Examples of mammography reports

All the mammography reports used in this study are in simplified Chinese. In Figure 1-3, we provide some examples of de-identified text reports and their English translation (automatically translated by Google Translate), together with the probabilistic labels predicted by our NLP model, namely, Mammo-RoBERTa.

Figure 1. Mammography report of Example #1 and its labels generated by Mammo-RoBERTa. Reference standard (Ground-truth): a mass is present in the left breast; no clear mass in the right breast. Our NLP labels indicate a confidence score of 0.9908 for a left mass and 0.0031 for a right mass, respectively.
### Findings:

Types of mammary glands: mixed type. 
Nipple: No depression, fixation, or deformation. 
Areola and skin: No abnormal thickening was seen. 
Subcutaneous fat: widened, the front edge of the gland is smooth, and the boundary is clear. 
Suspension ligament: No development. 
Glandular structure: The density of glands is uneven, and round and irregular dense shadows can be seen between them. 
Lump and signs of malignant transformation: There are 5 suspicious small calcifications in the lower left outer quadrant of the breast, without mass nodules. The cord-like duct shadow and the cluster-like dense area fuse into a cluster. Duct phase: no hyperplasia and expansion. 
Vascular phase: natural walking, no tortuosity, thickening, or deformation. 
Fat layer behind glands: clear; no abnormalities. 
Armpit: No swollen lymph nodes were seen.

### Impression:

1. There are 5 suspicious small calcifications in the lower outer quadrant of the left breast. Please review it in a short-term basis if necessary. 
2. Bilateral breast hyperplasia. Please follow up with the clinic.

<table>
<thead>
<tr>
<th>Breast</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prob. of mass</td>
<td>0.0100</td>
<td>0.0041</td>
</tr>
</tbody>
</table>

---

### Findings:

The parenchymal composition of the bilateral mammary glands was heterogeneous and dense (ACR type C); the nipples were not depressed or deformed; the areola and skin were not abnormally thickened; the subcutaneous fat was clear, and the suspensory ligament was thickened with a bull's-horn-like protrusion. The gland was basically symmetrical bilaterally, with a large amount of flocculent flakes and uneven density. An oval-shaped isointense mass was seen in the upper outer quadrant of the right breast, without lobulation, with partially clear and partially blurred margins and no clear calcification or burr; about 2.0×1.8 cm in size, and two tiny punctate and coarse calcifications in the right breast; two tiny punctate and coarse calcifications in the left breast, without clear mass; no abnormal density foci were seen in the posterior fat layer of the gland. There was no abnormal density foci in the posterior fat layer. The blood vessels were evenly distributed bilaterally. No enlarged lymph nodes were seen in the bilateral axillae.

### Impression:

1. Bilateral breast parenchymal structure: uneven and compact type (ACR type C). (2) A mass in the upper outer quadrant of the right breast is likely to be benign. It is recommended to closely combine with ultrasound. BIRADS 3 (3) Benign calcification of both breasts. BIRADS 2 (4) Bilateral breast hyperplasia. Please follow up with the clinic.

<table>
<thead>
<tr>
<th>Breast</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prob. of mass</td>
<td>0.3483</td>
<td>0.9543</td>
</tr>
</tbody>
</table>