**Data Instruction Document for BIDMC and MGH Note Sets**

\* Note: some of these files were not de-identified at the time of posting and are therefore not included. De-identified versions of these may be added at a later date.

## bidmc\_984

Description: BIDMC test set of notes. Contains 50% positive samples.

## bidmc\_2296

Description: BIDMC training set of notes.

## bidmc\_3280

Description: Complete set of all BIDMC notes.

## mgh\_680

Description: MGH test set of notes. Contains 50% positive samples.

## mgh\_1652

Description: MGH training set of notes.

## mgh\_2332

Description: Complete set of all MGH notes.

## merged\_mgh\_bidmc222

Description: Complete set of all MGH and BIDMC notes combined.

## mgh\_bidmc\_1664

Description: MGH + BIDMC test set of notes. Contains 50% positive samples.

## mgh\_bidmc\_3948

Description: MGH + BIDMC training set of notes.

## negative\_coefficients

Description: Negative coefficients from the MGH + BIDMC trained model.

## positive\_coefficients

Description: Positive coefficients from the MGH + BIDMC trained model.