

Bibliometrics

The bibliometric methods used by other fields as a measure of—or at least as a surrogate for the measure of—the impact of research can be divided into two general classes: journal-level metrics, and article-level metrics.

Journal-level metrics rank the journals in field of research. It is generally assumed that any article published in a journal with a high “impact-factor” is of high quality, and therefore researchers with publications in high impact-factor journals are doing high quality research.

Article-level metrics are more directly, and aggregate the data about all of an individual’s published articles into a single measure, in order to give an indication of the quality of that researcher’s work.

Journal-level metrics

Journal-level metrics are the earliest metrics to become widely used, and continue to be influential. They measure the impact of journals in a research field using citation counts between journals; the citation counts can be used either directly or indirectly. Journal-level metrics were originally developed to aid librarians in selecting what journals to which to subscribe and maintain in their collections.

But journal-level metrics have become a way to indicate the quality of a researcher’s publications. Presumably, the logic goes, a journal must publish the best articles in a field in order to have a high impact factor, so therefore any article published in a journal with a high impact factor must be of the best articles.

Journal Impact Factor All citation-based metrics can probably be traced back to a paper by Gross and Gross (1927); Archambault and Larivière (2009) take that article as the starting point for the Journal Impact Factor. Since the Journal Impact Factor is the best known and possibly most widely used citation-based metric for research, its origins can act as an example for all journal-level metrics in general.

Finding the “indispensable” journals The question posed by Gross and Gross (1927) was, “What... scientific periodicals are needed in a college library?” They wanted to identify to which journals a library should subscribe, and to do so they used quantitative methods to make comparisons between journals, and those methods eventually led to the Journal Impact Factor, along with other related bibliometrics.

Gross and Gross decided to use a quantitative method to evaluate journals, as opposed to merely asking an expert to compile a list of “indispensable” journals in order to avoid a list that was “seasoned too much by the needs, likes and dislikes of the compiler.” In other words, they used quantitative methods to minimize any single individual’s biases from the evaluation of journals.

Counting citations to sort journals The basic method pioneered by Gross and Gross, and repeated later by other researchers, was to select a pre-eminent journal or a reference that was generally accepted as key to the field at hand, and then compile and quantify the sources cited in that keystone reference/journal.

This is an extremely simplified example: if the journal that was selected as the central reference for a field, “Field”, contained five citations to a “Journal A”, ten citations to a “Journal B”, and four citations to a “Journal C”, then for that field the journals would be ranked and reported something like this:

- Leading periodicals in Field
 - Journal B – 10
 - Journal A – 5
 - Journal C – 4

These early citation metrics were inherently field-specific, since they took data from a central reference to some particular field.

Compiling citations for multiple fields Soon, though researchers began to cross-compile citation information from multiple journals, and then to include journals from multiple fields in those compilations. For example, Gregory (1937) compiled more than 27,000 citations from across 27 subfields in medicine.

Gregory gives a concise explanation of the purpose of her metrics:

The foregoing Tables (I-27) answer primarily the needs of the specialist in his attempt to keep adequately abreast of the literature in his field. Two further Tables have been compiled, to indicate to the medical library and the librarian (A), the indispensable periodicals for al the fields consulted; (B), a short list of essential periodicals in general medicine which cover a large amount of material. This latter list is designed primarily for the individual and for the small library.

There were 27 “foregoing Tables”, one for each of the subfields, yet in Gregory’s Herculean compilation, there were no comparisons made between those fields. This is a feature that would continue in similar metrics for decades: when multiple fields were compiled and evaluated at the same time, the results for each field were reported separately. Presumably this was because the intended audience for these metrics had no need for cross-field comparisons, since libraries are generally tasked to select the best journals for some particular field, and are not given carte blanche to choose which fields to support, and researchers are interested in their field, not all fields.

The compilation of citations from Martyn and Gilchrist (1968) is cited by Archambault and Larivière (2009) as having a large influence on methods used in the Journal Impact Factor, and it too continues the practice of reporting of metrics for separate fields separately.

But eventually, when journal impact factors become used as a way to judge the quality of research, this lack of cross-field comparisons will become an issue as institutions try to equitably evaluate the quality of research done by their faculty and staff in an array of fields. Ways to address this problem are offered by some journal-level as refinements to the methodology of the Journal Impact Factor. (More details below.)

Reporting ratios instead of counts Another feature of the Martyn and Gilchrist (1968) compilation that was carried forward into the Journal Impact Factor, and other bibliometrics, was first proposed by Hackh (1936): reporting a ratio instead of a count. Specifically, the ratio of the number of citations to the number of pieces that might be cited.

Consider another extremely simplified example: if there were five citations to articles published in “Journal A” in the same year that “Journal A” published a total of 20 articles, the citation ratio would be “5:20”, or “0.25”.

We can extend our earlier simplified example to include reporting ratios instead counts, and hopefully see why this is the preferred method:

Journal	Citation count	Citable pieces	Ratio
Journal A	5	20	0.25
Journal B	10	20	0.50
Journal C	4	10	0.40

- Leading periodicals in Field
 - Journal B – 0.50
 - Journal C – 0.40
 - Journal A – 0.25

Despite having fewer citations than Journal A, Journal C has I higher ratio citations to citable pieces, and so is ranked higher. Using that ratio to rank journals is a way to prevent quantity from overwhelming quality.

And thus the Journal Impact Factor This, then, is the core of how the Journal Impact Factor and a number of similar metrics work. Gather citation counts and citable pieces for each journal, calculate the ratios, and rank the journals.

However, there are a number of potential problems with this method. Those and some others problems with the Journal Impact Factor are most easily explained by looking at a some other journal-level metrics that use refined versions of this same basic methods.

References

- Archambault, Éric, and Vincent Larivière. 2009. “History of the Journal Impact Factor: Contingencies and Consequences.” *Scientometrics* 79 (3): 635–49. doi:10.1007/s11192-007-2036-x.
- Gregory, Jennie. 1937. “An Evaluation of Medical Periodicals.” *Bulletin of the Medical Library Association* 25 (3): 172–88.
- Gross, P. L. K., and E. M. Gross. 1927. “College Libraries and Chemical Education.” *Science* 66 (1713): 385–89.
- Hackh, Ingo. 1936. “The Periodicals Useful in the Dental Library.” *Bulletin of the Medical Library Association* 25 (1-2): 109–12.
- Martyn, John, and Alan Gilchrist. 1968. *An Evaluation of British Scientific Journals*. London: Aslib.