Aspects of labels data to consider

This document contains summary of the content in the different label files. These files contain all possible outcome variables we want to predict against. There are in total 3 tables, listed below.

Table1:

DIAG and PROZ:

Description:

Contains all diagnosis/procedures and dates diagnosis/procedures were given on two different granular levels.

Fields:

1. **Datum_der_Diagnose_DAT (Many per DKAT (can give same diagnose many times)**
2. **Katalog_des_Diagnosecode_DKAT (One per diagnose date)**
3. **Diagnosecode_DKEY (Many per DKAT)**
4. **Lokalisation_des_Diagnose_LOK (Many per DKEY)**
5. Date_from_Pr (**Many per LOK (can give same prozedure many times)**
6. **Lokalisation_des_Prozedure_LOK (many per**
7. **Prozedure_code_ICPMK (many per DKEY date)**
8. **ID_des_OP_Katalog_ICPML (many per ICPMK)**

Outcome variables:

DKAT, DKEY; LOK des DKEY, ICPMK, ICPML, LOK des ICPMK

Comments:

1. Study date and diagnosis date not always same (will not align with xml data)
2. One study date -> one diagnosis, later study on same patient lead to different diagnosis (how should we treat the images as data records?)
3. One patient -> one-many study -> one-one (I think) Diagnosis date-> one-many DKAT -> many-many DKEY
4. One patient – One-many study -> one-one Proz date-> one - many ICPMK -> many -many ID des ICPMK

Table2:

IOL:

Description:

Contains for every patient the eye length for each eye

Fields:

1. **eye_length_AL (many per patient)**

2. **laterality_AL (many per patient)**

Comments:

1. Should we use as feature or outcome variable?
2. How to handle missing values´

Outcome variables:

1. **eye_length_AL, laterality_AL**

Table3:

VISUS

Description:

Contains the information about every patients visual acuity. Current and original to track progression as well as how it was measured. Per measurement it tracks measure date and which eye is measured.

Fields:

2. **laterality_AUGE**
3. **date_of_messe_DAT**
4. **visual_acuity_VISUS**
5. **messart_visus_VISART**
6. **usprüngliche_visus_angabe_VISORIG**

Comments:

3. Should we use as feature or outcome variable?
4. How to handle missing values
5. Completely separate prediction task?

Outcome variables:

-