Data Analysis

Aaron R. Caldwell

07 July, 2021

Introduction

In this study we collected data on 300 sport and exercise science research articles (100 from 3 journals). Based on the work of Büttner et al. (2020), we anticipated at least 150 (50%) of the articles would include a hypothesis that was tested. Based on the work of Fanelli (2010), Scheel, Schijen, and Lakens (2020), and Büttner et al. (2020) we hypothesized that the percentage of articles that find support for their hypothesis was greater than 80%.

Hypothesis

For this study, we hypothesized that the rate of positive results (i.e., studies that find at least partial support for their hypothesis) was greater than 80%. Therefore, the null hypothesis (H_0) was that the proportion of positive results was less than .8 and our alternative was greater than .8. There was no other effect being estimated in this study therefore the intercept of the model is what will be tested.

$H_0 : Intercept \leq 0.8$

$H_1 : Intercept > 0.8$

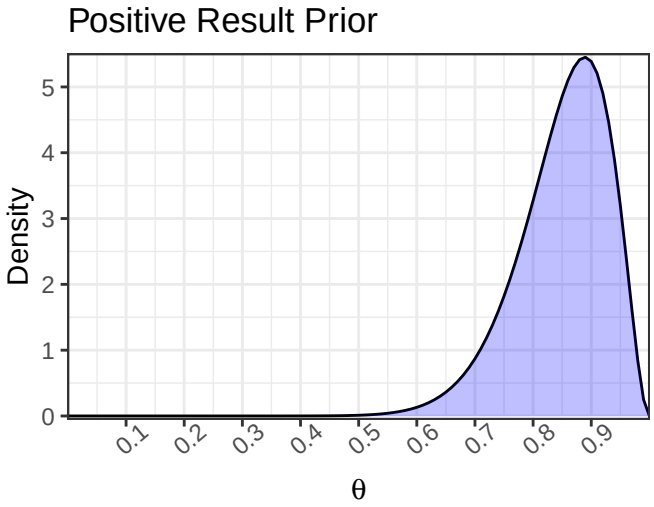
We also hypothesized that more than 60% of studies would test a hypothesis.

$H_0 : Intercept \leq 0.6$

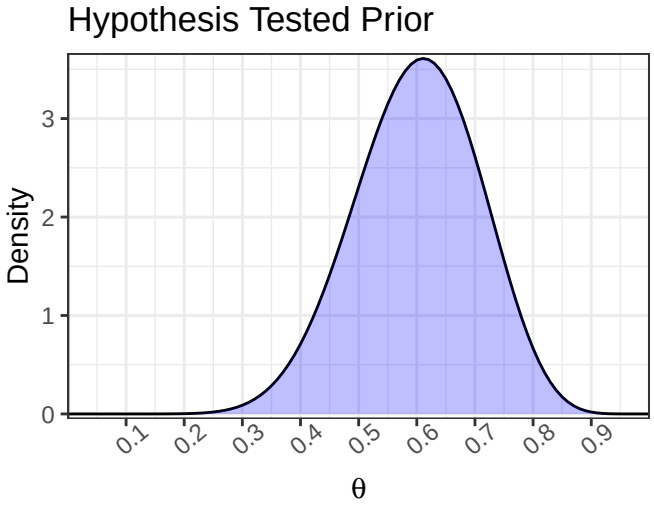
$H_1 : Intercept > 0.6$

Prior Choice

The prior we selected for this analysis was informed by the previous studies assuming the true positive rate is approximately 85% (Fanelli 2010). However, we would like to avoid "spiking" the prior in favor of our hypothesis and therefore want a skeptical prior. Based on the work of Scheel, Schijen, and Lakens (2020) and Büttner et al. (2020) the estimated positive rates in original research investigations ranged from 82%-92%, and even some fields included in the survey by Fanelli (2010) observed rates as low as ~70%. Therefore, we selected a prior of $\beta(17, 3)$, and is visualized it below. This prior is centered around .85, but includes the possibility of higher (.9) and much lower (.7) proportions as compatible parameter estimates.



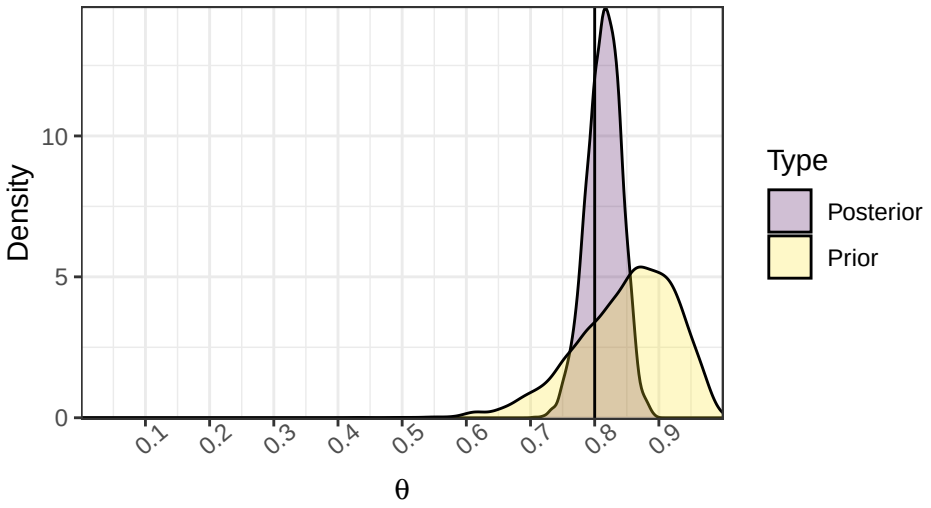
Similarly, we used prior centered at 60% for the secondary hypothesis test.



Hypothesis Test Analyses

Positive Result Rate Analysis

Below, we have incorporated this prior (`prior_1`) and then analyzed this with the `brm` function (saved as `m_final`).



In addition, the hypothesis was tested with the `hypothesis` function and the posterior compatibility intervals (C.I.).

Table 1: Hypothesis Test

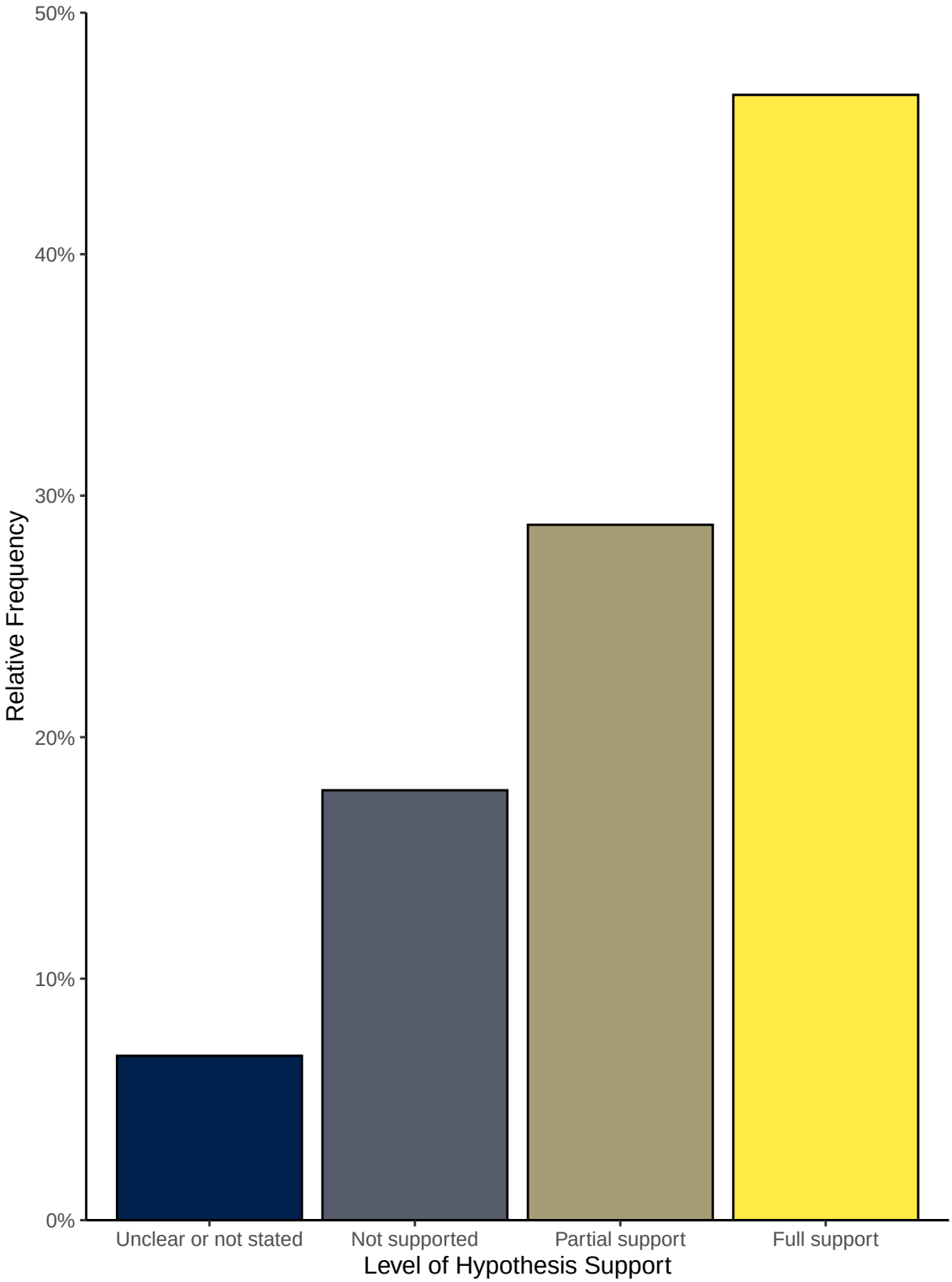
Hypothesis	Estimate	Est.Error	CI.Lower	CI.Upper	Evid.Ratio	Post.Prob	Star
(Intercept)-(0.8) > 0	0.0142668	0.0269525	-0.0319772	0.0568162	2.427592	0.70825	

Table 2: 95% Posterior C.I.

	2.5%	97.5%
b_Intercept	0.7577946	0.8629953
prior_b	0.6781015	0.9662099
lp___	-5.5422325	-3.1629022

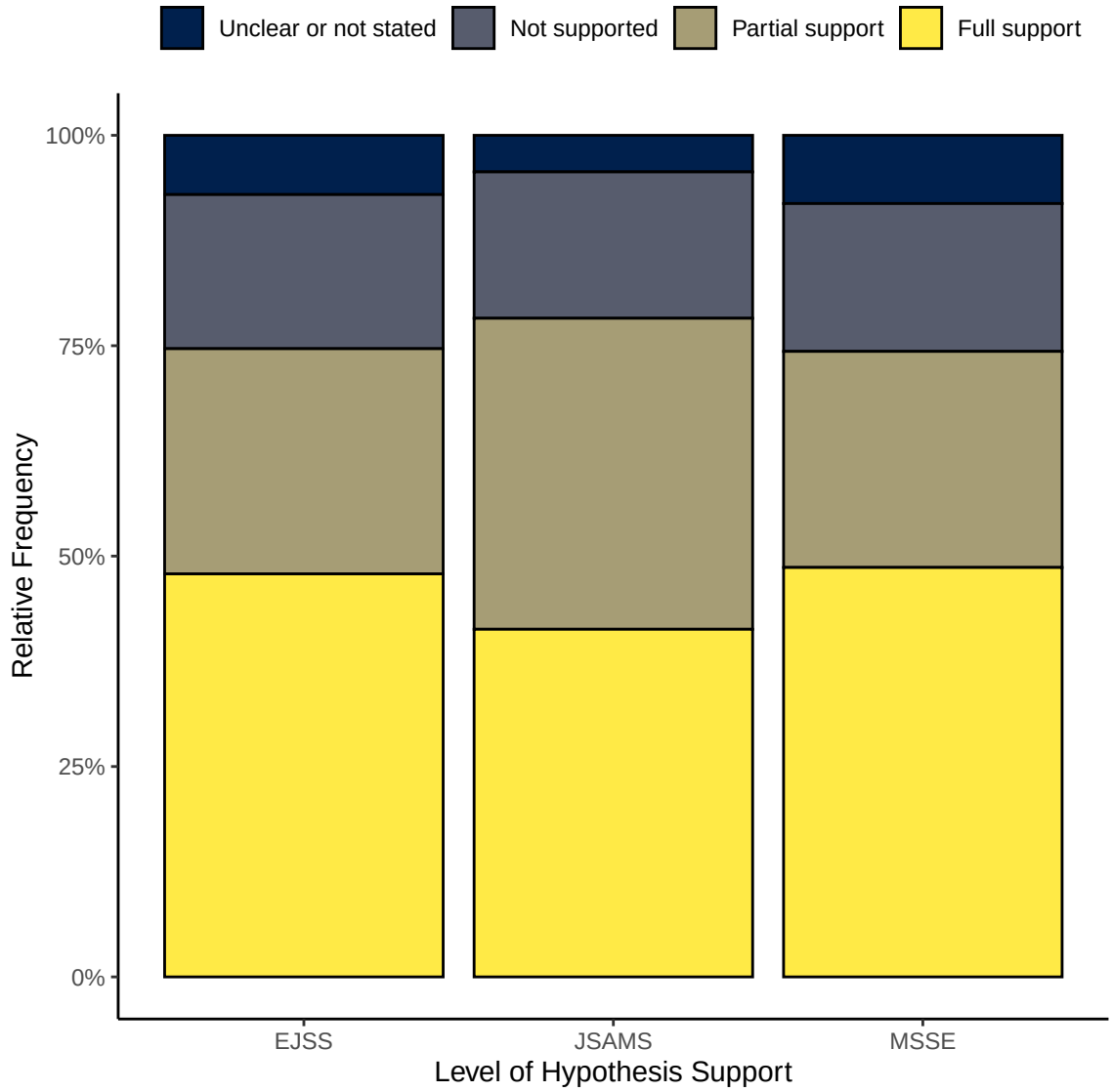
From the simulated scenario we find that given the data the hypothesis that the true positive result rate is greater than 80% is 2.43 times more likely than the true value being less than 80%.

Level of Support Reported



Positive Results by Journal

We can also break down the positive results by journal.



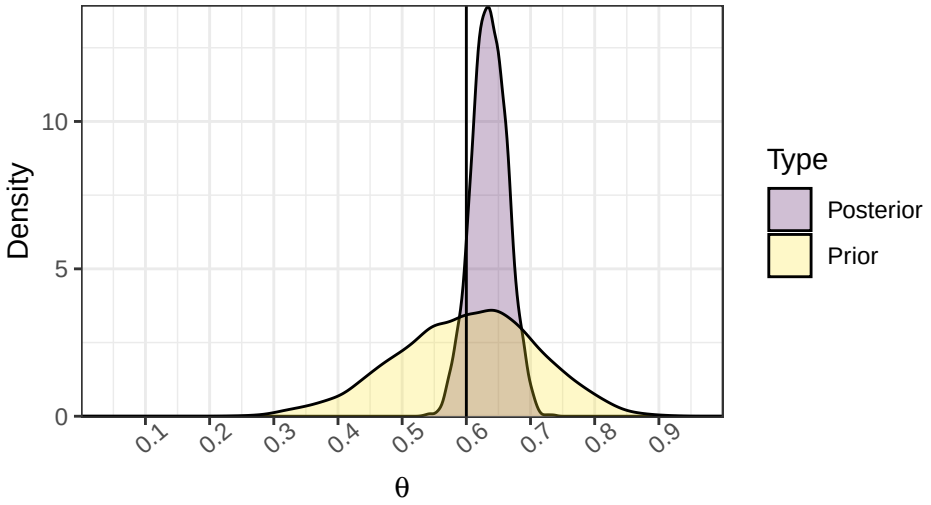
We can also tabulate the "support" by journal.

	Unclear or not stated	Not supported	Partial support	Full support
EJSS	5	13	19	34
JSAMS	2	8	17	19
MSSE	6	13	19	36

Overall, there were no discernible differences in the degree of support reported in the three journals, $\chi^2(6) = 2.4$, $p = 0.88$.

Hypothesis Tested Analysis

Below, we have incorporated this prior (`prior_1`) and then analyzed this with the `brm` function (saved as `m_final`).



In addition, the hypothesis was tested with the `hypothesis` function and the posterior compatibility intervals (C.I.).

Table 4: Hypothesis Test

Hypothesis	Estimate	Est.Error	CI.Lower	CI.Upper	Evid.Ratio	Post.Prob	Star
(Intercept)-(0.6) > 0	0.0358125	0.0276545	-0.0107656	0.0814775	9.723861	0.90675	

Table 5: 95% Posterior C.I.

	2.5%	97.5%
b_Intercept	0.5811895	0.6897217
prior_b	0.3766667	0.8015812
lp___	-5.9064835	-3.2418816

From the simulated scenario we find that given the data the hypothesis that the true positive result rate is greater than 80% is 9.72 times more likely than the true value being less than 80%.

Comparing Hypothesis Testing by Journal

While there were no discernible differences in the degree of support reported in the three journals, there was a difference in the rate at which hypothesis testing was reported.

	No	Yes
EJSS	29	71
JSAMS	54	46
MSSE	26	74

$\chi^2(2) = 20.43$, $p = 3.7 \times 10^{-5}$

Descriptives Statistics

Hypothesis Testing and Statistics

Contingency tables and binomial probabilities can be created for other data collected in this study of published studies.

Table 7: Statistics on Hypothesis Testing

		Hypothesis Tested	
	All	No	Yes
Total	278	93 (100)	185 (100)
Significance Testing			
No	8 (2.9)	4 (4.3)	4 (2.2)
Yes	270 (97.1)	89 (95.7)	181 (97.8)
Effect Size Reported			
No	49 (17.6)	19 (20.4)	30 (16.2)
Yes	229 (82.4)	74 (79.6)	155 (83.8)

Table 8: Statistics on Significance Testing

		Significant p-value	
	All	No	Yes
Total	268	4 (100)	264 (100)
Effect Size Reported			
No	46 (17.2)	2 (50)	44 (16.7)
Yes	222 (82.8)	2 (50)	220 (83.3)
p-value Type			
A mix: some exact and relative p-value(s)	149 (55.6)	1 (25)	148 (56.1)
Exact (e.g., p = .049)	91 (34)	2 (50)	89 (33.7)
Relative (e.g., p < .05)	28 (10.4)	1 (25)	27 (10.2)

- Effect sizes were reported often with 79.33% [74.3, 83.77] of manuscripts reported some estimate of effect size.
- 98.51% [96.22, 99.59] of manuscripts reported a "significant" p-value.
- 33.7% [28.09, 39.68] of manuscripts reported exact p-values (p = .045) versus relative p-values (p < .05).
- 89.63% [85.36, 93] of manuscripts reported at least *some* exact p-values (e.g., p = .045) versus relative p-values (e.g., p < .05).

Preregistration

Recently, there has been an emphasis on trial preregistration. This should be a requirement for clinical trials and any randomized control trial (RCT), and is highly encouraged for animal studies. Overall, we estimate that most manuscripts are out of compliance for clinical trials and no animal studies included preregistration information

Table 9: Preregistration Reported

		Preregistration	
	All	No	Yes
Total	300	273 (100)	27 (100)
Clinical Trial			
No	260 (86.7)	256 (93.8)	4 (14.8)
Yes	40 (13.3)	17 (6.2)	23 (85.2)
RCT			
No	236 (78.7)	227 (83.2)	9 (33.3)
Yes	64 (21.3)	46 (16.8)	18 (66.7)
Animal Study			
No	291 (97)	264 (96.7)	27 (100)
Yes	9 (3)	9 (3.3)	0 (0)

- 9% [6.01, 12.82] of manuscripts reported preregistration or clinical trial registration information.

Sample Size Information

All current guidelines require sample size reporting and encourage sample size justification. However, very few studies actually included a sample size justification. We should note that a sample size justification *does not* mean a power analysis was reported.

Table 10: Sample Size Justification

		Journal		
	All	EJSS	JSAMS	MSSE
Total	300	100 (100)	100 (100)	100 (100)
Sample Size Justification				
No	232 (77.3)	81 (81)	86 (86)	65 (65)
Yes	68 (22.7)	19 (19)	14 (14)	35 (35)
Sample Size Reported				
No (sample size(s) unclear/unreported)	5 (1.7)	0 (0)	4 (4)	1 (1)
Partial (some information missing)	2 (0.7)	0 (0)	1 (1)	1 (1)
Yes (e.g., total and group sample sizes provided)	293 (97.7)	100 (100)	95 (95)	98 (98)

- 22.67% [18.05, 27.83] of manuscripts reported some form of sample size justification.
- 97.67% [95.25, 99.06] of manuscripts reported all the required sample size information.

Table 11: Sample Sizes by Category

sci_cat	median_n	IQR_n	max	min	count
Applied exercise physiology (human)	21	31	3272	4	89
Basic physiology (animal and cell physiology)	33	25	60	12	9
Biomechanics	20	23	222	2	46
Clinical research	40	63	1202	1	29
Environmental physiology (heat, cold, & altitude)	16	15	36	5	10
Epidemiology	920	3619	44078	63	26
Motor learning/control/behavior	47	170	539	13	19
Other	112	717	4937	10	27
Sport performance	36	84	8182	9	28
Sport/exercise psychology	40	130	1768	12	12

Table 12: ANOVA: Compare Sample Size

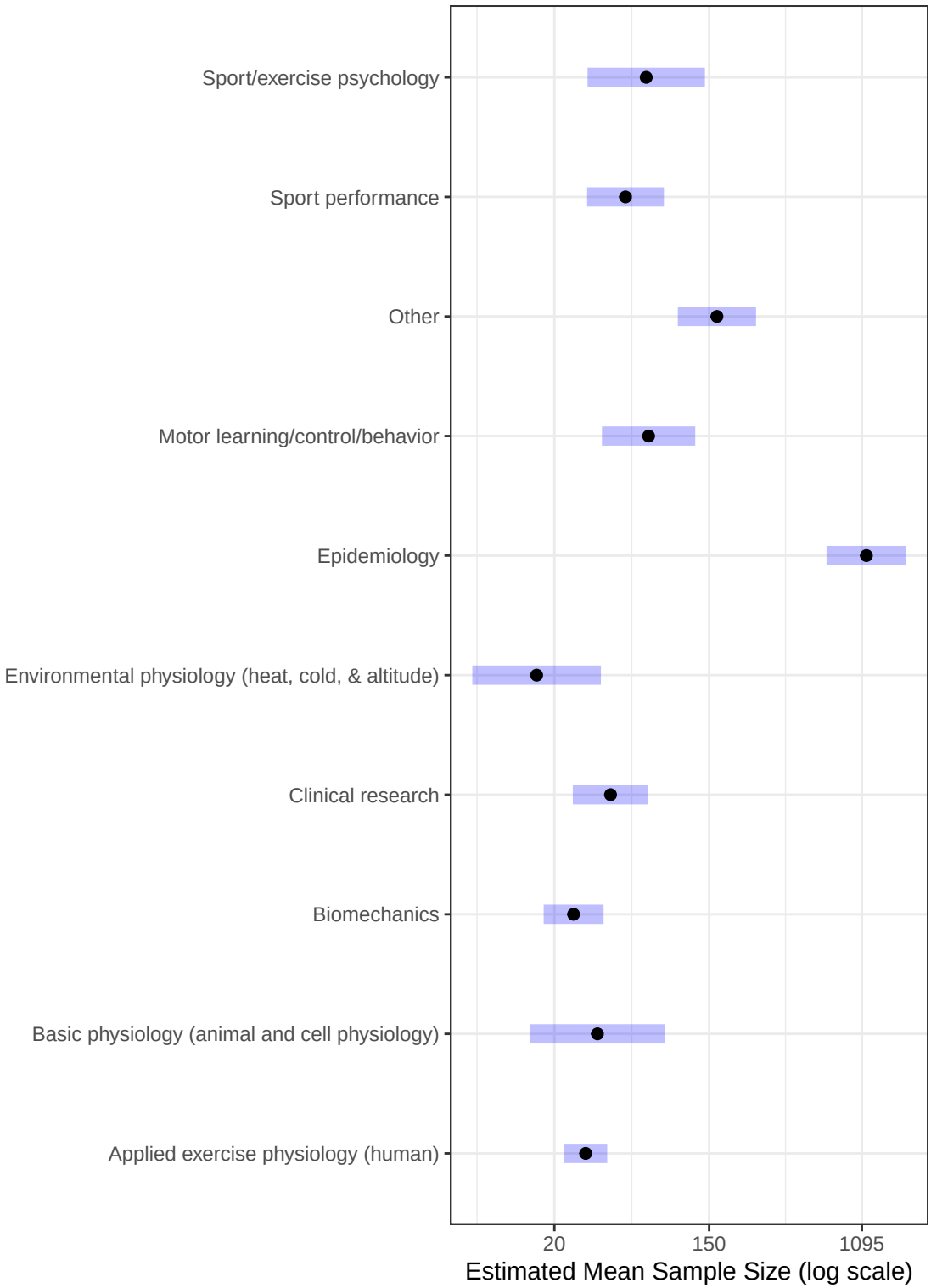
	num Df	den Df	MSE	F	ges	Pr(>F)
sci_cat	9	285	1.8086	21.811	0.4079	0

Further, we can illustrate the range of sample sizes with plots of sample sizes reported sample size and compare the sample sizes by sub-discipline. Overall, from the given data we would conclude that the sample sizes have a large range, and unsurprisingly epidemiology has the largest median sample size.

An ANOVA on the log transformed values of sample size indicate significant differences between disciplines in sample size.

Table 13: Estimated Means: Compare Sample Size

sci_cat	response	SE	df	lower.CL	upper.CL
Applied exercise physiology (human)	30	4	285	23	40
Basic physiology (animal and cell physiology)	35	16	285	14	85
Biomechanics	26	5	285	17	38
Clinical research	42	10	285	25	68
Environmental physiology (heat, cold, & altitude)	16	7	285	7	37
Epidemiology	1162	306	285	691	1952
Motor learning/control/behavior	68	21	285	37	125
Other	166	43	285	100	276
Sport performance	50	13	285	31	83
Sport/exercise psychology	66	26	285	31	142


Sport/exercise psychology
Sport performance
Other
Motor learning/control/behavior
Epidemiology
Environmental physiology (heat, cold, & altitude)
Clinical research
Biomechanics
Basic physiology (animal and cell physiology)
Applied exercise physiology (human)
20
150
1095
Estimated Mean Sample Size (log scale)

Other Open Science Practices

We can estimate the proportion of studies that report some type of open data or data sharing in their manuscript.

- 0.67% [0.08, 2.39] of manuscripts reported some form of data sharing or open data.

There is some concern that the replicability of kinesiology is not known so we can also report the proportion of studies that explicitly stating they are replicating previous work.

- 0% [0, 1.22] of manuscripts explicitly stated they were replicating a previous study.

References

Büttner, Fionn, Elaine Toomey, Shane McClean, Mark Roe, and Eamonn Delahunt. 2020. "Are Questionable Research Practices Facilitating New Discoveries in Sport and Exercise Medicine? The Proportion of Supported Hypotheses Is Implausibly High." *British Journal of Sports Medicine*, July, bjsports-2019-101863. https://doi.org/10.1136/bjsports-2019-101863.

Fanelli, Daniele. 2010. "'Positive' Results Increase down the Hierarchy of the Sciences." Edited by Enrico Scalas. *PLoS ONE* 5 (4): e10068. https://doi.org/10.1371/journal.pone.0010068.

Scheel, Anne M., Mitchell Schijen, and Daniel Lakens. 2020. "An Excess of Positive Results: Comparing the Standard Psychology Literature with Registered Reports," February. https://doi.org/10.31234/osf.io/p6e9c.