

**Assessing the proportion of supported hypotheses in sport and exercise medicine research:  
a cross-sectional, meta-research study.**

**Supplemental File**

**P-Hacking through the Garden of Forking Paths (without leaving a trace):**

A multiple regression model containing 20 independent binary variables (i.e., two levels per variable) will produce >1 million effect estimates by including the values of each variable in all possible combinations:

$$2^{20} = 1,048,576$$

**Cherry-picking in the orchard of statistically significant outcomes:**

The probability of one outcome variable demonstrating a statistically significant result ( $p < 0.05$ ) by chance is 5%. However, the probability that one outcome variable out of thirteen tested outcome variables will achieve statistically significant result, by chance, is 49% if not corrected for multiplicity:

$$1-(1-0.05)^{13} = 49\%$$

**METHODS**

**Search strategy:**

We searched PubMed to identify all content that was published (in print or online ahead of print) in *BJSM*, *Sports Medicine*, *AJSM*, and *JOSPT* between January 1<sup>st</sup>, 2019 and May 31<sup>st</sup>, 2019 using the following search strategy:

*"Br J Sports Med" [Journal] OR Sports Medicine [Journal] OR "Am J Sports Med" [Journal] OR "J Orthop Sports Phys Ther" [Journal]*

We applied a 'Publication Dates' filter in PubMed to include content published (in print or online ahead of print) from January 1<sup>st</sup>, 2019 until May 31<sup>st</sup>, 2019:

*"Br J Sports Med"[Journal] OR ("Sports Med"[Journal] OR "Curr Sports Med Rep"[Journal]) OR "Am J Sports Med"[Journal] OR "J Orthop Sports Phys Ther"[Journal] AND ("2019/01/01"[PDAT] : "2019/05/31"[PDAT])*

**Identification:**

The research field of Sport and Exercise Medicine is ambiguously defined, which complicates the identification of relevant research articles. We identified original research studies published in the three highest-ranked Impact Factor Sport and Exercise Medicine journals as per the Journal Citation Reports 'Sports Sciences' Category. We also identified the highest-ranked Impact Factor specialist Sports Physical Therapy and Sports Physiotherapy journal in the Journal Citation Reports 'Sports Sciences' Category.

The *British Journal of Sports Medicine* has a 2018 Journal Impact Factor of 11.645 and publishes 2 issues per month. *Sports Medicine* has a 2018 Journal Impact Factor of 7.583 and publishes 1 issue per month. The *American Journal of Sports Medicine* has a Journal Impact Factor of 6.053 and publishes 1 issue per month. The *Journal of Orthopaedic & Sports Physical Therapy* has a 2018 Journal Impact Factor of 3.058 and publishes 1 issue per month.

Identified records were exported from PubMed and imported into a project folder in the reference management software, *Zotero*.

#### **Screening:**

Two assessors independently screened the title and abstracts of identified records, using the systematic review software *Rayyan*, to determine whether a record is an original research study. We excluded any study design or publication type that did not fulfil the criteria of an original research study (e.g., systematic review +/- meta-analysis, narrative review, educational review, consensus statement, editorial, commentary, research update). We also excluded case-reports, qualitative studies, animal studies, cadaveric studies, and cell studies.

#### **Inclusion:**

We screened full-text articles of original research studies to include original research studies that reported at least one study hypothesis. Original research studies that did not state a study hypothesis were excluded from data extraction.

#### **Data extraction:**

Two study assessors independently extracted data from included studies. A third investigator independently arbitrated between-assessor disagreement.

We extracted meta-data relating to the number of stated study aims, and the number and nature of reported study hypotheses. If studies reported at least one study hypothesis, we extracted meta-data relating to the following;

- 1) the prevalence and number of stated study aims;
- 2) the number of stated reported study hypothesis;
- 3) primary study hypothesis – this was the first appearing study hypothesis;
- 4) whether the primary study hypothesis contained multiple (i.e.,  $\geq 2$ ) sub-hypotheses. We identified sub-hypotheses in primary study hypotheses to determine whether partial or full support was generated by the data for authors' hypotheses. We determined a primary study hypothesis to contain two or more sub-hypotheses if it reported more than one study hypothesis. For example, a study hypothesis may report more than one independent and/or dependent variables within the primary study hypothesis.
- 5) the number of sub-hypotheses within the primary study hypothesis;
- 6) whether the primary study hypothesis was an alternative or null study hypothesis;
- 7) whether the alternative primary study hypothesis was directional or not;
- 8) whether the study reported that the primary study hypothesis was "supported" or rejected by the results, and;

- 9) whether the supported study hypothesis was fully or partially supported by study results. Full support indicates that all components of the primary study hypothesis, including all study outcomes (if more than one), time-points (if more than one), and participant groups (if more than one), were supported by the data/results. Partial support indicates that at least one, but not all, components of the primary study hypothesis were supported by the data/results.

The levels of independent variable are detailed in a pre-specified study protocol that we uploaded to the Open Science Framework (<https://osf.io/u582r/>) prior to data collection. Shortly after initiating study data extraction, we eliminated the 'diagnostic accuracy' and 'predictive accuracy' categories from the study design variable due to substantial overlap of these study designs with observational study designs.

#### **Analysis:**

We present data as absolute frequencies and relative proportions of the denominator of relevant categories. We estimate the proportion of original research studies reporting at least one study hypothesis by dividing the proportion of original research studies that report at least one study hypothesis by all original research studies. We calculate the proportion of supported hypotheses by dividing the number of studies with supported hypotheses by all studies reporting at least one study hypothesis. We calculate 95% Confidence Intervals around each proportion value to estimate the probability that the true underlying population proportion parameter of supported hypotheses lies in a specified range. We constructed upper and lower 95% Confidence Interval boundaries around proportion estimates using the following formula:

$$95\%CI = p \pm z \times \sqrt{\frac{p(1-p)}{n}}$$

...where  $p$  is the proportion estimate of interest,  $n$  is the numeric denominator value of the calculated fraction, and  $z$  represents the  $z$ -value for a two-tailed 95% Confidence Interval (i.e., 1.96).

#### **Results:**

We identified 669 items that were published in *BJSM*, *Sports Medicine*, *AJSM*, and *JOSPT* between January 1<sup>st</sup>, 2019 and May 31<sup>st</sup>, 2019. We present the number of and proportion of identified published items and included original research studies with and without at least one reported hypothesis in Supplemental Table 1. We proportion of study designs of included original research studies reporting at least one study hypothesis in Supplemental Table 2. The number and proportion of stated study aims in original research studies reporting at least one study hypothesis are reported in Supplemental Table 3. The prevalence of sub-hypotheses reported in primary study hypotheses are presented in Supplemental Table 4.

#### **Limitations:**

Journals may or may not impose upon authors, through editorial guidelines, the requirement to specify hypotheses, thereby influencing estimates reported in the current study.

**Supplemental Table 1. Number of identified published items, and included original research studies with and without at least one reported hypothesis**

Journal	Studies identified, k (%)	Original research studies included, k (%)	Original research studies reporting $\geq 1$ study hypothesis, k (%)
<i>BJSM</i>	263 (39.3%)	51 (23.7%)	12 (9.3%)
<i>Sports Medicine</i>	95 (14.2%)	6 (2.8%)	3 (2.3%)
<i>AJSM</i>	241 (36.0%)	131 (60.9%)	97 (75.2%)
<i>JOSPT</i>	70 (10.5%)	27 (12.6%)	17 (13.2%)
<b>Total</b>	<b>669</b>	<b>215</b>	<b>129</b>

**Supplemental Table 2. Study design of included original research studies**

Study Characteristics	k	Proportion (95% CI)
<b>Original research studies:</b>	<b>129</b>	-
Randomized controlled trial	17	13.2% (7.3% - 19.0%)
Quasi-randomized controlled trial	4	3.1% (0.1% - 6.1%)
Observational (prospective cohort)	27	20.9% (13.9% - 28.0%)
Observational (retrospective cohort)	29	22.5% (15.3% - 29.7%)
Observational (case-control)	7	5.4% (1.5% - 9.3%)
Observational (cross-sectional)	22	17.1% (10.6% - 23.5%)
Case-series	21	16.3% (9.9% - 22.6%)
Other	1	0.8% (0.7% - 2.3%)
Unclear	1	0.8% (0.7% - 2.3%)

**Supplemental Table 3. Number of stated study aims in original research studies that report at least one hypothesis**

Study Characteristics	k	Proportion (95% CI)
<b>Original research studies reporting <math>\geq 1</math> study hypothesis</b>	<b>129</b>	-
Number of aims stated:		
One	97	75.2% (67.7% - 82.6%)
Two	23	17.8% (11.2% - 24.4%)
Three	6	4.7% (1.0% - 8.3%)
Four	2	1.6% (0.1% - 3.7%)
Five	1	0.8% (0.1% - 2.3%)

**Supplemental Table 4. Prevalence of sub-hypotheses in primary study hypothesis**

Study Characteristics	k	Proportion (95% CI)
<b>Original research studies reporting <math>\geq 1</math> study hypothesis</b>	<b>129</b>	-
Primary study hypothesis containing no sub-hypotheses	61	47.3% (38.7% - 55.9%)
Primary study hypothesis containing $> 1$ sub-hypothesis	68	52.7% (44.1% - 61.3%)
Two sub-hypotheses	37	54.4% (42.6% - 66.2%)
Three sub-hypotheses	22	32.4% (21.2% - 43.5%)
Four sub-hypotheses	5	7.4% (1.1% - 13.6%)
Five sub-hypotheses	3	4.4% (0.1% - 9.3%)
$> 5$ sub-hypotheses	1	1.5% (0.1% - 4.3%)

## Supplemental Box 1. Flow-diagram of study results

