# Mediation

*Adam Garber*

Norwegian University of Science and Technology - A Course in `MplusAutomation`

June 01, 2021

---

## Lab preparation

---

## Load packages

```
library(mediation)
library(tidyverse)
library(MplusAutomation)
library(here)
library(gt)
library(gtsummary)
```

---

## Lab outline

1. Estimate a mediation model using the {`mediation`} package
2. Estimate the same model using the Structural Equation Modeling (SEM) framework with {`MplusAutomation`}
3. For the second empirical example, estimate parallel models using the `mediation` and `SEM` methods

---

## A quick detour - Equivalent models

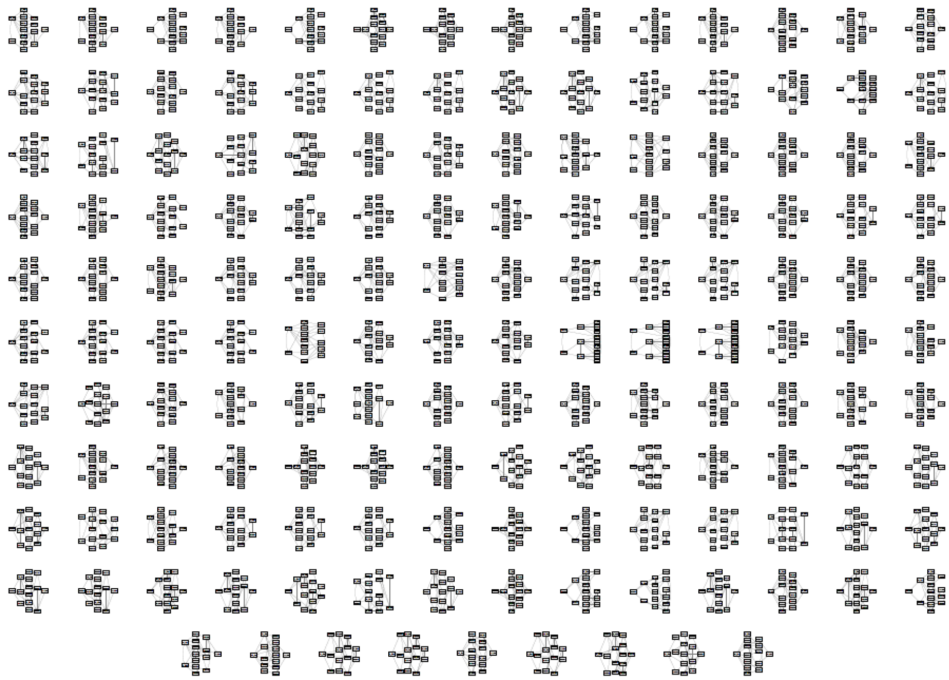


*Figure.* Picture adapted from SEM slides by Sacha Epskamp http://sachaepskamp.com/files/SEM22019/SEM2_2019_Week2_slides.pdf

---

## The empirical examples of mediation used in this exercise are from the following article

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R package for causal mediation analysis.

https://cran.r-project.org/web/packages/mediation/vignettes/mediation.pdf

---

## Data source for example 1

Brader T, Valentino NA, Suhat E (2008). **What Triggers Public Opposition to Immigration? Anxiety, Group Cues, and Immigration.** American Journal of Political Science, 52(4), 959–978.

https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-5907.2008.00353.x

To see metadata run - `?framing`

---

Read in the `framing` dataset

```
set.seed(4212020)

data("framing", package = "mediation")

framing <- droplevels(framing) %>%  # drop factor levels with frequency zero
  mutate(emo = emo - 2)
```

---

Take a look at variables used in the mediation model

| Name | Labels |
|---|---|
| emo | Measure of subjects' negative feeling during the experiment (1-10). 1 indicates the most negative feeling. |
| treat | Framing codition interaction term. News story with conditions tone (Negative/Positive) and ethnic identity |
| cong_mesg | Whether subjects requested sending an anti-immigration message to Congress on their behalf. |
| age | Age of subject (18-85) |
| educ | Education (1-4) |
| gender | Gender (Male/Female) |
| income | Subjects' income, measured as a 19-point scale. |

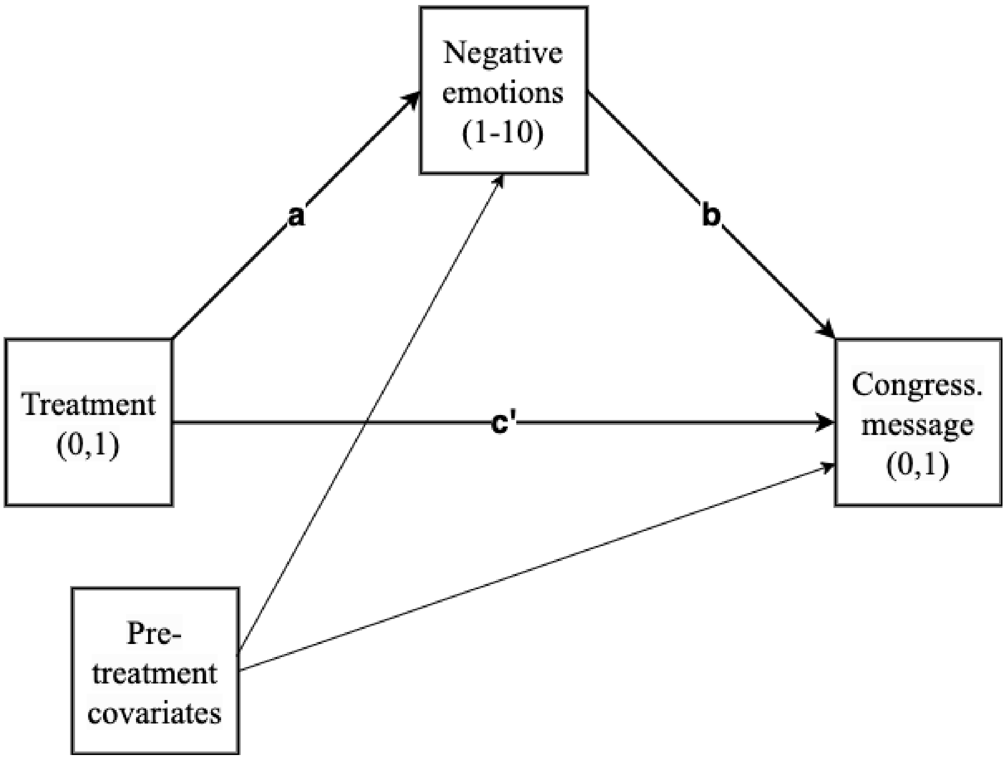
---

Look at descriptives table for the framing dataset using {`gtsummary`}

```
table_data <- framing %>%
  select(emo, treat, cong_mesg, age, educ, gender, income)

table1 <- tbl_summary(table_data,
                      statistic = list(all_continuous() ~ "{mean} ({sd})"),
                      missing = "no") %>%
  bold_labels()

table1
```

| **Characteristic** | **N = 265** |
|---|---|
| **emo** | 4.97 (2.77) |
| **treat** | 68 (26%) |
| **cong_mesg** | 88 (33%) |
| **age** | 48 (16) |
| **educ** | |
| less than high school | 20 (7.5%) |
| high school | 92 (35%) |
| some college | 70 (26%) |
| bachelor's degree or higher | 83 (31%) |
| **gender** | |
| male | 126 (48%) |
| female | 139 (52%) |
| **income** | 11 (4) |

---




---

### Estimate a mediation model in R using {`mediation`}

step 1: fit a linear model of the mediator (`emo`) regressed on treatment (`treat`) and pre-treatment covariates

```
med_fit <- lm(emo ~ treat + age + educ + gender + income,
              data = framing)
```

---

step 2: fit a general linear model (`glm`) with the binary outcome variable `cong_mesg` regressed on treatment (`treat`), mediator, and pre-treatment covariates

```
out_fit <- glm(cong_mesg ~ emo + treat + age + educ + gender + income,
               data = framing,
               family = binomial("probit"))
```

---

step 3: estimate the mediation effects with bias corrected bootstrapped confidence intervals

```r
med_out <- mediate(med_fit, out_fit, treat = "treat", mediator = "emo",
                   boot = TRUE, boot.ci.type ="bca", sims = 100)

summary(med_out)
```

```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the BCa Method
##
##                          Estimate 95% CI Lower 95% CI Upper p-value
## ACME (control)             0.0812       0.0238         0.13  <2e-16 ***
## ACME (treated)             0.0822       0.0238         0.14  <2e-16 ***
## ADE (control)              0.0112      -0.0921         0.12    0.72
## ADE (treated)              0.0122      -0.1031         0.13    0.72
## Total Effect               0.0934      -0.0200         0.26    0.16
## Prop. Mediated (control)   0.8698     420.8309       668.91    0.16
## Prop. Mediated (treated)   0.8804     367.5789       584.02    0.16
## ACME (average)             0.0817       0.0228         0.13  <2e-16 ***
## ADE (average)              0.0117      -0.0976         0.13    0.72
## Prop. Mediated (average)   0.8751     394.2049       626.46    0.16
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 265
##
##
## Simulations: 100
```

---

### Run mediation model 1 using the Structural Equation Modeling framework with {MplusAutomation}

---

```r
m1_mediate  <- mplusObject(
  TITLE = "m1 mediate framing",
  VARIABLE =
   "usevar =
    cong_mesg emo treat age
    educ gender income;

    categorical = cong_mesg; ! outcome is binary",

  ANALYSIS =  "bootstrap = 500; ! set number of bootstrap samples (500 for example purposes)" ,

  MODEL =
   "emo on treat age educ gender income;            ! mediator linear regression
    cong_mesg on emo treat age educ gender income;  ! outcome GLM regression
```

```
    Model indirect:
    cong_mesg ind treat;" ,

  OUTPUT =
    "sampstat standardized cinterval (bcbootstrap); ! bias-corrected bootstrap",

  PLOT = "type=plot2;",

  usevariables = colnames(framing),
  rdata = framing)

m1_mediate_fit <- mplusModeler(m1_mediate,
                    dataout=here("15-mediation", "mplus_files", "framing.dat"),
                    modelout=here("15-mediation", "mplus_files", "m1_mediate_framing.inp"),
                    check=TRUE, run = TRUE, hashfilename = FALSE)
```

---

**Model 1 Mplus output**

```
STANDARDIZED TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS

                    Estimate       S.E.  Est./S.E.    P-Value

Effects from TREAT to CONG_MES

  Total                0.109      0.075      1.453      0.146
  Total indirect       0.101      0.031      3.253      0.001

  Specific indirect 1
    CONG_MES
    EMO
    TREAT              0.101      0.031      3.253      0.001

  Direct
    CONG_MES
    TREAT              0.008      0.071      0.119      0.905
```

---

## Data source for example 2

Vinokur AD, Price RH, Schul Y (1995). **Impact of the JOBS Intervention on Unemployed Workers Varying in Risk for Depression.** American Journal of Community Psychology, 23(1), 39–74.

https://link.springer.com/content/pdf/10.1007/BF02506922.pdf

To see metadata run - `?jobs`

---

*Note:* For this example we will ignore the issue of non-compliance addressed in Tingley et al. (2014) as this causal inference topic is beyond the scope of this course.

Read in the data from the job search intervention study (`jobs`)

```
data("jobs", package = "mediation")
```

Take a look at variables used in the mediation model

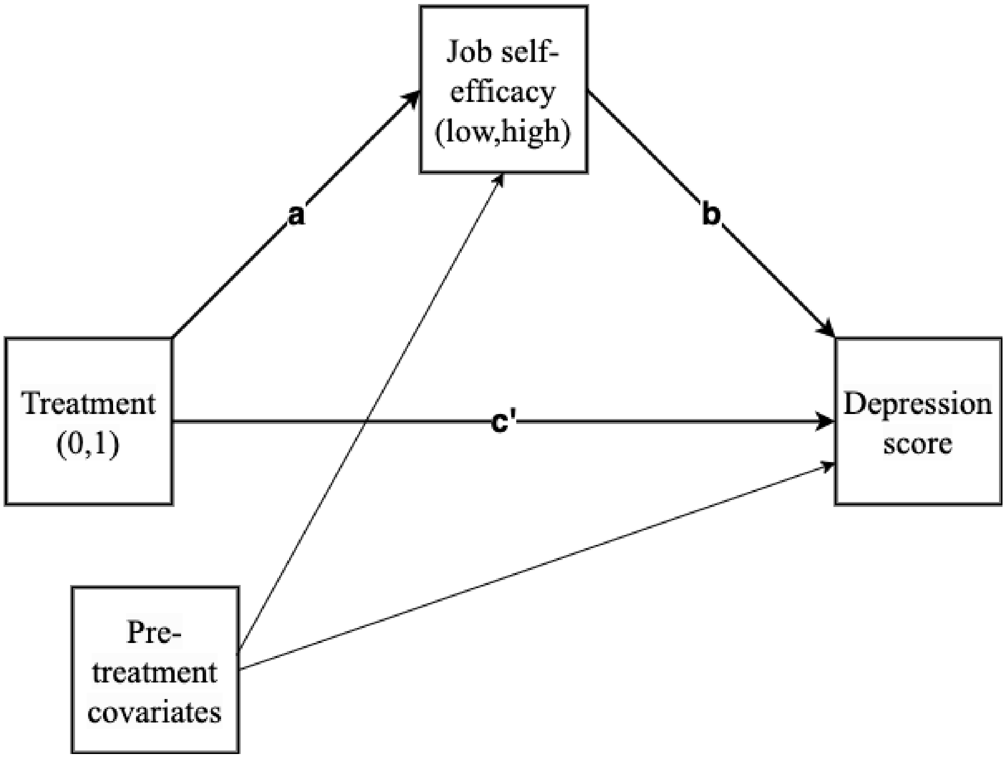
| Name | Label |
|---|---|
| depress2 (Y) | Measure of depressive symptoms post-treatment. |
| treat (X) | Indicator variable for whether participant was randomly selected for the JOBS II training program. 1 = as |
| job_dich (Z) | The job_seek measure recoded into two categories of high and low. 1 = high job search self-efficacy. |
| sex | Indicator variable for sex. 1 = female |
| age | Age in years. |
| marital | Factor with five categories for marital status. |
| nonwhite | Indicator variable for race. 1 = nonwhite. |
| educ | Factor with five categories for educational attainment. |
| income | Factor with five categories for level of income. |

---

Look at descriptives of the framing dataset using {`gtsummary`}

```
jobs_desc <- jobs %>%
  dplyr::select(depress2, job_dich, treat, sex, age, marital, nonwhite, educ, income)

tablej <- tbl_summary(jobs_desc,
                      statistic = list(all_continuous() ~ "{mean} ({sd})"),
                      missing = "no" ) %>%
  bold_labels()

tablej
```

---




---

step 1: fit a binomial logist model using `glm` with the binary mediator (`job_dich`) regressed on treatment (`treat`) and pre-treatment covariates

```
jmed_fit <- glm(job_dich ~ treat + sex + age + marital +
                  nonwhite + educ + income,
                data = jobs, family = binomial)
```

---

step 2: fit a linear model with depression score (`depress2`) regressed on treatment, mediator, and pre-treatment covariates

```
jout_fit <- lm(depress2 ~ job_dich + treat +
                 sex + age + marital + nonwhite + educ + income,
               data = jobs)
```

---

step 3: Estimate the mediation effects with bias corrected bootstrapped confidence intervals.

```
jmed_out <- mediate(jmed_fit, jout_fit, treat = "treat", mediator = "job_dich",
                    boot = TRUE, boot.ci.type ="bca", sims = 100)

summary(jmed_out)
```

```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the BCa Method
##
##                 Estimate 95% CI Lower 95% CI Upper p-value
## ACME             -0.0148      -0.0536         0.00    0.02 *
## ADE              -0.0306      -0.1032         0.04    0.56
## Total Effect     -0.0454      -0.1303         0.03    0.28
## Prop. Mediated    0.3257      -1.2982         4.91    0.26
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 899
##
##
## Simulations: 100
```

---

### Run mediation model 2 as a SEM model with {MplusAutomation}

---

```
m2_jmediate  <- mplusObject(

  TITLE = "m2 jobs mediate",

  VARIABLE =
   "usevar = treat sex
    age marital nonwhite
    educ income depress2 job_dich;

    categorical = job_dich; ! moderator is binary",

  ANALYSIS =
   "bootstrap = 500; ! set number of bootstrap draws (500 for example purposes)" ,

  MODEL =
   "job_dich on treat sex age marital nonwhite educ income;

    depress2 on job_dich treat sex age marital nonwhite educ income;

    Model indirect:
    depress2 ind treat;" ,

  OUTPUT =
    "sampstat standardized cinterval (bcbootstrap); ! bias-corrected bootstrap",

  PLOT = "type=plot2;",

  usevariables = colnames(jobs),
  rdata = jobs)
```

```r
m2_jmediate_fit <- mplusModeler(m2_jmediate,
                    dataout=here("15-mediation", "mplus_files", "jobs.dat"),
                    modelout=here("15-mediation", "mplus_files", "m2_mediate_jobs.inp"),
                    check=TRUE, run = TRUE, hashfilename = FALSE)
```

---

**Model 2 Mplus output**

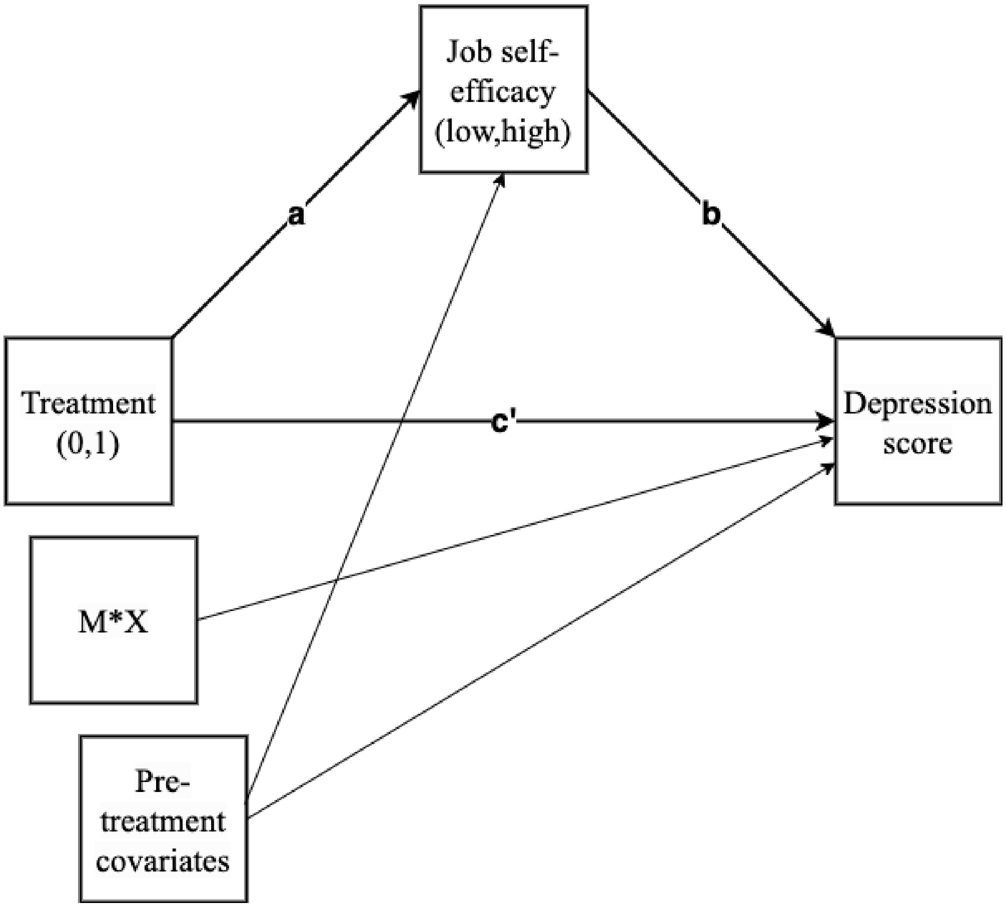
STANDARDIZED TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS

```
                    Estimate       S.E.  Est./S.E.    P-Value

Effects from TREAT to DEPRESS2

  Total               -0.043      0.033     -1.306      0.192
  Total indirect      -0.029      0.012     -2.385      0.017

  Specific indirect 1
    DEPRESS2
    JOB_DICH
    TREAT             -0.029      0.012     -2.385      0.017

  Direct
    DEPRESS2
    TREAT             -0.015      0.033     -0.440      0.660
```

---

## Run model 3 including the mediator*treatement interaction (potential outcomes framework)

For further reading on this topic see chapter 3 of *Regression and mediation analysis using Mplus* (Muthen et al., 2017)




```
m3_jmed  <- mplusObject(

  TITLE = "m3 MX jobs mediate",

  VARIABLE =
   "usevar =
    treat sex age marital nonwhite
    educ income depress2 job_dich mx; ",

  DEFINE = "mx = job_dich*treat;",

  ANALYSIS =  "bootstrap = 500; ",

  MODEL =
   "job_dich on treat sex age marital nonwhite educ income;
    depress2 on job_dich treat mx sex age marital nonwhite educ income;

    Model indirect:
    depress2 MOD job_dich mx treat; ",

  OUTPUT =
    "sampstat cinterval(bootstrap); ",
```

```
  usevariables = colnames(jobs),
  rdata = jobs)

m3_jmed_fit <- mplusModeler(m3_jmed,
                    dataout=here("15-mediation", "mplus_files", "jobs.dat"),
                    modelout=here("15-mediation", "mplus_files", "m3_mediate_jobs.inp"),
                    check=TRUE, run = TRUE, hashfilename = FALSE)
```

---

**Model 3 Mplus output**

TOTAL, INDIRECT, AND DIRECT EFFECTS BASED ON COUNTERFACTUALS (CAUSALLY-DEFINED EFFECTS)

```
Effects from TREAT to DEPRESS2

                   Estimate       S.E.  Est./S.E.    P-Value

  Tot natural IE     -0.026      0.011     -2.357      0.018
  Pure natural DE    -0.022      0.055     -0.401      0.688
  Total effect       -0.048      0.055     -0.878      0.380

 Other effects

  Pure natural IE    -0.023      0.012     -1.938      0.053
  Tot natural DE     -0.026      0.052     -0.494      0.621
  Total effect       -0.048      0.055     -0.878      0.380
```

---

## References


Brader T, Valentino NA, Suhat E (2008). What Triggers Public Opposition to Immigration? Anxiety, Group Cues, and Immigration. American Journal of Political Science, 52(4), 959–978.

Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. Structural equation modeling: a multidisciplinary journal, 25(4), 621-638.

Ingels, S. J., Pratt, D. J., Herget, D. R., Burns, L. J., Dever, J. A., Ottem, R., … & Leinwand, S. (2011). High School Longitudinal Study of 2009 (HSLS: 09): Base-Year Data File Documentation. NCES 2011-328. National Center for Education Statistics.

Muthén, B. O., Muthén, L. K., & Asparouhov, T. (2017). Regression and mediation analysis using Mplus. Los Angeles, CA: Muthén & Muthén.

Muthén, L.K. and Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R package for causal mediation analysis.

Vinokur AD, Price RH, Schul Y (1995). Impact of the JOBS Intervention on Unemployed Workers Varying in Risk for Depression. American Journal of Community Psychology, 23(1), 39–74.

Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686