

Statistical / Data Science Methods in Clinical Research

Thomas T. John, PhD
Assistant Teaching Professor
Department of Mathematics
Syracuse University

Abstract

- This presentation will go through some typical statistical analyses used in scientific research, in particular clinical research, as well as the results of such analyses typically presented in current clinical/medical journals.
- Using some actual health related datasets, analyses and examples will be carried out live, including graphical visualization of the data.
- Along with the interpretations of the analyses, the discussion will include some of the related nuances that researchers collecting and recording data should keep in mind.

A Joke

- Once upon a time there was a shepherd tending his sheep at the edge of a country road.
- A brand-new Range Rover screeches to a halt next to him.
- The driver, a young man dressed in a Brioni suit, Cerrutti shoes, Ray-Ban glasses, Jovial Swiss wristwatch and a Bhs tie gets out and asks the shepherd: 'If I guess how many sheep you have, will you give me one of them?'
- The shepherd looks at the young man, then looks at the sprawling field of sheep and says: 'Okay.'

Joke (continued)

- The young man parks the SUV, connects his notebook and wireless modem, enters a NASA site, scans the ground using his GPS, opens a database and 60 Excel tables filled with algorithms, then prints a 150-page report on his high-tech mini printer.
- He then turns to the shepherd and says: 'You have exactly 1,586 sheep here.'
- The shepherd answers: 'That's correct, you can have your sheep.'

Joke (continued)

- The young man takes one of the animals and puts it in the back of his vehicle.
- The shepherd looks at him and asks: 'Now, if I guess your profession, will you pay me back in kind?'
- The young man answers: 'Sure.'
- The shepherd says: 'You are a consultant.'
- 'Exactly! How did you know?' asked the young man with bewilderment

Joke (continued)

- 'Very simple,' answers the shepherd.
- '*First*, you came here without being called.
- *Second*, you charged me a fee to tell me something I already knew.
- *Third*, you do not understand anything about my business and
- 'I'd really like to have my dog back.'

Comparison Between Groups

Typically called:
2-Independent-Samples Comparison

https://www.npr.org/sections/health-shots/2023/06/26/1184390543/intermittent-fasting-effective-weight-loss

SHOTS - HEALTH NEWS

Intermittent fasting is as effective as counting calories, new study finds

June 26, 2023 · 5:00 PM ET

By Will Stone

Annals of Internal Medicine

ORIGINAL RESEARCH

Time-Restricted Eating Without Calorie Counting for Weight Loss in a Racially Diverse Population

A Randomized Controlled Trial

Shuhao Lin, MS, RD; Sofia Cienfuegos, PhD; Mark Ezpeleta, PhD; Kelsey Gabel, PhD, RD; Vasiliki Pavlou, MS, RD; Andrea Mulas, MS, RD; Kaitie Chakos, MS, RD; Mara McStay, MS, RD; Jackie Wu, MS, RD; Lisa Tussing-Humphreys, PhD, RD; Shaina J. Alexandria, PhD; Julianne Sanchez, MD; Terry Unterman, MD; and Krista A. Varady, PhD

Background: Time-restricted eating (TRE), without calorie counting, has become a popular weight loss strategy, yet long-term randomized trials evaluating its efficacy are limited.

Objective: To determine whether TRE is more effective for weight control and cardiometabolic risk reduction compared with calorie restriction (CR) or control.

Design: 12-month randomized controlled trial. (ClinicalTrials.gov: NCT04692532)

Setting: University of Illinois Chicago from January 2021 to September 2022.

Participants: 90 adults with obesity.

Intervention: 8-hour TRE (eating between noon and 8:00 p.m. only, without calorie counting), CR (25% energy restriction daily), or control (eating over a period of 10 or more hours per day). Participants were not blinded.

Measurements: Change in body weight, metabolic markers, and energy intake by month 12.

Results: Seventy-seven persons completed the study. Mean age was 40 years (SD, 11), 33% were Black, and 46% were

Hispanic. Mean reduction in energy intake was -425 kcal/d (SD, 531) for TRE and -405 kcal/d (SD, 712) for CR. Compared with the control group, weight loss by month 12 was -4.61 kg (95% CI, -7.37 to -1.85 kg; $P \leq 0.01$) (-4.87% [CI, -7.61% to -2.13%]) for the TRE group and -5.42 kg (CI, -9.13 to -1.71 kg; $P \leq 0.01$) (-5.30% [CI, -9.06% to -1.54%]) for the CR group, with no statistically significant difference between TRE and CR (0.81 kg [CI, -3.07 to 4.69 kg; $P = 0.68$]) (0.43% [CI, -3.48% to 4.34%]).

Limitation: Not blinded, not powered to detect relatively large differences in weight loss, and lack of adjustment for multiple comparisons.

Conclusion: Time-restricted eating is more effective in producing weight loss when compared with control but not more effective than CR in a racially diverse population.

Primary Funding Source: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases.

Figure 2. Change in body weight and energy intake between groups over 12 months.

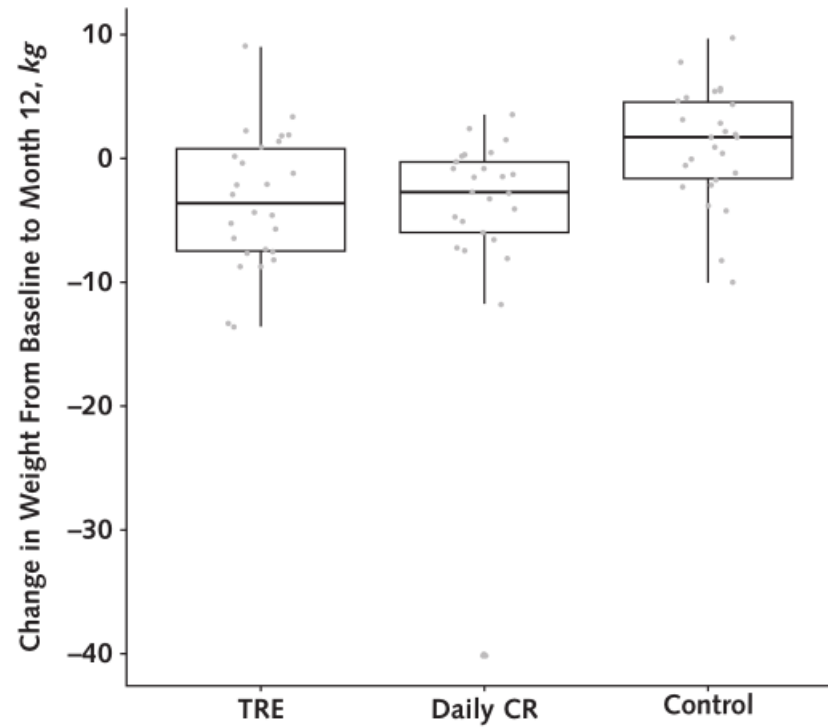
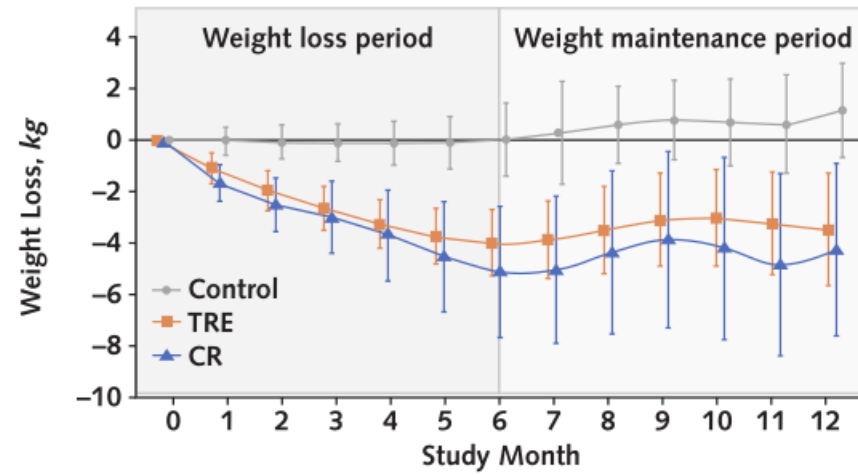


Table 2. Change in Body Weight and Body Composition From Baseline and Between Intervention Groups*

Variables	Participants, <i>n</i>	Change From Baseline (95% CI)			Difference Between Groups (95% CI)		
		TRE	Daily CR	Control	TRE vs. CR	TRE vs. Control	CR vs. Control
Primary outcome							
Body weight, kg							
12 mo	77	-3.49 (-5.65 to -1.32)	-4.30 (-7.63 to -0.96)	1.12 (-0.69 to 2.94)	0.81 (-3.07 to 4.69) <i>P</i> = 0.68	-4.61 (-7.37 to -1.85) <i>P</i> ≤ 0.01	-5.42 (-9.13 to -1.71) <i>P</i> ≤ 0.01
Secondary outcomes							
Body weight, kg							
6 mo	66	-4.00 (-5.31 to -2.70)	-5.14 (-7.66 to -2.62)	0.00 (-1.40 to 1.40)	1.14 (-1.63 to 3.91)	-4.00 (-5.87 to -2.13)	-5.14 (-7.95 to -2.33)
Body weight, %							
6 mo	66	-4.27 (-5.69 to -2.84)	-5.06 (-7.56 to -2.57)	-0.03 (-1.47 to 1.40)	0.80 (-2.01 to 3.61)	-4.23 (-6.21 to -2.26)	-5.03 (-7.84 to -2.22)
12 mo	77	-3.76 (-5.89 to -1.64)	-4.20 (-7.59 to -0.80)	1.11 (-0.72 to 2.94)	0.43 (-3.48 to 4.34)	-4.87 (-7.61 to -2.13)	-5.30 (-9.06 to -1.54)
Fat mass, kg							
6 mo	68	-2.68 (-3.75 to -1.61)	-2.25 (-4.62 to 0.13)	-0.13 (-1.33 to 1.08)	-0.43 (-2.96 to 2.10)	-2.55 (-4.12 to -0.98)	-2.12 (-4.71 to 0.46)
12 mo	76	-2.20 (-3.88 to -0.52)	-2.61 (-5.97 to 0.74)	0.57 (-1.14 to 2.27)	0.42 (-3.24 to 4.07)	-2.77 (-5.10 to -0.43)	-3.18 (-6.85 to 0.49)

Risk Ratios

From Logistic Regression

COVID during pregnancy may alter brain development in boys

April 18, 2023 · 5:00 AM ET

Heard on [Morning Edition](#)



Jon Hamilton



Original Investigation | Psychiatry

Sex-Specific Neurodevelopmental Outcomes Among Offspring of Mothers With SARS-CoV-2 Infection During Pregnancy

Andrea G. Edlow, MD, MSc; Victor M. Castro, MS; Lydia L. Shook, MD; Sebastien Haneuse, PhD; Anjali J. Kaimal, MD, MAS; Roy H. Perlis, MD, MSc

Abstract

IMPORTANCE Prior studies using large registries have suggested a modest increase in risk for neurodevelopmental diagnoses among children of mothers with immune activation during pregnancy, and such risk may be sex-specific.

OBJECTIVE To determine whether in utero exposure to SARS-CoV-2 is associated with sex-specific risk for neurodevelopmental disorders up to 18 months after birth, compared with unexposed offspring born during or prior to the COVID-19 pandemic period.

DESIGN, SETTING, AND PARTICIPANTS This retrospective cohort study included the live offspring of all mothers who delivered between January 1 and December 31, 2018 (born and followed up before the COVID-19 pandemic), between March 1 and December 31, 2019 (born before and followed up during the COVID-19 pandemic), and between March 1, 2020, and May 31, 2021 (born and followed up during the COVID-19 pandemic). Offspring were born at any of 8 hospitals across 2 health systems in Massachusetts.

EXPOSURES Polymerase chain reaction evidence of maternal SARS-CoV-2 infection during pregnancy.

MAIN OUTCOMES AND MEASURES Electronic health record documentation of *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* diagnostic codes corresponding to neurodevelopmental disorders.

Key Points

Question Is in utero exposure to maternal SARS-CoV-2 infection associated with greater rates of neurodevelopmental disorder diagnoses in male or female offspring, compared with controls with no such exposure?

Findings This cohort study of 18 355 infants delivered after February 2020 found that male but not female offspring born to mothers with a positive SARS-CoV-2 polymerase chain reaction test result during pregnancy were more likely to receive a neurodevelopmental diagnosis in the first 12 months after delivery, even after accounting for preterm delivery.

Meaning These findings suggest that male offspring exposed to SARS-CoV-2 in utero may be at increased risk for neurodevelopmental disorders.

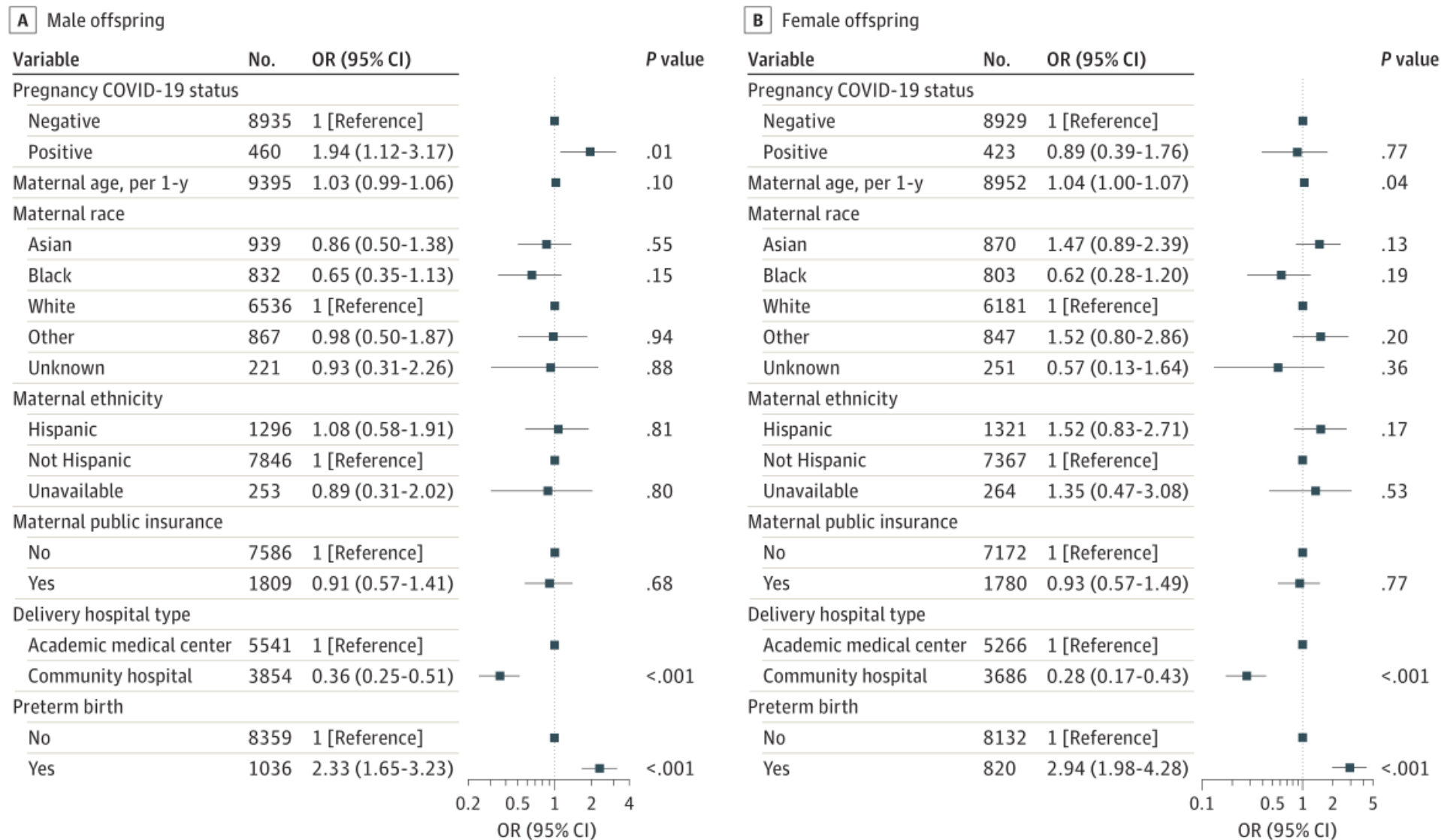
RESULTS The COVID-19 pandemic cohort included 18 355 live births (9399 boys [51.2%]), including 883 (4.8%) with maternal SARS-CoV-2 positivity during pregnancy. The cohort included 1809 Asian individuals (9.9%), 1635 Black individuals (8.9%), 12 718 White individuals (69.3%), and 1714 individuals (9.3%) who were of other race (American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, more than 1 race); 2617 individuals (14.3%) were of Hispanic ethnicity. Mean maternal age was 33.0 (IQR, 30.0-36.0) years. In adjusted regression models accounting for race, ethnicity, insurance status, hospital type (academic center vs community), maternal age, and preterm status, maternal SARS-CoV-2 positivity was associated with a statistically significant elevation in risk for neurodevelopmental diagnoses at 12 months among male offspring (adjusted OR, 1.94 [95% CI 1.12-3.17]; $P = .01$) but not female offspring (adjusted OR, 0.89 [95% CI, 0.39-1.76]; $P = .77$). Similar effects were identified using matched analyses in lieu of regression. At 18 months, more modest effects were observed in male offspring (adjusted OR, 1.42 [95% CI, 0.92-2.11]; $P = .10$).

CONCLUSIONS AND RELEVANCE In this cohort study of offspring with SARS-CoV-2 exposure in utero, such exposure was associated with greater magnitude of risk for neurodevelopmental diagnoses among male offspring at 12 months following birth. As with prior studies of maternal

(continued)

 **Open Access.** This is an open access article distributed under the terms of the CC-BY License.

Figure 2. Multiple Logistic Regression Forest Plot of Neurodevelopmental Diagnosis Within 12 Months



Data Collection/Analysis

Perspectives/Cautions/Concerns/...

Data Collection/Analysis: Perspectives/Cautions/Concerns/...

- Statistical analysis: Overall idea
- Random sample > Inference about the population
 - Who is the population we want to understand?
 - Is our sampling reaching enough for the scope?
- “Random”:
 - Biased sample > Biased results
- Measurement:
 - How accurately are we measuring, realistically?
- All other variables/factors in play
 - Balance out as best as possible
- Causation
 - We can infer causation only from controlled experiments

(if time allows)

Linear Trend

From Simple Linear Regression



SCIENCE

LISTEN & FOLLOW



Humans traveled less during COVID restrictions. Animals traveled more

June 8, 2023 · 2:04 PM ET



Nell Greenfieldboyce



Mountain goats roamed the streets of Llandudno, Wales in March of 2020, as the COVID-19 lockdown kept people and tourists away.

Christopher Furlong/Getty Images

ANIMAL MOVEMENT

Behavioral responses of terrestrial mammals to COVID-19 lockdowns

Marlee A. Tucker^{1*}, Aafke M. Schipper¹, Tempe S. F. Adams², Nina Attias^{3,4}, Tal Avgar⁵, Natarsha L. Babic⁶, Kristin J. Barker⁷, Guillaume Bastille-Rousseau⁸, Dominik M. Behr^{9,10}, Jerrold L. Belant¹¹, Dean E. Beyer Jr.¹²,

movement speed or habitat select roads (7), and altered diurnal pattern itat use (8). In addition to these landscape modification effects, animals can react to the presence and activity of humans. These often are perceived as a risk (9) can lead to changes in habitat use due avoidance of areas heavily used by

COVID-19 lockdowns in early 2020 reduced human mobility, providing an opportunity to disentangle its effects on animals from those of landscape modifications. Using GPS data, we compared movements and road avoidance of 2300 terrestrial mammals (43 species) during the lockdowns to the same period in 2019. Individual responses were variable with no change in average movements or road avoidance behavior, likely due to variable lockdown conditions. However, under strict lockdowns 10-day 95th percentile displacements increased by 73%, suggesting increased landscape permeability. Animals' 1-hour 95th percentile displacements declined by 12% and animals were 36% closer to roads in areas of high human footprint, indicating reduced avoidance during lockdowns. Overall, lockdowns rapidly altered some spatial behaviors, highlighting variable but substantial impacts of human mobility on wildlife worldwide.

In 2020, governments around the world introduced lockdown measures in an attempt to curb the spread of the novel severe acute respiratory syndrome coronavirus 2 (SARS CoV-2) virus. This resulted in a drastic reduction in human mobility including human confinement to living quarters, closure of recreation and protected areas, and reductions in the movement of vehicles and their associated by-products (e.g., noise and pol-

lutants) (1). This “anthropause” provides a unique opportunity to quantify the effects of human mobility on wildlife by decoupling these from landscape modification effects (e.g., roads) (2, 3). It is established that anthropogenic landscape modifications affect how animals use habitats (4) and interact with each other (5). For example, human infrastructure may induce various behavioral responses in animals, including avoidance (6), shifts in

Fig. 4. Changes in animal distance to roads during the COVID-19 lockdowns.

Decreasing distance to roads in response to the human footprint index (HFI). Colored points represent individuals ($n = 2160$), with point size proportional to the inverse sampling variance of the response ratio for each individual. The black line is the predicted effect size (response ratio; RR). The shaded area indicates 95% CI, and the dashed gray line at zero illustrates no change. Negative values indicate closer proximity to roads during the early 2020 lockdowns, whereas positive values indicate increased distance from roads during the lockdowns.

