CHEST RADIOLOGY

Immune checkpoint inhibitor-related pneumonitis and COVID-19: a case-matched comparison of CT findings

Riccardo Picasso[1] · Andrea Cozzi[2] · Virginia Picasso[3] · Federico Zaottini[1] · Federico Pistoia[1] · Sara Perissi[4] · Carlo Martinoli[1,4]

Abstract

Objectives To compare the radiological findings of immune checkpoint inhibitor-related pneumonitis (IRP) and COVID-19 pneumonia, evaluating the potential of the CO-RADS score to differentiate between them.

Methods Two readers blindly reviewed chest CTs from age- and sex-matched groups of 33 patients with IRP and 33 patients with COVID-19 pneumonia. Each examiner evaluated the presence of 13 CT features, semiquantitatively scored lung involvement, and assigned a CO-RADS score. Inter-reader reliability in the assessment of CT features and CO-RADS categories was evaluated with Cohen's κ. Distribution differences between groups were evaluated with the χ^2, Fisher's, and Mann–Whitney U tests.

Results Substantial or higher inter-reader reliability was found in CO-RADS assignments ($\kappa = 0.664$) and in the evaluation of CT features ($\kappa \geq 0.638$), among which the sole feature found to significantly differentiate IRP from COVID-19 pneumonia was unilateral presentation ($p < 0.001$). Lung involvement semiquantitative scores and CO-RADS scores were significantly higher ($p < 0.001$) in COVID patients (median involvement score 4, IQR 4–6; median CO-RADS score 5, IQR 4–5) than in IRP patients (median involvement score 2.5, IQR 2–4; median CO-RADS score 3, IQR 3–4) but exploratory analysis of CO-RADS specificity revealed comparatively low values, ranging between 51.5% (Reader 1) and 54.6% (Reader 2).

Conclusions CT features of IRP and COVID-19 pneumonia frequently overlap, save for the extent of lung involvement and bilaterality. In the current SARS-CoV-2 pandemic, the low specificity of the CO-RADS score for the differential diagnosis of COVID-19 pneumonia and IRP may prompt to reconsider the role of imaging in IRP work-up.

Keywords COVID-19 pneumonia · Computed tomography (CT) · CO-RADS · Checkpoint inhibitor-related pneumonitis (IRP)

✉ Federico Zaottini
federico.zaottini.fz@gmail.com

Riccardo Picasso
riccardo.picasso@gmail.com

Andrea Cozzi
andrea.cozzi@gmail.com

Virginia Picasso
virgi.picasso@gmail.com

Federico Pistoia
federicopistoia1@gmail.com

Sara Perissi
saraperissi@gmail.com

Carlo Martinoli
carlo.martinoli@unige.it

[1] Unit of Radiology, IRCCS Ospedale Policlinico San Martino, Via Largo Rosanna Benzi, 10, 16132 Genoa, Italy

[2] Imaging Institute of Southern Switzerland (IIMSI), Ente Ospedaliero Cantonale (EOC), Via Tesserete 46, 6900 Lugano, Switzerland

[3] Unit of Radiology, Ospedale Lavagna, Via Don Giovanni Battista Bobbio 25, 16033 Lavagna, Italy

[4] Dipartimento Di Scienze Della Salute (DISSAL), Università Degli Studi Di Genova, Via Alberti L.B 2, 16132 Genoa, Italy

⚪ Springer

Introduction

Since the approval of ipilimumab for the treatment of metastatic melanoma in 2011 [1], immune checkpoint inhibitors (ICI) revolutionized cancer therapy and came to represent the standard of care for the treatment of several solid and haematological malignancies [2]. Nevertheless, the hyperstimulation of the immune system engendered by ICI may determine adverse events—driven by an anomalous response against normal tissues—such as ICI-related pneumonitis (IRP), mainly related to drugs targeting the programmed cell death protein 1 (PD-1) [3]. While IRP represents a rather frequent, clinically relevant, and potentially lethal condition, there is no available dedicated test to obtain a confirmed diagnosis of IRP, which is therefore suspected in patients receiving immunotherapy who develop pneumonia-like symptoms and exhibit computed tomography (CT) findings suggestive for this condition, after having excluded other causes of pneumonia [4–6]. In addition, although several CT patterns of lung damage associated with IRP have been described, the most common findings of IRP are rather non-specific and may include ground-glass opacities, septal thickening, and traction bronchiectasis, which broadly overlap with typical signs of viral pneumonia and interstitial lung disease [7].

During the COVID-19 pandemic, several authors reported the close clinical and imaging similarity of SARS-CoV-2 pneumonia and IRC, pointing out the rising challenge in distinguishing between the two conditions in patients receiving ICI [8–10]. Although preliminary data suggested a potential role of artificial intelligence and perfusion CT in this task [11], the current epidemiological transition of COVID-19 towards an endemic status, the frequency of undetermined ground-glass opacities detected during routine follow-up CT studies of oncological patients without pneumonia symptoms, and the potential use of CT to disclose signs of COVID-19 pneumonia in patients with negative reverse transcription polymerase chain reaction (RT-PCR) tests in high prevalence scenarios [12] are all likely to sizably increase the burden of doubtful cases. Indeed, the distinction between IRP and COVID-19 pneumonia is critical, not only for implications related to disease contention and patients' isolation in SARS-CoV-2 infections, but also because a diagnosis of IRP requires a prompt suspension of immunotherapy and, in selected cases, the administration of corticosteroids.

CO-RADS is a standardized assessment scheme which provides a level of suspicion for pulmonary involvement of COVID-19 based on the features seen at unenhanced chest CT. The system has been developed in a moderate to high prevalence setting with the aim of facilitating the recognition of COVID-19 infection in patients with clinically evident pneumonia and unavailable positive RT-PCR test [13]. Therefore, this study aimed to retrospectively compare radiological findings of IRP and COVID-19 pneumonia in an age- and sex-matched cohort with closely comparable clinical presentation, also evaluating the potential of the CO-RADS score to discriminate between these two conditions.

Methods

Approval for this monocentric study, performed at IRCCS Ospedale Policlinico San Martino (Genoa, Italy), was obtained from the competent Ethics Committee (Comitato Etico Regione Liguria, protocol code 12,306, approved on 02/05/2022). Informed consent to participate to the study was waived due to the retrospective nature of the study; however, all the patients signed the informed consent to undergo the diagnostic examinations and have their data used for research purpose unless emergency situations have occurred.

Study design and population

We screened our institutional imaging database and electronic medical charts of the Radiology Unit of IRCCS Ospedale Policlinico San Martino to identify cancer patients who developed symptoms suspicious for pneumonia during ICI therapy, were referred for CT and this demonstrated findings compatible with IRP, and were ultimately diagnosed with this condition after clinical exclusion of alternative etiologies in the five-year period from March 1, 2016, to February 28, 2021. Due to the retrospective nature of the study, the clinical work-up of patients diagnosed with IRP comprised various combinations of blood tests, bacterial cultures, and bronchioalveolar lavage. Patients with examinations performed after the start of the COVID-19 pandemic in Italy (February 21, 2020) were included in this study only if results of reverse transcriptase–polymerase chain reaction test negative for SARS-CoV-2 infection—performed within ± 3 days of their CT examination—were retrievable in institutional electronic medical charts. After recording their demographic and clinical characteristics (age, sex, cancer type, ICI target), enrolled patients were matched for sex and age (tolerance: ± 6 months) with patients who had a diagnosis of COVID-19 pneumonia confirmed by a positive reverse transcriptase–polymerase chain reaction test and had performed chest CT—both within 48 h from hospital admission—in a period (March 17, 2020, to November 27, 2020) encompassing the first and second wave of the COVID-19 pandemic in Italy before the start of vaccination campaigns.

Image acquisition and analysis

All chest CT examinations were conducted on dual-source 128×2 slices CT scanner (Somatom Definition flash, Siemens, Germany). Acquisition parameters were as follows: slice thickness 2.0 mm, 120 kVp, mAs according to patient body size, spiral pitch factor 0.98, and collimation width 0.625. The CT acquisition protocol was adapted to the clinical question (e.g. with contrast medium if the examination was aimed to confirm or rule-out pulmonary embolism or in patients during oncological follow-up; without contrast medium if lung infiltrates or pneumonia were suspected); multiplanar and high-resolution reconstructions were made, as per institutional protocol.

Chest CT examinations of all included patients were reviewed independently and in a random order by two board-certified radiologists (F.Z., Reader 1, and R.P., Reader 2) with 5 and 4 years of clinical experience in thoracic imaging and with 1 year of experience each in the application of the CO-RADS classification [13]. Both readers had access to standard chest CT interpretation settings and tools provided by our institutional PACS system but were blinded to exam identifiers, patients' names, and medical history.

First, both readers performed a semiquantitative assessment of the extent of lung involvement according to a five-category scheme [14] (0%, 0; 1–25%, 1; 26–50%, 2; 51–75%, 3; over 75%, 4) applied on each lung, with a maximum involvement score of 8. Then, 13 CT features were evaluated according to the Fleischner Society Glossary of Terms for Thoracic Imaging [15]: craniocaudal (lower, upper, or mixed) and axial (peripheral, central, or mixed) distributions of lung findings, their laterality (unilateral or bilateral), the presence of ground-glass opacities and their appearance (round/circumscribed, patchy, or diffuse), interlobular septal thickening (present or absent), crazy paving (present or absent), consolidations (present or absent), air bronchogram (present or absent), tree-in-bud nodules (present or absent), traction bronchiectasis (present or absent), linear band-like/strip-like opacities (present or absent), mediastinal lymphadenopathy (present or absent [16]), and pleural effusion. Finally, the readers assigned a CO-RADS category to each case.

Statistical analysis

The Shapiro–Wilk test was used to perform distribution analysis. Consequentially, normal distributions were reported using median ± standard deviation and non-normal distributions were reported as median with their interquartile range (IQR). Inter-reader reliability in the assessment of lung involvement extent, of CT features, and of the CO-RADS categories was evaluated with Cohen's κ, reported with its 95% confidence interval (CI) and

interpreted according to the Landis and Koch classification [17]. Analyses of distribution differences were performed on the whole number of observations with the application of patient clustering: distribution differences of ordinal items (lung involvement and CO-RADS scores) between the IRP group and the COVID-19 pneumonia group were evaluated with the Wilcoxon–Mann–Whitney U test, whereas the χ^2 and the Fisher's tests were used to evaluate distribution differences of all remaining nominal items.

The diagnostic performance of the CO-RADS classification in distinguishing IRP and COVID-19 pneumonia was explored separately for the two readers by dichotomizing CO-RADS categories either as "positive for COVID-19" (CO-RADS scores 4 and 5) or "negative for COVID-19" (CO-RADS scores 1–3). Considering the case–control nature of this study and its potential implications on diagnostic performance indexes [18], evaluations of the two readers were reported in a descriptive fashion, while only a preliminary evaluation of specificity was performed.

All analyses were performed with SPSS v.26.0 (IBM Corp., Armonk, N.Y., USA), the p value significance threshold being lowered to $p < 0.003$ after applying the Bonferroni–Holm correction to account for multiple statistical testing.

Results

Study population

A total of 33 patients with IRP (24 males, 73%) with an average age of 68.1 ± 11.9 years were retrieved, according to the database search described above, in the period between January 11, 2017, and February 02, 2021. A diagnosis of melanoma had been made in 11/33 patients (33%), while the remaining 22/33 (67%) had been diagnosed with non-small-cell lung cancer. Anti-PD-1 monotherapy was administered to 31/33 patients (94%), one patient with non-small-cell lung cancer (3%) received anti-PD-L1 monotherapy, and a combined anti-PD-1/anti-CTLA-4 therapy was administered to another patient (3%) with non-small-cell lung cancer. These 33 patients with IRP were matched for sex and age to 33 patients with an RT-PCR confirmed diagnosis of COVID-19 pneumonia (average age 68.2 ± 11.4 years). As detailed in Table 1, no statistically significant difference was found between the two groups in terms of length of symptoms presence before chest CT, peripheral oxygen saturation before supplemental oxygen administration, need of supplemental oxygen administration, and composite adverse patient outcome (i.e. need of intensive care unit admission or death at 30 days from hospitalization).

Table 1 Patient characteristics according to pneumonia type

	IRP	COVID-19	*p* value
Average age	68.1 ± 11.9 years	68.2 ± 11.4 years	0.974
Median length of symptoms before CT	6 days (IQR 0–11.5)	10 days (IQR 3–14)	0.035
Median SpO$_2$ before supplemental O$_2$ administration	91% (IQR 81.5–96%)	94% (IQR 92–96%)	0.094
Need of supplemental O$_2$ administration*			
Yes	10 (45.5%)	23 (69.7%)	0.095
No	12 (54.5%)	10 (30.3%)	
Adverse outcome[‡,§]			
Yes	9 (32.1%)	15 (45.5%)	0.121
No	19 (67.9%)	18 (54.5%)	

IQR interquartile range, *IRP* immune checkpoint inhibitor-related pneumonitis

*Only retrievable for 22 patients in the IRP group

‡Need of intensive care unit admission or death at 30 days from hospitalization

§Only retrievable for 28 patients in the IRP group

Table 2 Assessment of inter-reader reliability in the evaluation of the 13 chest CT features

Chest CT feature	Cohen's κ (95% confidence interval)
Predominant craniocaudal distribution	0.833 (0.716–0.949)
Predominant axial distribution	0.638 (0.467–0.808)
Laterality	1.000 (0.759–1.000)
Ground-glass opacities presence and appearance	0.706 (0.564–0.847)
Interlobular septal thickening	0.709 (0.533–0.884)
Crazy paving	0.649 (0.462–0.837)
Consolidations	0.714 (0.547–0.880)
Air bronchogram	0.655 (0.472–0.839)
Tree-in-bud nodules	0.743 (0.503–0.982)
Traction bronchiectasis	0.725 (0.554–0.897)
Linear band-like opacities	0.937 (0.852–1.000)
Mediastinal lymphadenopathy	0.698 (0.507–0.888)
Pleural effusion	0.956 (0.870–1.000)

Chest CT features

The analysis of inter-reader reliability (Table 2) for the evaluation of the 13 chest features showed a substantial agreement for 9 features, ranging from $\kappa = 0.638$ (95% CI 0.467–0.808) for the assessment of axial distribution to $\kappa = 0.743$ (95% CI 0.503–0.982) for the evaluation of the presence of tree-in-bud nodules. The remaining 4 features had an almost perfect agreement ranging from $\kappa = 0.833$ (95% CI 0.716–0.949) for evaluation of craniocaudal distribution of pulmonary findings to $\kappa = 1.000$ (95% CI 0.759–1.000) for the assessment of the unilateral or bilateral distribution of findings. The semiquantitative evaluation of lung involvement extent (with a visual score ranging from 1 to 8) also showed a high inter-reader agreement (75.8%, 95% CI 64.2–84.5%) and substantial or higher inter-reader reliability, with a non-weighted $\kappa = 0.683$ (95% CI 0.558–0.809) and a linear-weighted $\kappa = 0.805$ (95% CI 0.705–0.905).

As detailed in Table 3, among the 13 chest CT features, a statistically significant association with COVID-19 pneumonia or IRP was observed only for the laterality of findings, with a unilateral presentation being observed in 21.2% of IRP cases and in no COVID-19 cases ($p < 0.001$). As shown in Fig. 1, IRP patients also had a significantly lower total extent of lung involvement (median 2.5, IQR 2–4) compared to patients with COVID-19 pneumonia (median 4, IQR 4–6, Mann–Whitney U 1174, $p < 0.001$).

CO-RADS assessment

As expected, overall CO-RADS scores were significantly higher (Mann–Whitney U 980, $p < 0.001$) in the COVID-19 pneumonia group (median 5, IQR 4–5) than in the IRP group (median 3, IQR 3–4), as depicted in Fig. 2. Table 4 details category-specific CO-RADS scores assigned by the two readers, highlighting a 77.3% agreement (51/66 cases, 95% CI 65.8–85.7%) with an overall substantial inter-reader reliability ($\kappa = 0.664$, 95% CI 0.512–0.814) that had its lowest category-wise value in the CO-RADS 3 category ($\kappa = 0.536$, 95% CI 0.295–0.778). Consequently, inter-reader agreement in the dichotomized CO-RADS classification improved marginally to 84.8% (95% CI 74.3–91.6%, 54 of 66 cases) and inter-reader reliability slightly lowered to $\kappa = 0.651$ (95% CI 0.453–0.848).

The distribution of CO-RADS scores in the two groups is listed in Table 5: Reader 1 correctly identified 17/33 IRP patients (52%) with a CO-RADS 1–3 score, while the remaining 16/33 patients were incorrectly assigned a CO-RADS 4 (9/33, 27%) or a CO-RADS 5 score (7/33, 21%), for a resulting 51.5% specificity (95% CI 33.5–69.2%). Reader

Table 3 Distribution of CT features according to pneumonia type

CT feature	IRP	(%)	COVID-19		χ^2	p value
Predominant craniocaudal distribution						
None	32	48.5%	31	47.0%	5.309	0.07
Lower	15	22.7%	25	37.9%		
Upper	19	28.8%	10	15.1%		
Predominant axial distribution						
Central	1	1.5%	0	0.0%	1.09	0.58
Peripheral	25	37.9%	27	40.9%		
Mixed	40	60.6%	39	59.1%		
Laterality						
Monolateral	14	21.2%	0	0.0%	15.661	<0.001
Bilateral	52	78.8%	66	100.0%		
Ground-glass opacities presence and appearance						
None	2	3.0%	0	0.0%	3.196	0.362
Patchy	29	43.9%	25	37.9%		
Diffuse	18	27.3%	18	27.3%		
Mixed	17	25.8%	23	34.8%		
Interlobular septal thickening						
Absent	31	47.0%	18	27.3%	5.485	0.03
Present	35	53.0%	48	72.7%		
Crazy paving						
Absent	46	69.7%	35	53.0%	3.866	0.073
Present	20	30.3%	31	47.0%		
Consolidations						
Absent	24	36.4%	25	37.9%	0.032	1
Present	42	63.6%	41	62.1%		
Air bronchogram						
Absent	39	59.1%	40	60.6%	0.032	1
Present	27	40.9%	26	39.4%		
Tree-in-bud nodules						
Absent	53	80.3%	61	92.4%	4.117	0.074
Present	13	19.7%	5	7.6%		
Traction bronchiectasis						
Absent	49	74.2%	41	62.1%	2.235	0.135
Present	17	25.8%	25	37.9%		
Linear band-like opacities						
Absent	46	69.7%	32	48.5%	6.142	0.021
Present	20	30.3%	34	51.5%		
Mediastinal lymphadenopathy						
Absent	43	65.2%	53	80.3%	3.819	0.078
Present	23	34.8%	13	19.7%		
Pleural effusion						
Absent	55	83.3%	48	78.0%	2.165	0.141
Present	11	16.7%	18	22.0%		

IRP immune checkpoint inhibitor-related pneumonitis

2 showed a closely comparable 54.6% specificity (95% CI 36.4–71.9%), having correctly assigned a CO-RADS 1–3 score to 18/33 IRP patients (55%) and having incorrectly assigned a CO-RADS 4 score to 7/33 patients (21%) and a CO-RADS 5 score to 8/33 patients (24%) (Fig. 3).

Discussion

The introduction of ICI in cancer care during the last decade is widely recognized as a major milestone in cancer research and treatment. Currently, more than two-third of

Fig. 1 Distribution of the extent of lung involvement across the two pneumonia groups. *IRP* Immune checkpoint inhibitor-related pneumonitis
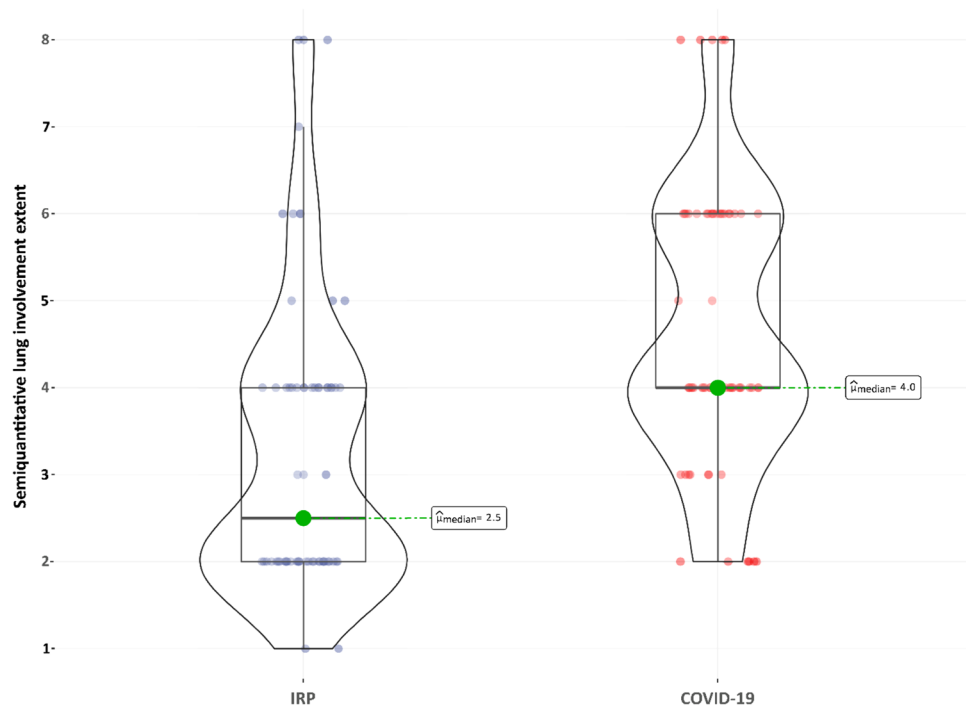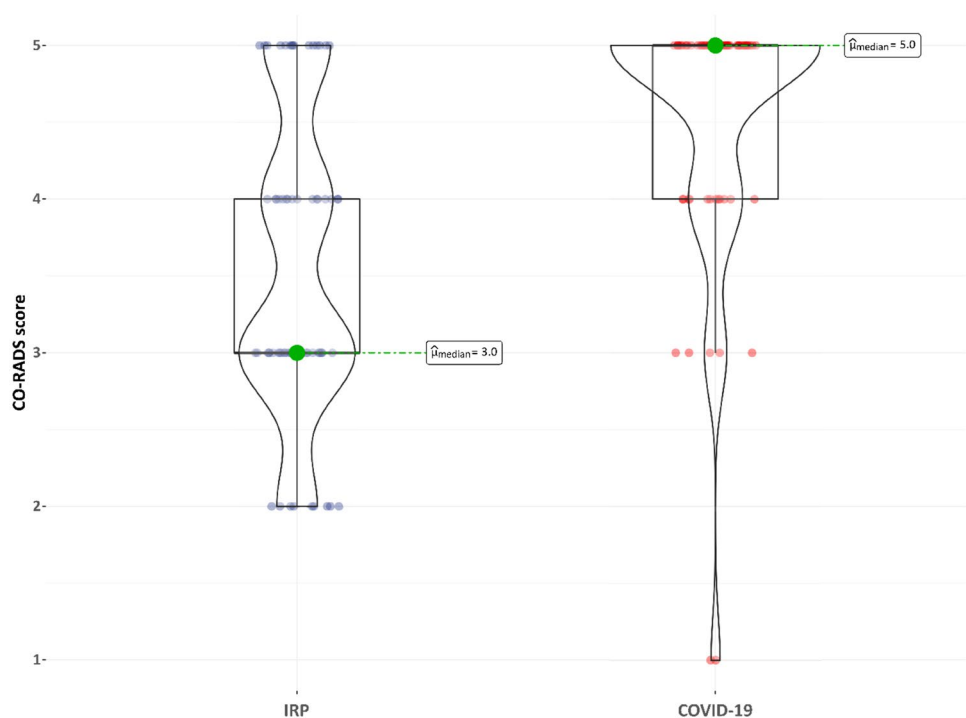

Fig. 2 Distribution of CO-RADS scores across the two pneumonia groups. *IRP* Immune checkpoint inhibitor-related pneumonitis

drug-related trials in oncology regard ICI and further growth of clinical indications to single or combined ICI therapy is easily foreseeable in the next future [19]. However, IRP represents a frequent adverse effect of ICI therapy: a recent meta-analysis [20] found an overall 2.7% incidence of IRP in patients treated with anti-PD-1 molecules, rising to 10% when ICI are part of combination therapies [21]. Although

its pathogenesis remains largely unknown, IRP is widely hypothesized to be a multi-layered autoimmune process, including abnormal T-cells reaction against self-peptides, production of autoantibodies, overexpression of inflammatory cytokines, and development of complement-mediated inflammation [22]. Bronchoalveolar lavage specimens from IRP patients demonstrated a decrease of T regulatory cells,

Table 4 Cross-tabulation of CO-RADS score assignments between readers

| | Reader 2 | | | | | | Agreement (95% confidence interval) |
	CO-RADS 1	CO-RADS 2	CO-RADS 3	CO-RADS 4	CO-RADS 5	Total	
Reader 1							
CO-RADS 1	1	0	0	0	0	1	100% (20.7–100.0%)
CO-RADS 2	0	4	0	0	0	4	100% (51.1–100.0%)
CO-RADS 3	0	1	10	0	5	16	62.5% (38.6–81.5%)
CO-RADS 4	0	0	4	10	2	16	62.5% (38.6–81.5%)
CO-RADS 5	0	0	1	2	26	29	89.7% (73.6–96.4%)
Total	1	5	15	12	33	66	77.3% (65.8–85.7%)
κ (95% confidence interval)	1.000 (0.759–1.000)	0.881 (0.640–1.000)	0.536 (0.295–0.778)	0.637 (0.396–0.879)	0.696 (0.455–0.937)	0.664 (0.512–0.814)	

Table 5 Distribution of CO-RADS scores among pneumonia groups according to each reader assessment

|		IRP	COVID-19
Reader 1			
Negative for COVID-19	CO-RADS 1	0	1
	CO-RADS 2	4	0
	CO-RADS 3	13	3
Positive for COVID-19	CO-RADS 4	9	7
	CO-RADS 5	7	22
Reader 2			
Negative for COVID-19	CO-RADS 1	0	1
	CO-RADS 2	5	0
	CO-RADS 3	13	2
Positive for COVID-19	CO-RADS 4	7	5
	CO-RADS 5	8	25

IRP immune checkpoint inhibitor-related pneumonitis

proliferation of CD8 lymphocytes, and CD4/CD8 ratio inversion [23, 24], but to date alveolar lavage and lung biopsy are not routinely indicated due to the absence of specific pathological findings [25]. The increasing diffusion of ICI-based therapies, the relatively high incidence of IRP in treated patients, and the current COVID-19 pandemic are posing serious challenges in the interpretation of chest CTs of patients who develop pneumonia-like symptoms during ICI treatments. This holds particularly true when considering that IRP patients may be silent carriers of SARS-COV-2 infection, that imaging features of IRP and COVID-19 pneumonia can substantially overlap, and that classification systems devised to aid the interpretation of CT scans in suspected COVID-19 patients were developed in a medium-to-high prevalence scenario. The current transition to an endemic SARS-CoV-2 circulation warrants further investigations about the potential of these scores in the differential diagnosis of COVID-19 and other interstitial lung diseases such as IRP.

In this retrospective study, two radiologists blindly reviewed 66 CT examinations from two groups of age- and sex-matched patients with COVID-19 pneumonia and IRP: of note, the two groups did not significantly differ in major clinical characteristics. For every CT examination, each reader assigned a CO-RADS score and evaluated the presence of a series of predetermined descriptive CT features. Finally, the extension of lung involvement was graded through a semiquantitative scale consisting of five consecutive classes. While we observed a high overall inter-reader agreement (77.3%) and overall substantial inter-reader reliability (Cohen's $\kappa = 0.664$) in CO-RADS assignments, class-specific reliability analysis showed only a moderate reliability for the CO-RADS 3 category (Cohen's $\kappa = 0.536$), reflecting uncertainties in the differential diagnosis of COVID-19 pneumonia in less-than typical cases. Again, the two readers had at least substantial agreement in the evaluation of descriptive CT features, with high agreement in the semiquantitative estimation of lung involvement. Aside from the bilateral presentation found in all COVID-19 patients but only in 80% of IRP patients, no other descriptive CT feature was significantly associated with one of the two groups, highlighting the close overlap of CT appearance of the two conditions. However, our results also enable us to hypothesize that in larger samples the presence of linear band-like opacities and of interlobular septal thickening may

Fig. 3 Six examples of CO-RADS system application in COVID-19 pneumonia and IRP cases. **a** COVID-19 pneumonia classified as CO-RADS 5, true positive; **b** IRP classified as CO-RADS 5, false positive; **c** IRP classified as CO-RADS 4, false positive; **d** IRP classified as CO-RADS 2, true negative; **e** IRP classified as CO-RADS 2, true negative; **f** COVID-19 pneumonia classified as CO-RADS 3, false negative
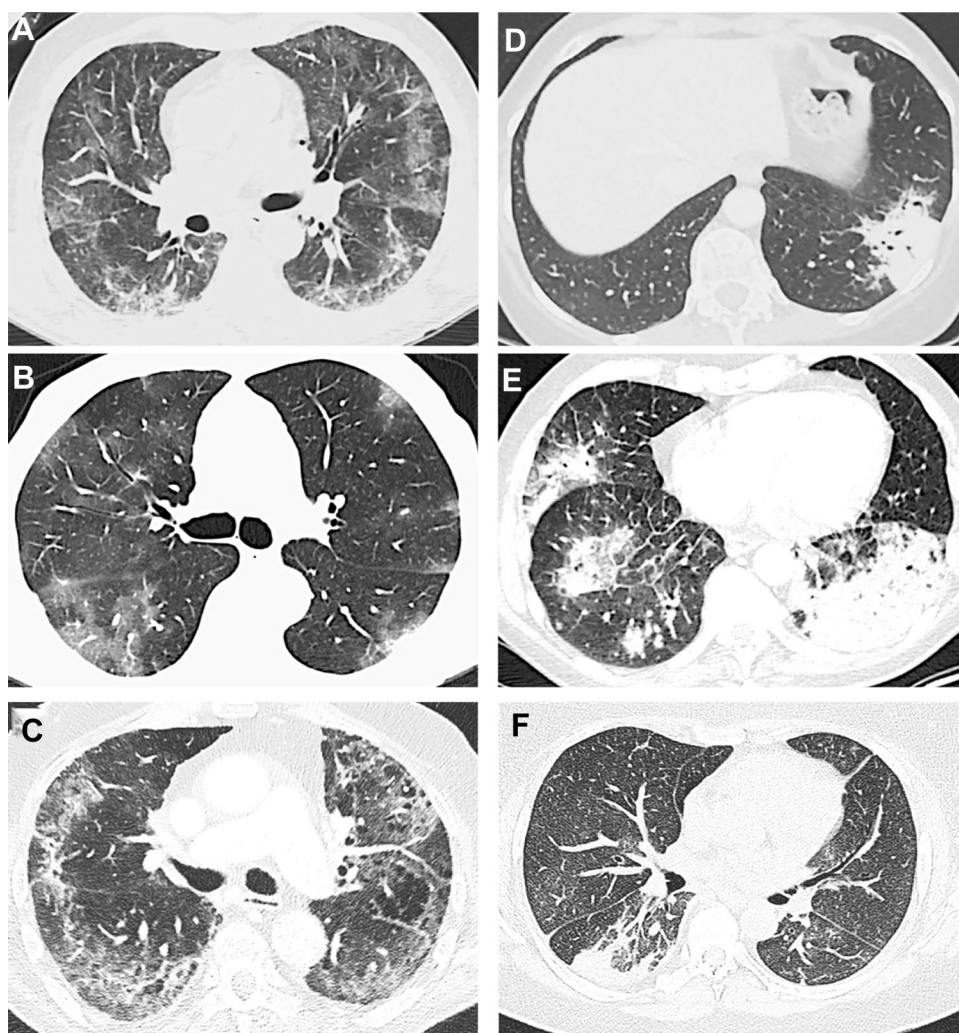
turn out significantly more present in COVID-19 patients, also considering that COVID-19 patients had a significantly higher extent of lung involvement ($p < 0.001$).

Even if—as expected—the overall CO-RADS scores were significantly higher in the COVID-19 pneumonia group, exploratory analysis of specificity revealed medium-to-low estimates: among IRP patients, a CO-RADS score ≥ 4 was incorrectly assigned in at least 45% of patients (48% for Reader 1 and 45% for Reader 2, respectively). In addition, no IRP cases which were assigned a CO-RADS ≥ 4 score by one reader had a CO-RADS < 3 score assigned by the other, underlining the challenging nature of this differential diagnosis. Post hoc case revision revealed that "patchy" ground-glass opacities and septal thickening were the features most commonly driving the reader towards erroneous diagnosis of COVID-19 pneumonia. This appears in line with the fact that in the CO-RADS system, the detection of multifocal bilateral ground-glass opacities is pivotal in shifting from CO-RADS 3 to higher classes: however, a comparatively high number of IRP patients in our study demonstrated that specific feature, suggesting that its relative diagnostic weight should be reconsidered when interpreting CT scans of patients receiving ICI therapy.

A recent meta-analysis [26] showed that from 2 to 58% of SARS-COV-2 patients (mean summary estimate 12%) may have an initial false negative RT-PCR test and that in these patients CT currently plays a pivotal role as a complementary tool to diagnose COVID-19 pneumonia: thus, evidence of a low performance of CO-RADS in the differential diagnosis of IRP and COVID-19, coupled to the absence of significantly associated CT features—save for semiquantitatively scored lung involvement and bilateral presentation—may prompt a reconsideration of the diagnostic pathway of ICI patients during the current pandemic.

Our study presents several limitations, chiefly related to its single-centre nature and relatively small sample size: however, we tried to mitigate the effects of these limitations with a case-matched analysis of IRP patients—who did not have a significantly different clinical presentation compared to COVID-19 patients—and by building

a highly self-consistent group of IRP patients, all with a clinically confirmed IRP diagnosis which is essentially a diagnosis of exclusion and is reached in a minority of patients. In addition, the retrospective nature of the study prevented any investigation on the impact of chest CT findings in modifying the clinical course of IRP patients.

In conclusion, in our case–control analysis, significant differences in chest CT appearance of IRP and COVID-19 pneumonia were represented by bilateral presentation and lung involvement extent, while 12 other CT features variously overlapped between the two groups. The CO-RADS score exhibited a medium-to-low differential diagnostic potential, suggesting that in the current SARS-CoV-2 pandemic context clinical and imaging findings may not be sufficient to appropriately reach a differential diagnosis when evaluating patients receiving.

ICI and developing pneumonia-like symptoms. Extensive multimodal investigation including blood and culture tests, bronchoalveolar lavage analysis, and CT should be warranted in doubtful cases and therapeutic decision making (e.g. suspending ICI and starting steroids) should be based on a careful and personalized cost–benefit analysis that must consider the probability of IRP and individual patient's conditions. Further studies are warranted to investigate if the development of advanced imaging techniques and innovative CT-based biomarkers may assist the differential diagnosis between IRP and COVID-19 pneumonia.

Declarations

Conflict of interests The authors have no relevant financial or non-financial interest to disclose.

Ethical approval This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the competent Ethics Committee of IRCCS San Martino (Comitato Etico Regione Liguria, protocol code 12306, approved on 02/05/2022).

Consent to participate Informed consent was obtained from all individual participants included in the study to undergo the diagnostic examination and to have their data used for research purpose.

Consent to publish The authors affirm that human research participants provided informed consent to have their data used for research purpose.

References

1. Hodi FS, O'Day SJ, McDermott DF et al (2010) Improved survival with Ipilimumab in patients with metastatic melanoma. N Engl J Med 363:711–723. https://doi.org/10.1056/NEJMoa1003466

2. Wilky BA (2019) Immune checkpoint inhibitors: the linchpins of modern immunotherapy. Immunol Rev 290:6–23. https://doi.org/10.1111/imr.12766

3. Nishino M, Sholl LM, Hatabu H, Ramaiya NH, Hodi FS (2015) Anti-PD-1–relatedpneumonitis during cancer immunotherapy. N Engl J Med 373:288–290. https://doi.org/10.1056/NEJMc1505197

4. Thompson JA, Schneider BJ, Brahmer J et al (2019) Management of immunotherapy related toxicities, Version 1.2019, NCCN clinical practice guidelines in oncology. J Natl Compr Cancer Netw 17:255–289. https://doi.org/10.6004/jnccn.2019.0013

5. Schneider BJ, Naidoo J, Santomasso BD et al (2021) Management of immune-related adverse events in patients treated with immune checkpoint inhibitor therapy: ASCO guideline update. J Clin Oncol 39:4073–4126. https://doi.org/10.1200/JCO.21.01440

6. Takeshi J, Lee KS, Nishino M et al (2021) Chest CT diagnosis and clinical management of drug-related pneumonitis in patients receiving molecular targeting agents and immune checkpoint inhibitors: a position paper from the fleischner society. Radiology 298:550–566. https://doi.org/10.1148/radiol.2021203427

7. Kalisz KR, Ramaiya NH, Laukamp KR, Gupta A (2019) immune checkpoint inhibitor therapy–related pneumonitis: patterns and management. Radiographics 39:1923–1937. https://doi.org/10.1148/rg.2019190036

8. Dumoulin DW, Gietema HA, Paats MS, Hendriks LEL, Cornelissen R (2020) Differentiation of COVID-19 pneumonitis and ICI induced pneumonitis. Front Oncol 10:577696. https://doi.org/10.3389/fonc.2020.577696

9. Russano M, Citarella F, Napolitano A et al (2020) COVID-19 pneumonia and immune-related pneumonitis: critical issues on differential diagnosis, potential interactions, and management. Expert Opin Biol Ther 20:959–964. https://doi.org/10.1080/14712598.2020.1789097

10. Sullivan RJ, Johnson DB, Rini BI et al (2020) COVID-19 and immune checkpoint inhibitors: initial considerations. J Immunother Cancer 8:e000933. https://doi.org/10.1136/jitc-2020-000933

11. Brendlin AS, Mader M, Faby S et al (2021) AI Lung segmentation and perfusion analysis of dual-energy CT can help to distinguish COVID-19 infiltrates from visually similar immunotherapy-related pneumonitis findings and can optimize radiological workflows. Tomography 8(1):22–32. https://doi.org/10.3390/tomography8010003

12. Rona G, Arifoğlu M, Voyvoda N, Batırel A (2021) Should CT be used for the diagnosis of RT-PCR-negative suspected COVID-19 patients? Clin Respir J 15:491–498. https://doi.org/10.1111/crj.13332

13. Prokop M, van Everdingen W, van Rees VT et al (2020) CORADS: a categorical CT assessment scheme for patients suspected of having COVID-19—definition and evaluation. Radiology 296:E97–E104. https://doi.org/10.1148/radiol.2020201473

14. Chung M, Bernheim A, Mei X et al (2020) CT imaging features of 2019 NovelCoronavirus (2019-nCoV). Radiology 295:202–207. https://doi.org/10.1148/radiol.2020200230

15. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Müller NL, Remy J (2008) Fleischner society: glossary of terms for thoracic imaging. Radiology 246:697–722. https://doi.org/10.1148/radiol.2462070712

16. Sardanelli F, Cozzi A, Monfardini L et al (2020) Association of mediastinal lymphadenopathy with COVID-19 prognosis. Lancet Infect Dis 20:1230–1231. https://doi.org/10.1016/S1473-3099(20)30521-1

17. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174. https://doi.org/10.2307/2529310

18. Rutjes AWS, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PMM (2005) Case–control and two-gate designs in diagnostic accuracy studies. Clin Chem 51:1335–1341. https://doi.org/10.1373/clinchem.2005.048595

19. Robert C (2020) A decade of immune-checkpoint inhibitors in cancer therapy. NatCommun 11:3801. https://doi.org/10.1038/s41467-020-17670-y

20. Nishino M, Giobbie-Hurder A, Hatabu H, Ramaiya NH, Hodi FS (2016) Incidence of programmed cell death 1 inhibitor-related pneumonitis in patients with advanced cancer. JAMA Oncol 2:1607. https://doi.org/10.1001/jamaoncol.2016.2453

21. Naidoo J, Wang X, Woo KM et al (2017) Pneumonitis in patients treated with anti-programmed death-1/programmed death ligand 1 therapy. J Clin Oncol 35:709–717. https://doi.org/10.1200/JCO.2016.68.2005

22. Ando H, Suzuki K, Yanagihara T (2021) Insights into potential pathogenesis andtreatment options for immune-checkpoint inhibitor-related pneumonitis. Biomedicines 9:1484. https://doi.org/10.3390/biomedicines9101484

23. Suresh K, Naidoo J, Zhong Q et al (2019) The alveolar immune cell landscape is dysregulated in checkpoint inhibitor pneumonitis. J Clin Invest 129:4305–4315. https://doi.org/10.1172/JCI128654

24. Strippoli S, Fucci L, Negri A et al (2020) Cellular analysis of bronchoalveolar lavage fluid to narrow differential diagnosis of checkpoint inhibitor-related pneumonitis in metastatic melanoma. J Transl Med 18:473. https://doi.org/10.1186/s12967-020-02650-z

25. Brahmer JR, Lacchetti C, Schneider BJ et al (2018) Management of immune-relatedadverse events in patients treated with immune checkpoint inhibitor therapy: American Society of Clinical oncology clinical practice guideline. J Clin Oncol 36:1714–1768. https://doi.org/10.1200/JCO.2017.77.6385

26. Pecoraro V, Negro A, Pirotti T, Trenti T (2022) Estimate false-negative RT-PCR rates for SARS-CoV-2: a systematic review and meta-analysis. Eur J Clin Invest 52:e13706