Smart scanning: automatic detection of superficially located lymph nodes using ultrasound – initial results

Smart Scanning: Automatisches Erfassen oberflächlicher Lymphknoten im Ultraschall – erste Ergebnisse

Authors
Maximilian Rink1, Julian Künzel1, Christian Stroszczynski2, Friedrich Jung3, Ernst Michael Jung2

Affiliations
1 Department of Otorhinolaryngology, University Hospital Regensburg, Regensburg, Germany
2 Department of Radiology, University Hospital Regensburg, Regensburg, Germany
3 Institute of Biotechnology, Brandenburg University of Technology Cottbus-Senftenberg, Senftenberg, Germany

Keywords
artificial intelligence, ultrasound, multimodal ultrasound, lymphatic, lymph node ultrasound, technical aspects

received 30.1.2024
accepted after revision 14.5.2024
published online 2024

Bibliography
Fortschr Röntgenstr
DOI 10.1055/a-2331-0951
ISSN 1438-9029
© 2024. Thieme. All rights reserved.
Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Correspondence
Maximilian Rink
Department of Otorhinolaryngology, University Hospital Regensburg, Regensburg, Germany
maximilian.rink@ukr.de

Supplementary Material is available at https://doi.org/10.1055/a-2331-0951.

ABSTRACT

Purpose Over the last few years, there has been an increasing focus on integrating artificial intelligence (AI) into existing imaging systems. This also applies to ultrasound. There are already applications for thyroid and breast lesions that enable AI-assisted sonography directly on the device. However, this is not yet the case for lymph nodes.

Materials and Methods The aim was to test whether already established programs for AI-assisted sonography of breast lesions and thyroid nodules are also suitable for identifying and measuring superficial lymph nodes. For this purpose, the two programs were used as a supplement to routine ultrasound examinations of superficial lymph nodes. The accuracy of detection by AI was then evaluated using a previously defined score. If available, a comparison was made with cross-sectional imaging.

Results The programs that were used are able to adequately detect lymph nodes in the majority of cases (78.6%). Problems were caused in particular by a high proportion of echogenic fat, blurred differentiation from the surrounding tissues and the occurrence of lymph node conglomerates. The available cross-sectional images did not contradict the classification of the lesion as a lymph node in any case.

Conclusion In the majority of cases, the tested programs are already able to detect and measure superficial lymph nodes. Further improvement can be expected through specific training of the software. Further developments and studies are required to assess risk of malignancy.

Key Points
- The inclusion of AI in imaging is increasingly becoming a scientific focus.
- The detection of lymph nodes is already possible using device-integrated AI software.
- Malignancy assessment of the detected lymph nodes is not yet possible.

Citation Format

ZUSAMMENFASSUNG

Hintergrund In den letzten Jahren rückt die Integration künstlicher Intelligenz (KI) in bestehende Bildgebungen zunehmend in den Fokus. Dies gilt auch für die Sonografie. Für Läsionen der Schilddrüse sowie der Mamma existieren bereits Anwendungen, die unmittelbar am Gerät eine KI-assistierte Sonografie ermöglichen. Für Lymphknoten ist dies bisher nicht der Fall.

Material und Methoden Getestet wurde, ob bereits etablierte Programme zur KI-assistierten Sonografie von Läsionen der Brust beziehungsweise Schilddrüsenknoten sich grundsätzlich auch dazu eignen, oberflächliche Lymphknoten...
Introduction

Ultrasound is a broadly available imaging technique with minimal side effects, and it offers a wide range of applications, including the detection and assessment of lymph nodes. Indications for the sonographic assessment of lymph nodes include, for example, the detection of cancer spread, re-staging after oncological treatment or, as a complement to clinical examination, the differential diagnosis of unclear lymph node enlargement [1, 2, 3, 4, 5, 6].

Once the examiner has detected the lymph nodes as part of the routine sonographic examination, the status of the lymph nodes must be assessed. When assessing lymph nodes, the shape, the ratio of cortex to medulla, the presence of a hilar sign, echogenicity, the pattern of vascularization, the extent of indurcation, and the perfusion all play a decisive role [7]. Measurement on at least two planes is required for complete documentation. This is particularly necessary if sonographic follow-up examinations are scheduled. Automated detection and measurement with the help of algorithms based on artificial intelligence (AI) immediately after the examination could be helpful here.

Modern ultrasound technology using high-performance equipment and multi-frequency probes can be used to combine detailed B-scan technology, high-resolution color-coded duplex procedures, ultrasound elastography techniques, and contrast-enhanced ultrasound (CEUS) with perfusion analysis [8, 9]. This allows the examiner to obtain a multi-parametric “overall impression” of the lymph node and use this as a basis for assessment [10]. AI also offers promising potential regarding the generation and time-consuming evaluation of this overall impression. For example, initial studies have already shown that a combination of AI and CEUS can provide a more accurate status assessment of thyroid nodules [11].

Research on the additional assessment of sonographic findings using AI has been steadily increasing in recent years. Applications have been described in particular for the assessment of breast lesions [12, 13], the thyroid gland [14, 15, 16, 17], as well as for cardiac ultrasound [18]. One particular difficulty for the successful application of AI in the assessment of ultrasound images is the high variability in the quality and settings of the respective findings [19]. However, modern AI programs work via pattern recognition, which is made considerably more difficult due to this variability. With regard to the input data, the basic rule for AI is: “garbage in, garbage out”.

AI-supported sonographic assessment of superficial lymph nodes has only been minimally investigated. Initial studies show, for example, that there is potential for improved identification of axillary lymph node metastases in patients with breast cancer [20] or for sonographic differential diagnosis of the various causes of enlarged cervical lymph nodes [21]. Until now, however, the ultrasound images have usually been evaluated afterwards using special software external to the actual ultrasound device. Nevertheless, current ultrasound devices already offer the integrated option of AI-supported identification and evaluation of lesions on the ultrasound device (onboard application). Initial studies are investigating the quality of these programs for the differential diagnosis of various lesions [17, 22, 23]. As far as we know, there are no corresponding smart scanning programs specifically for lymph nodes. The aim of this study was, therefore, to examine the extent to which already established AI programs could also be used for superficially located lymph nodes. This is of interest from our point of view, as these programs can already be used in everyday clinical practice.

The primary goal of the study was to determine the suitability of the two AI programs for the automated identification and measurement of lymph nodes. Ideally, once a node has been correctly identified, the AI program should automatically and accurately capture the edge and perform a measurement on both planes of the respective image. The determined scores formed the basis for answering this question.

Materials and Methods

The study included superficial lymph nodes regardless of their location (e.g., cervical, inguinal, or axillary) and regardless of the patient’s basic disease. All examinations were performed in the interdisciplinary ultrasound center of our hospital as part of the clinical routine. All examinations were performed by an experienced examiner (DEGUM III, more than 20 years of experience in ultra-
sound, more than 3000 examinations per year) and stored digitally in the Picture Archiving and Communication System (PACS) to record the findings. The PACS files were then analyzed retrospectively. A positive ethics vote from the local ethics committee was available for the retrospective evaluation. The AI program was used for patient lymph nodes that were documented with an image as part of the clinical question and not for every visible lymph node of the patient. Smart thyroid was available for a considerably shorter period of time than smart breast. A comparison of the two programs was not the aim of our study.

Two different ultrasound devices were used for the examinations: Resona 7 (Mindray, Shenzhen, China) and Resona R9 (Mindray). An AI program for the detection and assessment of breast lesions (smart breast) was used on the Resona 7, while an AI program for thyroid nodules (smart thyroid) was used on the Resona R9. The detection program for breast tumors (smart breast, Resona 7, Mindray) was used for 52 lymph nodes in 34 patients. The more recent AI program for thyroid tumors (smart thyroid, Resona R9, Mindray) was used in 4 patients with 4 lymph nodes.

Smart breast was originally developed for the sonographic assessment and status evaluation of breast lesions. The BI-RADS criteria required for this are stored in the software and are automatically analyzed after the lesion has been detected, although manual corrections can be made by the examiner [23]. The software relies on the best possible image settings by the examiner, as the analysis is based on this data. The program cannot perform a fully automated tomographic examination with scanning of the lesion, data set generation, and interpretation. The same applies to smart thyroid, where the TIRADS criteria are stored accordingly.

The examination to identify lymph nodes using smart breast was carried out according to the following procedure:
1. Adjustment of a superficially located lymph node on the B-scan
2. Start of the “smart breast” program
3. Definition of the axis orientation in the ultrasound image (sagittal or transversal)
4. Start of the software-supported lesion search. The program then identifies the lymph node and circles its edge.
5. Start of the automatic measurements (the two diameters that can be measured within the selected plane are measured here, for example the ventrodorsal and mediolateral diameters in the transverse section)
6. Saving of the findings

Modifications can be made to the smart breast program at various points in the examination, such as correcting the automatically drawn edge. However, this was not done in the study, as the aim was to find out how well the automatic definition of the margin and the measurement points works. The procedure in the smart thyroid program was equivalent.

Both smart breast and smart thyroid are programs that have been evaluated by the software manufacturer and are available to buy for routine users to assist in the differential diagnosis of breast lesions and thyroid nodules, respectively. Guiblan et al. investigated the agreement in the assessment of breast lesions according to BI-RADS between the examiner and smart breast [23]. Before being used for lymph node sonography, smart breast was tested at our hospital on 100 breast lesions with a diameter of at least 10 mm in order to check the functionality of the program as described. Incorrect markings or identifications were corrected manually. For the smart thyroid program, an equivalent procedure was carried out with 100 thyroid nodules with a diameter of at least 8 mm.

Following this first phase, an attempt was made to automatically detect 10 lymph nodes, each with a diameter of at least 10 mm. In this second step, any errors that occurred were also corrected manually; the results up to this point were not used for this study or everyday clinical practice. The software was then set up on-site and used as an addition to the routine examination according to the standard. Therapy-relevant decisions regarding the examined patients were never dependent on the result of the AI program. AI errors were no longer corrected manually, and no further training in any other form took place, as the aim was to evaluate whether the two programs in their current form were already suitable for detecting and measuring lymph nodes. A total of three linear probes were used, one of which has a somewhat wider frequency spectrum than the other two (1 × 3–15 MHz and 2 × 3–9 MHz).

The evaluation assessed whether the AI software was able to recognize the lymph nodes and correctly determine their size. For this purpose, a score from 0 (lymph node not detected at all) to 5 (lymph node fully detected, complete margin accurately captured and without any artifact formation) was created. The levels of the score were determined before the images were analyzed. The different score levels are shown in Table 1, representative examples are shown in Fig. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5. The score was determined in consensus by two of the authors. A score of 4 or 5 was considered “good”, a score of 3 was considered “adequate”. All lower scores meant an “inadequate” result. The AI did not assess whether the findings were malignant or benign. This is currently not possible with the AI programs used.

Wherever possible, cross-sectional imaging performed close to the time of the ultrasound examination (maximum interval of 4 weeks) was used for comparison in order to verify the evaluation of a sonographic finding as a lymph node. The reports produced by radiologists (CT, MRI) or specialists in nuclear medicine (PET-CT) as part of their routine clinical practice were used as a reference.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Presentation of the score. LN = lymph node.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>LN not recognized</td>
</tr>
<tr>
<td>1</td>
<td>LN is recognized, but rough misinterpretation of the limits</td>
</tr>
<tr>
<td>2</td>
<td>LN detected, but significant errors when circling the edge or setting the measuring points/no setting of measuring points</td>
</tr>
<tr>
<td>3</td>
<td>LN recorded, and edge largely captured correctly</td>
</tr>
<tr>
<td>4</td>
<td>Edge of the LN only almost completely correctly circled or measurement only in one dimension with completely correctly detected edge; no relevant artifact</td>
</tr>
<tr>
<td>5</td>
<td>Complete edge of the LN precisely captured and no artifact; measurement of both axes mandatory</td>
</tr>
</tbody>
</table>
Following the first evaluation of the images, all findings with a score of 2 or lower were analyzed again to identify possible causes for the inferior detection by the AI software. During the first evaluation, possible factors that were noticed were the close presence of low-echo structures such as vessels, the presence of a lymph node conglomerate, a blurred demarcation from the surrounding area, and a high proportion of echogenic fat. For this reason, specific attention was paid to these criteria during the review, as well as to artifacts in the image and an overall unfavorable image quality.

The evaluations in this study were retrospective and purely exploratory.

Results

Table 2 summarizes the scores achieved by both programs.

In total, 30 of the 56 lymph nodes (53.6 %) achieved a score of 4 or 5 and thus a “good” result, while another 25 % achieved an “adequate” result with a score of 3. Only 12 of the 56 lymph nodes (21.4 %) had a score of less than 3 and were therefore assessed as “inadequate”.

As described above, the images with a score of 2 or lower were subsequently analyzed again. The exact individual analysis for the follow-up of the images with a score of 2 or lower is shown in Table 3. Overall, it can be seen that the AI program had problems delimiting the margin, particularly in lymph nodes with a high proportion of echo-rich fat (Fig. 3). The division of a con-
glomerate into individual lymph nodes was also not successful (▶ Fig. 1).

In the case of the unrecognized lymph node with a score of 0 (corresponding to lymph node 9 in ▶ Table 3), the AI program mistook a neighboring vessel for a lymph node (▶ Fig. 6).

In none of the cases did the cross-sectional imaging – if performed – contradict the sonographic evaluation of the lesion by the AI as a lymph node. It should be noted that cross-sectional imaging was only available for 22 lymph nodes for comparison. In total, an MRI examination was available for 2 nodes, a CT scan for 13, and a PET-CT scan for 7 for comparison. In the remaining cases, no cross-sectional imaging was available for comparison due to the retrospective approach of this study. A detailed list can be found in Supplement 1. Even though the cases with cross-sectional imaging as a comparison did not show any false-positive results, the case presented in ▶ Fig. 6 demonstrates that false-positive results do occur.

Overall, it should be noted that only certain regions were examined sonographically in this study, all of which were in a superficial location. “Good” identification and measurement of the lymph node by AI was only possible in slightly more than half of the cases (53.6 % score 4 and 5) and still requires the best possible visualization by the examiner. Good delineation of the node from surrounding structures also improves the accuracy of the AI program.

### Discussion

The aim of this pilot study was to test whether an onboard AI application for diagnostic breast or thyroid workup can also be used to identify lymph nodes. After a very short initial training period as described in the materials and methods section, the system can recognize lymph nodes and record their size. This can make the procedure easier and faster for the examiner.

Research on the integration of AI into ultrasound as well as other imaging methods is increasingly becoming a focus of various disciplines [24, 25]. In the field of ultrasound, most studies analyze whether the integration of AI can improve the differential diagnosis of lesions, for example of the thyroid gland [14, 15, 16, 17] or the breast [12, 13].

Determining the status of lymph nodes using ultrasound is a frequent and highly relevant clinical question. In principle, there is great potential for improving the diagnostic workup with the additional assessment by AI. The first step is always the correct identification and marking of a lesion as a lymph node. This can

### Table 2

Summary of the scores achieved with the respective AI programs and their incidence rates. The percentages are rounded to the nearest whole number.

<table>
<thead>
<tr>
<th></th>
<th>Score 0</th>
<th>Score 1</th>
<th>Score 2</th>
<th>Score 3</th>
<th>Score 4</th>
<th>Score 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smart Breast, Resona 7</td>
<td>1 (2%)</td>
<td>2 (4%)</td>
<td>9 (17%)</td>
<td>14 (27%)</td>
<td>16 (31%)</td>
<td>10 (19%)</td>
</tr>
<tr>
<td>Smart Thyroid, Resona R9</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>3 (75%)</td>
<td>1 (25%)</td>
</tr>
</tbody>
</table>

### Table 3

Presentation of possible influencing factors for lymph nodes detected with a score of 2 or worse. The numbers refer to the consecutive numbering of the lymph nodes as part of the data collection. LN: lymph node.

<table>
<thead>
<tr>
<th>LN 2</th>
<th>Proximity to echo-poor structures</th>
<th>Conglomerate</th>
<th>Blurred demarcation from surroundings</th>
<th>High fat content</th>
<th>Artifacts</th>
<th>Unfavorable image quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>LN 9</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>LN 12</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>LN 15</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>LN 17</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>LN 18</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>LN 21</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>LN 32</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>LN 42</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>LN 47</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>LN 48</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>LN 51</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>
be done by the examiner, with subsequent assessment of the lymph node by the AI program. However, automatic detection of the node with direct subsequent analysis by the AI software would be more time-efficient. The examiner would then only have to select a suitable image and everything else would be done automatically. As a first step in this study, we focused on whether lymph nodes can be reliably detected and their extent determined by the AI programs available on-site.

In our study, a score of 3 or higher was achieved for 78.6% of the lymph nodes. Considering that the programs that we used were not primarily developed for the identification of lymph nodes, we consider this to be a promising result. In particular, it underlines the high potential of AI to improve the identification and interpretation of lymph nodes using onboard AI applications, possibly with special programs in the future. An increase in accuracy appears possible through higher case numbers and thus increased training of the AI.

As far as we know, this is the first study examining the AI-assisted identification and measurement of lymph nodes right at the ultrasound machine using onboard AI software.

Despite these encouraging results, some limitations of our study must be taken into account. The images were acquired as part of routine clinical practice, not in a prospective study setting. Therefore, image settings such as orientation, gain, or frequency are not standardized but rather were defined individually for each patient. The evaluation was performed by two examiners in consensus, so there was no blinding to each other’s assessment and one of the two examiners was also the original investigator. As a pilot study, the aim was to show that the method works in principle. The number of cases is therefore still small. The comparison with cross-sectional imaging is limited by the fact that cross-sectional imaging from an external radiology/nuclear medicine department may exist for the nodes without comparative imaging, which was not accessible due to the retrospective nature of the study. Due to the retrospective approach, it was only possible to evaluate images collected and stored in the everyday clinical routine. Therefore, no statement can be made about the susceptibility of the AI software to false-positive results. The example in Fig. 6 shows that false-positive findings can occur (i.e., misinterpretation of a structure other than a lymph node). Further prospective studies are required for a more precise statement on this.

For the future aim of improving the assessment of lymph node status by adding AI, one important limitation must be mentioned. Currently, only a selected section through the lymph node is analyzed by the mentioned programs. However, an assessment of the entire lymph node in all planes would be necessary for a status assessment. The fact that AI can in principle be used to assess the status of circumscribed lesions in real time has already been demonstrated for the thyroid gland [17] and liver masses [22], for example.

Developers and manufacturers of ultrasound AI software should develop lymph node specific applications and these programs should then be validated in prospective studies. A second step would be the use of tomographic ultrasound images for AI-supported assessment of status.

**Conclusion**

Our study takes a first step by showing that AI programs can, in principle, detect and measure lymph nodes directly on the device in a sectional plane. An assessment of lymph node status is not yet possible with the programs used. Further developments and studies are required for this.

**Conflict of Interest**

Julian Künzel received fees for lectures from GE Healthcare (Chicago, IL, USA).

**References**

