Diagnostic Performance and Reliability of Non-Enhanced Imaging Characterization Quotients for the Differentiation of Infectious and Malignant Pulmonary Nodules in Hematological Patients Using 3T MRI

Diagnostische Genauigkeit und Zuverlässigkeit nativer Bildgebungsscharakterisierungsquotingnten zur Differenzierung infektiöser und maligner pulmonaler Herdbefunde in hämatologischen Patienten im 3T-MRT

Authors
Sebastian Niko Nagel1, Damon Kim1, Tatjana Wylutzki1, Ingo G. Steffen1, Stefan Schwartz2, Tobias Pienkofer1, Bernd Hamm1, Thomas Elgeti1, 3

Affiliations
1 Department of Radiology, Charité – Universitätsmedizin Berlin, Germany
2 Department of Hematology and Oncology, Charité – Universitätsmedizin Berlin, Germany
3 Department of Nuclear Medicine, Charité – Universitätsmedizin Berlin, Germany

Key words
thorax, diagnostic radiology, infection, lymphoma, MR imaging

received 28.03.2019
accepted 19.08.2019

Bibliography
DOI https://doi.org/10.1055/a-1005-7424
Published online: 2019
Fortschr Röntgenstr
© Georg Thieme Verlag KG, Stuttgart - New York
ISSN 1438-9029

Correspondence
Dr. Sebastian Niko Nagel
Department of Radiology, Charité-Universitätsmedizin Berlin, Hindenburgdamm 30, 12203 Berlin, Germany
Tel.: +49/30/450 62 77 33
sebastian.nagel@charite.de

ZUSAMMENFASSUNG

Ziel
Beurteilung der diagnostischen Genauigkeit und Reliabilität nichtkontrastmittelverstärkter Charakterisierungszquotienten (engl. „non-enhanced imaging characterization quotients“, kurz NICQs) in der Magnetresonanztomografie (MRT) zur differenzialdiagnostischen Einordnung pulmonaler Herdbefunde in hämatologischen Patientinnen und Patienten.

Material und Methoden
Es wurden insgesamt 83 Läsionen in 45 konsekutiven hämatologischen Patienten analysiert (10 bakterielle Pneumonien, 16 Pilzpnemonien, 19 pulmonale Lymphom-Manifestationen). Das MRT-Protokoll bestand aus T2-gewichteten Single-Shot Fast-Spin-Echo-(FSE) und T1-gewichteten Gradientenecho (GRE)-Sequenzen. Der jeweils T2-basierte T2-NICQmean und T2-NICQ90th wurde aus der Signalintensität der Läsion, der Muskulatur und des Fettgewebes errechnet ((SILäsion -SIMuskulatur) / (SIFettgewebe-SIMuskulatur) *100), der simple T1-basierte Quotient T1-Qmean aus der Signalintensität der Läsion und der Muskulatur (SILäsion /SIMuskulatur). Die Bildauswertung erfolgte durch einen Radiologen mit > 7 Jahren und eine Radiologin mit 1 Jahr Erfahrung. Für die statistische Auswertung kamen der Kruskal-Wallis- oder Mann-Whitney-U-Test, die Receiver-Operating-Characteristic (ROC)-Analyse mit Berechnung der Area Under the Curve (AUC) sowie die Berechnung von Intraclass-Korrelationskoeffizienten (ICCs) zum Einsatz.

Ergebnisse
Die Medianwerte beider T2-NICQs unterschieden sich signifikant zwischen infektiösen Veränderungen und Lymphom-Manifestationen im Allgemeinen (T2-NICQmean 20,33 vs. 10,14; T2-NICQ90th 34,96 vs. 25,52) sowie Pilzinfiltrationen und Lymphom-Manifestationen im Besonderen (T2-NICQmean 19,00 vs. 10,14; T2-NICQ90th 34,49 vs. 25,25). Die AUCs auf Patientenebene für die T2-NICQs lagen dabei zwischen 0,73 und 0,79. Die ICCs betrugen mindestens >0,85, mit Ausnahme der Bewertung der Intrarater-Reliabilität des T2-NICQ90th (0,79).

Schlussfolgerung

Nagel SN et al. Diagnostic Performance and... Fortschr Röntgenstr

Published online: 24.10.2019
Kernaussagen:
- Nichtkontrastmittelverstärkte Charakterisierungquotienten bieten einen schnellen und pragmatischen Ansatz für die Beurteilung pulmonaler Herde in hämatologischen Patientinnen und Patienten.
- Die diagnostische Leistung von nichtkontrastmittelverstärkten Charakterisierungquotienten ist engemessen, um infektiöse und pilzpneumonische Infiltrate von Lymphom-Manifestationen zu unterscheiden.
- Nichtkontrastmittelverstärkte Charakterisierungquotienten zeigen eine gute bis exzellente Intra- und Interrater-Reliabilität.

ABSTRACT

Purpose To evaluate the diagnostic performance and reliability of non-enhanced imaging characterization quotients (NICQs) derived from magnetic resonance imaging (MRI) in the differential diagnosis of pulmonary nodules in hematological patients.

Materials and Methods A total of 83 lesions in 45 consecutive hematological patients were analyzed (10 bacterial pneumonias, 16 fungal pneumonias, 19 pulmonary lymphoma manifestations). The MRI protocol included T2-weighted single-shot fast spin echo (FSE) and T1-weighted gradient echo (GRE) sequences. T2-based T2-NICQmean and T2-NICQ90th were calculated from signal intensities measured in the lesion, muscle, and fat ((SIlesion - SIMuscle)/(SIfat - SIMuscle) * 100), and simple T1-based T1-Qmean from signal intensities of the lesion and muscle. Images were read by one radiologist with > 7 years and one with 1 year of experience. For statistical evaluation the Kruskal-Wallis or Mann-Whitney U-test, receiver operating characteristic (ROC) analysis with calculation of areas under the curve (AUC), and intraclass correlation coefficients (ICCs) were used.

Introduction
The differential diagnosis of pulmonary findings can be a challenge for radiologists in hematological patients, especially when the distinction between an infection and pulmonary involvement of an underlying disease is required [1]. Morphologic features, such as the halo sign in invasive bronchopulmonary aspergillosis on computed tomography scans, can be of help, but are nonspecific [2, 3]. In addition, the results of other diagnostic procedures, such as the galactomannan test, may also be inconclusive. In a study by Cao et al., for instance, this test identified only 73.5% of patients with aspergillosis [4]. If the diagnosis remains unclear, appropriate treatment may be delayed, which is known to be associated with higher morbidity and mortality in the case of fungal infections [5, 6]. To establish the correct diagnosis, invasive diagnostic procedures such as bronchoscopic sampling may become necessary [7]. Taking these considerations into account, a fast and noninvasive diagnostic approach that helps in the diagnostic workup of unclear cases is desirable.

Results
Medians of T2-NICQs differed significantly when comparing infectious lesions and lymphoma manifestations in general (T2-NICQmean 20.33 vs. 10.14; T2-NICQ90th 34.96 vs. 25.52) or fungal lesions and lymphoma manifestations in particular (T2-NICQmean 19.00 vs. 10.14; T2-NICQ90th 34.49 vs. 25.25). The AUCs for T2-NICQs on the per-patient level ranged from 0.73 to 0.79. ICCs were at least > 0.85, except for intrarater testing of T2-NICQ90th (0.79).

Conclusion
The overall diagnostic performance of T2-NICQs is adequate for differentiating infectious and fungal lesions from lymphoma manifestations. The results show good to excellent intra- and interrater agreement. We therefore consider NICQs helpful in the diagnostic workup of pulmonary nodules in hematological patients.

Key points:
- Non-enhanced Imaging Characterization Quotients provide a fast and pragmatic approach for assessing pulmonary lesions in hematological patients.
- The diagnostic performance of Non-enhanced Imaging Characterization Quotients is adequate for differentiating infectious and fungal infiltrates from lymphoma manifestations.
- Non-enhanced Imaging Characterization Quotients show good to excellent intra- and interrater agreement.

Citation Format

Magnetic resonance imaging (MRI) provides excellent soft-tissue contrast and different studies using either diffusion-weighted imaging (DWI) or dynamic contrast-enhanced imaging (DCE) showed that it can help to differentiate between benign and malignant pulmonary nodules [8–11]. However, these methods may not be suitable for patients who do not tolerate long examination times or who cannot receive intravenous contrast medium. Against this background, encouraging results have been obtained in a study using T1- and T2-weighted images acquired with a fast MRI protocol at 3 T and defining scaled signal intensities as non-enhanced imaging characterization quotients (NICQs) [12]: combining T2-NICQmean with T1-Qmean yielded 77% sensitivity and 95% specificity to differentiate benign and malignant pulmonary nodules. The advantage of this approach lies in its simplicity, both in terms of image acquisition and image analysis. To the best of our knowledge, NICQs have not been assessed in a larger collective of hematological patients.

The aim of the present study was therefore twofold: first, to assess the diagnostic performance of NICQs in differentiating bac-
terial and fungal infections from lymphoma manifestations in hematological patients and, second, to assess the reliability of NICQs in this setting.

Patients and Methods

Patients

The local ethics committee approved this prospective study (EA4/017/14). Written consent was obtained from all participants. The inclusion criterion was the presence of at least one solid pulmonary lesion in a recent, clinically indicated chest X-ray or computed tomography (CT) scan. Patients with general contraindications to MRI were excluded.

The EORTC/MSG diagnostic criteria [13] served as the standard of reference in fungal infections (at least probable). Histopathologic proof in lymphomas. Response clearly attributable to treatment was considered the standard of reference if the aforementioned information was not available.

MRI technique

All MRI examinations were performed on a 3 T MRI scanner (Magnetom Skyra, Siemens Healthineers, Erlangen, Germany) using a scan time-optimized protocol by Nagel et al. that was already successfully tested in a similar setting [12, 14] and originally derived from Biederer et al. and Attenberger et al. [15, 16]. The following three sequences were acquired with a surface coil positioned on the chest: axial and coronal T2-weighted (T2w) single-shot fast spin echo sequences (TR/TE 500/27, flip angle 160, matrix 256 × 320, slice 5 mm) and axial T1-weighted (T1w) gradient echo sequences (TR/TE 5.39/2.04, flip angle 9, matrix 180 × 320, slice 3 mm), all using multiple breath-hold regimens.

Image analysis

All images were read independently by two radiologists with different levels of experience (S.N. with > 7 years and T.W. with 1 year of MRI experience) and blinded to the clinical data. Regions of interest (ROIs) were drawn in the lesion, muscle, and fat on T2w images. At least one representative layer was chosen for the ROI, but if the lesion was well visible on multiple layers, then also multiple ROIs were allowed. Only ROIs with at least 10 voxels were considered for the final analysis. The ROIs in muscle and fat were placed as close as possible to the lesion and preferably at the same height in the phase-encoding direction. A sample set of ROIs is shown in ▶ Fig. 1.

To test for interrater reliability, S.N. and T.W. both read all datasets; to test for intrarater reliability, T.W. repeated the reading of the datasets from the first 19 consecutive patients. Details of the location of the lesions were available to correlate repeated readings.

Statistical analysis

Scaled signal intensities, defined as T2-NICQmean and T2-NICQ90th, were calculated for T2w images ((SI lesion – SI Muscle) / (SI Fat – SI Muscle)) *100) and as a simple quotient, T1-Qmean, for T1w images (SI lesion / SI Muscle) [12, 18]; calculations were done using the mean and 90th percentile of signal intensity within the lesion for T2-NICQmean and T2-NICQ90th, respectively.

Categorical parameters are given as frequencies. All metric data were tested for normal distribution using the Shapiro-Wilk test. For normally distributed data, descriptive statistics are given as mean and standard deviation. If no normal distribution was found, the median and interquartile range are provided.

Statistical testing included the Kruskal-Wallis or Mann-Whitney U-test, receiver operating characteristic (ROC) analysis, and calculation of intraclass correlation coefficients (ICCs). Unless otherwise stated, results are given on a per-lesion basis. Statistical analysis was done using R (Version 3.5.1) [19] or SPSS (SPSS Statistics, Version 25.0, IBM Corp., Armonk, NY, USA).

Results

A total of 83 lesions in 45 consecutive hematological patients with pulmonary nodules or masses diagnosed by routine clinical imaging were included in this study (24–76 years, median 59 years,
Lesion analysis

Data were not normally distributed. Median values of T1-Q\textsubscript{mean} and T2-NICQ\textsubscript{mean} differed significantly between the entities that were evaluated (p < 0.05). Median values of T2-NICQ\textsubscript{90th} were close to statistical significance (p = 0.054). Results are summarized in Table 1 and presented as boxplots in Fig. 3a–c.

In the analysis of infectious lesions in general vs. lymphoma manifestations, the median values differed significantly for T2-NICQ\textsubscript{mean} and T2-NICQ\textsubscript{90th}, but not for T1-Q\textsubscript{mean}. For bacterial pneumonias vs. lymphoma manifestations, the median values differed significantly for T2-NICQ\textsubscript{mean} and T2-NICQ\textsubscript{90th}, but not for T1-Q\textsubscript{mean}. On the per-patient level, the median values differed significantly for T2-NICQ\textsubscript{90th} only. For fungal lesions vs. lymphoma manifestations, the median values differed significantly for T1-Q\textsubscript{mean} and T2-NICQ\textsubscript{mean}, but not for T2-NICQ\textsubscript{90th}. On the per-patient level, the median values differed significantly for T2-NICQ\textsubscript{mean} and T2-NICQ\textsubscript{90th}, but no longer for T1-Q\textsubscript{mean}. Details are provided in Table 1.

T2-NICQ\textsubscript{mean} and T2-NICQ\textsubscript{90th} showed the best areas under the curve (AUCs) for comparing infectious and fungal lesions vs. lymphoma manifestations. When comparing the per-lesion and the per-patient analysis, the results for T1-Q\textsubscript{mean} differed only slightly while the AUCs of T2-NICQ\textsubscript{mean} and T2-NICQ\textsubscript{90th} clearly increased on the per-patient level. A summary of the analysis is provided in Table 2.

Inter- and intrarater reliability

The interrater reliability was consistently excellent (ICC above 0.94). The intrarater reliability in general was lower, especially for T2-NICQs (T2-NICQ\textsubscript{mean} 0.85, T2-NICQ\textsubscript{90th} 0.79). However, the results were still good. Detailed results are compiled in Table 3.

Discussion

The current data confirm that NICQs are suitable parameters for differentiating infections from lymphoma manifestations, as already suggested by the preliminary results by Nagel et al. [12]. Moreover, T1-Q\textsubscript{mean}, T2-NICQ\textsubscript{mean}, and T2-NICQ\textsubscript{90th} show almost consistently excellent inter- and intrarater reliability.

The results underline that especially T2-NICQ\textsubscript{mean} and T2-NICQ\textsubscript{90th} can be of help in distinguishing infectious lesions in general or fungal infiltrates in particular from pulmonary lymphoma manifestations. If findings still remain unclear, the quotients at least may help in the workup by providing early guidance for further diagnostics. Of note again is the simplicity of the approach, using fast standard sequences and easy to calculate parameters in the image assessment. The total scan time of approx. 12 minutes makes the protocol suitable for imaging critically ill patients [14].

Compared to a previous study in 29 patients that considered infectious vs. various malignant pulmonary nodules in general [12], the AUC of T1-Q\textsubscript{mean}, T2-NICQ\textsubscript{mean}, and T2-NICQ\textsubscript{90th} on a per-lesion basis is lower in the present study when considering infections vs. lymphoma manifestations (T1-Q\textsubscript{mean} 0.62 vs. 0.72, T2-NICQ\textsubscript{mean} 0.66 vs. 0.73, T2-NICQ\textsubscript{90th} 0.65 vs. 0.82). Nevertheless, regarding the critical issue of discriminating infections from pulmonary lymphoma manifestations in hematological patients, T2-NICQ\textsubscript{mean} and T2-NICQ\textsubscript{90th} showed promising results with an AUC of up to 0.76 with a specificity of up to 1.0 on a per-patient basis. Thus, the present findings confirm adequate diagnostic performance. It has to be noted that the number of lesions and patients evaluated mostly exceeds the numbers described in imaging-based thoracic diagnostic studies using MRI [8–11].

The concept of scaled signal intensities in MRI in general seems promising in the evaluation of pulmonary nodules. For example, ...
T1 and T2 values normalized to muscles correlated significantly with the standard uptake values (SUV) from PET/CT images in a study by Koo et al. that considered the differentiation of benign and malignant nodules [20]. In another study by Li et al., the qualitative assessment of the T1 and T2 signal intensity of nodules in relation to muscle and fat was used in the follow-up after cryoablation of lung tumors [21].

Future research might focus on the use of diffusion-weighted imaging (DWI), as it has been shown to be helpful in differentiating benign from malignant lesions in various studies [8, 9, 22–26]. Although this would prolong scanning time contrary to the concept of a strongly speed-optimized protocol, DWI also does not need iv contrast agents. Combining T1-Qmean and T2-NICQs with DWI should therefore be evaluated, especially when motion correction for free-breathing acquisition is available to image critically ill patients.

The second aim of our study was to assess the reliability of NICQ measurement as a diagnostic tool. The intra- and interrater reliability was good to excellent with T1-Qmean consistently showing the best results. Data on the inter- and intraobserver variability of NICQs have not been published before, so we can only speculate why ICCs are highest for T1-based measurement. One possible reason is that only two measurements are required and errors arising from a third measurement are avoided. In addition, a signal loss towards the center of the image was observed especially in T2w images. A possible explanation for this could be the use of surface coils and differences in the MR sequences: Gradient echo sequences are less susceptible to B1 inhomogeneities than fast spin echo sequences, which becomes more apparent at higher field strengths [27].

Thus, the obtained T1w images seem to provide more flexibility with respect to placing the ROI, which is represented by a
Infectious lesions in general and fungal infiltrates in particular were found to have higher T2 and lower T1 signal intensities. T2-NICQmean and T2-NICQ90th ((SILesion–SIMuscle)/(SIFat–SIMuscle) * 100); T1-Qmean (SILesion/SIMuscle).

Analysis of T1-Qmean, T2-NICQmean, and T2-NICQ90th on a per lesion and per patient level. AUCs, 95% confidence intervals and p-values are provided. AUC rating: 0.6–0.7 poor, 0.7–0.8 fair, 0.8–0.9 good, > 0.9 excellent. T2-NICQmean and T2-NICQ90th ((SILesion–SIMuscle)/(SIFat–SIMuscle) * 100); T1-Qmean (SILesion/SIMuscle). Optimal cutoffs were determined using Youden-Index; sensitivity and specificity are provided below. For T1-Qmean, values below the cutoff and for T2-NICQmean and T2-NICQ90th values above the cutoff indicate an infection. Please note the high specificity of T2-NICQs. ROC: receiver operator characteristic; AUC: area under the curve; SI: signal intensity.
smaller interquartile range of T1-Qmean. Nevertheless, the T2w images of the scan time-optimized protocol used here are also suitable for reliable evaluation, especially for the identification of lymphoma lesions with low T2-NICQ.

In general, the use of a higher field strength when scanning the lung may lead to susceptibility effects at air-tissue interfaces [28]. In this regard, Fink et al. could show that the imaging characteristics did not substantially differ between 1.5 T and 3 T and further, a better spatial resolution and a higher signal-to-noise ratio can be expected [29]. The latter can be seen as a major advantage of higher field strengths, as the amount of protons to produce a signal is relatively low in the lungs [28].

Interestingly, the intrarater agreement was slightly inferior to the interrater agreement. A possible explanation may be the lower experience level of the second reader. Nevertheless, also the less experienced radiologist achieved at least good agreement in repeated measurements. Data on the inter- and intrarater variability of scaled signal intensity are currently not available. Theoretically, the inter- and intrarater variability might be enhanced by the use of mapping techniques, which would overcome the use of scaling regions [30]. However, using mapping techniques in a pulmonary setting might be limited by breathing artifacts resulting from the need to acquire different echo times for one image.

This study has some limitations. First, histopathologic proof was not available in all patients. Thus, clinical response to treatment was the only standard of reference available in some cases. We tried to compensate for this drawback by only including patients in whom the therapeutic response could be clearly attributed to the initiation of a new medication. Second, imaging quotients were evaluated without further consideration of clinical information. Taking additional data into account, e. g., laboratory results, may further enhance overall diagnostic performance. Thirdly, no follow-up studies were considered, although these are helpful in assessing pulmonary changes, especially infections. In addition, repeated scans in general could be one major strength of pulmonary MRI, as they do not require ionizing radiation.

In conclusion, the overall diagnostic performance of T2-NICQs is adequate for differentiating infectious and fungal lesions from lymphoma manifestations. Results further show good to excellent intra- and interrater agreement. We therefore consider NICQs helpful in the diagnostic workup of pulmonary nodules in hematological patients.


