

Table for G*Power manuscript

Robert Thibault

today

Table 1. Estimated number of articles published between 1 Jan 2017 and 31 May 2022 that reference the G*Power software.

	PubMed Central (95% CI)	PubMed (95% CI)
Any power calculation	21000 (20000 - 22000)	47000 (46000 - 49000)
Sample size calculation	14000 (13000 - 16000)	32000 (28000 - 36000)
ANOVA sample size calculation	4000 (2000 - 5000)	8000 (5000 - 11000)

We excluded articles that discuss G*Power, but do not report using this software for a power calculation. The table includes rows for publications that report using G*Power for any power calculation related to any statistical test (Any power calculation), a power calculation for any statistical test that solves for sample size (Sample size calculation), and a power calculation for an ANOVA that solves for sample size (ANOVA sample size calculation). The total number of articles in each database from 1 Jan 2017 to 31 May 2022 is: PubMed Central 3,285,893; PubMed 7,318,980. Numbers are rounded to the nearest thousand to avoid suggesting a higher level of precision than this estimation provides.

Table 2. Types of power calculations.

Power calculation solving for	N=140	Percentage	(95% CI)
Sample size	94	67%	(59 - 75%)
Power	29	21%	(14 - 28%)
Effect size	13	9%	(5 - 15%)
Sample size (after completing the study)	5	3%	(1 - 7%)
Unsure	3	2%	(<1 - 5%)

The sum of the first column is greater than n = 140 because four articles performed multiple power calculations that solved for different elements. We only coded *Sample size (after completing the study)* when an article unambiguously described a calculation for a future study.

Table 3. Reproducibility of sample size calculations from a random sample of publications referencing G*Power

Transparency element	N=94	Percent	(95% CI)
Reproducible			
Yes, without assumptions	2	2%	(<1 - 6%)
Likely, with assumptions	27	29%	(20 - 38%)
No	65	69%	(59 - 78%)
All 6 elements reported†			
Yes	13	14%	(8 - 21%)
No	81	86%	(78 - 92%)
Alpha			
0.05	80	85%	(77 - 92%)
Other	2	2%	(<1 - 6%)
Not reported	12	12%	(7 - 20%)
Power			
0.80	52	55%	(45 - 65%)
0.95	16	17%	(10 - 25%)
Other	18	19%	(12 - 28%)
Not reported	8	8%	(4 - 15%)
Effect size type			
d	14	15%	(8 - 23%)
f	15	16%	(9 - 24%)
Non-standardized	11	11%	(6 - 19%)
Other	8	8%	(4 - 15%)
Not reported	46	49%	(39 - 59%)
Effect size value			
Reported	75	80%	(71 - 87%)
Not reported	19	20%	(13 - 29%)
Statistical test			
ANOVA	15	16%	(9 - 24%)
t-test	12	13%	(7 - 20%)
Other	6	6%	(2 - 12%)
Not reported	61	65%	(55 - 74%)
Sample size‡			
Reported	90	96%	(91 - 99%)
Not reported	4	4%	(1 - 9%)

*We use the term 'likely' because we cannot be certain that all of our assumptions were correct. †These six elements are: alpha, power or beta, effect size type, effect size value, statistical test, and sample size. We considered statistical test reported if they named the general test, even if details were missing (e.g., reporting an ANOVA, but not what type of ANOVA). ‡The median sample size was 55 (IQR: 28 to 105).

Supplementary Table 1. Article characteristics

Article characteristic	N=94	Percent
Unit of study		
Human	84	89%
Non-human animal	10	11%
Year of publication		
2017	4	4%
2018	7	7%
2019	20	21%
2020	29	31%
2021	22	23%
2022	12	13%
Protocol		
Yes	8	9%
No	86	91%
Multiple sample size calculations		
Yes	4	4%
No	90	96%

The median journal impact factor was 2.5 (range: 1.4 to 11). 18 articles were published in journals that did not have an impact factor.

Table 4. Measures of quality of the sample size calculations

Quality measure	N=94	Percent	(95% CI)
Analysis match			
Yes	29	31%	(22 - 40%)
No	13	14%	(8 - 21%)
Unsure	45	48%	(38 - 58%)
NA (protocol)	7	7%	(3 - 14%)
Error			
Yes	8	8%	(4 - 15%)
No	15	16%	(9 - 24%)
Unsure	71	76%	(66 - 84%)
Adjusted for multiple comparisons			
Yes	0	0%	(NA)
No, and multiple analyses are performed	71	76%	(66 - 84%)
No, but a single outcome was identified	21	22%	(15 - 31%)
Unsure	2	2%	(<1 - 6%)
Justification for chosen effect size			
Previously published research	24	25%	(17 - 35%)
Effect size conventions	24	25%	(17 - 35%)
General reference to another study	9	9%	(5 - 16%)
Pilot data	5	5%	(2 - 11%)
Effect size of interest	4	4%	(1 - 9%)
Other	3	3%	(1 - 7%)
No justification reported	32	34%	(25 - 44%)

Some articles provided more than one justification for their chosen sample size, and thus the sum of the percentages exceeds 100%.

Supplementary Table 2. Inter-rater agreement

Variable	Cohen's k	Percent agreed	n disagreed	n total	n categories
id	1.00	100%	0	146	146
pmcid	1.00	100%	0	146	146
protocol	0.65	97%	5	146	2
include	0.72	98%	3	146	2
participants	0.82	96%	5	140	2
journal	0.73	74%	37	140	142
pub_year	0.95	96%	6	140	7
impact_factor	0.82	83%	24	140	112
power_calc_type	0.68	84%	23	140	10
multiple	0.78	89%	15	140	3
version	0.95	98%	2	92	3
version_text	0.99	99%	1	92	13
power	0.92	99%	1	92	3
power_text	0.90	94%	5	85	13
alpha	0.84	97%	3	92	3
alpha_text	1.00	100%	0	80	4
sample_size	0.21	92%	7	92	3
sample_size_text	0.88	88%	10	82	70
effect_size_type	0.84	89%	10	92	10
effect_size_value	0.80	93%	6	92	3
effect_size_value_text	0.95	95%	3	65	34
stat_test	0.70	85%	14	92	7
reproducible	0.24	71%	27	92	4
justification	0.48	61%	36	92	15
just_previous	0.40	82%	17	92	2
just_pilot	0.48	96%	4	92	2
just_convention	0.75	90%	9	92	2
just_mcid	0.39	97%	3	92	2
just_none	0.56	78%	20	92	2
just_ref	0.21	88%	11	92	2
just_other	-0.02	96%	4	92	2
mult_compare	0.24	66%	31	92	4
anova_within_between	0.53	68%	29	92	6
match	0.46	66%	31	92	5
error	0.25	72%	26	92	4

The variables are listed as they appear in the open data. See the data dictionary for a description of each variable. Cohen's kappa is mostly irrelevant for variables with a large number of categories, and can be ignored. Not all variables were relevant for all articles we coded; thus, 'n total' differs among the variables. 'justification' was coded as a radio question with 7 options. We re-coded this variable into 7 binary variables and calculate the inter-rater agreement for each one.

Table 100. ANOVA outcomes

Option used	N=37	Percent	(95% CI)
Non-default option	3	7%	(2 - 19%)
Default option	18	49%	(33 - 64%)
Unsure	16	43%	(28 - 59%)

temp

Table 101. REC results

Flowchart item	US (n)	UK (n)
Institutions sampled	20	15
Public forms identified	3	4
Institutions contacted	17	10
Public forms identified after contact	4	4
Private forms provided	3	2
Forms indicated as not available to share	4	4
No response	6	0
Sample size	9	10
Justification	6	6
Calculation	2	1
Calculation (indirectly)	2	1

temp