

# Is Peer-Reviewing Worth the Effort?

Kenneth Church, Raman Chandrasekar, John E. Ortega and Ibrahim Said Ahmad

Northeastern University: {k.church, r.chandrasekar, j.ortega, i.ahmad}@northeastern.edu

<https://github.com/kwchurch/is-peer-reviewing-worth-the-effort>

## Abstract

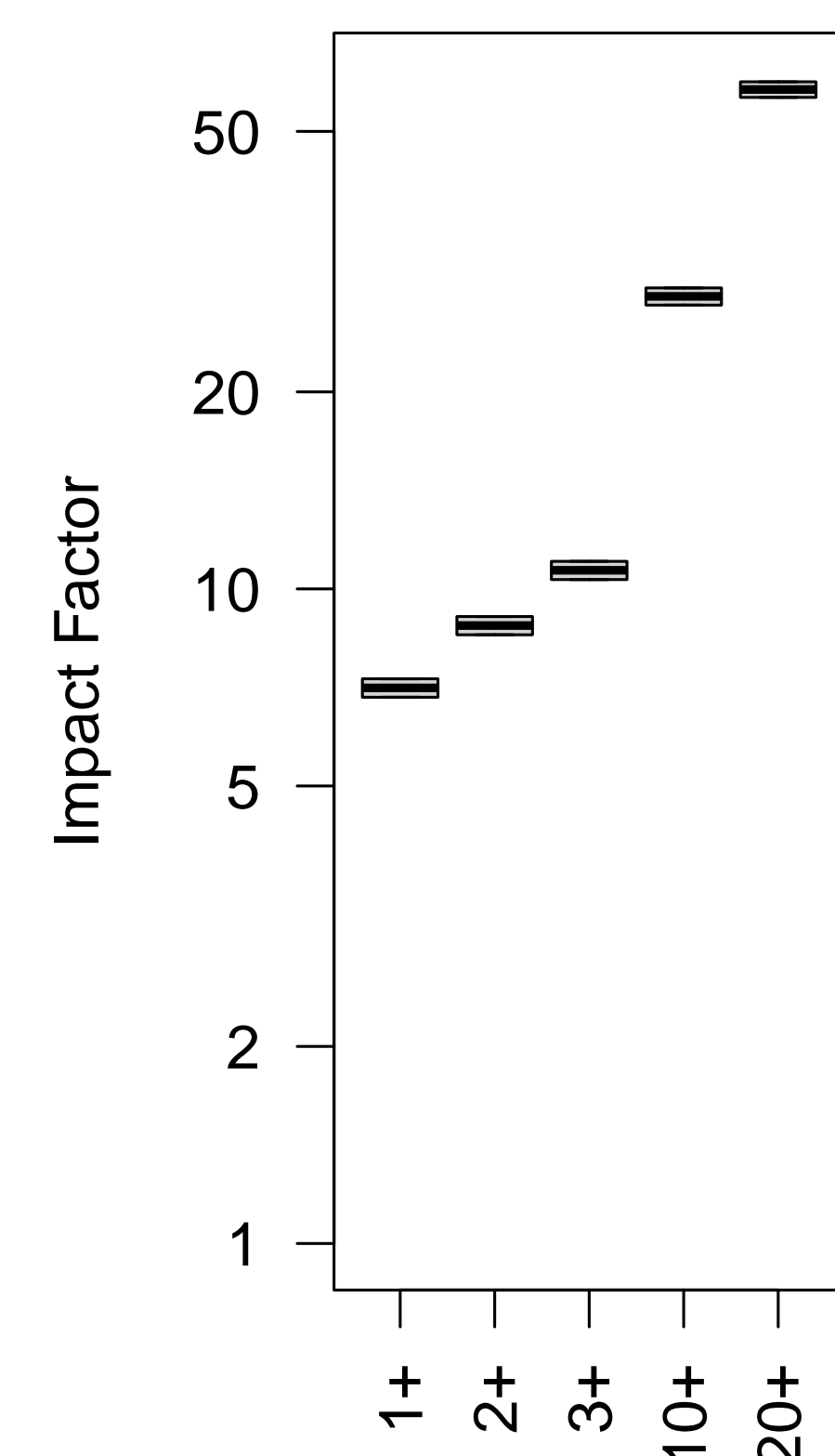
- How effective is peer-reviewing?  
Goal: Identify important papers
- Predict future citations from:
  - 1 Venue (conf, journal) and
  - 2 Citations soon after publication
- Conclusion: Early citations  $\gg$  Venue
- End with constructive proposal

3710 Papers in ACL Anthology

	2016	2017	2018	2019
ACL Conf	0.140	0.136	0.096	0.068
EMNLP	0.031	0.116	0.103	0.084
TACL	0.069	0.111	0.130	0.120
SemEval	0.036	-0.005	-0.026	-0.024
Workshops	-0.110	-0.104	-0.094	-0.077
1,121,081 Papers in PubMed, ArXiv or ACL				
ACL	0.0086	0.0109	0.016	0.015
ArXiv	0.0255	0.0088	0.024	0.021
PubMed	-0.0212	-0.0012	-0.014	-0.013

Table 3: These  $\rho$  are smaller than  $\rho$  in Table 2; it is easier to predict citations in one year from another, than from venue.

Early Citations



PubMed Venues

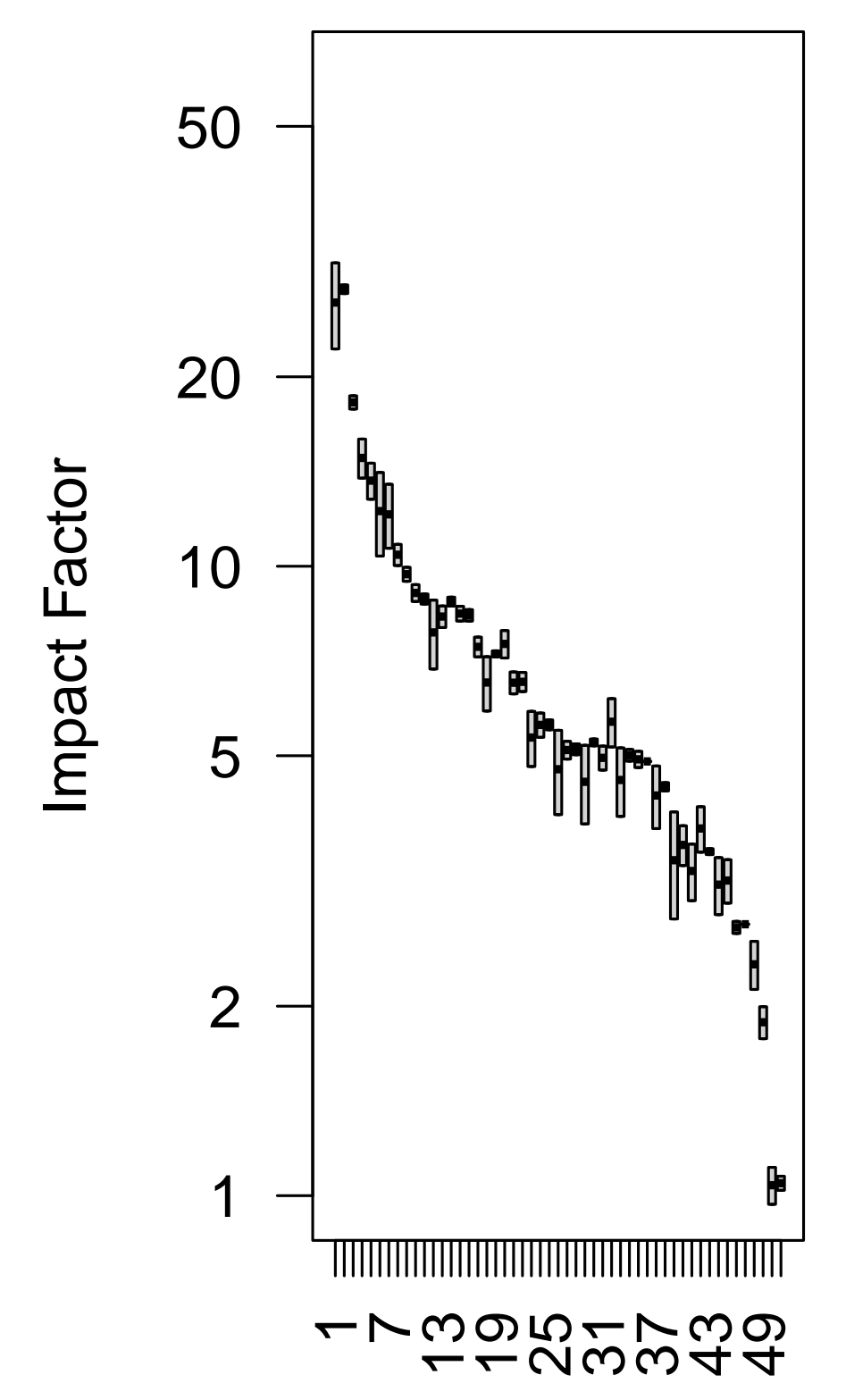
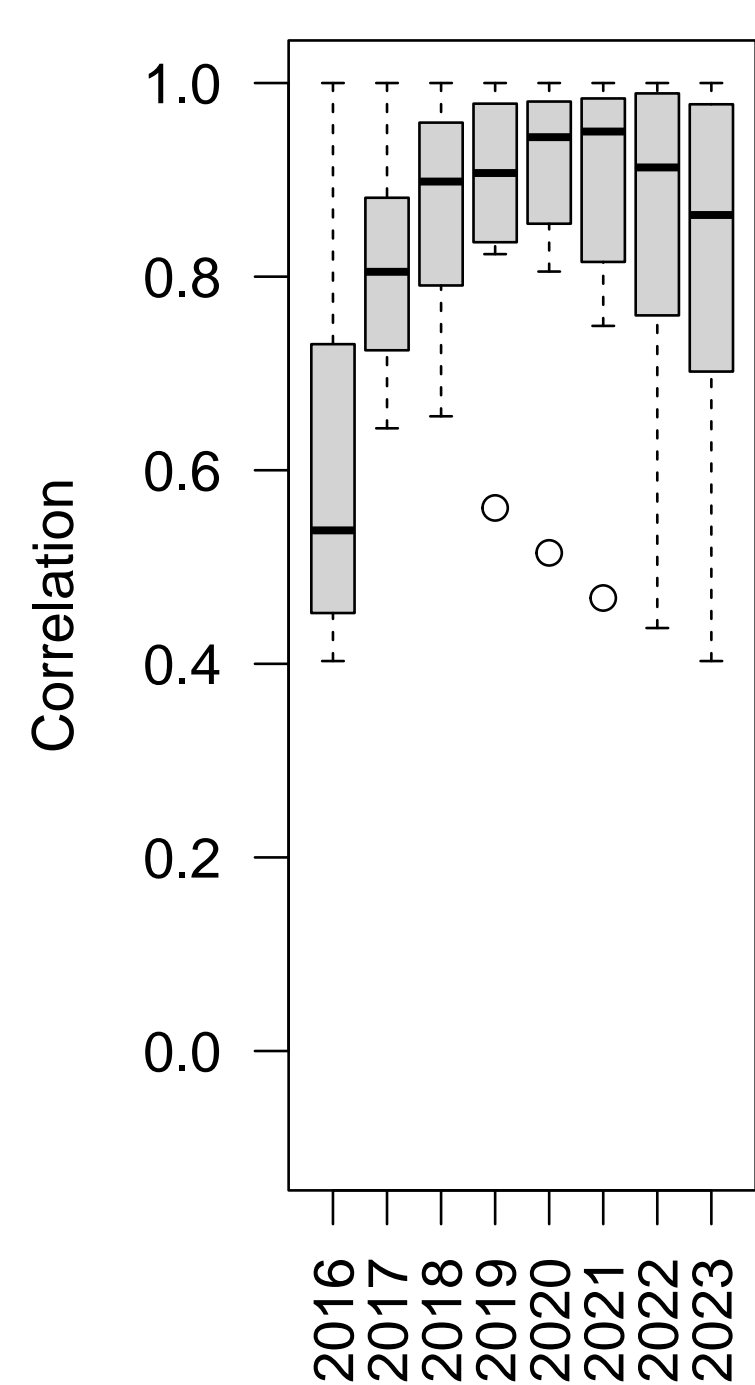


Figure 2: Left:  $\mu$  (impact) by early citations; Right:  $\mu$  (impact) by 50 venues in PubMed. Simple rule of thumb: for predicting future citations ( $\mu$  in future), it is better to select papers with 1+ early citations ( $\mu \approx 6.9$ ) than most of the 50 venues. Selecting papers with 20+ early citations ( $\mu \approx 61$ ) is better than the top venue ( $\mu$  for *Science*  $\approx 30$ ).

Large Correlations



Small Correlations

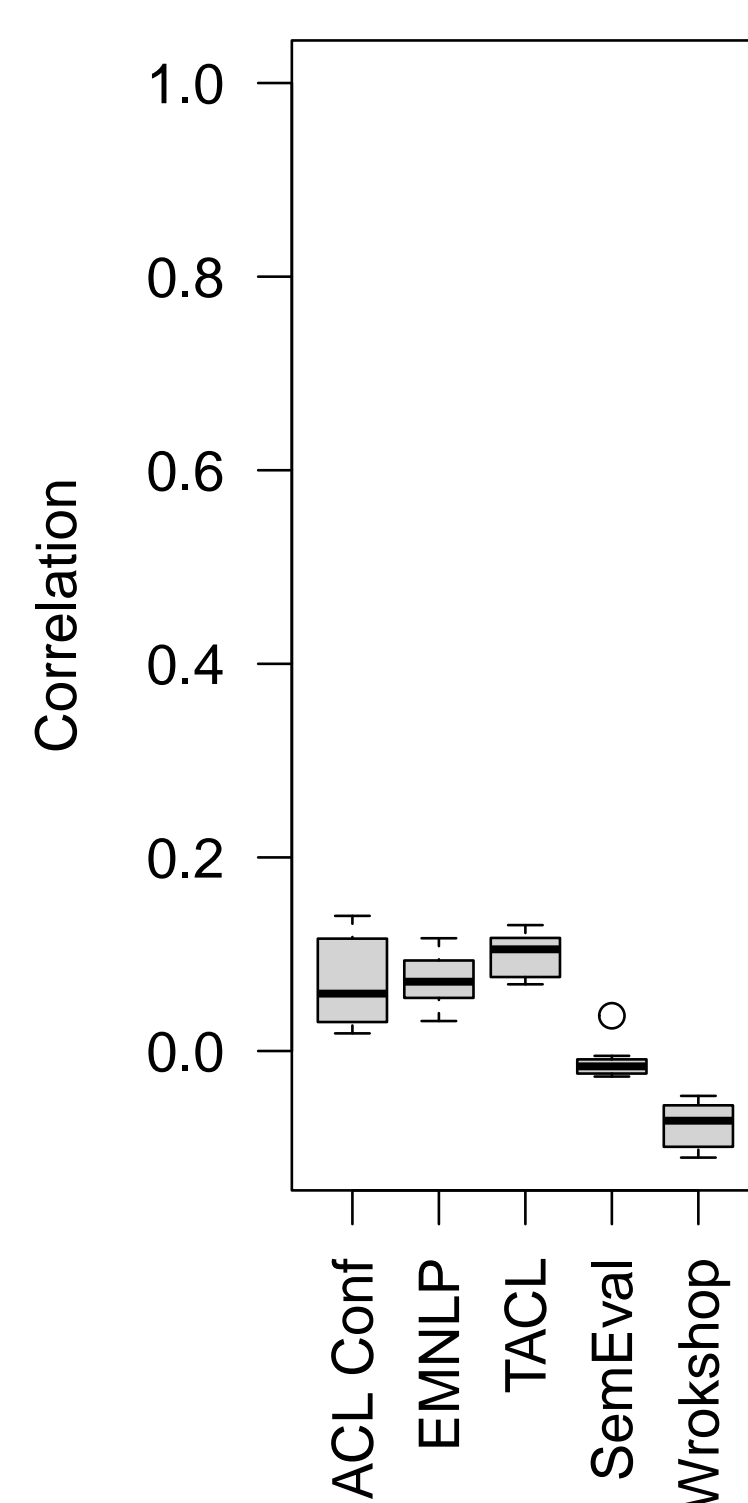


Figure 1: Early Returns (left)  $\gg$  Venue (right), based on  $\text{cor}(\rho)$  from Tables 2-3. Data: Semantic Scholar.

Venue	Id in S2	2016	2017	2018	2019	2020	2021
NAACL	9724599	5	7	5	1	3	1
LREC	12260053	0	0	0	1	0	0
LREC	28309452	2	8	4	10	7	7
EMNLP	1380793	0	2	16	19	17	19
COLING	18649702	0	1	2	1	3	1
SemEval	17378758	0	0	0	2	0	0

Table 1: Citation counts for a few ACL papers

	2016	2017	2018	2019	2020	2021
Data: 3710 ACL Papers Pub. in 2016						
2016	1.00	0.80	0.66	0.56	0.51	0.47
2017	0.80	1.00	0.92	0.85	0.81	0.75
2018	0.66	0.92	1.00	0.98	0.94	0.88
2019	0.56	0.85	0.98	1.00	0.98	0.93
2020	0.51	0.81	0.94	0.98	1.00	0.98
2021	0.47	0.75	0.88	0.93	0.98	1.00
Data: 1,026,798 PubMed Papers Pub. in 2016						
2016	1.00	0.77	0.64	0.55	0.50	0.45
2017	0.77	1.00	0.90	0.82	0.75	0.68
2018	0.64	0.90	1.00	0.94	0.89	0.83
2019	0.55	0.82	0.94	1.00	0.94	0.90
2020	0.50	0.75	0.89	0.94	1.00	0.95
2021	0.45	0.68	0.83	0.90	0.95	1.00

Table 2: It is easy to predict citations for a paper in year  $i$  from citations in year  $j$ . These  $\text{cor}(\rho)$  are based on Table 1. For papers pub. in 2016, we compute a vector of citations they had in year  $i$  and another vector for citations in year  $j$ .  $\rho$  is large, especially when  $i \approx j$ .

Group	h median	$\mu$	$\sigma$	N
PubMed, ArXiv and ACL Anthology				
0 citations	48	1	1.3	292,566
1+ citations	345	3	6.9	828,515
2+ citations	345	5	8.7	604,536
3+ citations	345	6	10.6	448,541
10+ citations	343	17	28.6	88,490
20+ citations	341	37	61.4	23,593
ACL Anthology	73	2	9.9	3710
ArXiv	236	2	6.4	101,176
PubMed	292	2	5.4	1,026,798

Deep Dive into ACL Anthology

0 citations	9	0	1.0	1.8	953
1+ citations	73	3	13.0	68.2	2757
2+ citations	73	4	17.1	79.2	2025
3+ citations	73	6	21.6	90.0	1550
10+ citations	73	23	57.0	155.6	481
20+ citations	71	54	114.9	235.6	190
ACL Main Conf.	42	5	18.7	45.2	377
EMNLP	41	6	25.8	78.9	269
TACL	17	11	70.5	280.3	45
SemEval	15	1	4.7	16.3	230
Workshops	24	1	3.8	10.1	1111

Table 4: A few early citations compare favorably to most venues. Early citations are based on first year after publication; scores ( $h, \mu$ ) are based on 4th year after pub. Note ArXiv is better than PubMed in terms of  $\mu$ .

## Conclusions

Early citations are more predictive of future citations than venue. Consequences:

- 1 **Exclusivity:** Better  $\mu$  (impact) if we select by early citations than current baseline (standard reviewing by PCs).
- 2 **Inclusivity:** More papers ( $N$ ) have early citations than accepted by venues.
- 3 **Robustness:** Results were replicated over types of papers and pub. dates.

## Constructive Proposal

- Challenges for Peer-Reviewing:
  - Too many submissions and too few qualified reviewers
- One proposed alternative reviewing process:
  - 1 Authors post papers on ArXiv
  - 2 Papers qualify for review if cited or nominated
- Qualifications for nominators: prof. (or equiv. in industry)
  - Nominations include suggestions for reviewers as well as a promise to review four papers per nomination
- Reviewers should be familiar with submission, e.g., someone suggested in a nomination or an author that cited the submission