FDA BI Research Weekly Report

Yage Wang, Zheng Gong 20170716

**Question 1. How many BI records are found?**

Answer: 26,584 BI related records are found. 25,787 records have matched report number. [File: 1\_Final\_BI\_Number\_List.txt]

Method 1: Use text mining to identify

Data source: MAUDE database [foitext.zip, foitextadd.zip, foitextchange.zip, foitextthru1995.zip, foitext(1996-2016).zip] (25 files in total), use column “FOI\_TEXT”

Word List 1: ['breast pain', 'breast swelling', 'breast cyst', 'breast calcification', 'capsular contracture', 'lymph node enlargement', 'firmness of breast', 'breast implant', 'breast augmentation', 'breast enlargement', 'mammary implant', 'mammary augmentation', 'mammary enlargement', 'mammary swelling', 'breast prosthesis', 'mammary prosthesis']

Word List 2: ['hematoma', 'mass', 'lump', 'rupture', 'deflated', 'infection', 'abscess', 'leukopenia', 'nodules', 'skin discoloration', 'skin lesion', 'seroma', 'effusion', 'fluid']

Fuzzy Match Rate: 95%

If one record contains at least one word in Word List 1, or it contains “breast/mammary” **and** at least one word in Word List 2, then it would be considered as BI related record.

In this method, 10,981 BI related records are found.

Method 2: Use product code to identify

Data source: MAUDE database [foidev.zip, foidevadd.zip, foidevchange.zip, foidevthru1997.zip, foidev(1998-2016).zip] (23 files in total), use column “DEVICE\_RECORD\_PRODUCT\_CODE”

Product Code List: [‘FWM’, ‘FTR’]

If one record whose product code equals to one of the list, it would be considered as BI related record.

In this method, 26,604 BI related records are found. (FWM = 11097, FTR = 15507)

Combination of two methods

The union of two methods has 27,475 distinct BI records.

If time criteria set to 01/01/1994 – 02/01/2017, 26,584 records are found. (The data from MAUDE are from 1992 to 2017/06)

Finally, 25787 records have matched report number.

**Question 2. How many ALCL records are found?**

Answer: 197 ALCL related records are found. 121 records have matched report number. [File: 2\_ALCL\_Number\_List.txt]

151 records are included in BI records. 105 records have matched report number. [File: 3\_ALCL\_BI\_Number\_List.txt]

Method: Use text mining to identify

Data source: MAUDE database [foitext.zip, foitextadd.zip, foitextchange.zip, foitextthru1995.zip, foitext(1996-2016).zip] (25 files in total), use column “FOI\_TEXT”

Word List: [‘ALCL’, ‘anaplastic large cell lymphoma’]

Fuzzy Match Rate: 90%

If one record contains at least one word in Word List, it would be considered as ALCL related record.

In this method, 197 ALCL records are found. 151 of 197 records are included in Question 1 BI records.

**Question 3. Explain tables and columns used in Q1 and Q2.**

There are two kinds of tables in MAUDE database: “foidev” and “foitext”.

“foidev” contains 45 columns before 2008, 28 columns after 2009. Please see [File: 4\_FDA\_BI\_Table\_Column.xlsx]

**Question 4. Present the variety of Brand Name and Model Number in BI records**

For brand name, 3,280 distinct values are found out of 25,843 names (not null).

For model name, 1,795 distinct values are found out of 18,215 names (not null).

The distribution of all values could be finished in 1 or 2 days if needed.

1. Separated list of text mining and product code (today)
2. Distribution
3. Mdrfoi statistic report
4. word frequency analysis of method 2 (product code)
5. due: Monday 9 am