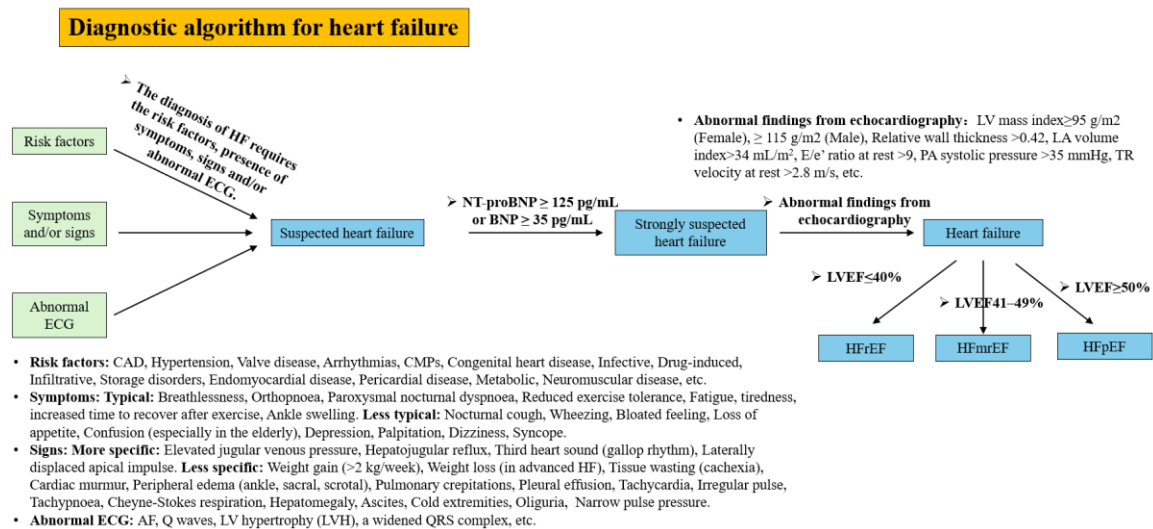


## Task Requirements

Extract the original fields of disease symptoms, descriptions, and values from the clinical notes of the report to infer the given disease. The extraction and inference process needs to follow the given disease diagnosis flowchart. The original file of each report is stored in a text file, and the part that needs to be annotated is stored in a JSON file (we did not provide the source JSON file and text file in DiReCT).

	1	4/8/2024 11:35 AM
	1	4/3/2024 6:04 PM
	1	4/8/2024 1:14 PM
	1	4/3/2024 6:04 PM
	1	4/8/2024 11:03 AM
	1	4/3/2024 6:04 PM

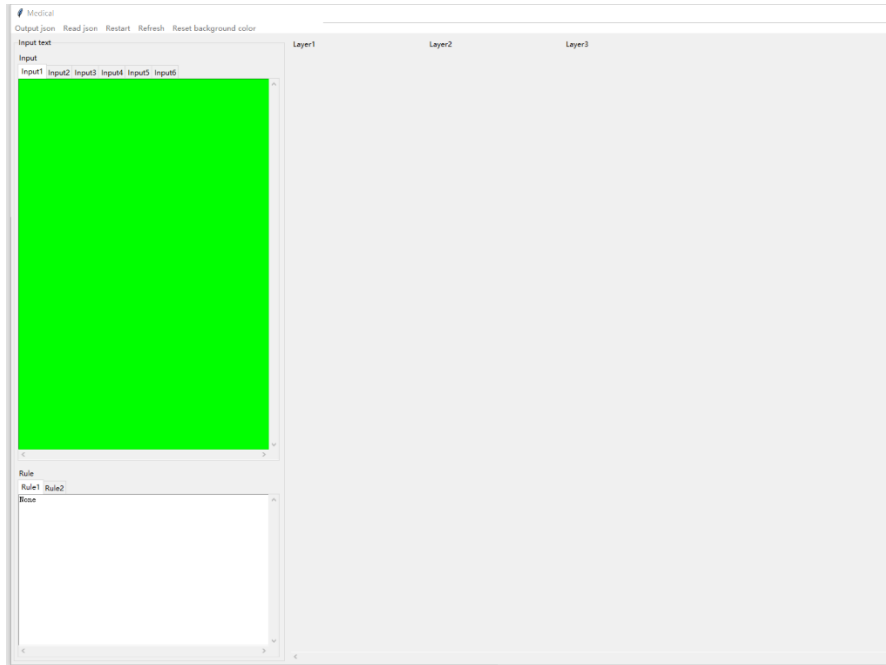
Please refer to the slides attached to each disease for the **diagnostic process** (same as the diagnostic graph shown on GitHub). Below is an example of Heart Failure.



Each blue box here represents each step of the diagnostic status in the disease diagnosis process. They may be suspicious of a certain disease or a large category of diseases. The risk factors, symptoms, and the values or requirements on the arrows in the green boxes are necessary conditions to reach each diagnostic status. This requires you to find the corresponding description in the report as evidence. It is particularly important to note that the content in the green box can only reach the suspicion of a certain disease. The subsequent status often requires some values, images (electrocardiogram), or some test results.

## Annotation Software

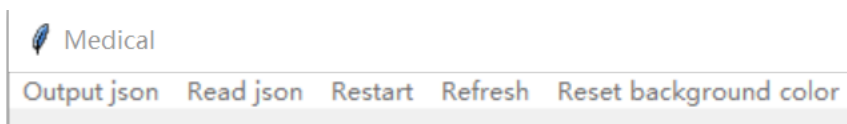
The software does not need to be installed, but it requires a Windows system. It will prompt you to perform a security check at the beginning, please ignore it. The software interface is as follows:



After opening, a config file will be generated. Adjusting the values in it can change the interface structure.

```
[Input]
inputtextnumber = 6
inputrulenumner = 2
inputtextwidth = 60
inputtextheight = 5
inputrule_1_name = Rule 1
inputrule_1 = None
```

By changing the inputtextwidth, you can change the display width of the green part on the left.



The Read json button can read the json file shown in Task Requirements and load the report content.

The Restart button will restart the software and refresh all interfaces.

The Output json button is used to save the annotated data.

Input

Input1

Input2

Input3

Input4

Input5

Input6

Admission Exam:  
VITALS: 97.5 PO 127 / 65 51 18 100 RA  
EYES: Anicteric, pupils equally round  
ENT: Ears and nose without visible erythema, masses, or trauma. Oropharynx without visible lesion, erythema or exudate. Mucous membranes moist  
CV: Heart regular, no murmur  
RESP: Lungs clear to auscultation with good air movement bilaterally  
GI: Abdomen soft, \_\_\_ to palpation  
MSK: Neck supple, moves all extremities, strength grossly full and symmetric bilaterally in all limbs  
SKIN: No rashes or ulcerations noted  
NEURO: Alert, oriented, face symmetric, gaze conjugate with EOMI, speech fluent, moves all limbs  
PSYCH: pleasant, appropriate affect

Input1-6 represents some records in the report. Click read json and read the file to automatically display the content. They are:

Input1: Chief Complaint

Input2: History of Present Illness

Input3: Past Medical History

Input4: Family History

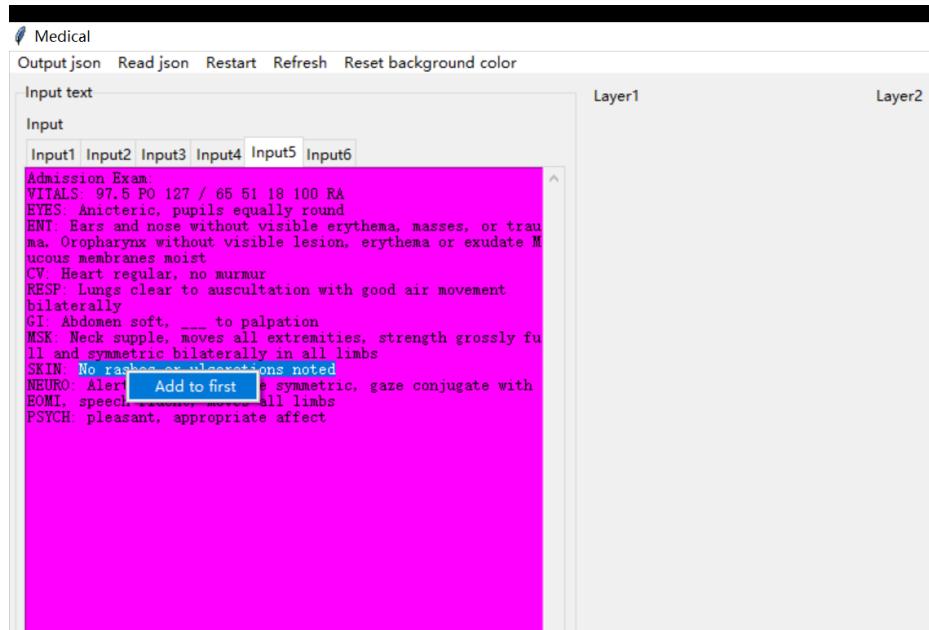
Input5: Physical Exam

Input6: Pertinent Results

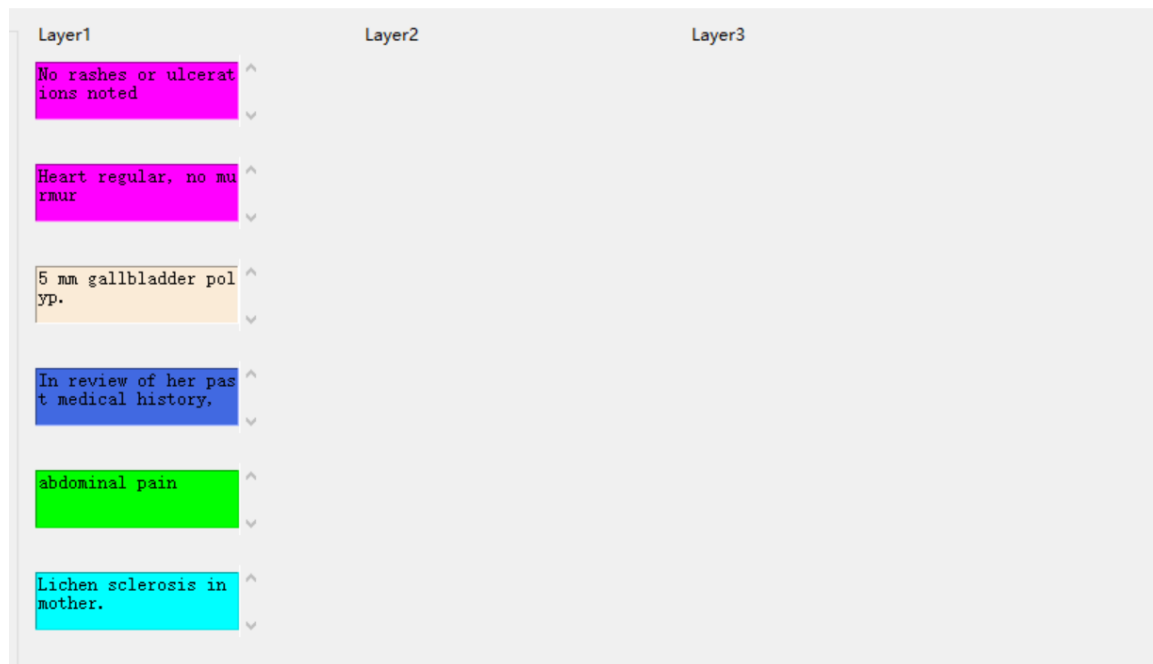
We manually copied all the required data from the raw text file and saved them in the JSON file for annotation.

## Data Annotation

First, read input1-input6 in sequence and find the original text description (can be a paragraph, a word, or a sentence) that causes the given disease as evidence. Hold down the left mouse button and drag to select the text you need, then right-click Add to first.



At this time, the layer1 on the right will show the content you selected

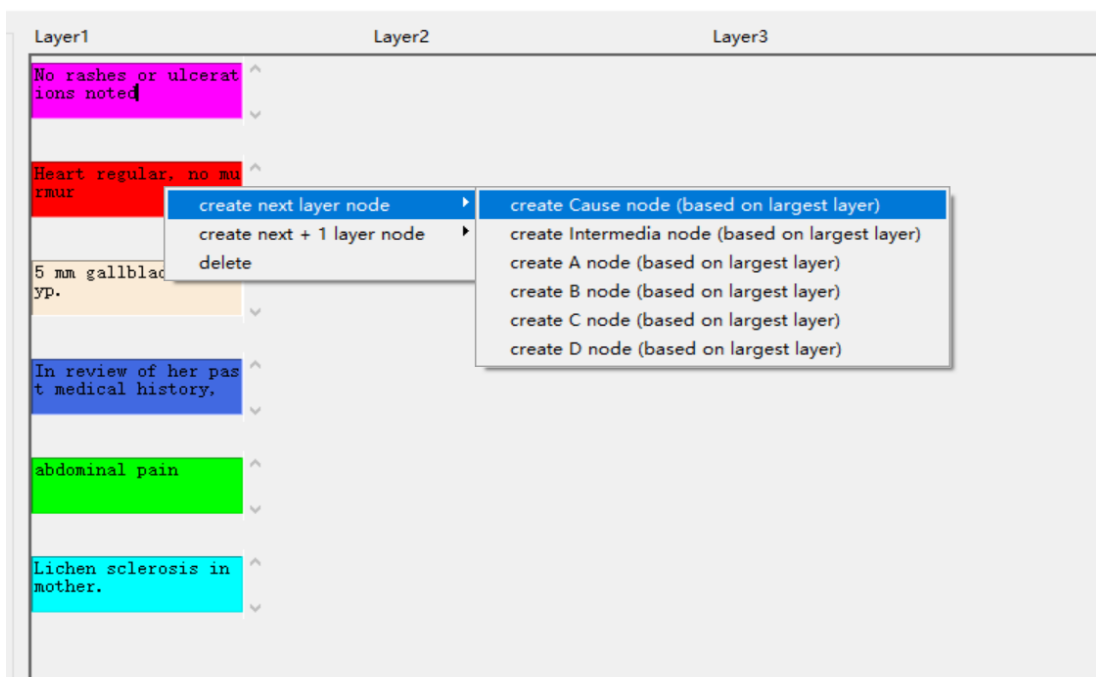


In this step, please try to find the text or numerical description that can achieve each diagnostic state as comprehensively as possible to facilitate the subsequent annotation. In

addition, it is recommended to add the evidence of the diagnostic state on the left side of the diagnostic diagram first.

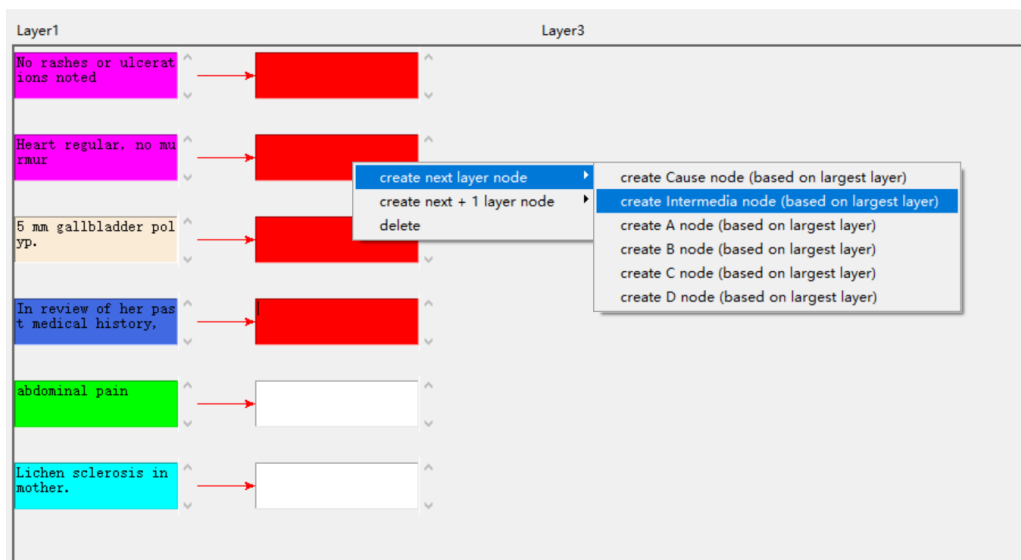
Next, please infer the diagnostic chain based on the evidence you found. You also need to give your explanation for each step of reasoning, which comes from the diagnostic flowchart (that is, explain why a piece of evidence can reach a state). Right-click a box and click delete to delete it.

First, click to select evidence (or multiple pieces of evidence if a diagnostic state requires multiple pieces of evidence to exist at the same time). After selecting, it will turn red (click again to cancel the selection). Then right-click and select create next layer node, and then select create cause node. This will add a new white box to fill in the reason.

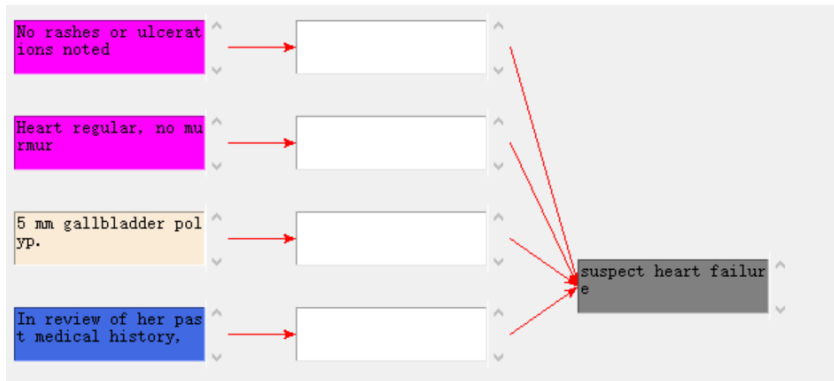




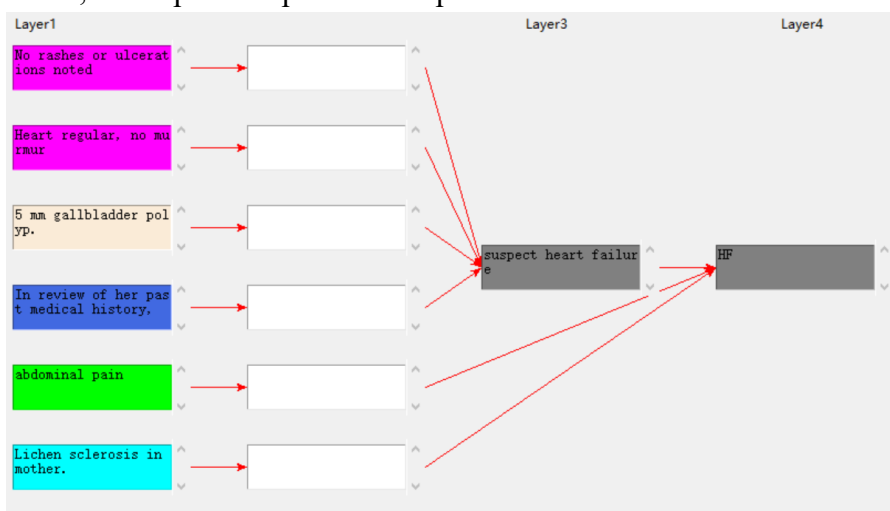
Due to the limitation of the software, the degree of freedom is low. It is recommended to add the reasons for the corresponding diagnosis status to all evidences first. When adding, add them in order from top to bottom.



Then, add the first layer of diagnostic states (mostly suspect) according to the diagnostic map, and click to select all the reasons that can reach this state. Next, right-click on a selection to create next layer node, and select create intermedia node. A gray box will appear, and you need to fill in the name of this state (consistent with the diagnostic map).



Next, continue to deduce for the given disease, select evidence and current diagnostic status, and repeat the previous step.



Please repeat the above steps until the given disease state is derived.

Pay attention to finding all the evidence that is useful for deriving each disease state as much as possible. But you should strictly follow the provided **diagnostic process** for the corresponding disease. If you find an observation that is not recorded in the diagnostic process please add it to the slide.

Pay attention to the fact that when looking for evidence, the unit selected for the text description in the clinical note should be at least a complete statement, and for the numerical value, you need to choose the numerical value and the name before the numerical value, such as proBNP-3843. Evidence similar to the image below only requires the sentence after the colon, and the title "-CXR\_\_" does not need to be included.

- CXR \_\_: "New mild-to-moderate interstitial abnormality, most suggestive of congestive heart failure."

Pay attention to the fact that if there is no evidence for a certain diagnosis status in the derivation, please create a piece of evidence according to the status requirements of the **diagnosis process** and add it to the appropriate part of the input.

Pay attention to the fact that one piece of evidence can only point to one state, and the state closer to the right of the diagnosis diagram has a higher priority. That is, if a piece of evidence points to Heart Failure, it can no longer point to Suspected Heart Failure.

Pay attention to some information that will leak the final diagnosis of the disease, such as some of the doctor's own inferences, and directly appear the name of the final diagnosis. Please remove them.

After the annotation is completed, the last step is to save. Please click Output json in the upper left corner to save the file to a new folder, and keep the file name consistent with the json file read in.