Some meeting notes and conclusion in the Yesterday’s meeting (Euma, Marta and Kyle correct me if I am wrong anywhere):

**For the Echo Notes**:

We are looking at the differences between patients who have CHF with the patients who don’t, given that this patient had developed Sepsis during the ICU stay (From the Angus script). The Angus script in summary - pulls out patients who developed sepsis during the ICU stay(/hospital stay – NEED to CONFIRM) based on the ICD codes during the discharge.

* Look at the patients with admission notes which consist of the keywords listed below. The note is considered admission note if the description column of that note in the noteevents table consists of the word ‘admission’(all cases). In other words if the description contains *admission* then the note is considered to be an admission note.
  + HFrEF
  + heart failure with reduced ejection
  + heart failure with preserved ejection fraction
  + HFpEF
  + Congestive heart failure
  + diastolic heart failure
  + systolic heart failure
  + CHF
* I will come up with a similarity metric on a scale of 0-1 (this can be implemented based on rule-based Levinstein distance (aka edit distance) or similarity-based metric using word vectors/embeddings) or other technique. I will give Kyle 5 sample reports for each decimal (for example for 0.4 - 5 documents and for 0.5 - 5 documents) to select the threshold we set for the similarity score of the documents. Based on this threshold we would consider that a patient was thought as having CHF(in the hospital staff opinion) or not.

**Doubts:**

* If a patient had sepsis in the past (before the admission to the hospital) and does not develop any sepsis during the ICU stay or the hospital stay. Would this patient have ICD code of sepsis during the discharge.

**ToDo:**

* Go through the Angus script.
* SQL table
* Read the documents Euma sent me

**See the facebook post for the similarity**

1. See count of hamd\_ids which admission note
2. See the admission description also for the CHF symptoms

* Check Euma Paper
* Euma Slack inputs
* Get more patients
* Get more criterion and take more terms from the paper
* Check the intersection with Sepsis Patients from the script Euma gave or Marta sent email initiallally or from the slack

**What is p-value in statistics**