

# P-Index, A Novel Algorithm for Bibliometric Analysis

Stuart Spore, NYU Health Sciences Libraries, New York University School of Medicine



The chart at the center illustrates relationships among publication numbers, H-index, and P-index values for five age-based cohorts of faculty at the NYU School of Medicine. From the left, members of cohort A are 60 or more years old, cohort B, 50-60 years old, cohort C, 40-50 years old, cohort D, 30-40 years old, and cohort E, less than 30 years old.

### What is the H-index?

The H-index is a widely-used measure of an author's number of publications and the number of times that each of those publications is cited elsewhere. The value of h is equal to the number of papers (N) by a given author that are cited N or more times in other publications. For example, the publication set below generates an H-Index of 6; counting down from the top item #6 is first item where item # is greater than the number of times cited.

| Item # | Citation                            | Times Cited |
|--------|-------------------------------------|-------------|
| 1      | Journal of Cement v.3(4):35         | 23          |
| 2      | Journal of Concrete v.12(1):3       | 20          |
| 3      | Journal of Hard Substances v.1(1):1 | 16          |
| 4      | Cement Weekly v.30(52):25           | 10          |
| 5      | Journal of Cement v.2(1):12         | 6           |
| 6      | Journal of Asphalt v.33(12):6       | 3           |
| 7      | Journal of Cement v.1(2):3          | 0           |
| 8      | Cement Weekly v.16(1):1             | 0           |
| 9      | Journal of Asphalt v.2(3):12        | 0           |
| 10     | Journal of Concrete v.6(6):6        | 0           |

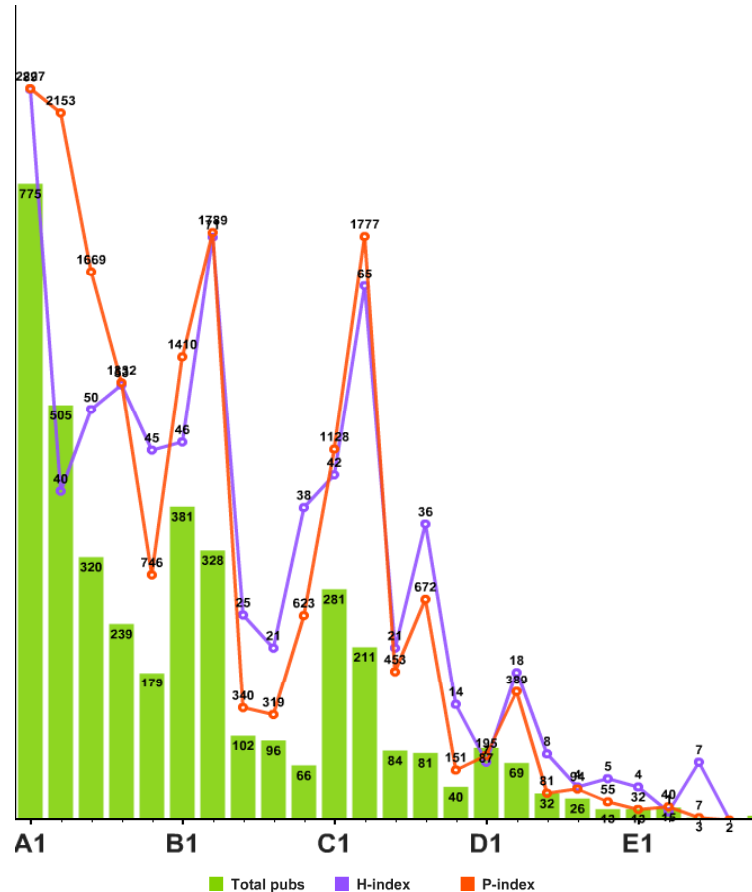
### What is the P-Index?

The P-index is a metric developed to measure faculty publishing productivity at NYU School of Medicine. Each publication is assigned a value of the impact factor of the host journal plus 1. Articles published in journals with no impact factor are arbitrarily given a value of 0.05. The P-index is the sum of all the values, for example:

| Journal                          | Impact Factor | P-Value |
|----------------------------------|---------------|---------|
| Lancet neurology                 | 14.270        | 15.270  |
| Neurology forum                  | n/a           | 0.050   |
| Pediatric neurology              | 1.505         | 2.505   |
| Reviews in Neurological Diseases | n/a           | 0.050   |
| Revue neurologique               | 0.508         | 1.508   |
| P-Index                          |               | 19.383  |

Books, chapters, and minor journal articles (e.g. meeting abstracts, letters, editorials) are excluded from P-index calculations.

Publishing and P-Index data for this poster were derived from the NYU School of Medicine Faculty Bibliography, an online database curated by the NYU Health Sciences Libraries. H-index values were calculated from ISI Web of Knowledge®.



### Do P-index and H-index agree?

The answer is a (very) rough yes. Both P-index and H-index values rise and fall together, although the numbers themselves obviously diverge. The range of possible values is much wider for P-index than H-index, (0-2250 versus 0-85), making differences appear more pronounced in the P-index. For example, in cohort B, calculating degree of change from right to left, we find:

|         | B5 | B4   | B3   | B2   | B1   |
|---------|----|------|------|------|------|
| P-Index | 0  | 0.51 | 1.06 | 5.26 | 0.79 |
| H-Index | 0  | 0.55 | 1.19 | 2.84 | 0.65 |

While P-Index measures a greater change between B3 and B2 than H-Index does, there is an overall harmony. The causes behind the B3 to B2 difference are unclear, but probably relate to different data sets used to generate the metrics.

### Does bibliometrical analysis have an age bias?

Simple volume plays a significant role in driving P-index and H-index values. More publications mean higher scores and, as one would expect, senior faculty tend to have more publications than younger faculty. However, that effect is less pronounced for those over 40 years old.

Moreover volume alone is not a guarantee of higher P-index and H-index scores. Compare D1 with D2 and B1 with B2, for example:

| Cohort Member | Publications | P-Index | H-Index |
|---------------|--------------|---------|---------|
| D1            | 87           | 195     | 7       |
| D2            | 69           | 389     | 18      |
| B1            | 381          | 1410    | 46      |
| B2            | 328          | 1789    | 71      |

Still senior faculty are more likely than younger to publish in high prestige venues and may be more likely to be cited by their colleagues. Overall, age does tend to correlate with higher bibliometric scores; both volume and citation frequency play a part in that effect.

### P-index versus H-index?

P-index and the H-index calculations are ultimately grounded in citation analysis, but H-index is calculated directly while P-index is calculated obliquely (via ISI Impact Factors). If both measurements count the same things in similar ways, why choose one over the other? The most significant advantage of the P-index is that its calculation can be readily scaled over large numbers of faculty or academic/clinical departments. A recent project which entailed generation of P-index values for 1,887 faculty and 30 academic departments for 3 separate years (5,751 calculations total) required less than 25 minutes to complete. H-Index calculation, on the other hand, requires by citation by citation comparison. Even with good enabling software, H-index is difficult to scale for large projects. H-Index does have the advantage of being more widely used which can make extra-mural comparisons easier.