

# MatchSubjects

Michael Caiola (Michael.Caiola@fda.hhs.gov) and Meijun Ye (Meijun.Ye@fda.hhs.gov)

Matching HEA subjects to gender and then age

```
function [HEA2,STR2,TBI] = MatchSubjects()
```

First, open up master Excel sheet

```
load("PredictionDatabase.mat","T");
T.Sex = categorical(T.Sex);
TBI = T(T.Category == "TBI",:);
HEA = T(T.Category == "Normal",:);
HEA_ind = 1:height(HEA);
match_ind = [];
age_diff = [];
for i = 1:height(TBI)
    Q = HEA(HEA_ind,:);
    switch TBI.Sex(i)
        case "female"
            Q = Q(Q.Sex=="female",:);
        case "male"
            Q = Q(Q.Sex=="male",:);
    end
    Q.Match = abs(Q.Age-TBI.Age(i));
    if min(Q.Match) > 5
        continue
    end
    age_diff(end+1) = min(Q.Match);
    in = find(Q.Match == age_diff(end));
    in = in(randi(length(in))); %randomize if mulitple matches
    match_ind(end+1) = Q.FileID(in);
    HEA_ind(HEA_ind == find(HEA.FileID==match_ind(end))) = [];
end
HEA2 = HEA(ismember(HEA.FileID,match_ind),:);
```

```
STR = T(T.Category == "Stroke",:);
STR_ind = 1:height(STR);
match_ind = [];
age_diff = [];
for i = 1:height(TBI)
    Q = STR(STR_ind,:);
    switch TBI.Sex(i)
        case "female"
            Q = Q(Q.Sex=="female",:);
        case "male"
            Q = Q(Q.Sex=="male",:);
    end
```

```

Q.Match = abs(Q.Age-TBI.Age(i));
if min(Q.Match) > 5
    continue
end
age_diff(end+1) = min(Q.Match);
in = find(Q.Match == age_diff(end));
in = in(randi(length(in))); %randomize if mulitple matches
match_ind(end+1) = Q.FileID(in);
STR_ind(STR_ind == find(STR.FileID==match_ind(end))) = [];
end
STR2 = STR(ismember(STR.FileID,match_ind),:);

```

## Output

```

disp("# HEA subjects/sessions: " + numel(unique(HEA2.Subject)) + "/" + height(HEA2))
disp("# STR subjects/sessions: " + numel(unique(STR2.Subject)) + "/" + height(STR2))
disp("# TBI subjects/sessions: " + numel(unique(TBI.Subject)) + "/" + height(TBI))
figure;
ages = 18:2:65;
histogram(HEA2.Age,ages);
hold on
histogram(STR2.Age,ages);
histogram(TBI.Age,ages);
legend("HEA","STR","TBI")

```

## Disclaimer

This software and documentation (the "Software") were developed at the Food and Drug Administration (FDA) by employees of the Federal Government in the course of their official duties. Pursuant to Title 17, Section 105 of the United States Code, this work is not subject to copyright protection and is in the public domain. Permission is hereby granted, free of charge, to any person obtaining a copy of the Software, to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, or sell copies of the Software or derivatives, and to permit persons to whom the Software is furnished to do so. FDA assumes no responsibility whatsoever for use by other parties of the Software, its source code, documentation or compiled executables, and makes no guarantees, expressed or implied, about its quality, reliability, or any other characteristic. Further, use of this code in no way implies endorsement by the FDA or confers any advantage in regulatory decisions. Although this software can be redistributed and/or modified freely, we ask that any derivative works bear some notice that they are derived from it, and any modified versions bear some notice that they have been modified.