

Longitudinal Analysis

Francesco Beghini, Jackson Pullman, Marcus Alexander, Shivkumar Vishnempet Shridhar,
Drew Prinster, Adarsh Singh, Rigoberto Matute Juárez,
Edoardo M. Airoidi, Ilana L. Brito, Nicholas A. Christakis

First regression model

Model the strain-sharing rate at timepoint 2 using the Mahalanobis distance as covariate

```
lmer_first_model <- lmer(T2 ~ T1 + relationship + maha_dist + (1|village_code_ego) + (1|ego),  
                          sharing_rate_only_followup)  
summary(lmer_first_model)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [  
## lmerModLmerTest]  
## Formula: T2 ~ T1 + relationship + maha_dist + (1 | village_code_ego) +  
##      (1 | ego)  
##      Data: sharing_rate_only_followup  
##  
## REML criterion at convergence: 37840.5  
##  
## Scaled residuals:  
##      Min      1Q  Median      3Q      Max  
## -2.6204 -0.6339 -0.0743  0.5078 19.8980  
##  
## Random effects:  
##      Groups      Name      Variance Std.Dev.  
##      ego      (Intercept) 0.159680 0.39960  
##      village_code_ego (Intercept) 0.008861 0.09413  
##      Residual      0.794220 0.89119  
## Number of obs: 14268, groups:  ego, 290; village_code_ego, 4  
##  
## Fixed effects:  
##              Estimate Std. Error      df t value Pr(>|t|)  
## (Intercept)   4.357e-02  5.621e-02  2.532e+00   0.775  0.50413  
## T1            1.034e-01  8.467e-03  1.388e+04  12.206 < 2e-16 ***  
## relationshipes 2.525e-01  3.686e-02  1.413e+04   6.852  7.6e-12 ***  
## maha_dist     -3.132e-02  9.171e-03  1.407e+04  -3.415  0.00064 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Correlation of Fixed Effects:  
##              (Intr) T1      rltnsh
```

```
## T1          -0.014
## relatnshpys -0.046 -0.220
## maha_dist    0.017  0.022  0.066
```

Model the strain-sharing rate at timepoint 2 using the separate sociodemographic variables

```
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient

## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: T2 ~ T1 + relationship + gender.mm + gender.ff + indigenous.both +
##   indigenous.one + age_difference_abs + average_age + DDS_average +
##   DDS_difference_abs + education_difference_abs + average_education +
##   religion_same + building_same + wealth_difference_abs + average_wealth +
##   average_bristol + different_bristol + watersource_same +
##   same_usage_painkillers + same_usage_antibiotics + same_usage_antidiarrheal +
##   same_usage_antiparasite + same_usage_vitamins + same_usage_zinc +
##   same_usage_antifungal + (1 | village_code_ego) + (1 | ego)
## Data: sharing_rate_only_followup
##
## REML criterion at convergence: 30672.6
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.9751 -0.6339 -0.0783  0.5132 19.6027
##
## Random effects:
##   Groups             Name                Variance Std.Dev.
##   ego                (Intercept) 0.16725   0.4090
##   village_code_ego    (Intercept) 0.03589   0.1894
##   Residual                        0.76022   0.8719
## Number of obs: 11692, groups:  ego, 261; village_code_ego, 4
##
## Fixed effects:
##
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)    5.304e-01  1.657e-01  1.896e+01   3.200 0.004719
## T1              7.194e-02  9.327e-03  1.163e+04   7.713 1.33e-14
## relationshipyes  1.163e-01  4.377e-02  1.156e+04   2.657 0.007893
## gender.mmTRUE   -3.921e-02  3.090e-02  1.152e+04  -1.269 0.204437
## gender.ffTRUE    7.732e-02  1.984e-02  1.146e+04   3.897 9.79e-05
## indigenous.bothTRUE  1.765e-01  4.256e-02  3.325e+03   4.148 3.44e-05
## indigenous.oneTRUE  8.106e-02  3.564e-02  4.318e+03   2.274 0.022996
## age_difference_abs  1.344e-03  7.611e-04  1.165e+04   1.766 0.077490
## average_age     -1.895e-03  1.391e-03  1.037e+04  -1.362 0.173203
## DDS_average      2.023e-02  1.264e-02  9.867e+03   1.600 0.109595
## DDS_difference_abs -2.891e-02  8.339e-03  1.160e+04  -3.467 0.000528
## education_difference_abs -1.570e-02  4.557e-03  1.162e+04  -3.444 0.000575
## average_education -5.238e-02  8.775e-03  9.485e+03  -5.970 2.46e-09
## religion_sameTRUE -1.297e-02  2.334e-02  1.083e+04  -0.556 0.578543
## building_sameTRUE  5.156e-01  8.148e-02  1.155e+04   6.327 2.59e-10
## wealth_difference_abs -6.606e-03  9.017e-03  1.165e+04  -0.733 0.463814
## average_wealth    5.693e-02  1.516e-02  9.752e+03   3.756 0.000174
## average_bristol    7.538e-02  5.656e-02  1.164e+04   1.333 0.182664
## different_bristol  1.181e-02  7.217e-03  1.166e+04   1.637 0.101716
## watersource_sameTRUE  7.761e-02  2.139e-02  1.127e+04   3.628 0.000287
## same_usage_painkillersTRUE  2.792e-02  1.645e-02  1.151e+04   1.698 0.089605
## same_usage_antibioticsTRUE  1.832e-02  2.802e-02  1.104e+04   0.654 0.513258
```

```

## same_usage_antidiarrhealTRUE -2.790e-01 4.297e-02 9.372e+03 -6.494 8.81e-11
## same_usage_antiparasiteTRUE -1.994e-03 4.352e-02 9.908e+03 -0.046 0.963447
## same_usage_vitaminsTRUE -6.009e-02 2.205e-02 1.135e+04 -2.725 0.006436
## same_usage_antifungalTRUE -3.311e-01 7.063e-02 1.105e+04 -4.688 2.79e-06
##
## (Intercept) **
## T1 ***
## relationshipyes **
## gender.mmTRUE
## gender.ffTRUE ***
## indigenous.bothTRUE ***
## indigenous.oneTRUE *
## age_difference_abs .
## average_age
## DDS_average
## DDS_difference_abs ***
## education_difference_abs ***
## average_education ***
## religion_sameTRUE
## building_sameTRUE ***
## wealth_difference_abs
## average_wealth ***
## average_bristol
## different_bristol
## watersource_sameTRUE ***
## same_usage_painkillersTRUE .
## same_usage_antibioticsTRUE
## same_usage_antidiarrhealTRUE ***
## same_usage_antiparasiteTRUE
## same_usage_vitaminsTRUE **
## same_usage_antifungalTRUE ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## Correlation matrix not shown by default, as p = 26 > 12.
## Use print(x, correlation=TRUE) or
##      vcov(x)      if you need it

## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient

```

Second regression model

Model each species sharing status at timepoint 2 using the Mahalanobis distance as covariate

```
lmer_species_rand <- glmer(data = xcov, T2_t ~ T1_t + relationship + maha_dist + (1|species) + (1|village_code_ego), data = xcov, family = binomial)
summary(lmer_species_rand)
```

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 0) [glmerMod]
##   Family: binomial ( logit )
## Formula: T2_t ~ T1_t + relationship + maha_dist + (1 | species) + (1 |
##   village_code_ego) + (1 | ego)
##   Data: xcov
##
##           AIC          BIC    logLik deviance df.resid
##  94239.5   94312.0 -47112.8   94225.5    229778
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.6126 -0.2635 -0.1923 -0.1330  18.6394
##
## Random effects:
##   Groups                Name                Variance Std.Dev.
##   species                (Intercept)  0.730136  0.85448
##   ego                    (Intercept)  0.207429  0.45544
##   village_code_ego (Intercept)  0.006726  0.08201
## Number of obs: 229785, groups: species, 544; ego, 294; village_code_ego, 4
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -3.107451   0.076908 -40.405 < 2e-16 ***
## T1_tTRUE       1.963549   0.026470  74.180 < 2e-16 ***
## relationshipTRUE 0.287753   0.038224   7.528 5.15e-14 ***
## maha_dist     -0.001421   0.001223  -1.162   0.245
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) T1_TRU rlTRUE
## T1_tTRUE      -0.064
## rltnshpTRUE  -0.075 -0.046
## maha_dist     -0.262  0.008  0.098
```