Diagnostic concordance of thoracic images for Covid-19 disease

DOI: 10.5377/alerta.v7i2.18157

Neftali Alexander Granados Flores1,2, Marlon Alexander Lazo Ferrufino2


*Correspondence
neftaligranados.dr@gmail.com

Abstract
Introduction. Reverse transcription polymerase chain reaction (RT-PCR) is the gold standard method for diagnosing SARS-CoV-2 disease. However, due to limited accessibility to this test during the pandemic, diagnostic imaging was used to support diagnostic suspicion and avoid delays in medical care. Objective. Determine the accuracy of diagnostic imaging (chest X-ray and computed tomography) in diagnosing SARS-CoV-2 infection, compared to RT-PCR result. Methodology. An analytical cross-sectional study was conducted. The imaging reports of 138 patients were compared with their RT-PCR results to calculate sensitivity, specificity, positive predictive value, and negative predictive value for both chest X-ray and computed tomography. Concordance between the imaging results and RT-PCR was measured using Cohen’s Kappa index and Bayes factor. Results. Computed tomography showed a sensitivity of 92.9%, a specificity of 64%, a positive predictive value of 92.1%, and a negative predictive value of 66.7%. On the other hand, X-rays showed a sensitivity of 86%, a specificity of 52.9%, a positive predictive value of 92.9%, and a negative predictive value of 34.6%. Conclusion. Computed tomography showed moderate diagnostic concordance and is particularly useful in cases of moderate to high clinical suspicion, diagnostic discrepancy, or the need to confirm complications. On the other hand, X-rays showed low diagnostic concordance and should be used in combination with RT-PCR for a definitive diagnosis, especially in cases of high clinical suspicion.

Keywords
SARS-CoV-2, Sensitivity and Specificity, Tomography, X-rays, Predictive Value of Tests.

Introduction
COVID-19, a disease caused by SARS-CoV-2, a virus belonging to the Coronaviridae family, is characterized by its high contagiousness, morbidity and mortality. This disease triggered the most important health emergency of the modern era, causing the collapse of health systems in many countries due to the consumption of human and financial resources, claiming the direct or indirect lives of 14.9 million people in its first two years.1

Resumen
Introducción. La reacción en cadena de la polimerasa de transcripción inversa (RT-PCR) es el estándard de oro para el diagnóstico de enfermedad por SARS-CoV-2. En el contexto de la pandemia con accesibilidad limitada a esta prueba, las imágenes diagnósticas aportaron hallazgos que sustentan la sospecha diagnóstica, evitando retrasos en atención médica. Objetivo. Determinar la sensibilidad, especificidad, valor predictivo positivo y negativo de las imágenes diagnósticas y su concordancia respecto al resultado de RT-PCR. Metodología. Estudio transversal analítico. Se comparó el resultado del reporte por imágenes con los resultados de RT-PCR en 138 pacientes. Se calculó la sensibilidad, especificidad, valor predictivo positivo y negativo, para los rayos X de tórax y tomografía computarizada para el diagnóstico de infección por SARS-CoV-2. Se utilizó el Índice Kappa de Cohen y el factor de Bayes para medir la concordancia y fuerza de asociación entre las variables. Resultados. La tomografía computarizada presentó una sensibilidad de 92,9%, una especificidad del 64%, un valor predictivo positivo de 92,1% y un valor predictivo negativo de 66,7%; mientras que, los rayos X presentaron una sensibilidad del 86%, una especificidad del 52,9%, un valor predictivo positivo de 92,9% y un valor predictivo negativo del 34,6%. Conclusión. La tomografía mostró concordancia diagnóstica moderada, su utilidad es mayor en casos de sospecha clínica moderada-alta, discrepancia diagnóstica o confirmación de complicaciones. Los rayos X mostraron concordancia diagnóstica baja, este método es de utilidad en casos de alta sospecha clínica, pero necesita comprobación con RT-PCR para un diagnóstico definitivo.

Palabras clave
SARS-CoV-2, Sensibilidad y Epecificidad, Tomografía, Rayos X, Valor Predictivo de las Pruebas.

Conflict of interests:
The author declare there are not conflict of interests.

© 2024 by the authors.
This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).
In El Salvador, the first reported case of SARS-CoV-2 was on March 18, 2020. Since then, the incidence spread exponentially throughout the national territory, reaching 555 confirmed cases, with 12 deaths and 26 accumulated recovered cases up to May of the same year. This generated the need for accurate and timely diagnoses, so the Ministry of Health indicated the performance of 29,931 RT-PCR tests until May 2020, i.e., an average of 599 tests per day nationwide. RT-PCR is known to be highly specific but has the disadvantage of variable sensitivity (60-70% to 95-97%), depending on the stage of the disease at which the test is performed. False negatives are more likely in the early stages, with a 100% probability if performed on the first day of exposure to the virus, 38% on the day of symptom onset, and 20% on the third day of symptoms. This fact, together with the need for structured equipment for its handling, interpretation time, and transmission of results to local operational levels, complicated its systematic application in the face of overcrowded emergency units and the need to provide rapid patient care. Therefore, they began to rely on available imaging tests to confirm the diagnosis, classify the severity of the disease, evaluate complications and rule out other simulating entities, avoiding delays in the initiation of treatment.

Although definitive diagnosis of SARS-CoV-2 infection is based on any of the three types of laboratory tests (polymerase chain reaction, antigen detection tests, or antibody detection tests), the World Health Organization (WHO) recommended performing thoracic imaging studies in the following scenarios: symptomatic patients when RT-PCR is not available, results are delayed or there is high clinical suspicion with negative laboratory test. In addition, it suggests performing imaging studies complementary to the clinical and laboratory approach in suspected patients, confirmed non-hospitalized or with mild symptoms to decide on hospital admission, or with moderate or severe symptoms to help determine the type of care and therapeutic management.

Several imaging techniques are available for the management of patients with clinical suspicion of SARS-CoV-2 infection, such as radiography, computed tomography (CT), and lung ultrasound. These methods are complementary to each other and offer options for the evaluation of the different organs and systems according to the evolutionary stage of the disease; imaging tests are also important for the timely identification of acute complications of the disease and chronic sequelae in the pulmonary parenchyma.

The objective of the research was to determine the sensitivity, specificity, positive and negative predictive value of diagnostic images and their concordance concerning the RT-PCR result in patients who consulted with suspected SARS-CoV-2 infection at the General Hospital of the Salvadoran Social Security Institute (HG-ISSS) from March to September 2020.

**Methodology**

An observational study was conducted at the General Hospital of the Salvadoran Social Security Institute (HG-ISSS) from March to September 2020. The study evaluated clinical records, diagnostic imaging studies (X-rays and CT scans), and RT-PCR results for SARS-CoV-2 in patients with suspected SARS-CoV-2.

The study universe included patients who consulted from March to September 2020 at the HG-ISSS with respiratory symptoms suspicious of SARS-CoV-2 disease, regardless of the time of evolution and severity, totaling 1309 patients identified. This study included 214 patients with suspected SARS-CoV-2 who underwent X-rays, CT scans, and RT-PCR tests during the research period.

The sample size of 214 patients was calculated using the finite population equation, with a sampling error of 0.05 and a 95% confidence interval. The calculated prevalence of SARS-CoV-2 disease is 1.14%, based on data from the Salvadoran population and the report of the Office of the Resident Coordinator and the Office for the Coordination of Humanitarian Affairs (OCHA) of May 11, 2020. The theoretical sensitivity and specificity of RT-PCR were used to obtain a sample size of 138 patients.

The sample was selected by simple random sampling, including patients who met the inclusion criteria: clinical suspicion of SARS-CoV-2 disease, imaging studies, and RT-PCR test, and excluding those with incomplete clinical record data. A data collection instrument was used, which consisted of a self-completion sheet divided into three sections. The first section included patient demographic variables such as sex, age, medical history, and symptoms described in the clinical history of the emergency unit. The second part focused on the radiological and tomographic reports in 64 slices, in high-resolution single phase with reconstruction in pulmonary and soft tissue window. These reports were classified according to the Coronavirus Classification...
System (CO-RADS), which classifies the findings as typical, atypical, indeterminate, or negative. They were interpreted by institutional radiologists with an average experience in diagnostic imaging of five years.

In the case of Chest X-rays, these were performed on stationary digital equipment and interpreted by ten radiology and imaging resident physicians in their third (five residents) and fourth (five residents) year of training, who have at least two years of previous medical specialty (surgery, internal medicine or gynecology). The same CO-RADS classification categories adapted to chest radiography were applied. The segmentation of the interpreters according to the imaging study modality into radiologist and radiology resident physicians is stipulated by the administrative organization of the Radiology and Imaging Department HG-ISSS, so it was adopted in the methodological design of this research.

The third section addressed the results of RT-PCR tests obtained from the internal databases of the HG-ISSS epidemiology department, considering only those reported with positive or negative results.

Double-entry tables were constructed with the data collected to determine sensitivity, specificity, and positive and negative predictive values, using Microsoft Excel 365® and the Epi Info™ automatic calculators for analysis. Statistical analysis was performed with the STATA 14 program, calculating Cohen’s Kappa index (k) to evaluate the concordance of the instruments of the categorical measures, considering the value of one as perfect concordance, 0 to 0.99 as weak concordance, 0 to -0.99 as weak discordance, -1 as total discordance and > 0.75 as acceptable concordance. Confidence intervals (CI) were taken into consideration. In addition, the Bayes factor (BF) was calculated to measure the strength of association based on Jeffrey’s value classification scheme: weak (1.1 - 3), moderate (3.1 - 10), strong (10 - 30), very strong (30 - 100) and extreme (>100).

The study was conducted by the principles of the Declaration of Helsinki, with prior authorization from the ethics committee of the Salvadoran Social Security Institute for the review of clinical records and the generation of a protected database, guaranteeing the confidentiality and exclusive scientific use of the participants’ information.

Results

A total of 138 patients with suspected SARS-CoV-2 disease were studied. Of the total, 58 % were men, with a mean age of 58 ± ten years. Most of them were economically active (contributors), resided in urban areas, having completed basic education (9th grade). The most affected age group was 61-65, with 26 patients representing 19 % of the total cases (Figure 1).

Comorbidity was present in 73.9 % of the patients. Arterial hypertension was the most frequent, with 61.6 %, followed by diabetes mellitus with 39.1 % and obesity with 28.3 %. These categories were not mutually exclusive; therefore, a patient could present more than one comorbidity. There was no comorbidity in 26.1 % of the patients. In 93.5 % of the cases, symptoms were reported as the reason for consultation (Table 1), the most common being fever (77.5 %), cough (68.8 %), general malaise (56.5 %), and dyspnea (45.6 %). Only 6.5 % of patients were asymptomatic.

When evaluating diagnostic imaging, 62 % and 63 % of patients were categorized with typical findings on radiographs and CT, respectively (Table 2).

The main radiographic findings were “patchy opacities” (43 %) and alveolar-interstitial infiltrates (22 %), predominantly bilateral (40 %) and peripheral (38 %), mainly affecting the lower third in almost half of the cases. On CT, the findings included "ground-glass opacity" (43 %) and alveolar-interstitial infiltrates (22 %), identical to those reported by X-ray (Table 3). Furthermore, CT lesions were predominantly peripherally distributed (59 %), predominantly affecting the lower lobes (69.6 %) with no predilection for laterality, followed by the middle lobe (58.7 %) and left upper lobe (35.5 %).

Specific signs were identified in the air bronchogram in some cases, including fibrous bands and perivascular thickening; septal thickening (interlobular) and “crazy paving” pattern (ground glass associated with septal thickening) were reported less frequently. 33.3 % of the cases did not report specific signs in the tomographic report.

Extrapulmonary manifestations included pleural effusion in 10 % and enlarged lymph nodes in 1 % of cases (Table 3).

Of the total number of patients evaluated, 82 % had a positive RT-PCR test result for SARS-CoV-2, used as the gold standard for evaluating imaging methods.

Cohen’s Kappa index was calculated to evaluate the diagnostic agreement between the imaging methods: 0.32 (95 % CI 0.17 - 0.47) for X-rays, indicating a low diagnostic agreement (0.2 - 0.4), and 0.58 (95 % CI 0.42 - 0.73) for CT, representing a moderate diagnostic agreement (0.4 - 0.6). In addition, values were obtained to specify the degree of probative strength, obtaining...
for X-rays the BF of 1.8 corresponding to a weak probative strength (1.1 - 3) and CT the BF of 3.6 concerning a moderate probative strength (3.1 - 10) (Table 4).

**Discussion**

The sensitivity and specificity of different imaging methods for detecting SARS-CoV-2 infection show mixed evidence. For X-rays, reports indicate relatively lower sensitivity (15.5 - 69 %) but higher specificity (89 - 97 %). The positive predictive value (PPV) is 83.8 %, and the negative predictive value (NPV) is 50.1 %. These data appear to conflict with those reported in the study. However, it is important to note that the international data were obtained from studies conducted at the outset of the pandemic, when the clinical

---

**Table 1.** Characterization of patients with suspected SARS-CoV-2

<table>
<thead>
<tr>
<th>Variables</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>138</td>
</tr>
<tr>
<td>Average age in years</td>
<td>58 ± 10</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>58%</td>
</tr>
<tr>
<td>Female</td>
<td>42%</td>
</tr>
<tr>
<td>Initial evaluation</td>
<td></td>
</tr>
<tr>
<td>Symptomatic</td>
<td>93.5%</td>
</tr>
<tr>
<td>Asymptomatic</td>
<td>6.5%</td>
</tr>
<tr>
<td>Clinical history</td>
<td></td>
</tr>
<tr>
<td>Comorbidities</td>
<td>73.9%</td>
</tr>
<tr>
<td>No comorbidities</td>
<td>26.1%</td>
</tr>
</tbody>
</table>

**Table 2.** Diagnostic classification of thoracic imaging for SARS-CoV-2

<table>
<thead>
<tr>
<th>Classification (Interpretation)</th>
<th>X-ray (Resident physician)</th>
<th>CT (Radiologist)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical findings</td>
<td>62 %</td>
<td>63 %</td>
</tr>
<tr>
<td>Atypical findings</td>
<td>12 %</td>
<td>15 %</td>
</tr>
<tr>
<td>Indeterminate findings</td>
<td>7 %</td>
<td>5 %</td>
</tr>
<tr>
<td>Negative findings</td>
<td>19 %</td>
<td>17 %</td>
</tr>
</tbody>
</table>
and imaging behavior of SARS-CoV-2 infection was not yet fully understood. This could explain the high sensitivity and positive predictive value currently observed. The low specificity observed in the study, in comparison to international studies (89–97 %)\textsuperscript{,vi–viii} may be attributed to technical factors that influence the image, as well as the level of experience of the medical personnel who corroborated the interpretation. In particular, the interpretations were mainly performed by personnel in their last year of training in diagnostic imaging.

Based on this, X-rays, being cheaper and more accessible, could be most useful in cases of high clinical suspicion to confirm positivity. Some studies suggest its usefulness as a diagnostic alternative in emergency departments, especially when access to RT-PCR tests is limited or in situations where false negatives are suspected.\textsuperscript{,ix}

It is important to note that the absence of X-ray findings does not rule out the presence of the disease, especially in the early stages. Some studies report that up to 38.8 % of symptomatic patients with SARS-CoV-2 have a normal chest X-ray at the time of diagnosis, which suggests a low diagnostic sensitivity in the early stages of the disease, so its use as a screening method is not recommended.\textsuperscript{,ix–xvi}

The most frequently reported findings on chest radiographs are similar to those of other viral pneumonias: focal and ground-glass opacities, which can range from very subtle to affecting both lungs, with a characteristic peripheral distribution.\textsuperscript{,vii–xv} In initial case series studies in Wuhan, China, where the first cases were documented, they reported that 75 % of patients showed bilateral and 25 % unilateral findings.\textsuperscript{xv} Other case series studies, such as that of Wang et al. reported that out of a total of 138 patients, 100 % showed bilateral findings.\textsuperscript{xvi} In the present study, a lower proportion of bilateral and unilateral findings was found, which could be attributed to temporal and geographic differences in the evolution of the virus and its clinical and imaging manifestations.\textsuperscript{xvii}

Chest CT is reported to be the most sensitive and specific method for detecting SARS-CoV-2 infection compared to chest X-ray.\textsuperscript{xvii} However, CT may be normal on admission, even with confirmed infection by RT-PCR.\textsuperscript{xviii} Some studies report that 2–56 % of patients may have a normal CT scan between day 0–3 ± 3 from symptom onset, despite a positive RT-PCR.\textsuperscript{xviii}

The most frequently reported findings on chest radiographs are similar to those of other viral pneumonias: focal and ground-glass opacities, which can range from very subtle to affecting both lungs, with a characteristic peripheral distribution.\textsuperscript{,vii–xv} In initial case series studies in Wuhan, China, where the first cases were documented, they reported that 75 % of patients showed bilateral and 25 % unilateral findings.\textsuperscript{xv} Other case series studies, such as that of Wang et al. reported that out of a total of 138 patients, 100 % showed bilateral findings.\textsuperscript{xvi} In the present study, a lower proportion of bilateral and unilateral findings was found, which could be attributed to temporal and geographic differences in the evolution of the virus and its clinical and imaging manifestations.\textsuperscript{xvii}

Chest CT is reported to be the most sensitive and specific method for detecting SARS-CoV-2 infection compared to chest X-ray.\textsuperscript{xvii} However, CT may be normal on admission, even with confirmed infection by RT-PCR.\textsuperscript{xviii} Some studies report that 2–56 % of patients may have a normal CT scan between day 0–3 ± 3 from symptom onset, despite a positive RT-PCR.\textsuperscript{xviii}

Table 3. Diagnostic Thoracic Imaging Findings

<table>
<thead>
<tr>
<th>Findings</th>
<th>X-Ray</th>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types of lesions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patchy opacities</td>
<td>43%</td>
<td>49%</td>
</tr>
<tr>
<td>Consolidation</td>
<td>16%</td>
<td>20%</td>
</tr>
<tr>
<td>Alveolar-interstitial infiltrates</td>
<td>22%</td>
<td>24%</td>
</tr>
<tr>
<td>Others</td>
<td>19%</td>
<td>7%</td>
</tr>
<tr>
<td>Lesion distribution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unilateral</td>
<td>2%</td>
<td>0%</td>
</tr>
<tr>
<td>Bilateral</td>
<td>40%</td>
<td>14%</td>
</tr>
<tr>
<td>Peripheral</td>
<td>38%</td>
<td>59%</td>
</tr>
<tr>
<td>Central</td>
<td>10%</td>
<td>12%</td>
</tr>
<tr>
<td>None/other</td>
<td>10%</td>
<td>15%</td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper third</td>
<td>14%</td>
<td>26.1%</td>
</tr>
<tr>
<td>Middle third</td>
<td>28%</td>
<td>58.7%</td>
</tr>
<tr>
<td>Lower third</td>
<td>48%</td>
<td>69.6%</td>
</tr>
<tr>
<td>No lesions</td>
<td>10%</td>
<td>35.5%</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>69.6%</td>
</tr>
</tbody>
</table>

Table 4. Sensitivity, specificity, positive predictive value, negative predictive value of thoracic imaging for SARS-CoV-2

<table>
<thead>
<tr>
<th>Diagnostic test</th>
<th>(S)</th>
<th>(S)</th>
<th>(PPV)</th>
<th>(NPV)</th>
<th>(K)</th>
<th>(BF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest X-ray</td>
<td>86%</td>
<td>52.9%</td>
<td>92.9%</td>
<td>34.6%</td>
<td>0.32</td>
<td>1.8</td>
</tr>
<tr>
<td>Chest CT</td>
<td>92.9%</td>
<td>64%</td>
<td>92.1%</td>
<td>66.7%</td>
<td>0.58</td>
<td>3.6</td>
</tr>
</tbody>
</table>

Findings: (S) Sensitivity. (S) Specificity. (PPV) Positive predictive value. (NPV) Negative predictive value. (K) Kohen Kappa index. (BF) Bayes factor.
In the study, CT showed moderate diagnostic concordance, with high sensitivity and PPV but moderate specificity and NPV; these values are similar to those reported internationally. For example, Sharma et al. reported a sensitivity of up to 98 %,44 Kim et al. 94 % (95 % CI) for chest CT, and Herpe et al. reported a specificity of 88 %, sensitivity of 80 %, PPV 89 % and NPV 79 %.45

CT shows a higher yield in cases of moderate clinical suspicion, diagnostic discrepancy, or for the detection of complications. Common CT findings include peripherally distributed, bilateral, multifocal ground-glass areas associated with subsegmental patchy consolidations, mainly affecting lower lobes and posterior segments.46

A case series in China reported that "ground glass" was the most common finding at 45-67 %, similar to the data obtained in the investigation.47 Consolidative lesions were approximately 30-60 %, a percentage almost three times higher than that found in the investigation, which could be attributable to differences in the disease at the time of evaluation, circulating strains, and lineages of the virus.

The investigation also documented other specific findings, such as the thickening of interlobular septa and the crazy-paving pattern, findings similar to those reported in Mexican case series.48-50 Because CT can be normal in the initial periods of the disease, it is not recommended as a screening or early diagnostic tool.51-53 Although the sensitivity of RT-PCR varies considerably among studies, e.g., Smith et al. reported a sensitivity of 30 % to 91 %,54 when evaluating imaging studies concerning RT-PCR results for SARS-CoV-2, it was found that “positive” categories correlated well with RT-PCR results but categories considered “negative” showed low negative predictive values due to their limited ability to rule out disease in the absence of imaging findings, independent of the study modality analyzed.

Regarding the clinical characterization of the patients, most cases were symptomatic, with fever being the main symptom of SARS-CoV-2 infection, data consistent with international studies that report a high percentage of fever (98.6 %), fatigue (69.6 %) and dry cough (59.4 %) as the main symptoms in the first cases reported in China.55 While, other case series mention that fever (83 %), cough (82 %), and respiratory distress (31 %) were the main symptoms.56

A limitation of this study was the lack of consistency in the imaging reports, leading to variability in the terminology used by different radiologists. Additionally, two different groups of professionals interpreted the images based on the imaging method used: radiologist physicians interpreted computed tomography (CT) scans, while radiology resident physicians interpreted chest radiographs (X-rays). This difference in interpretation could have introduced biases related to the experience of the group interpreting the radiographs, potentially affecting the accuracy of the evaluations. Furthermore, since each imaging method has unique characteristics that influence its sensitivity and specificity in detecting disease, the results obtained are not directly comparable with each other, regardless of the interpreter.

It is recommended to standardize the information in reports by adopting institutional templates. Additionally, expanding the use of Picture Archiving Communication Systems (PACS) and Integrated Health Services Networks (RISS) for storing and transmitting diagnostic images is advised. This will enhance the coordination of care services, making it easier for medical specialties to collaborate in order to implement timely treatments based on imaging results.

**Conclusion**

Computed tomography showed high sensitivity and positive predictive value, with moderate specificity, negative predictive value, and diagnostic concordance, making it a viable alternative in cases of moderate-high clinical suspicion, diagnostic discrepancy, or to rule out complications. X-rays showed high sensitivity and positive predictive value, but low specificity, negative predictive value, and low diagnostic concordance. They will be useful in cases of high clinical suspicion but need to be complemented with RT-PCR for a definitive diagnosis.

**Funding**

No external funds were available.

**References**


Diagnostic concordance of thoracic images for Covid-19 disease


