Supplemental Table 1. Diagnostic yields of grayscale mode in differentiating benign from malignant LNs.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Mode | Sensitivity | | Specificity | PPV | NPV | Accuracy |
| Grayscale 5 | |  |  |  |  |  |
| 0 | 1.9% | | 81.0% | 11.8% | 37.8% | 35.4% |
| 1+ | 98.1% | | 19.0% | 62.2% | 88.2% | 64.6% |
| 2+ | 86.5% | | 62.0% | 75.6% | 77.2% | **76.1%** |
| 3+ | 56.3% | | 87.3% | 85.8% | 59.5% | 69.4% |
| 4+ | 37.7% | | 94.3% | 90.0% | 52.7% | 61.7% |
| 5+ | 8.4% | | 99.4% | 94.7% | 44.4% | 46.9% |
| Grayscale 4 |  | |  |  |  |  |
| 0 | 1.9% | | 81.0% | 11.8% | 37.8% | 35.4% |
| 1+ | 98.1% | | 19.0% | 62.2% | 88.2% | 64.6% |
| 2+ | 86.5% | | 62.7% | 75.9% | 77.3% | **76.4%** |
| 3+ | 48.4% | | 88.6% | 85.3% | 55.8% | 65.4% |
| 4+ | 34.9% | | 94.9% | 90.4% | 51.7% | 60.3% |
| Grayscale 3 |  | |  |  |  |  |
| 0 | 1.9% | | 81.0% | 11.8% | 37.8% | 35.4% |
| 1+ | 98.1% | | 19.0% | 62.2% | 88.2% | 64.6% |
| 2+ | 86.5% | | 62.7% | 75.9% | 77.3% | **76.4%** |
| 3+ | 47.0% | | 88.6% | 84.9% | 55.1% | 64.6% |
| Grayscale 2 |  | |  |  |  |  |
| 0 | 7.0% | | 50.6% | 16.1% | 28.6% | 25.5% |
| 1+ | 93.0% | | 49.4% | 71.4% | 83.9% | **74.5%** |
| 2+ | 50.2% | | 87.3% | 84.4% | 56.3% | 66.0% |

Grayscale 5 refers that five grayscale features with the highest accuracies (absence of CHS, heterogeneity, short axis＞10mm, round shape, presence of necrosis) are used, as well as grayscale 4(absence of CHS, heterogeneity, short axis＞10mm, round shape), 3(absence of CHS, heterogeneity, short axis＞10mm) and 2(absence of CHS, heterogeneity) . The items of 1+, 2+, 3+, 4+ and 5+ represented at least 1, 2, 3, 4 and 5 kinds of sonographic features were present in each mode, respectively. Bold indicates that the best accuracy of each method in differentiating benign from malignant LNs.

CHS, central hilar structure; LN, lymph node; NPV, negative predictive value; PPV, positive predictive value

Supplemental Table 2. Diagnostic efficiency of EBUS single and multimodal imaging in the model group with three grayscale features.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Mode | Sensitivity | Specificity | PPV | NPV | Accuracy |
| **G** |  |  |  |  |  |
| 0 | 1.9% | 81.0% | 11.8% | 37.8% | 35.4% |
| 1+ | 98.1% | 19.0% | 62.2% | 88.2% | 64.6% |
| 2+ | 86.5% | 62.7% | 75.9% | 77.3% | **76.4%** |
| 3+ | 47.0% | 88.6% | 84.9% | 55.1% | 64.6% |
| **F** | 85.6% | 49.4% | 69.7% | 71.6% | **70.2%** |
| **E** | 82.8% | 79.8% | 84.8% | 77.3% | **81.5%** |
| **G+F** |  |  |  |  |  |
| 0 | 1.4% | 84.2% | 10.7% | 38.6% | 36.5% |
| 1+ | 98.6% | 15.8% | 61.5% | 89.3% | 63.5% |
| 2+ | 91.2% | 46.8% | 70.0% | 79.6% | 72.4% |
| 3+ | 81.9% | 67.7% | 77.5% | 73.3% | **75.9%** |
| 4+ | 45.6% | 89.2% | 85.2% | 54.7% | 64.1% |
| **G+E** |  |  |  |  |  |
| 0 | 1.4% | 81.7% | 9.4% | 37.8% | 35.4% |
| 1+ | 98.6% | 18.4% | 62.2% | 90.6% | 64.6% |
| 2+ | 93.5% | 55.7% | 74.2% | 86.3% | 77.5% |
| 3+ | 82.3% | 80.4% | 85.1% | 77.0% | **81.5%** |
| 4+ | 40.0% | 95.6% | 92.5% | 53.9% | 63.5% |
| **F+E** |  |  |  |  |  |
| 0 | 4.7% | 53.2% | 11.9% | 29.1% | 25.2% |
| 1+ | 95.4% | 46.8% | 70.9% | 88.1% | 74.8% |
| 2+ | 73.0% | 82.3% | 84.9% | 69.2% | **76.9%** |
| **G+F+E** |  |  |  |  |  |
| 0 | 0.9% | 84.8% | 7.7% | 38.6% | 36.5% |
| 1+ | 99.1% | 15.2% | 61.4% | 92.3% | 63.5% |
| 2+ | 95.4% | 45.6% | 70.5% | 87.8% | 74.3% |
| 3+ | 88.4% | 62.0% | 76.0% | 79.7% | 77.2% |
| 4+ | 78.6% | 80.4% | 84.5% | 73.4% | **79.4%** |
| 5+ | 38.6% | 96.2% | 93.3% | 53.5% | 63.0% |

Three grayscale features (absence of CHS, heterogeneity and short axis > 10mm) were used in each multimodal combination, and elastography had the highest diagnostic accuracy of 81.5%, higher than that of any multimodal combination. Absence of hilar vascularity was used in blood flow Doppler and grading score 4-5 was used in elastography. The items of 1+, 2+, 3+, 4+ and 5+ represented at least 1, 2, 3, 4 and 5 kinds of sonographic features were present in each mode, respectively. Bold indicates that the best accuracy of each method in differentiating benign from malignant LNs.

CHS, central hilar structure; E, elastography; F, blood flow Doppler; G, grayscale; NPV, negative predictive value; PPV, positive predictive value

Supplemental Table 3. Diagnostic efficiency of EBUS single and multimodal imaging in the validation group with three grayscale features.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Mode | Sensitivity | Specificity | PPV | NPV | Accuracy |
| **G** |  |  |  |  |  |
| 0 | 1.3% | 79.7% | 7.7% | 37.6% | 34.8% |
| 1+ | 98.7% | 20.3% | 62.4% | 92.3% | 65.2% |
| 2+ | 88.6% | 61.0% | 75.3% | 80.0% | **76.8%** |
| 3+ | 43.0% | 91.5% | 87.2% | 54.6% | 63.8% |
| **F** | 86.1% | 50.9% | 70.1% | 73.2% | **71.0%** |
| **E** | 87.3% | 78.0% | 84.2% | 82.1% | **83.3%** |
| **G+F** |  |  |  |  |  |
| 0 | 1.3% | 83.1% | 9.1% | 38.6% | 36.2% |
| 1+ | 98.7% | 17.0% | 61.4% | 90.9% | 63.8% |
| 2+ | 93.7% | 50.9% | 71.8% | 85.7% | **75.4%** |
| 3+ | 82.3% | 64.4% | 75.6% | 73.1% | 74.6% |
| 4+ | 41.8% | 91.5% | 86.8% | 54.0% | 63.0% |
| **G+E** |  |  |  |  |  |
| 0 | 1.3% | 83.1% | 9.1% | 38.6% | 36.2% |
| 1+ | 98.7% | 17.0% | 61.4% | 90.9% | 63.8% |
| 2+ | 94.9% | 55.9% | 74.3% | 89.2% | 78.3% |
| 3+ | 82.3% | 81.4% | 85.5% | 77.4% | **81.9%** |
| 4+ | 41.8% | 96.6% | 94.3% | 55.3% | 65.2% |
| **F+E** |  |  |  |  |  |
| 0 | 5.1% | 57.6% | 13.8% | 31.2% | 27.5% |
| 1+ | 94.9% | 42.4% | 68.8% | 86.2% | 72.5% |
| 2+ | 78.5% | 86.4% | 88.6% | 75.0% | **81.9%** |
| **G+F+E** |  |  |  |  |  |
| 0 | 1.3% | 86.4% | 11.1% | 39.5% | 37.7% |
| 1+ | 98.7% | 13.6% | 60.5% | 88.9% | 62.3% |
| 2+ | 97.5% | 47.5% | 71.3% | 93.3% | 76.1% |
| 3+ | 89.9% | 61.0% | 75.5% | 81.8% | 77.5% |
| 4+ | 76.0% | 83.1% | 85.7% | 72.1% | **79.0%** |
| 5+ | 41.8% | 96.6% | 94.3% | 55.3% | 65.2% |

Three grayscale features (absence of CHS, heterogeneity and short axis > 10mm) were used in each multimodal combination, and elastography had the highest diagnostic accuracy of 83.33%, higher than that of any multimodal combination. Absence of hilar vascularity was used in blood flow Doppler and grading score 4-5 was used in elastography. The items of 1+, 2+, 3+, 4+ and 5+ represented at least 1, 2, 3, 4 and 5 kinds of sonographic features were present in each mode, respectively. Bold indicates that the best accuracy of each method in differentiating benign from malignant LNs.

CHS, central hilar structure; E, elastography; F, blood flow Doppler; G, grayscale; NPV, negative predictive value; PPV, positive predictive value