Cross-Checking for Plagiarism

Human knowledge has traditionally been expanded by the oral and written word (today, audio and video also play a role). In the oral tradition, plagiarism is a common and accepted practice as previous concepts are constantly being incorporated into current knowledge. Lack of tangible pre-existing documents at different points in time makes checking for plagiarism impossible in the oral tradition. Conversely, in the scientific literature, the availability of previous documents makes checking for plagiarism easier. One feature of scientific literature is its constant need for originality and honesty. Although lack of originality may be caused by many factors, in science we are always concerned that it may be because of plagiarism or piracy. Plagiarism involves the use of unauthorized portions of one's own work (self-plagiarism or duplicate publication) or that of others, or both. Piracy always refers to appropriation of ideas, data, or both, without permission or acknowledgment of the original source. There is no sin in the use of someone else's ideas or data as long as permissions have been granted and the sources of this information are clearly acknowledged when an article is initially submitted.

In 1989, the Department of Health and Human Services Office of Research Integrity noted that 12% of the misconduct cases it investigated were because of plagiarism.¹ In 2005, the National Science Foundation revealed that 66% of the cases it investigated were related to plagiarism.¹ This infraction occurs in all walks of life, but, in a somewhat paradoxical fashion, codes of honor tend to be stronger for high school and college students compared with physicians and scientists. Tremendous external pressure, particularly on non-English-speaking faculty to publish in English-language scientific journals, is probably the single most important factor contributing to all types of plagiarism (thus, unfortunately, most authors accused of plagiarism reside outside of the United States). Dr. Loren Greene from New York University coined the term *text offenders* and supports the creation of a data base containing their names.¹ In 23 university campuses surveyed by Rutgers University, 38% of undergraduate students acknowledged being involved in at least 1 act of "cut and paste" in the previous 2 years, and nearly half of them considered this behavior as trivial and unimportant.² Approximately 5% of college students downloaded at least 1 entire paper. This behavior did not change once students were warned not to plagiarize. If these frightening statistics reflect people who are being "sworn" into strict codes of honor, what can be expected of those already practicing science?

How do medical and scientific journals detect plagiarism? First, as editors we rely on the integrity of those submitting manuscripts. Second, our reviewers detect most instances of plagiarism, generally with the help of data from PubMed. Until now, only a small number of journals routinely have used more advanced computerized methods to detect plagiarism.

Screening software, as expected, is better at detecting plagiarism than solely human effort. When 1 set of assignments was manually studied, plagiarism was detected in 3%, but when the same batch was subjected to Turnitin (iParadigms, Oakland, Calif), the plagiarism rate increased to 13%.³ Other popular plagiarism detection services found on the Web and used by high schools and colleges include EVE (Essay Verification Engine; CaNexus.com), iThenticate (iParadigms), and Plagiarism.org. Google also offers free plagiarism detection software. The problem for those involved in academic publishing is that these methods were not designed to specifically evaluate the scientific and medical literature.

CrossRef.org and iParadigms (the creators of Turnitin) have designed an umbrella plagiarism detection program that will use iThenticate to verify the originality of material against a vast data base of proprietary content that, to date, includes the archives of 8 major scientific publishers as well as open Web sources. The software that will be used for this activity is called CrossCheck. Because the *American Journal of Neuroradiology* (AJNR) will be using CrossCheck from now on, we will attempt to briefly describe this process. CrossCheck is a paid service that compares the text of a submitted manuscript with those found in its data base (which includes articles published by Elsevier, Wiley-Blackwell, Taylor and Francis, BMJ Publishing Group, and the *New England Journal of Medicine*, among others). When textual overlap is detected, the system automatically produces an "originality" report that will be used by the editorial board of the AJNR to determine if a true plagiarism situation exists. Conversely, AJNR will make its electronic data base available for incorporation into the CrossCheck plagiarism detection system. The impact that CrossCheck will have on our manuscript flow is negligible; thus, we are not expecting any delays in the publishing of accepted manuscripts.

As stated in a previous communication, we at AJNR take the issue of plagiarism seriously and are proud of the originality of our content. Plagiarism is an intellectually and academically dishonest behavior for which we will, now more than ever, be on the lookout. We believe that the routine use of a system such as CrossCheck will help assure that our articles continue to meet the highest scientific standards.

References

1. Roig M. **Plagiarism in the sciences: conference highlights.** *Sci Editor* 2006;29:48–49
2. Muha D. **New study confirms internet plagiarism is prevalent.** Available at: <http://ur.rutgers.edu/medrel/viewArticle.html?ArticleID=3408>. Accessed April 18, 2008
3. Jocoy C, DiBiase D. **Plagiarism by adult learners online: A case study in detection and remediation.** *Int Rev Res Open Distance Learn* 2006;7 Available at: <http://www.irrodl.org/index.php/irrodl/article/view/242/466>. Accessed May 16, 2008

M. Castillo
Editor-in-Chief
K. Halm
Managing Editor

DOI 10.3174/ajnr.A1178