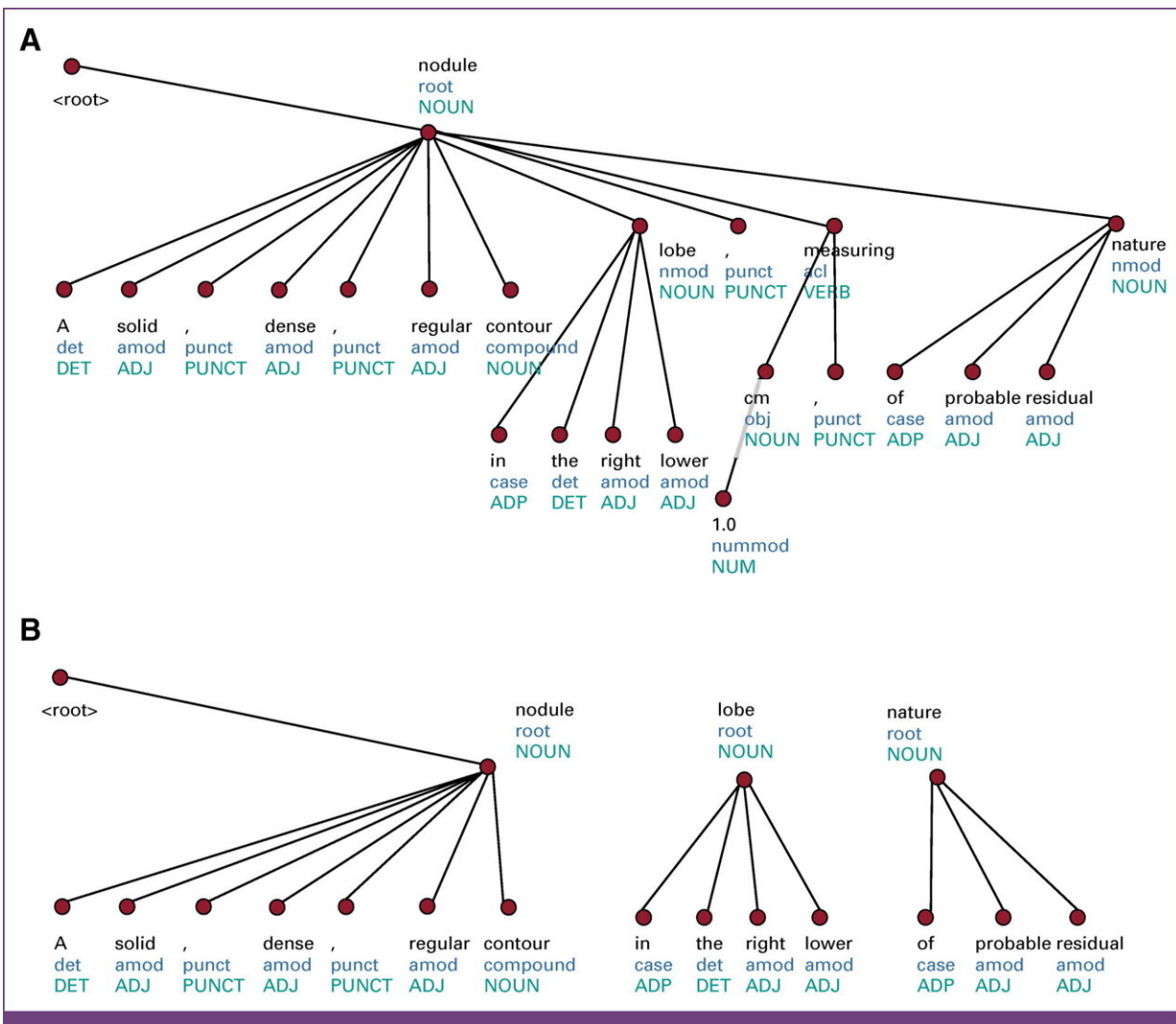


AI is capable of detecting incidental lung cancer in written medical reports

October 10 2023



An example of our information extraction algorithm. (A) Initial dependency tree. (B) Final information chunks. Credit: *JCO Global Oncology* (2023). DOI: 10.1200/GO.23.00191

Recently published in [JCO Global Oncology](#), a new article presents the development of artificial intelligence (AI) for the accurate detection of potentially cancerous nodules described in computed tomography (CT) reports, conducted outside of the cancer screening context.

The technology is a [natural language processing](#) (NLP) tool, trained to analyze chest CT reports. The study was conducted by the D'Or Institute for Research and Education (IDOR), in partnership with the Federal University of Health Sciences of Porto Alegre (UFCSPA), and the Universities of Florida and Stanford in the United States.

Natural language processing advancements in artificial intelligence have frequently made headlines in recent years, and while it may still be a topic of futuristic narratives, the truth is that the use of this technology is already well-established and surrounds us daily. If you've ever used online banking chatbots, enabled automatic video captions, or issued commands to virtual assistants on your devices, you may have wondered how machines can understand human communication so well.

Understanding and responding to [human language](#) is the primary goal of natural language processing (NLP), one of several fields within artificial intelligence. NLP is used for several functions in many well-known tools, including chatbots and digital assistants like Siri and Alexa. However, its potential is also being explored in the medical field.

With this in mind, the current study sought to invest in the technology to develop a tool capable of detecting the possibility of lung [cancer](#) through CT reports that had been performed for reasons other than cancer screening.

"When we perform a chest CT, in the context of lung cancer, there are

two main indications. One is for cancer screening in a patient at risk, usually over 50 years old with a smoking history. The other context is a patient who has, for example, a suspicion of a pulmonary embolism and undergoes a CT scan to investigate that, and an incidental nodule appears in that exam."

"This is called an incidental finding. Our NLP tool operates in this latter situation because sometimes these patients may lose the opportunity to receive an [early diagnosis](#). The incidental finding is easier to overlook because the doctor is focused on another hypothesis and may not suspect those details in the initial interpretation," explains Dr. Rosana Rodrigues, a radiologist researcher at IDOR and one of the study's authors.

Lung nodules

When analyzing chest images, the presence of lung nodules is a relatively common finding, and the majority of them are benign. However, serious risks can hide in 1 to 3% of these cases. Due to their high prevalence, lung nodules are often overlooked in emergency hospital exams.

A prime example of this occurred during the COVID-19 pandemic, where CT scans were frequently performed in clinics and hospitals to identify lung involvement in the disease. In this scenario, many of these nodules were described in medical reports but were not adequately investigated for their cancer potential. Identifying the issue in its early stages provides an opportunity to apply more effective therapies, with a greater chance of curing patients.

"We had the idea for the tool during the pandemic because we were performing 30 to 40 CT scans per day on patients suspected of COVID-19. The doctors who were requesting the CT had a complete focus on the disease because they needed to know if the patient required hospitalization. When we were writing the reports, in addition to the

presence or absence of COVID-19 lung impairment, we also saw many lung nodules, some of which were suspicious for lung cancer," says Dr. Rodrigues.

"That raised a huge concern for us, because during the pandemic, no one would retrieve these reports, so this cancer diagnosis could be lost. That's when we started thinking about how we would recover all these suspicious exams. That's when we've planned the NLP tool," recalls the radiologist, who also works at 3 hospitals in Rio de Janeiro, including the public hospital of the Federal University of Rio de Janeiro (UFRJ).

The possibility of identifying lung cancer in patients who are not suspected of having the problem encouraged the article researchers to consider solutions for this diagnostic window. That's when they had the idea of developing an automated NLP tool capable of searching for suspicious nodules identified incidentally in chest CT medical reports.

Teaching the machine

Like us, artificial intelligence isn't born knowing. To train the NLP tool developed for the study, the radiologists on the team conducted a retrospective analysis of over 21,500 chest CT reports performed at a research-affiliated hospital between 2020 and 2021. Of these thousands of exams, 484 presented incidentally detected lung nodules with potential carcinogenicity, whose descriptions were used to train the NLP tool in identifying these lesions.

After training, the NLP underwent internal validation involving the assessment of over 300 chest CT reports, with 157 of them containing incidentally detected nodules suspected of malignancy and 148 serving as control group to calculate the tool's accuracy potential.

The NLP was taught to understand the text written in the reports,

without access to images. Researchers programmed it to report as suspicious any incidentally detected nodules not previously known in the patient's history, with a diameter greater than 4 mm and without clinical context associated with cancer, pneumonia, or small airway disease. The tool was also capable of categorizing higher-risk nodules, such as those larger than 8 mm or those with a solid component greater than 6 mm.

In the internal evaluation, the NLP tool achieved an accuracy of 98% in detecting the [nodules](#) of interest. The positive results led the researchers to conduct a second test in May 2022, this time analyzing over 900 chest CT reports, which were randomly selected from 57 different hospitals.

In this second test, the NLP demonstrated an even more impressive accuracy of 98.6%, which was further validated by a final check by radiologist doctors, establishing a gold standard for the tool's testing. These results reaffirmed the accuracy and competence of artificial intelligence for assistance in clinical applications.

Of the 484 incidental findings used for NPL training, 8 patients were diagnosed with lung cancer and were able to undergo early treatments. According to the study, 2 of these diagnoses could have been missed without the assistance of the NLP tool.

Considering that 70% of lung cancer cases are curable when treated in precocious stages, early detection of the disease significantly improves these chances for patients. The [artificial intelligence](#) was developed in a widely used programming language, Python, and its use is compatible with several institutions and hospitals where Portuguese is the official language. The tool may contribute significantly to the early recognition of [lung](#) cancer, especially in patients treated in emergency services and outside specialized oncology centers.

More information: Rodrigo Basilio et al, Natural Language

Processing for the Identification of Incidental Lung Nodules in Computed Tomography Reports: A Quality Control Tool, *JCO Global Oncology* (2023). [DOI: 10.1200/GO.23.00191](https://doi.org/10.1200/GO.23.00191)

Provided by D'Or Institute for Research and Education

Citation: AI is capable of detecting incidental lung cancer in written medical reports (2023, October 10) retrieved 28 April 2024 from <https://medicalxpress.com/news/2023-10-ai-capable-incident-lung-cancer.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.