ABSTRACT

Background Despite current recommendations, there is no recent scientific study comparing the influence of CT reconstruction kernels on lung pattern recognition in interstitial lung disease (ILD).

Purpose To evaluate the sensitivity of lung (i70) and soft (i30) CT kernel algorithms for the diagnosis of ILD patterns.

Materials and Methods We retrospectively extracted between 15–25 pattern annotations per case (1 annotation = 15 slices of 1 mm) from 23 subjects resulting in 408 annotation stacks per lung kernel and soft kernel reconstructions. Two subspecialized chest radiologists defined the ground truth in consensus. 4 residents, 2 fellows, and 2 general consultants in radiology with 3 to 13 years of experience in chest imaging performed a blinded readout. In order to account for data clustering, a generalized linear mixed model (GLMM) with random intercept for reader and nested for patient and image and a kernel/experience interaction term was used to analyze the results.

Results The results of the GLMM indicated, that the odds of correct pattern recognition is 12% lower with lung kernel compared to soft kernel; however, this was not statistically significant (OR 0.88; 95%-CI, 0.73–1.06; p = 0.187). Furthermore, the consultants’ odds of correct pattern recognition was 78% higher than the residents’ odds, although this finding did not reach statistical significance either (OR 1.78; 95%-CI, 0.62–5.06; p = 0.283). There was no significant interaction between the two fixed terms kernel and experience. Intrarater agreement between lung and soft kernel was substantial (κ = 0.63 ± 0.19). The mean inter-rater agreement for lung/soft kernel was κ = 0.37 ± 0.17/κ = 0.38 ± 0.17.
**Introduction**

High-resolution computed tomography (CT) has been established as a cornerstone in the diagnosis and follow-up of interstitial lung disease (ILD) [1]. Interstitial lung disease presents with a distinct array of CT imaging patterns in a certain distribution [2, 3].

CT is an essential diagnostic test in the diagnosis of suspected idiopathic pulmonary fibrosis (IPF). With the release of novel recommendations by the Fleischner Society as well as the ATS/ERS/JRS/ALAT, the role of CT became even more apparent [1, 4, 5]. In the appropriate clinical context and the presence of the usual interstitial pneumonia (UIP) pattern on chest CT, a confident diagnosis of IPF can be achieved [6]. Growing evidence supports the assumption that also a “probable UIP” pattern, while lacking the typical honeycombing appearance, predicts the presence of IPF with substantial likelihood [1, 5]. Therefore, the stratification into the diagnostic categories of UIP based on CT are a crucial task for radiologists. Consequently, an utmost critical step in the diagnosis of ILD is the correct identification of the CT patterns present on a particular scan.

To achieve optimized spatial resolution, CT scans of the lungs are reconstructed with edge-enhancing reconstruction algorithms maximizing contrast. Using lung kernels for the assessment and measurement of lung nodules as well as for the diagnosis of ILD has been recommended by several publications [1, 7, 8]. However, the lung kernel simultaneously increases image noise [9]. Previously published results showed that computer-aided detection (CAD) applications exhibit better results when used along with soft reconstruction kernels [10]. We noticed that these soft kernels, which are designed to elevate tissue contrast while reducing noise, are increasingly applied in the clinical routine to assess the lung parenchyma. Various studies have shown that reconstruction kernels and other technical parameters have a substantial impact on image perception [11–16].

Despite the wide endorsement of using lung kernel reconstruction algorithms for the diagnosis of chest CT patterns, there is no
recent evidence supporting those recommendations. Results published prior to the introduction of helical CT in the early 1990s indicated the beneficial use of lung kernel reconstructions on chest CT [9, 17–19]. Despite the advances in CT technique, to our knowledge no systematic investigation of image reconstruction algorithms and their impact on lung parenchymal assessment has been conducted.

The purpose of this study was to directly compare the radiologists’ sensitivity for the classification of image patterns on chest CT using lung- and soft kernels. We used a subset of CT patterns linked to the UIP pattern. In order to reflect a broad range of clinical scenarios, we involved radiologists with various levels of experience.

Materials and Methods

The local ethics review board (IRB) approved this retrospective study. Twenty-three chest CT scans were retrospectively selected from the local ILD database from our tertiary care center. All included patients had a multidisciplinary team (MDT) diagnosis of IPF or NSIP (nonspecific interstitial pneumonia, including idiopathic and secondary NSIP) on CT in patients with connective tissue disease) between January 2018 and May 2018. CT scans were acquired on a 128-MDCT scanner (Somatom Definition Flash, Siemens Healthcare) with iterative reconstruction (IRIS, Siemens Healthcare) and an integrated detector system (Stellar, Siemens Healthcare). Chest CT scans were performed with 100 kVp and 100 Reference mAs. FOV was kept constant at 32 cm. The slice reconstruction thickness was 1 mm. For raw dataset processing, level three iteration was applied. Axial slices were reconstructed with two different kernels (i30 and i70) at a slice thickness of 1 mm. For raw dataset processing, level three iteration was applied. Axial slices were reconstructed with two different kernels (i30 and i70) at a slice thickness of 1 mm for interpretation. Multiplane reconstructions were not performed.

The following inclusion criteria were applied: 1) Patients with ILD-board approved usual interstitial pneumonia (UIP) or nonspecific interstitial pneumonia (NSIP); 2) The scans included soft and lung kernel raw data reconstructions; 3) Diagnostic image quality (no image artifacts). In total, we included 23 consecutive patients with interstitial lung disease (13 patients with UIP pattern, 10 patients presenting with NSIP pattern). The CT scans were anonymized and transferred to a separate database for further pattern assessment.

Standard of Reference

Two subspecialized chest radiologists (AC, LE) with 12 and 20 years of experience in chest imaging reviewed the CT scans with soft kernels as well as lung kernels. Window level and width were set at of −500 HU and 1500 HU, respectively. On these scans, the chest radiologists annotated eight typical CT patterns deemed critical: healthy lung parenchyma, ground glass opacification, reticulations, consolidation, honeycombing, reticulations in combination with ground glass opacification, bronchiectasis, and emphysema (▶ Fig. 1). The open-source software Slicer 4.10.2 (http://www.slicer.org) was used for color annotating each slice. 15–25 areas per case were selected. The continuous annotated areas were used to create a stack of 15 consecutive slices for radiological evaluation (1 stack = 15 contiguous slices of 1 mm), resulting in a database containing 816 image stacks (408 soft kernel, 408 lung kernel). Within each stack, the predefined CT pattern was marked for further readout. After the completion of the annotation, a consensus panel among the chest radiologists resolved discrepancies. The amount of the selected patterns matched the frequencies of the patterns in the real UIP and NSIP cases.

Image Analysis

Eight independent readers blinded to the patient diagnosis and annotated patterns performed the readout. The readers were two board-certified radiologists with 12 and 13 years of experience, two radiology fellows (one thoracic radiology fellow, one fellow in abdominal imaging) with four and five years of experience, respectively, and four residents with three to four years of training. Each reader received an individually randomized database of the CT cases. It should be noted that the presentation sequence of the soft and the lung kernel was also randomly assigned. The readout was performed using Slicer software. On each pre-annotated CT stack, readers were asked to determine the particular CT pattern highlighted. Readers recorded the patterns. A synopsis of the study design is depicted in ▶ Fig. 2.
**Statistics**

In order to account for data clustering, a generalized linear mixed model (GLMM) with random intercept for reader and nested for patient and image with interaction term for the fixed effects kernel and experience was used to analyze the results. The investigated binary endpoint was the correct pattern recognition (0 = no, 1 = yes). The model accounts for the repeated measures of reader, patient and image. The analysis was performed using R Software version 4.0.4. Intra-rater (lung versus soft kernel) agreement and inter-rater agreement were calculated according to the weighted Kappa statistic for the lung kernel and the soft kernel separately. The following classifications of Kappa levels were used: slight ($\kappa \leq 0.20$), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), and almost perfect agreement ($\kappa \geq 0.81$). Kappa levels were compared using correlation coefficient testing (z-statistic). The significance level was set to 0.05. MedCalc version 15.0 (MedCalc Software, Ostend, Belgium) and SPSS (IBM, Armonk, New York, USA) were utilized.

**Results**

The eight radiologists scored in total 80 healthy lung parenchyma regions, 22 ground glass patterns, 67 regions with reticulations, 29 areas presenting with consolidation, 86 cases of honeycombing, 65 cases showing reticulations and ground glass, as well as 23 annotations of bronchiectasis and 36 emphysema cases (3264 classifications in lung and 3264 classifications in soft kernel images) (**Fig. 3**).

**Individual Reader Analysis**

Individual pattern sensitivities of all readers are shown in **Table 1**. Individual per reader analysis did not demonstrate significant differences in pattern sensitivities comparing lung kernel versus soft kernel, except for one resident, who detected significantly more emphysema pattern using soft kernel images ($p = 0.0455$) and one consultant, who detected more reticulation with ground glass opacities on lung kernel images. The pooled results of all readers demonstrated an overall pattern sensitivity of 77.8% for lung kernel and 76.3% for soft kernel ($p = 0.0838$).

**Generalized Linear Mixed Model**

The results of the GLMM (**Table 2**) indicated, that the probability of correct pattern recognition is 12% lower with lung kernel compared to soft kernel; however, this finding did not reach statistical significance (OR 0.88; 95%-CI, 0.73–1.06; $p = 0.187$). Analysis per experience level showed highest frequencies of correct pattern detection by consultants, followed by residents and fellows. The consultants’ odds of correct pattern recognition was 78% higher than the residents’ odds, although this finding did not reach statistical significance either (OR, 1.78; 95%-CI, 0.62–5.06; $p = 0.283$). The results indicated no significant interaction between the fixed terms kernel and experience.

**Missed Cases**

The misclassified cases are depicted in **Fig. 4**. For the lung kernel reconstruction, readers tended to falsely classify emphysema as healthy lung parenchyma, honeycombing was rated as reticulation, and ground glass opacities in combination with reticulations were scored as showing reticulation or ground glass and vice versa (**Fig. 4A**). In soft kernel reconstructions, more accurate classification of emphysema was observed. The hyper- and hypoattenuation in the ground glass pattern and normal lung was mistaken for emphysema using the soft reconstruction kernel. Also, reticulations were scored as ground glass opacification and reticulation (**Fig. 4B**).

**Intra-Rater Agreement**

The overall intra-rater agreement between the lung kernel and the soft kernel was substantial ($\kappa = 0.63 \pm 0.19$). The consultants showed a significantly higher agreement ($\kappa = 0.89 \pm 0.18$) compared to the fellow radiologists ($\kappa = 0.56 \pm 0.08$; $p = 0.0001$). The moderate agreement among the fellow radiologists was significantly better than among the residents ($\kappa = 0.44 \pm 0.04$; $p = 0.003$).

The mean inter-rater agreement for all patterns on the lung kernel did not vary significantly from the agreement on the soft kernel images $\kappa = 0.38 \pm 0.17$ and $\kappa = 0.37 \pm 0.17$; $p = 0.64$), although, they differed among the three experience levels (**Table 3**).

**Discussion**

Chest CT with lung kernel reconstruction for ILD pattern detection is widely recommended. However, the effect of lung kernel reconstruction algorithms on pattern recognition has not recently been investigated.

Accurate pattern recognition on CT scans is a prerequisite in ILD diagnosis. Many patients with IPF are diagnosed with CT imaging only, avoiding the potential risks of invasive tissue sampling [1, 20–22]. This is especially important, as subjects with possible IPF are prone to develop an acute exacerbation associated with high mortality [23–25]. Therefore, optimization of interpretation parameters and imaging standards is essential.

The detection of honeycombing as a hallmark of the UIP pattern is an important marker for clinical decision-making. In a multidisci-
### Table 1 Individual pattern sensitivities of all readers.

<table>
<thead>
<tr>
<th>Pattern (ground truth)</th>
<th>Kernel</th>
<th>Total n</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>F1</th>
<th>F2</th>
<th>C1</th>
<th>C2</th>
<th>Sens all</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n</td>
<td>n %</td>
<td>n %</td>
<td>n %</td>
<td>n %</td>
<td>n %</td>
<td>n %</td>
<td>n %</td>
<td>n %</td>
<td></td>
</tr>
<tr>
<td>Normal lung</td>
<td>Lung</td>
<td>80</td>
<td>80</td>
<td>100</td>
<td>79</td>
<td>99</td>
<td>79</td>
<td>99</td>
<td>80</td>
<td>100</td>
<td>74</td>
</tr>
<tr>
<td>Normal lung</td>
<td>Soft</td>
<td>80</td>
<td>79</td>
<td>99</td>
<td>77</td>
<td>96</td>
<td>79</td>
<td>99</td>
<td>79</td>
<td>96</td>
<td>74</td>
</tr>
<tr>
<td>GGO</td>
<td>Lung</td>
<td>22</td>
<td>22</td>
<td>100</td>
<td>15</td>
<td>68</td>
<td>18</td>
<td>82</td>
<td>20</td>
<td>91</td>
<td>16</td>
</tr>
<tr>
<td>GGO</td>
<td>Soft</td>
<td>22</td>
<td>22</td>
<td>100</td>
<td>17</td>
<td>77</td>
<td>18</td>
<td>82</td>
<td>16</td>
<td>73</td>
<td>15</td>
</tr>
<tr>
<td>Reticulation</td>
<td>Lung</td>
<td>67</td>
<td>67</td>
<td>100</td>
<td>56</td>
<td>84</td>
<td>46</td>
<td>69</td>
<td>60</td>
<td>90</td>
<td>49</td>
</tr>
<tr>
<td>Reticulation</td>
<td>Soft</td>
<td>67</td>
<td>67</td>
<td>100</td>
<td>49</td>
<td>73</td>
<td>54</td>
<td>81</td>
<td>52</td>
<td>78</td>
<td>39</td>
</tr>
<tr>
<td>Consolidation</td>
<td>Lung</td>
<td>29</td>
<td>29</td>
<td>100</td>
<td>22</td>
<td>97</td>
<td>29</td>
<td>100</td>
<td>28</td>
<td>97</td>
<td>27</td>
</tr>
<tr>
<td>Consolidation</td>
<td>Soft</td>
<td>29</td>
<td>29</td>
<td>100</td>
<td>27</td>
<td>93</td>
<td>28</td>
<td>97</td>
<td>28</td>
<td>97</td>
<td>27</td>
</tr>
<tr>
<td>Honeycombimg</td>
<td>Lung</td>
<td>86</td>
<td>86</td>
<td>100</td>
<td>65</td>
<td>76</td>
<td>58</td>
<td>67</td>
<td>61</td>
<td>71</td>
<td>72</td>
</tr>
<tr>
<td>Honeycombimg</td>
<td>Soft</td>
<td>86</td>
<td>86</td>
<td>100</td>
<td>64</td>
<td>74</td>
<td>55</td>
<td>64</td>
<td>65</td>
<td>76</td>
<td>50</td>
</tr>
<tr>
<td>GGO+RET</td>
<td>Lung</td>
<td>65</td>
<td>65</td>
<td>100</td>
<td>39</td>
<td>60</td>
<td>20</td>
<td>31</td>
<td>28</td>
<td>43</td>
<td>27</td>
</tr>
<tr>
<td>GGO+RET</td>
<td>Soft</td>
<td>65</td>
<td>65</td>
<td>100</td>
<td>36</td>
<td>55</td>
<td>15</td>
<td>23</td>
<td>39</td>
<td>60</td>
<td>34</td>
</tr>
<tr>
<td>Bronchiectasis</td>
<td>Lung</td>
<td>23</td>
<td>23</td>
<td>100</td>
<td>22</td>
<td>96</td>
<td>20</td>
<td>87</td>
<td>21</td>
<td>91</td>
<td>21</td>
</tr>
<tr>
<td>Bronchiectasis</td>
<td>Soft</td>
<td>23</td>
<td>23</td>
<td>100</td>
<td>21</td>
<td>91</td>
<td>18</td>
<td>78</td>
<td>23</td>
<td>100</td>
<td>19</td>
</tr>
<tr>
<td>Emphysema</td>
<td>Lung</td>
<td>36</td>
<td>36</td>
<td>100</td>
<td>18</td>
<td>50</td>
<td>32</td>
<td>89</td>
<td>31</td>
<td>86</td>
<td>25</td>
</tr>
<tr>
<td>Emphysema</td>
<td>Soft</td>
<td>36</td>
<td>36</td>
<td>100</td>
<td>25</td>
<td>69</td>
<td>32</td>
<td>89</td>
<td>32</td>
<td>89</td>
<td>36</td>
</tr>
<tr>
<td>All</td>
<td>Lung</td>
<td>408</td>
<td>408</td>
<td>100</td>
<td>328</td>
<td>80</td>
<td>298</td>
<td>73</td>
<td>328</td>
<td>80</td>
<td>321</td>
</tr>
<tr>
<td>All</td>
<td>Soft</td>
<td>408</td>
<td>408</td>
<td>100</td>
<td>321</td>
<td>79</td>
<td>278</td>
<td>68</td>
<td>327</td>
<td>80</td>
<td>323</td>
</tr>
</tbody>
</table>

Note: R: resident#; F: fellow#; C: consultant#; GGO: Ground Glass Opacity; RET: reticulation.
### Table 2 Generalized Linear Mixed Model.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Terms</th>
<th>Odds Ratio</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT kernel</td>
<td>lung kernel vs. soft kernel</td>
<td>0.880</td>
<td>0.729–1.064</td>
<td>0.187</td>
</tr>
<tr>
<td>Reader experience</td>
<td>fellow vs. resident</td>
<td>0.641</td>
<td>0.227–1.809</td>
<td>0.401</td>
</tr>
<tr>
<td></td>
<td>consultant vs. resident</td>
<td>1.775</td>
<td>0.623–5.060</td>
<td>0.283</td>
</tr>
</tbody>
</table>

Note: CI: confidence interval.

### Fig. 4 Confusion matrix across all readers. Misclassified patterns are highlighted in light blue. A: soft kernel; B: lung kernel. NORM: normal lung; GGO: ground glass opacity; RET: reticulation; CONS: consolidation; HCOMB: honeycombing; BECT: bronchiectasis; EMPH: emphysema.

### Abb. 4 Wahrheitsmatrix. Falsch klassifizierte Muster sind hellblau hervorgehoben. A: Weichteilkern; B: Lungenkern. NORM: normales Lungenparenchym; GGO: Milchglasverdichtung; RET: Retikulation; CONS: Konsolidation; HCOMB: Honigwabenmuster; BECT: Bronchiektasen; EMPH: Emphysem.
pulmonary team setting, this allows for the confident diagnosis of IPF [26–28]. To diagnose a typical UIP pattern on CT requires the presence of honeycombing in a basal and subpleural distribution. Although the sensitivity for honeycombing in our study was moderate, this is in alignment with previous reports [29, 30]. In the recently published ATS/ERS/JRS/ALAT criteria for UIP CT patterns, the role of CT in the diagnosis of IPF was expanded substantially. This is based on growing evidence that the typical UIP pattern with honeycombing represents a late-stage disease and that the probability of microscopic UIP is high in patients showing a probable UIP pattern on CT. Recent literature showed that, in the right clinical context and MDT approach, those cases could also be diagnosed as IPF with high confidence with CT alone [1, 6]. We expected the benefit of a higher spatial resolution with lung kernel reconstructions to be superior to soft kernel reconstructions for the correct classification of interstitial lung patterns (▶Fig. 5). The results of the utilized model reveal that the odds for correct pattern identification using lung kernel reconstructions is in fact 12 % lower compared to tissue kernel without reaching statistical significance. However, these findings challenge current recommendations for the diagnosis of interstitial lung disease as they indicate non-superiority of the lung kernel over the soft kernel.

The experienced consultants demonstrated the highest intra-rater agreement which equals the highest consistency between the lung kernel and the soft kernel.

Several limitations warrant discussion. First, we built an artificial reading scenario with annotated image stacks, and we only evaluated the effect of the different kernel reconstructions on pattern recognition, not the effect on the final diagnosis when reading the whole CT scan. This might not reflect an actual reading environment. However, this allowed us to isolate the effect of different spatial resolution algorithms on pattern classification. Second, our readers were blinded, not only to the MDT discussion and distribution, but also to clinical background information (age, patient history, smoking status, etc.). Although this is also critical information, we did not target a clinical diagnosis. For determining the CT pattern sensitivity and the kernel-effect in isolation, the focused approach was deemed suitable. Third, we used a limited subset of CT patterns. In this evaluation, we focused on patterns that are most critical for UIP diagnosis. For this evaluation, we used kernels with relatively high frequency (i70). It should be noted that many institutions use more mid-frequency enhancement for lung reading. The impact of the mid-frequency lung kernel reconstruction was not investigated in the present study. Fourth, the same pattern was selected multiple times in a specific patient. This could represent potential clustering. However, a generalized linear mixed model was designed to account for this problem. Fifth, we had a relatively low count of 8 different readers, so only few observations in the 3 experience levels, which limits the meaningfulness of the per experience analysis.

### Conclusion

In conclusion, we found that there is no significant difference between lung and soft kernel reconstructed CT images for the correct pattern recognition in interstitial lung disease. There are non-significant trends that the use of soft kernels and a higher level of experience lead to a higher probability of correct pattern identification.

<table>
<thead>
<tr>
<th>INTER rater agreement</th>
<th>INTRA rater agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung kernel</td>
<td>Soft kernel</td>
</tr>
<tr>
<td>kappa</td>
<td>SD</td>
</tr>
<tr>
<td>R (residents):</td>
<td>0.46 ± 0.09</td>
</tr>
<tr>
<td>F (fellows):</td>
<td>0.25 ± 0.05</td>
</tr>
<tr>
<td>C (consultants):</td>
<td>0.05 ± 0.04</td>
</tr>
<tr>
<td>All:</td>
<td>0.38 ± 0.17</td>
</tr>
</tbody>
</table>

Note: SD: standard deviation; P: p-value according to comparison of correlation coefficient.
Clinical Relevance

Accurate identification of interstitial lung patterns on CT is pivotal in the diagnosis of interstitial lung disease. Our results show that soft kernel reconstructions are comparable to or maybe even better than lung kernel reconstructions for the correct identification of interstitial lung patterns. This finding could have major impact on clinical practice, as it questions the current recommendations and the routinely performed separate lung kernel reconstructions for lung parenchyma analysis. Provided that the observations are confirmed in further studies, these additional lung kernel reconstructions could be omitted, which could save large amounts of data storage space.

Conflict of Interest

The authors declare that they have no conflict of interest.

References


