Bootstrapping to evaluate accuracy of citation-based journal indicators

Andersen, Jens Peter; Haustein, Stefanie

Publication date:
2015

Document Version
Peer reviewed version

Citation for published version (APA):
Andersen, J. P., & Haustein, S. (2015). Bootstrapping to evaluate accuracy of citation-based journal indicators. Poster session presented at 15th International Society of Scientometrics and Informetrics Conference (ISSI 2015), Istanbul, Turkey.

Bootstrapping to evaluate accuracy of citation-based journal indicators

Jens Peter Andersen & Stefanie Haustein

Medical Library, Aalborg University Hospital, Aalborg, Denmark
Email: jepea@rn.dk Twitter: @ipoga

École de bibliothéconomie et des sciences de l'information, Université de Montréal, Montréal, Canada
Email: stefanie.haustein@umontreal.ca Twitter: @stefhaustein

Motivation

Despite criticism, ranking indicators are in demand.

Essential to provide estimates of indicator accuracy, robustness, stability and confidence.

This study uses bootstrapping to **test the stability of citation-based journal indicators** - recent as well as traditional.

Data

All clinical medicine records in WoS 2012:
34 NSF specialties -> 2,699 journals -> 362,556 records.

2-year citation window
c = raw citations
s = relative citations (specialty standardised)

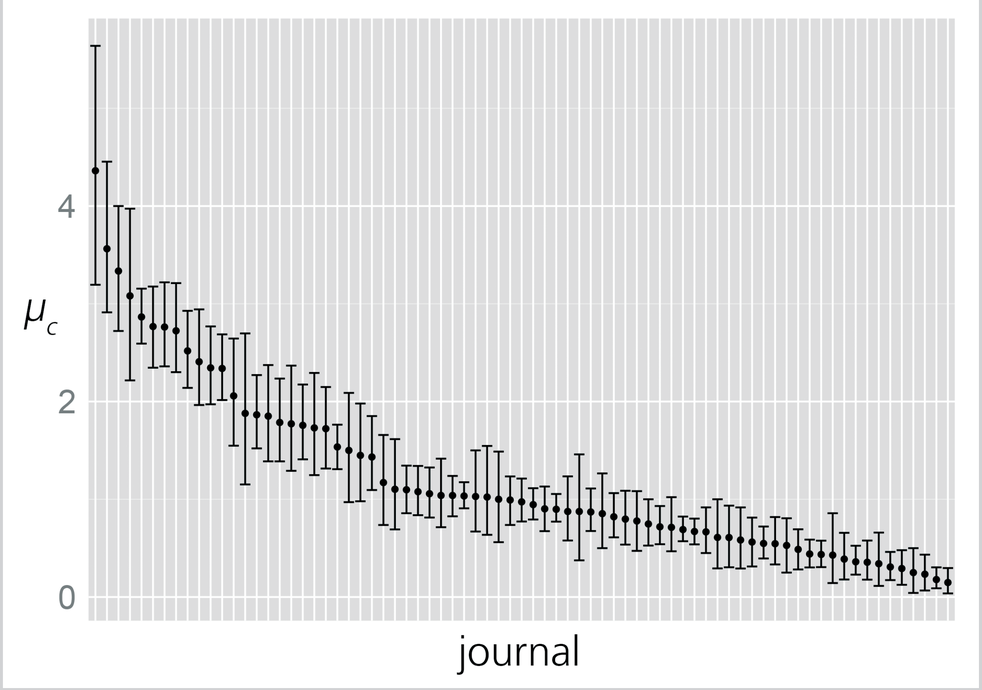

Figure 1: Mean raw citations per journal (data points) and bootstrapped stability intervals for dentistry journals.
Result: Bootstrapping identifies outlying scores. Stability intervals show the effect individual papers have on journal performance.

	All				≥50	
	Raw		Standardised			
Indicator	*mean*	*SD*	*mean*	*SD*	*mean*	*SD*
μ_c	2.321	3.897	1.000	1.679	1.052	1.261
M_c	1.477	2.278	1.000	1.543	1.079	1.471
μ_s	.835	1.107	1.000	1.326	1.053	1.076
M_s	.520	.717	1.000	1.381	1.075	1.297
N_{D10}	.081	.131	1.000	1.625	1.107	1.640
N_{SD10}	.078	.119	1.000	1.536	1.090	1.513

Table 1: Mean indicator values and standard deviations for all journals ("All") and journals publishing 50 or more papers ("≥50").
Result: All indicators are sensitive to sample sizes. N<50 journals have larger variance than N>50 journals.

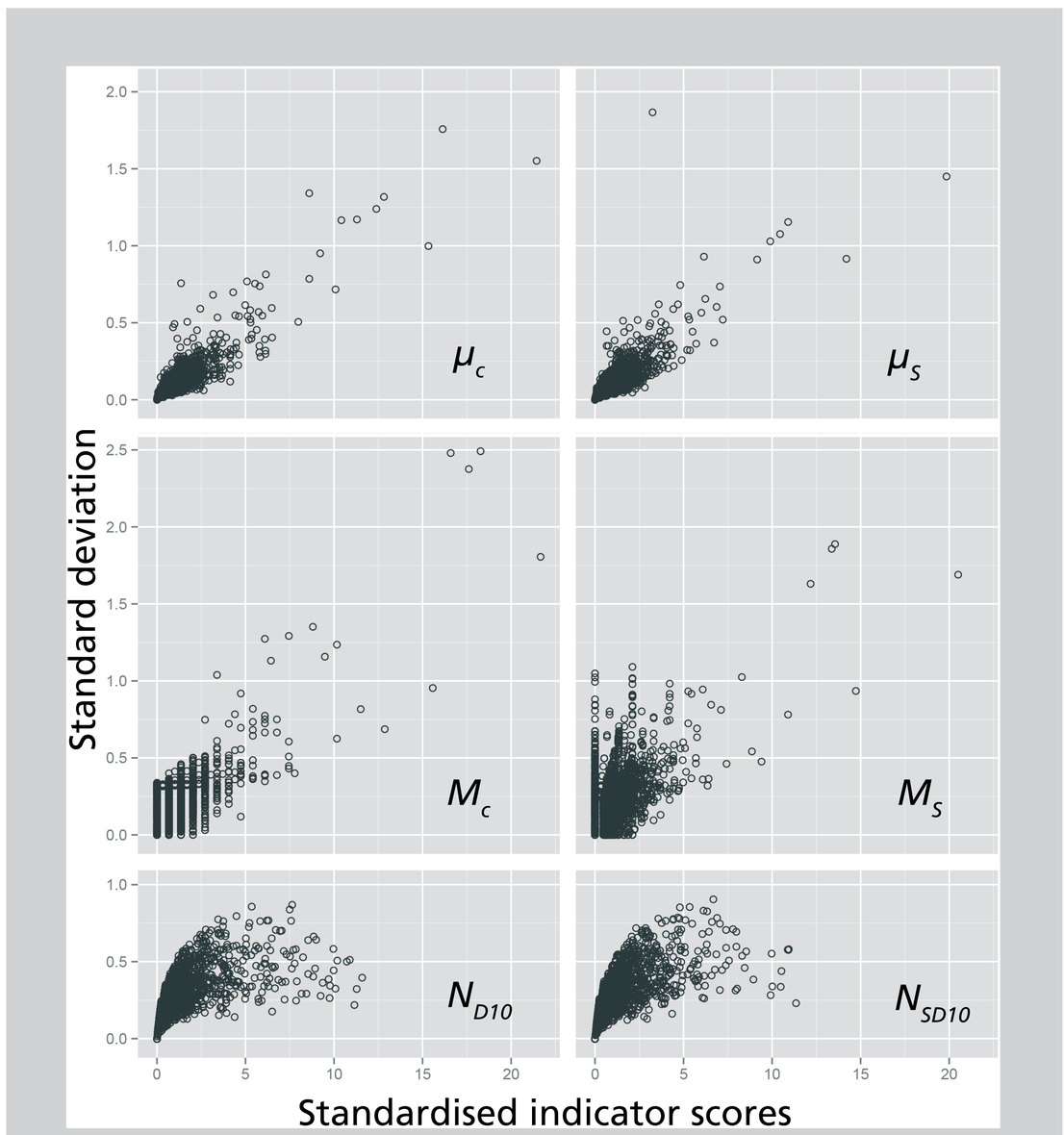

Figure 2: Standard deviation of standardised indicator scores per journal.
Result: Percentile-based indicators outperform mean- and median-based indicators with respect to stability. Median-based indicators perform worse than mean-based.

Indicators

μ_c and μ_s	mean raw and relative citations per paper.
M_c and M_s	median raw and relative citations per paper.
N_{D10} and N_{SD10}	top decile ratio of raw and relative citations.

Methods

Bootstrapping: Each sample (journal) is resampled 1,000 times, allowing calculation of stability data (95% confidence intervals).

Standardised (mean normalised) indicator scores used for comparison.

Results

See figure- and table-legends.

Further research

Additional indicators and specialty variations.

Université de Montréal

AALBORG UNIVERSITY HOSPITAL