Sample input for BioBert Model:

| Sample No | Sentences | Tags |
| --- | --- | --- |
| **1** | patients her2 hr positive mbc labc randomized ... | Intervention and Comparison (element I and C) |
| **2** | statistical significance defined p-value 0 05 | Not Relevant to the Evidence (Label N) |
| **3** | exclusion criteria included nephrotic proteinu... | Population/Problem (element P) |
| **4** | radial artery cannulation performed summarise ... | Intervention and Comparison (element I and C |

The Biobert model is a biomedical language representation model designed for biomedical text mining tasks such as biomedical named entity recognition, relation extraction, question answering, etc. So naturally it performs better than the regular Bert model.

Now, the model is able to predict if the sentence is among PICO labels. This is better as the context of the sentence would be considered and better prediction is made and is essentially a classification task on the sentences.

Whereas, the PICOBert model classifies each word in a sentence and hence cannot really well identify the PICO labels.

Currently creating a dataset in specific formats so that the model can be tested on the custom dataset and we can create a performance test to test different mode's accuracies and compare them.