, Wachsmann J, Chamarthy MR et al. Imaging of acute pulmonary embolism: An update. *Cardiovasc. Diagn. Ther* 2018; 8: 225–243
- [6] Brook OR, Brook A, Vollmer CM et al. HEALTH POLICY AND PRACTICE: Structured Reporting of Multiphasic CT for Pancreatic Cancer Brook et al. *Radiology* 2015; 274: 464–472
- [7] Schoeppe F, Sommer WH, Nörenberg D et al. Structured reporting adds clinical value in primary CT staging of diffuse large B-cell lymphoma. *Eur. Radiol* 2018; 28: 3702–3709
- [8] Jorg T, Heckmann JC, Mildenerberger P et al. Structured reporting of CT scans of patients with trauma leads to faster, more detailed diagnoses: An experimental study. *Eur. J. Radiol* 2021; 144: 109954
- [9] Sabel BO, Plum JL, Kneidinger N et al. Structured reporting of CT examinations in acute pulmonary embolism. *J. Cardiovasc. Comput. Tomogr* 2017; 11: 188–195
- [10] Jorg T, Halfmann MC, Arnhold G et al. Implementation of structured reporting in clinical routine: a review of 7 years of institutional experience. *Insights Imaging* 2023; 14: 61
- [11] Pinto dos Santos D, Kotter E. Structured radiology reporting on an institutional level – benefit or new administrative burden? *Ann. N. Y. Acad. Sci* 2018; 1434: 274–281
- [12] Sistro CL. Conceptual approach for the design of radiology reporting interfaces: The talking template. *J. Digit. Imaging* 2005; 18: 176–187
- [13] Dos Santos DP, Scheibl S, Arnhold G et al. A proof of concept for epidemiological research using structured reporting with pulmonary embolism as a use case. *Br. J. Radiol* 2018; 91: 1088
- [14] Jorg T, Halfmann MC, Rölz N et al. Structured reporting in radiology enables epidemiological analysis through data mining: urolithiasis as a use case. *Abdom. Radiol* 2023; 48: 3520–3529

- [15] Fink MA, Mayer VL, Schneider T et al. CT Angiography Clot Burden Score from Data Mining of Structured Reports for Pulmonary Embolism. *Radiology* 2022; 302: 175–184
- [16] Hempel JM, Pinto dos Santos D. Structured reporting and artificial intelligence. *Radiologe* 2021; 61: 999–1004
- [17] Pinto dos Santos D, Klos G, Kloeckner R et al. Development of an IHE MRRT-compliant open-source web-based reporting platform. *Eur. Radiol* 2017; 27: 424–430
- [18] Konstantinides S V, Meyer G, Becattini C et al. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). *Eur. Heart J* 2020; 41: 543–603
- [19] Raptis DG, Gourgoulianis KI, Daniil Z et al. Time trends for pulmonary embolism incidence in Greece. *Thromb. J* 2020; 18: 1–9
- [20] Keller K, Hobohm L, Ebner M et al. Trends in thrombolytic treatment and outcomes of acute pulmonary embolism in Germany. *Eur. Heart J* 2020; 41: 522–529
- [21] Jarman AF, Mumma BE, Singh KS et al. Crucial considerations: Sex differences in the epidemiology, diagnosis, treatment, and outcomes of acute pulmonary embolism in non-pregnant adult patients. *J. Am. Coll. Emerg. Physicians Open* 2021; 2: 1–9
- [22] Hall WB, Truitt SG, Scheunemann LP et al. The prevalence of clinically relevant incidental findings on chest computed tomographic angiograms ordered to diagnose pulmonary embolism. *Arch. Intern. Med* 2009; 169: 1961–1965
- [23] Hariharan P, Dudzinski DM, Rosovsky R et al. Relation Among Clot Burden, Right-Sided Heart Strain, and Adverse Events After Acute Pulmonary Embolism. *Am. J. Cardiol* 2016; 118: 1568–1573
- [24] Monreal M, Mahé I, Bura-Riviere A et al. Pulmonary embolism: Epidemiology and registries. *Presse Med* 2015; 44: e377–e383
- [25] Huhdanpaa HT, Tan WK, Rundell SD et al. Using Natural Language Processing of Free-Text Radiology Reports to Identify Type 1 Modic Endplate Changes. *J. Digit. Imaging* 2018; 31: 84–90
- [26] Casey A, Davidson E, Poon M et al. A systematic review of natural language processing applied to radiology reports. *BMC Med. Inform. Decis. Mak* 2021; 21: 179
- [27] Jungmann F, Kämpgen B, Mildenerger P et al. Towards data-driven medical imaging using natural language processing in patients with suspected urolithiasis. *Int. J. Med. Inform* 2020; 137: 104106
- [28] Adams LC, Truhn D, Busch F et al. Leveraging GPT-4 for Post Hoc Transformation of Free-Text Radiology Reports into Structured Reporting: A Multilingual Feasibility Study. *Radiology* 2023; 307 (4): 230725
- [29] Jorg T, Kämpgen B, Feiler D et al. Efficient structured reporting in radiology using an intelligent dialogue system based on speech recognition and natural language processing. *Insights Imaging* 2023; 14: 47
- [30] Das M, Wildberger JE. Diagnostik der akuten Lungenembolie: Ein Update. *Radiologe* 2016; 56: 543–556