**Date**: 8/30/2017

**Attendees**: Helen, Asiyah, Chul Ahn, Yage and Zheng

**Discussed**:

1. Analyzed quality control result, develop method to trade off false positive and false negative [Copy of 2\_Quality\_Control\_Data\_Asiyah.xlsx]
2. Reviewed updated word frequency and add new item to keyword list [3\_Keywords\_Word\_Frequency\_0830.xlsx]
3. Discussed solutions to conflicted records, some records need manually check in the future. [ 4\_Conflicted\_Records.xlsx]

**Action items**:

Yage & Zheng:

1. Update text mining method to identify different terms of keywords (e.g. breast pain -> painful breast)
2. Develop method to identify duplicate records
3. Develop method to deal with word – suspect, confirm, negation
4. Develop method to solve some text cell which is not split by space
5. Develop template for dictionary files
6. Update conflicted records, add product code in conflicted records
7. Update the word frequency based on updated methods and new keywords.
8. [Future] Detect journal articles from text content
9. [Future] To develop statistical analyses model, e.g. logistic regression for association analysis.

Helen & Asiyah:

1. Decide the tradeoff between false positive and false negative for each keyword sub-list
2. Update keywords list and send to Zheng & Yage (e.g. siltext for surface type)
3. Develop ontology and rules to combines features (e.g. CD30+ & CD30 positive should be combined)
4. Update manufacturer standardized name dictionary and send to Zheng & Yage
5. Offer journal markers of MAUDE database (the identifier of whether the text is from a journal) and send to Zheng & Yage

[Meeting adjourned.]