Trends in biochemical literature

Eugene Garfield

Any working biochemist intuitively knows that the literature is growing. However, quantifying this growth with any degree of precision is not an easy task. But after my interest in this problem was piqued by the late Robert Harte [1], I attempted to produce what is, I hope, an accurate and useful view of the biochemical literature.

Based on information extracted from the *Science Citation Index®* (*SCI®*) data base, I find that, in terms of articles published, the biochemical literature is still growing faster than the scientific literature as a whole. What's more, the number of references in a typical biochemical article is increasing as is the proportion of references to material more than five years older than the citing article.

For the purposes of this paper we have looked at 37 'core' primary journals. Our study encompassed the years 1968–1977 for all core journals (Table I), and also included 1962–1967 for the journals used by Harte*. I refer to the journals studied by Harte as the 'CEBJ journals', since their editors are full members of the Committee of Editors of Biochemical Journals of the IUB.

Of the 37 journals studied, 16 started publication in 1962 or later. Of these, eight started publication in 1970 or later. Thus, just in the number of journals considered important to biochemists, there has been a 76% increase in 16 years.

Table I shows that the number of articles per year produced by the core biochemistry journals increased from 9060 in 1968 to 14,418 in 1977. This amounts to an annual growth rate of 5.3% or a doubling time of 13.4 years. If we look at only the CEBJ journals for the same time period, we find that the number of articles they published annually increased from 6766 to 8491. This is an annual growth rate of 2.6%; however, non-CEBJ journals increased their output at an average annual rate of 11.1%.

The higher growth rate for the non-CEBJ journals is partly due to the birth of new journals. But it also results from the fact that the increase in the average number of items published per year was greater for non-CEBJ core journals. Table II shows that the average number of items published per year by a CEBJ journal increased from 615 in 1968 to 772 in 1977 – an increase of 26%. The average number of items published by a non-CEBJ core journal increased from 143 in 1968 to 228 in 1977 – an increase of 59%.

Over the longer period of 1962–1977 the CEBJ journals had an average annual growth rate of 5.1%. This growth rate conflicts with Harte's findings for the same journals for essentially the same time period (9.8%) but since detailed data were not provided in Harte's report, we have not been able to identify the reason for this difference.

For the earlier period of 1962–1967, the CEBJ journals had an average annual growth rate of 8.1% – about three times greater than the growth of the CEBJ journals during 1968–1977. These data clearly confirm the exponential growth that the literature experienced in the 1960s and the general slowdown which has occurred in the 1970s.

The 5.3% average annual growth rate observed for the core biochemical journals between 1968–1977 is slightly greater than the growth rate of the *SCI* data base (Table I). Increasing from 311,959 items in 1968 to 465,067 in 1977, the *SCI* had a 4.5% average annual growth rate. To the degree that the *SCI* data base represents the literature of science as a whole, we can say that the growth rate for the biochemistry literature was at least 18% higher.

Preliminary data from our unpublished studies on the literature of mathematics and botany allowed me to compare the growth of the biochemical literature to that of other fields. In striking contrast to biochemistry, the size of the core journals of pure mathematics remained almost constant during 1968–1977 and the number of botany articles increased by an annual growth rate of only 3%.

If we look at individual journals, Table I shows that the highest output of articles in 1977 came from *Biochim. Biophys. Acta* (2080), *J. Biol. Chem.* (1384), and *Biochem. Biophys. Res. Commun.* (1202). The largest average annual growth rates between 1968 and 1977 were shown by *Indian J. Biochem. Biophys.* (24.2%), *FEBS Lett.* (14.4%), and *Eur. J. Biochem.* (9.4%). These three journals all published a substantial number of articles and had a steady increase in articles over the years. The *Ital. J. Biochem.* also had a high annual growth rate (16.9%), but it published relatively few articles and its growth was erratic.

Earlier, I stated that the core biochemistry journals produced about 14,000 articles in 1977. One needs to remember, however, that biochemistry articles can appear in other than core journals – especially in multidisciplinary ones such as *Science, Nature,* and the *Proc. Natl. Acad. Sci. U.S.A.* Evidence of this can be seen in Tables III and IV. In these tables we have listed, for 1977, the 50 journals which were cited most by the core biochemistry journals and the 50 journals that cited the core journals the most. Each list contains a substantial number of journals that are not part of the biochemistry core.

To estimate how many biochemistry articles appear in non-core journals, we analysed the citation frequency between core and non-core journals. This indicated that non-core journals would contribute about 5000–10,000 additional biochemistry articles per year. Obviously, this is not very precise. But trying to measure the population of journal articles in a field like biochemistry is as elusive as measuring the ethnic or racial characteristics of a country like the U.S.A. where there is constant intermarriage. Nevertheless, when the estimated number of biochemistry articles published by non-core journals is added to the 14,000 articles published by core journals, it would seem that a minimum of 20,000–25,000 biochemistry articles were produced during 1977.

It should be noted here that since the *Proc. Natl. Acad. Sci. U.S.A.* is fourth among the journals most cited by the biochemistry core, an argument could have been made to include it as part of the core. But the same argument could be made, albeit somewhat less strongly, for *Nature, Science, J. Am. Chem. Soc., J. Bact., J. Cell Biol.,* and other important journals that are not devoted exclusively to biochemistry, but are highly cited in the core journals. Therefore, we felt it best to continue in this study our usual practice of defining the core journals as those which solely publish articles related to the field being examined.

Another 'growth' indicator within the biochemical literature is the increase in the average number of references contained in a typical article. To examine this factor I developed an 'R/S' value for each core

* To conserve space, some of the supporting data for this article have been omitted. These may be obtained from the author at the above address.

Eugene Garfield is President, Institute for Scientific Information® (ISI®), 3501 Market Street, University City Science Center, Philadelphia, PA 19104, U.S.A.