Lung CT Screening Reporting and Data System Speed and Accuracy Are Increased With the Use of a Semiautomated Computer Application

Toshimasa J. Clark, MD[a], Thomas F. Flood, MD, PhD[a], Suresh T. Maximin, MD[b], Peter B. Sachs, MD[a]

Abstract

Purpose: The Lung CT Screening Reporting and Data System (Lung-RADS™) is an algorithm that can be used to classify lung nodules in patients with significant smoking histories. It is published in table format but can be implemented as a computer program. The aim of this study was to assess the efficiency and accuracy of the use of a computer program versus the table in categorizing lung nodules.

Methods: The Lung-RADS algorithm was implemented as a computer program. Through the use of a survey tool, respondents were asked to categorize 13 simulated lung nodules using the computer program and the Lung-RADS table as published. Data were gathered regarding time to completion, accuracy of each nodule's categorization, users' subjective categorization confidence, and users' perceived efficiency using each method.

Results: The use of a computer program to categorize lung nodules resulted in significantly increased interpretation speed (80.8 ± 37.7 vs 156 ± 105 seconds, $P < .0001$), lung nodule classification accuracy (99.6% vs 76.5%, $P < .0001$), and perceived confidence and efficiency compared with the use of the table. There were no significant differences in accuracy when comparing thoracic radiologists with the remainder of the group.

Conclusions: Radiologists were both more efficient and more accurate in lung nodule categorization when using computerized decision support tools. The authors propose that other institutions use computerized implementations of Lung-RADS in the interests of both efficiency and patient outcomes through proper management. Furthermore, they suggest the ACR design future iterations of the Lung-RADS algorithm with computerized decision support in mind.

Key Words: Lung-RADS, decision support, lung cancer screening

J Am Coll Radiol 2015;12:1301-1306.

INTRODUCTION

The ACR's Lung CT Screening Reporting and Data System (Lung-RADS™) is an algorithm that can be used to classify lung nodules' risk for malignancy in patients with significant smoking histories [1]. Its results are generalizable to patients who meet the inclusion criteria for the National Lung Screening Trial, namely age 55 to 74 years with a history of cigarette smoking of at least 30 pack-years [2]. The Lung-RADS categories are defined in Figure 1. Each category is associated with a probability of malignancy that is derived from a simplification of the parsimonious logistic regression model described by McWilliams et al [3], and management recommendations follow from this inferred risk for malignancy.

As part of our institution's initiative to institute a lung cancer screening clinic, we elected to implement Lung-RADS as a computer program, positing that the use of this computer program to categorize lung nodules would increase efficiency and user satisfaction relative to the use of the published Lung-RADS table. We predicted no difference in the accuracy of nodule categorization between these methods.

METHODS

The Lung-RADS Algorithm as a Computer Application

The Lung-RADS algorithm was implemented as a computer program written in the JavaScript programming

[a]University of Colorado Denver, Aurora, Colorado.

[b]University of Washington and VA Puget Sound Health Care System, Seattle, Washington.

Corresponding author and reprints: Toshimasa J. Clark, MD, University of Colorado Denver, 12401 E 17th Avenue, Mail Stop L-954, Aurora, CO 80045; e-mail: toshimasa.clark@ucdenver.edu.

The authors have no conflicts of interest related to the material discussed in this article.

1546-1440/15/$36.00 ■ http://dx.doi.org/10.1016/j.jacr.2015.07.015

language. This application's source code is freely available as Open Source under the GNU Public License (GNU, Cambridge, Massachusetts), and the application is accessible on the Internet at http://raddecisionsupport.com/apps/lung_rads/index.html.

Empirical Survey

Two surveys were created in Google Forms (Google, Inc, Mountain View, California), reproduced in Online Supplemental Appendices 1 and 2. They were administered to rotating high school students, medical students, residents, fellows, and attending physicians at our institution. Each collected basic demographic information about the participant, such as training level and whether the participant was a subspecialized thoracic radiologist. The first presented the participant with three hypothetical nodules and the second with ten hypothetical nodules, described in text form without the use of protected health information. The characteristics of the three initial nodules were chosen arbitrarily, whereas those of the second set of nodules were generated such that the population incidence of solid, semisolid, and nonsolid nodules and the size distribution within each of these categories were approximated [4].

In both of these surveys, each participant categorized these nodules using the published Lung-RADS table (Fig. 1) and entered the time to completion as well as the derived Lung-RADS category and management recommendation. The participant then used the Lung-RADS computer application to categorize the same hypothetical lung nodules. Completion time, Lung-RADS category, and management recommendation were recorded in a linked Google Sheets spreadsheet upon form submission. Subjective data regarding the participant's perceived efficiency, confidence in characterization using the two methods, and general comfort level with technology were also recorded in the first survey using a 5-point Likert-type scale. The linked spreadsheets for these surveys were exported to Microsoft Excel format, with further statistical analysis performed in Excel using the Analysis ToolPak (Microsoft Corporation, Redmond, Washington). Continuous data (time to completion for each of these tasks) was analyzed using Student's t test (two tailed), with expectation of zero variance between groups. The 5-point Likert-type scale subjective data from the first survey and the discrete binary accuracy data were analyzed using χ^2 tests. P values less than .05 were considered to indicate statistical significance.

RESULTS

Twenty-one respondents completed the first survey, including six residents, seven body imaging attending radiologists, one body imaging fellow, six thoracic imaging attending radiologists, and one thoracic imaging fellow. Eighteen respondents completed the second survey, including one high school student, two medical students, five residents, four body imaging attending radiologists, one body imaging fellow, four thoracic imaging attending radiologists, and one thoracic imaging fellow. Expected categorization and accuracy for each of the methods for the 13 nodules are summarized in Table 1, with significant differences denoted with asterisks. It is of note that four of the nodules demonstrated ambiguity in their categorization because the Lung-RADS table has non-mutually exclusive category definitions, inconsistency in category definitions between the table and its footnotes, and a lack of defined order-of-operations rules regarding suspicious features and stability. For these nodules with ambiguous categorization, either category was accepted as a valid answer for both the table- and computer-based methods.

On a per nodule basis, there was overall categorization accuracy of 76.5% using the table-based method. There was no significant difference between thoracic radiologists and the remainder of the group (77.5% vs 74.2% accuracy, $P = .60$). Categorization accuracy using the computer-based method was 99.6% (a statistically significant difference relative to table-based method, $P < .0001$). Categorization accuracy specific to each nodule is described in Table 1, with significant differences denoted with asterisks.

For the table-based method, the mean time to categorize and extract management recommendations for each patient (the implicit patient in the first survey having three nodules and the three patients from the second survey having three, three, and four nodules, respectively) was 156 ± 105 seconds. For the computer-based method, the mean time to categorize and extract management recommendations for each patient was 80.8 ± 37.7 seconds (a statistically significant difference relative to the table-based method, $P < .0001$). There was no statistically significant between-group difference in time to characterize nodules using the table-based method when comparing thoracic radiologists with the remainder of the group ($P = .50$). There was a statistically significant difference when using the computer-based method, with thoracic radiologists requiring more time to use the computer application compared with the remainder of the group: 99.0 ± 48.5 versus