002 - Motivating Examples

EPIB 607 - FALL 2020

Sahir Rai Bhatnagar
Department of Epidemiology, Biostatistics, and Occupational Health
McGill University

`sahir.bhatnagar@mcgill.ca`

slides compiled on September 2, 2020

Case study 1: Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2

Case study 2: Comparison of Estimated Rates of Coronavirus Disease 2019 (COVID-19) in Border Counties in Iowa Without a Stay-at-Home Order and Border Counties in Illinois With a Stay-at-Home Order

Early phase COVID-19 vaccine trial[1]

Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial

Pedro M Folegatti*, Katie J Ewer*, Parvinder K Aley, Brian Angus, Stephan Becker, Sandra Belij-Rammerstorfer, Duncan Bellamy, Sagida Bibi, Mustapha Bittaye, Elizabeth A Clutterbuck, Christina Dold, Saul N Faust, Adam Finn, Amy L Flaxman, Bassam Hallis, Paul Heath, Daniel Jenkin, Rajeka Lazarus, Rebecca Makinson, Angela M Minassian, Katrina M Pollock, Maheshi Ramasamy, Hannah Robinson, Matthew Snape, Richard Tarrant, Merryn Voysey, Catherine Green*, Alexander D Douglas*, Adrian V S Hill*, Teresa Lambe*, Sarah C Gilbert*, Andrew J Pollard*, on behalf of the Oxford COVID Vaccine Trial Group†

Summary

Background The pandemic of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) might be curtailed by vaccination. We assessed the safety, reactogenicity, and immunogenicity of a viral vectored coronavirus vaccine that expresses the spike protein of SARS-CoV-2.

Methods We did a phase 1/2, single-blind, randomised controlled trial in five trial sites in the UK of a chimpanzee adenovirus-vectored vaccine (ChAdOx1 nCoV-19) expressing the SARS-CoV-2 spike protein compared with a meningococcal conjugate vaccine (MenACWY) as control. Healthy adults aged 18–55 years with no history of laboratory confirmed SARS-CoV-2 infection or of COVID-19-like symptoms were randomly assigned (1:1) to receive ChAdOx1 nCoV-19 at a dose of 5×10^{10} viral particles or MenACWY as a single intramuscular injection. A protocol amendment in two of the five sites allowed prophylactic paracetamol to be administered before vaccination. Ten participants assigned to a non-randomised, unblinded ChAdOx1 nCoV-19 prime-boost group received a two-dose schedule, with the booster vaccine administered 28 days after the first dose. Humoral responses at baseline and following vaccination were assessed using a standardised total IgG ELISA against trimeric SARS-CoV-2 spike protein, a multiplexed immunoassay, three live SARS-CoV-2 neutralisation assays (a 50% plaque reduction neutralisation assay [PRNT$_{50}$]; a microneutralisation assay [MNA$_{50}$, MNA$_{80}$, and MNA$_{90}$]; and Marburg VN), and a pseudovirus neutralisation assay. Cellular responses were assessed using an ex-vivo interferon-γ enzyme-linked immunospot assay. The co-primary outcomes are to assess efficacy, as measured by cases of symptomatic virologically confirmed COVID-19, and safety, as measured by the occurrence of serious adverse events. Analyses were done by group allocation in participants who received the vaccine. Safety was assessed over 28 days after

---

Phase 1/2 trial

- The focus in phase 1/2 trials is looking at what the vaccine does to the body and what the body does with the vaccine in *healthy* individuals

- Adults with no history of laboratory confirmed SARS-CoV-2 infection or of COVID-19-like symptoms were randomly assigned (1:1) to receive **ChAdOx1 nCoV-19** or **MenACWY** (Meningococcal) as a single intramuscular injection

- Convalescent plasma samples from adults with PCR-positive SARS-CoV-2 infection were obtained from symptomatic patients admitted to the hospitals to characterize the immunological properties of COVID-19[2]

- The enzyme-linked immunosorbent assay (ELISA) technique was used to detect antibodies (i.e. levels of immunity)

---

[2]Convalescent plasma is collected from someone who has recovered from a virus. When a person is infected with a virus, their body starts making antibodies to fight it. It is believed these antibodies could be the key ingredient for a treatment to help others with the same virus.
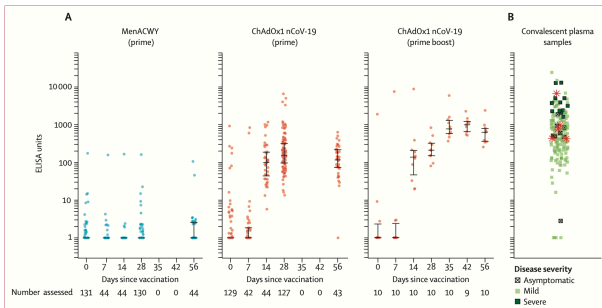
Figure 3: SARS-CoV-2 IgG response by standardised ELISA to spike protein in trial participants (A) and in 180 convalescent plasma samples from 172 patients with PCR-confirmed COVID-19 and eight asymptomatic health-care workers (B)

Error bars show median (IQR). Participants in the prime boost group received their second dose at day 28. Lower limit of quantification is 1 ELISA unit. Red stars in panel B show five samples also tested on the Marburg VN assay (see figure 4). MenACWY=meningococcal group A, C, W-135, and Y conjugate vaccine. SARS-CoV-2=severe acute respiratory syndrome coronavirus 2.

1. What levels of immunity are found in patients who have recovered from COVID-19? (panel B)
2. Relative to these what levels of immunity are found in persons who have received the ChAdOx1 nCoV-19 vaccine? Compare panel A (prime, 28 days) vs panel B.

What levels of immunity are found in patients who have recovered from COVID-19?[3]

```r
path <-
  "http://www.biostat.mcgill.ca/hanley/statbook/immunogenicityChAdOx1.nCoV-19vaccine.txt"
ds <- read.table(path)
head(ds)

##   RefIndexCategory IgGResponse.log10.ElisaUnits
## 1      Convalescent                         2.56
## 2      Convalescent                         2.74
## 3      Convalescent                         2.79
## 4      Convalescent                         3.32
## 5      Convalescent                         3.15
## 6      Convalescent                         2.35

str(ds)

## 'data.frame':^^I307 obs. of  2 variables:
##  $ RefIndexCategory          : Factor w/ 2 levels "Convalescent",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ IgGResponse.log10.ElisaUnits: num  2.56 2.74 2.79 3.32 3.15 2.35 2.72 2.95 2.42 2.64 ...

levels(ds$RefIndexCategory)

## [1] "Convalescent"          "Day28PostChAdOx1 nCoV-19"
```

---

[3] Data were (imperfectly) scraped from the Postscript file "behind" the pdf file by Dr. Hanley

What levels of immunity are found in patients who have recovered from COVID-19?

```
natural <- ds[ds$RefIndexCategory=="Convalescent",]
hist(natural$IgGResponse.log10.ElisaUnits,
     breaks = 20, col = "lightblue")
```

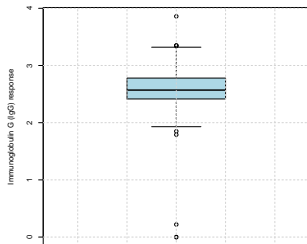**Histogram of natural$IgGResponse.log10.ElisaUnits**
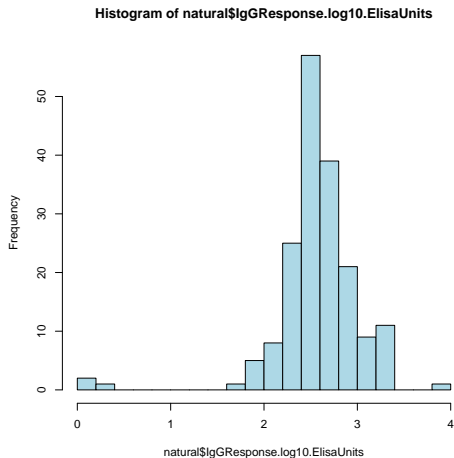
Three different methods of calculating the mean

```
summary(natural$IgGResponse.log10.ElisaUnits)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   2.417   2.570   2.577   2.780   3.860

boxplot(natural$IgGResponse.log10.ElisaUnits,
        col = "lightblue",
        ylab = "Immunoglobulin G (IgG) response")
grid(lty = "dashed")
```



```
t.test(natural$IgGResponse.log10.ElisaUnits)

## One Sample t-test with natural$IgGResponse.log10.ElisaUnits
## t = 75.0898, df = 179, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  2.509603 2.645064
## sample estimates:
## mean of x
##  2.577333

fit1 <- glm(IgGResponse.log10.ElisaUnits ~ 1, data = natural)
summary(fit1)

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.57733    0.03432   75.09   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.2120565)
##
##     Null deviance: 37.958  on 179  degrees of freedom
## Residual deviance: 37.958  on 179  degrees of freedom
## AIC: 234.65
##
## Number of Fisher Scoring iterations: 2

confint(fit1)

##    2.5 %   97.5 %
## 2.510061 2.644606
```

Naturally vs. vaccine-induced response levels

```r
p1 <- ggplot(data = ds, mapping = aes(x = RefIndexCategory, y = IgGResponse.log10.ElisaUnits,
    fill = RefIndexCategory)) + geom_jitter(alpha = 0.3) + theme_minimal() + theme(legend.position = "none")
p1 + geom_violin()
p1 + geom_boxplot()
```
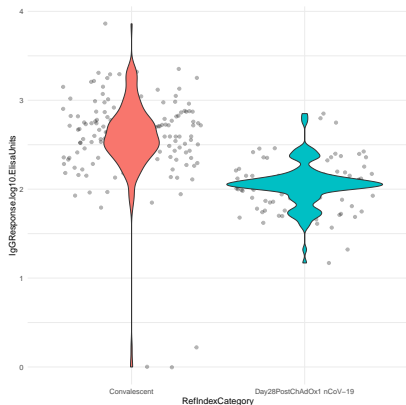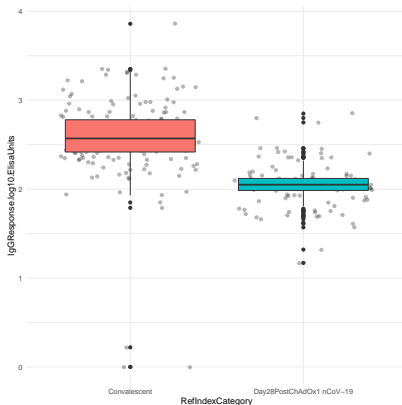

(a) Violin plot

(b) Boxplot

Comparing means using classic methods

1. Numerical summary

```
by(ds$IgGResponse.log10.ElisaUnits,ds$RefIndexCategory,summary)

## ds$RefIndexCategory: Convalescent
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   2.417   2.570   2.577   2.780   3.860
## --------------------------------------------------------------
## ds$RefIndexCategory: Day28PostChAdOx1 nCoV-19
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.170   1.985   2.050   2.047   2.120   2.850
```

Comparing means using classic methods

1. Numerical summary

```
by(ds$IgGResponse.log10.ElisaUnits,ds$RefIndexCategory,summary)

## ds$RefIndexCategory: Convalescent
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   2.417   2.570   2.577   2.780   3.860
## -----------------------------------------------------------
## ds$RefIndexCategory: Day28PostChAdOx1 nCoV-19
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.170   1.985   2.050   2.047   2.120   2.850
```

2. Another "dot" test

```
t.test(IgGResponse.log10.ElisaUnits ~ RefIndexCategory, data = ds)

## Welch Two Sample t-test with IgGResponse.log10.ElisaUnits by RefIndexCategory
## t = 13.1047, df = 284.781, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.4510720 0.6105238
## sample estimates:
##          mean in group Convalescent mean in group Day28PostChAdOx1 nCoV-19
##                            2.577333                               2.046535
```

Comparing means using regression

3. Regression

```
fit2 <- glm(IgGResponse.log10.ElisaUnits ~ RefIndexCategory, data = ds)
print(summary(fit2), signif.star = FALSE)


##
## Coefficients:
##                                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                                 2.57733    0.02874   89.67   <2e-16
## RefIndexCategoryDay28PostChAdOx1 nCoV-19   -0.53080    0.04469  -11.88   <2e-16
##
## (Dispersion parameter for gaussian family taken to be 0.1487187)
##
##     Null deviance: 66.339  on 306  degrees of freedom
## Residual deviance: 45.359  on 305  degrees of freedom
## AIC: 290.17
##
## Number of Fisher Scoring iterations: 2


confint(fit2)


##                                              2.5 %     97.5 %
## (Intercept)                               2.5209962  2.6336704
## RefIndexCategoryDay28PostChAdOx1 nCoV-19 -0.6183894 -0.4432064
```

Fitted regression line

```
plot(ds$RefIndexCategory, ds$IgGResponse.log10.ElisaUnits, pch=19, cex=0.5)
abline(h = seq(0,4,0.5),col = "lightblue")
lines(ds$RefIndexCategory, fit2$fitted.values, col = "red", lwd = 3)
```
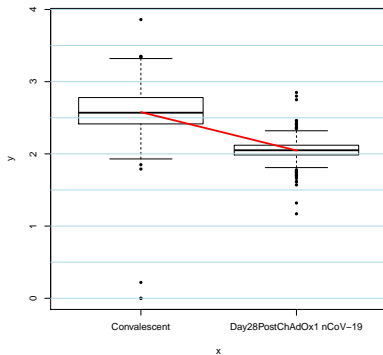

Figure: The red line is the fitted regression from the previous slide.

Case study 1: Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2

Case study 2: Comparison of Estimated Rates of Coronavirus Disease 2019 (COVID-19) in Border Counties in Iowa Without a Stay-at-Home Order and Border Counties in Illinois With a Stay-at-Home Order

Comparing Iowa and Illinois Cases[4]

Original Investigation | Public Health

Comparison of Estimated Rates of Coronavirus Disease 2019 (COVID-19) in Border Counties in Iowa Without a Stay-at-Home Order and Border Counties in Illinois With a Stay-at-Home Order

Wei Lyu, MS; George L. Wehby, PhD

Abstract

IMPORTANCE Iowa is 1 of 5 states in the US that have not issued a stay-at-home order during the coronavirus disease 2019 (COVID-19) pandemic. There is no empirical evidence on whether issuing a stay-at-home order in Iowa could have been associated with a reduced rate of COVID-19 infections in the state.

OBJECTIVE To compare COVID-19 cases in border counties in Iowa, which did not issue a stay-at-home order, with cases in border counties in Illinois, which did issue a stay-at-home order.

DESIGN, SETTING, AND PARTICIPANTS This cross-sectional study with a difference-in-differences design compared daily changes in COVID-19 cases per 10 000 residents in 8 Iowa counties bordering Illinois with those in the 7 Illinois counties bordering Iowa before and after Illinois issued a stay-at-home order on March 21, 2020. Additional sensitivity analyses were conducted to account for differences in timing of closing schools and nonessential businesses between the 2 states and differential trends in COVID-19 cases by county population density and poverty rates.

Key Points

Question Was the stay-at-home order in Illinois associated with different rates of coronavirus disease 2019 compared with Iowa, which did not issue a stay-at-home order?

Findings This cross-sectional study of border counties in Iowa and Illinois used difference-in-differences design and found an increase in estimated rates of COVID-19 cases per 10 000 residents in the border counties in Iowa compared with the border counties in Illinois after a stay-at-home order was implemented in Illinois but not in Iowa.

Case study 2: Comparison of Estimated Rates of Coronavirus Disease 2019 (COVID-19) in Border Counties in Iowa Without a Stay-at-Home Order and Border Counties in Illinois With a Stay-at-Home Order

14 / 22

Are the difference in curves real? Or just random variation?

- This study compared COVID-19 cases in border counties in Iowa, which did not issue a stay-at-home order, with cases in border counties in Illinois, which did issue a stay-at-home order.

Figure 1. Cumulative Coronavirus Disease 2019 (COVID-19) Cases per 10 000 Residents in Iowa and Illinois Border Counties
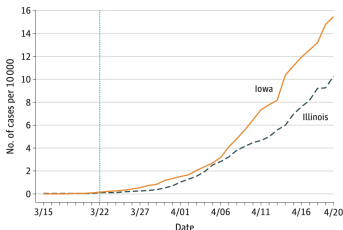
The vertical line represents the date on which the stay-at-home order took effect in Illinois.

May 15, 2020   3/7

Freely available county level data from NYTimes[5]

```r
library(covdata) # remotes::install_github("kjhealy/covdata")
library(dplyr); library(tidyr); library(ggplot2); library(readr)

# get population data from https://covid19.census.gov/datasets/
pop_county <- read_csv("https://opendata.arcgis.com/datasets/21843f238cbb46b08615fc53e19e0daf_1.csv") %>%
              dplyr::rename(fips = GEOID, population = B01001_001E, state = State) %>%
              dplyr::select(state, fips, population)

county_level <- nytcovcounty %>%
                dplyr::left_join(pop_county, by = c("state","fips")) %>%
                dplyr::mutate(cases.per.10k = cases/population * 1e4) %>%
                dplyr::filter(state %in% c("Iowa","Illinois")) %>%
                dplyr::group_by(county)

pop_state <- pop_county %>%
             dplyr::group_by(state) %>%
             dplyr::summarise(population = sum(population, na.rm = TRUE))

state_level <- county_level %>%
               dplyr::group_by(state, date) %>%
               dplyr::filter(date >= "2020-03-15") %>%
               dplyr::summarise(cases = sum(cases)) %>%
               dplyr::left_join(pop_state, by = "state") %>%
               dplyr::mutate(cases.per.10k = cases / population * 1e4, state = factor(state),
                             time = as.numeric(date - min(date)) + 1)
head(state_level)

## # A tibble: 6 x 6
## # Groups:   state [1]
##   state    date       cases population cases.per.10k  time
##   <fct>    <date>     <dbl>      <dbl>        <dbl> <dbl>
## 1 Illinois 2020-03-15    94   12821497       0.0733     1
## 2 Illinois 2020-03-16   104   12821497       0.0811     2
## 3 Illinois 2020-03-17   159   12821497       0.124      3
## 4 Illinois 2020-03-18   286   12821497       0.223      4
## 5 Illinois 2020-03-19   420   12821497       0.328      5
## 6 Illinois 2020-03-20   583   12821497       0.455      6
```
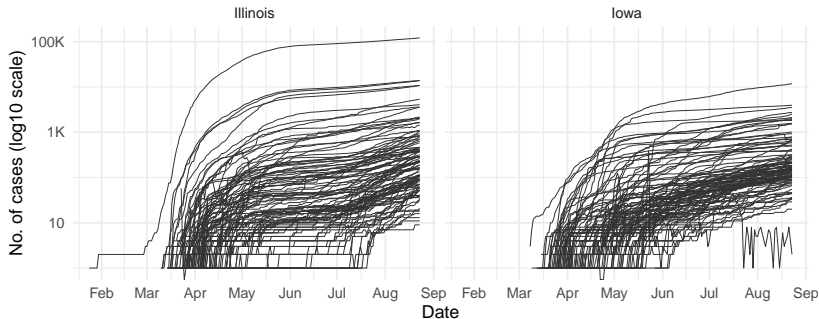
Case study 2: Comparison of estimated rate of Coronavirus Disease 2019 (COVID-19) in Border Counties in Iowa Without a Stay-at-Home Order and Border Counties in Illinois With a Stay-at-Home Order

County level cases for Iowa and Illinois - log10 scale

```
ggplot(data = county_level, mapping = aes(x = date, y = cases, group = county)) +
  geom_line(size = 0.25, color = "gray20") +
  scale_x_date(date_breaks = "1 month", date_labels = "%b")+
  scale_y_log10(labels = scales::label_number_si()) +
  guides(color = FALSE) + facet_wrap(~ state, ncol = 2) +
  labs(title = "COVID-19 Cases in Iowa and Illinois by County",
       x = "Date", y = "No. of cases (log10 scale)", caption = "Data: The New York Times") +
  theme_minimal()
```
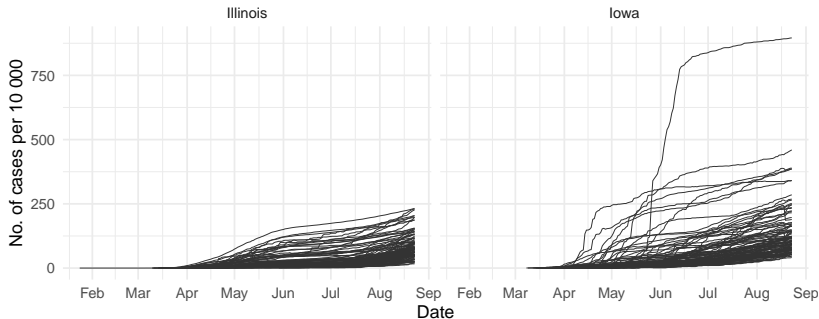

COVID−19 Cases in Iowa and Illinois by County

Data: The New York Times

Case study 2: Comparison of Estimated Rates of Coronavirus Disease 2019 (COVID-19) in Border Counties in Iowa Without a Stay-at-Home Order and Border Counties in Illinois With a Stay-at-Home Order

17 / 22

County level cases for Iowa and Illinois - per capita

```r
ggplot(data = county_level, mapping = aes(x = date, y = cases.per.10k, group = county)) +
  geom_line(size = 0.25, color = "gray20") +
  scale_x_date(date_breaks = "1 month", date_labels = "%b")+
  scale_y_continuous(labels = scales::label_number_si()) +
  guides(color = FALSE) + facet_wrap(~ state, ncol = 2) +
  labs(title = "COVID-19 Cases in Iowa and Illinois by County",
       x = "Date", y = "No. of cases per 10 000", caption = "Data: The New York Times") +
  theme_minimal()
```


COVID−19 Cases in Iowa and Illinois by County

Data: The New York Times

Case study 2: Comparison of Estimated Rates of Coronavirus Disease 2019 (COVID-19) in Border Counties in Iowa Without a Stay-at-Home Order and Border Counties in Illinois With a Stay-at-Home Order

18 / 22

State level cases for Iowa and Illinois - per capita

```
ggplot(data = state_level, mapping = aes(x = date, y = cases.per.10k, color = state)) +
  geom_line(size = 1) +
  scale_x_date(date_breaks = "1 month", date_labels = "%b")+
  scale_y_continuous(labels = scales::label_number_si()) +
  labs(title = "COVID-19 Cases in Iowa and Illinois",
       subtitle = "Cases since March 15, 2020",
       x = "Date", y = "No. of cases per 10 000", caption = "Data: The New York Times") +
  theme_minimal()
```


COVID−19 Cases in Iowa and Illinois
Cases since March 15, 2020
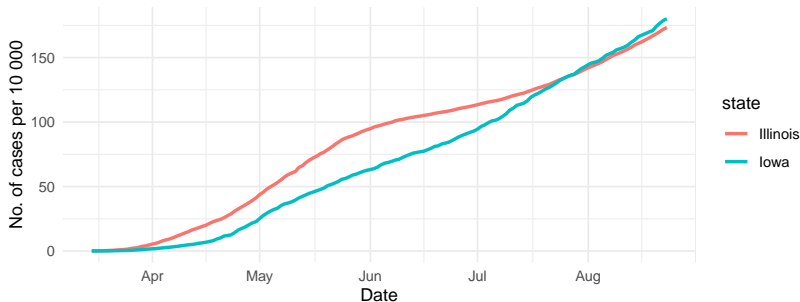
Data: The New York Times

Are the findings in the paper reproducible?

```
fit3 <- glm(cases.per.10k ~ state*time, data = state_level)
summary(fit3)

##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -7.07540    1.22153  -5.792 1.66e-08 ***
## stateIowa      -17.88124    1.72751 -10.351  < 2e-16 ***
## time            1.10890     0.01300  85.300  < 2e-16 ***
## stateIowa:time  0.06078     0.01838   3.306  0.00105 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 59.87398)
##
##     Null deviance: 953056  on 323  degrees of freedom
## Residual deviance:  19160  on 320  degrees of freedom
## AIC: 2251.3
##
## Number of Fisher Scoring iterations: 2
```

Model-based predictions

```
library(ggeffects)
ggeffects::ggpredict(fit3, terms = "state") %>%
  plot()
```


Predicted values of cases.per.10k
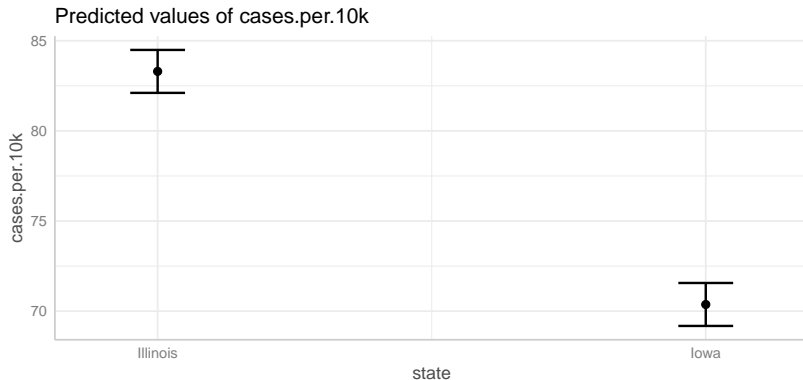
Session Info

```
R version 3.6.2 (2019-12-12)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Pop!_OS 19.10

Matrix products: default
BLAS:   /usr/lib/x86_64-linux-gnu/openblas/libblas.so.3
LAPACK: /usr/lib/x86_64-linux-gnu/libopenblasp-r0.3.7.so

attached base packages:
[1] tools      stats     graphics  grDevices utils     datasets  methods
[8] base

other attached packages:
 [1] ggeffects_0.14.1  covdata_0.4.4     NCStats_0.4.7     FSA_0.8.30
 [5] forcats_0.5.0     stringr_1.4.0     dplyr_1.0.2       purrr_0.3.4
 [9] readr_1.3.1       tidyr_1.1.2       tibble_3.0.3      ggplot2_3.3.2.9000
[13] tidyverse_1.3.0   knitr_1.29

loaded via a namespace (and not attached):
 [1] sjlabelled_1.1.3  tidyselect_1.1.0  xfun_0.16         haven_2.3.1
 [5] snakecase_0.11.0  colorspace_1.4-1  vctrs_0.3.4       generics_0.0.2
 [9] utf8_1.1.4        rlang_0.4.7       pillar_1.4.6      glue_1.4.2
[13] withr_2.2.0       DBI_1.1.0         dbplyr_1.4.2      modelr_0.1.5
[17] readxl_1.3.1      lifecycle_0.2.0   plyr_1.8.6        munsell_0.5.0
[21] gtable_0.3.0      cellranger_1.1.0  rvest_0.3.5       evaluate_0.14
[25] labeling_0.3      curl_4.3          fansi_0.4.1       highr_0.8
[29] broom_0.7.0       Rcpp_1.0.4.6      scales_1.1.1      backports_1.1.9
[33] formatR_1.7       jsonlite_1.7.0    farver_2.0.3      fs_1.3.2
[37] TeachingDemos_2.12 digest_0.6.25    hms_0.5.3         stringi_1.4.6
[41] insight_0.8.1     grid_3.6.2        cli_2.0.2         magrittr_1.5
[45] crayon_1.3.4      pkgconfig_2.0.3   ellipsis_0.3.1    MASS_7.3-51.5
[49] xml2_1.3.0        reprex_0.3.0      lubridate_1.7.4   assertthat_0.2.1
[53] httr_1.4.1        rstudioapi_0.11   R6_2.4.1          compiler_3.6.2
```