

Providing Choice & Value

Generic CT and MRI Contrast Agents





Screening for Self-Plagiarism in a Subspecialty-versus-General Imaging Journal Using iThenticate

A.U. Kalnins, K. Halm and M. Castillo

This information is current as of July 22, 2025.

AJNR Am J Neuroradiol 2015, 36 (6) 1034-1038 doi: https://doi.org/10.3174/ajnr.A4234 http://www.ajnr.org/content/36/6/1034

Screening for Self-Plagiarism in a Subspecialty-versus-General Imaging Journal Using iThenticate

A.U. Kalnins, K. Halm, and ^DM. Castillo

ABSTRACT

BACKGROUND AND PURPOSE: Self-plagiarism is a form of research misconduct that can dilute the credibility and reputation of a scientific journal, as well as the represented specialty. Journal editors are aware of this problem when reviewing submissions and use on-line plagiarism-analysis programs to facilitate detection. The *American Journal of Neuroradiology (AJNR)* uses iThenticate to screen several submitted original research manuscripts selected for review per issue and retrospectively assesses 3 issues per year. The prevalence of self-plagiarism in *AJNR* was compared with that in *Radiology*; the necessity and cost of more extensive screening in *AJNR* were evaluated.

MATERIALS AND METHODS: The self-duplication rate in *AJNR* original research articles was compared with that in *Radiology*, a general imaging journal that screens all submitted original research manuscripts selected for review by using iThenticate. The rate of self-duplication in original research articles from 2 randomly selected 2012 *AJNR* issues was compared with the rate in the prior year to gauge the need for more extensive screening. A cost analysis of screening all submitted original research manuscripts selected for review by using iThenticate was performed.

RESULTS: Using an empiric 15% single-source duplication threshold, we found that the rate of significant self-plagiarism in original research articles was low for both journals. While *AJNR* had more articles exceeding this threshold, most instances were insignificant. Analyzing 2 randomly chosen issues of *AJNR* for single-source duplication of >15% in original research articles yielded no significant differences compared with an entire year. The approximate annual cost of screening all submitted original research manuscripts selected for review was US \$6800.00.

CONCLUSIONS: While the rate of self-plagiarism was low in *AJNR* and similar to that in *Radiology*, its potential cost in negative impact on *AJNR* and the subspecialty of neuroradiology justifies the costs of broader screening.

ABBREVIATIONS: AJNR = American Journal of Neuroradiology; ORA = original research article; SORMSR = submitted original research manuscript selected for review

The appropriation of previously published ideas or words without proper credit is known as "plagiarism" and is considered a major breach of ethics in scientific publications. In 2005, the National Science Foundation identified plagiarism in nearly 66% of suspected cases of fraud.¹ The same institution implemented a plagiarism checking system and discovered that 1.5% of the nearly 8000 grants accepted in 2011 contained significant plagiarism.² The recognized significance of plagiarism has merited its inclu-

http://dx.doi.org/10.3174/ajnr.A4234

sion in the federal definition of "research misconduct," defined as, "fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results."³ While plagiarism may involve the use of the work of others, the use of one's own work without acknowledging its source is known as "redundant publication" or "self-plagiarism" and is no less significant.⁴ The pressure of academic productivity as measured by the number of one's publications or how often one is cited creates an incentive for authors to publish a greater volume of work, which may increase the likelihood of self-plagiarism.^{5,6}

The presence of sporadic phrases reproduced from an author's own prior work, particularly from "Materials and Methods" sections of manuscripts based on similar research methodologies, may be acceptable at the discretion of an editor and not necessarily suspected as intentional duplication.⁷ However, copying entire sections or illustrations of one's own prior work without appropriate attribution is not acceptable and constitutes self-plagia-

Received September 20, 2014; accepted after revision November 13.

From the Department of Radiology (A.U.K.), Stanford University Medical Center, Stanford, California; *American Journal of Neuroradiology* (K.H., M.C.), Oak Brook, Illinois; and Department of Radiology (M.C.), University of North Carolina at Chapel Hill, Chapel Hill, North Carolina.

Please address correspondence to Mauricio Castillo, MD, FACR, University of North Carolina at Chapel Hill, Department of Radiology, Room 3326, Old Infirmary Building, Manning Dr, Chapel Hill, NC 27599-75; e-mail: Castillo@med.unc.edu; @AleksKalnins

rism. While there are various different classifications for self-plagiarism, a well-accepted one developed by Miguel Roig defines 4 distinct types: recycling of text, copyright infringement, "salami slicing" or dividing 1 study into multiple publications with the intent of apparently increasing productivity, and duplicate publication.⁸⁻¹⁰

Editors of scientific journals are aware of the potential of selfplagiarism when reviewing submissions, and the use of electronic plagiarism analysis software has greatly facilitated its detection. Widely used plagiarism-analysis programs are iThenticate (http:// www.ithenticate.com) in scientific publications and Turnitin (http://turnitin.com) in general education.¹¹ These programs are similar in their comparison of selected documents with data bases of published or submitted articles including "CrossCheck" for iThenticate and "OriginalityCheck" for Turnitin.12,13 The iThenticate data base enables manuscripts to be compared with >37billion Web pages, >92 million published works from on-line and off-line research publications and data bases, and >37 million scholarly articles through the "CrossCheck" data base, the largest comparison data base of scientific, technical, and medical publications in the world.^{14,15} Analysis of documents by one of these programs generates an overall "similarity index" with works in the data base as well as an interactive summary of individual sources with which a document demonstrates sameness.^{16,17} In one study, plagiarism was detected in 3% of assignments submitted by adult learners when manually evaluated for significant overlap, but when evaluated with Turnitin, plagiarism was identified in 13%.¹⁸

Given the awareness of the practice of plagiarism and selfplagiarism as well as the ease of use and availability of electronic plagiarism analysis, it is convenient for scientific journals to use it for evaluation. At the time of this study, Radiology used iThenticate to screen all manuscripts for the presence of duplications since February 2013 (D. Levine, MD, personal e-mail communication, October 2014), and the American Journal of Neuroradiology (AJNR) selectively used the iThenticate service for approximately 3 submitted original research manuscripts selected for review (SORMSRs) per issue since 2012. AJNR publishes about 300-350 original articles per year, and of these, 3 issues are now randomly retrospectively selected for review with iThenticate to evaluate the presence of duplication. In addition, all manuscripts originating from authors who have been associated with previous fraudulent behavior are prospectively assessed with iThenticate before undergoing the peer review process, representing approximately 3 SORMSRs per issue.

When one analyzes manuscript submissions by using electronic plagiarism software, the acceptable percentage of similarity between scientific publications is not well-defined. According to the Turnitin Web site instructions for the interpretation of the "similarity index," there are no clear rules that define when plagiarism has taken place. However, the company previously disseminated guidelines to academic institutions suggesting a 15% "similarity index" for alerting to the possibility of plagiarism.^{16,19-22} Organizations involved in promoting ethics and best practices of scientific editing (International Committee of Medical Journal Editors, International Society of Managing and Technical Editors, Committee on Publication Ethics) do not define a specific numeric "cutoff" for the identification of plagiarism.²³⁻²⁵ While *AJNR* has adopted the 15% guideline, other journals set it as low as 10%.⁷

In this study, we first sought to evaluate the rate of duplications in published *AJNR* original research articles to see if it is comparable with that in the general radiology literature and to evaluate whether screening all SORMSRs versus a random sample is sufficient. We also performed an analysis of the costs of screening all SORMSRs for potential duplication with iThenticate to assess the economic impact on *AJNR*.

MATERIALS AND METHODS Part 1

We selected articles published in AJNR and Radiology in 2011 under the heading of "Original Research." Review articles, editorials, commentaries, vignettes, case reports, clinical reports, and so forth were excluded because these do not constitute most articles in an imaging journal and were thought to be more susceptible to differences in opinion regarding the significance of flagged content. All "Original Research" articles published in Radiology and AJNR in 2011 were uploaded into iThenticate in a .pdf format, and an analysis was run. iThenticate parameters were set to exclude quotations and bibliographic entries from comparison and to ignore single-source duplications of $\leq 3\%$, given the likelihood of random similarity at or below this threshold. The 15% single-source "similarity index" threshold was set for flagging manuscripts needing further evaluation.^{16,19-22} When an article exceeded this 15%, 2 individuals (Editor-in-Chief, Managing Editor) separately assessed the duplications and rated them as being significant or insignificant. The presence of plagiarism and selfplagiarism was evaluated by comparing the authors of the SORMSRs with the authors of prior publications that were identified by the iThenticate program as having >15% total similarity. If any author of the SORMSR was an author of the flagged prior publication, self-plagiarism was suspected and the manuscript was further assessed for significance. The presence of sporadic phrases from prior works was ignored, while the presence of large blocks of text was further judged as significant or insignificant.

A significant duplication was characterized as duplicate publication of the same article with no or minor modifications or unreasonable duplication of large or small blocks of text from any portion of the manuscript without appropriate modification or citation of the prior publication. This characterization was consistent with published recommendations for classification of significant (major or minor) duplications.^{26,27} When differences occurred between the 2 evaluators in the judgment of duplication as significant or insignificant, these were resolved by consensus. Total significant single-source duplications of original research articles (ORAs) in AJNR and Radiology were compared by using the Mann-Whitney U test, with the null hypothesis that there is no difference between the ranks of the significant single-source duplications of ORAs in AJNR and Radiology. Statistical analysis was performed by using MedCalc for Windows, Version 14.10.2, 64bit (MedCalc Software, Mariakerke, Belgium).

Part 2

We assessed the value of analyzing all versus a small sample of submitted manuscripts for the presence of duplication by evalu-

Table 1: Duplication in ORAs, AJNR 2011 and Radiology 2011

	AJNR 2011	Radiology 2011
Issues (No.)	11	12
ORAs (No.)	302	343
ORAs per issue (mean)	27.5	28.6
ORAs per issue (SD)	2.73	1.31
With $>$ 15% duplication (No.)	54	12
With $>$ 15% duplication (% of total)	17.9	3.5
With $>$ 15% duplication from a single source (No.)	12	2
With $>$ 15% duplication from a single source (% of total ORAs)	4	0.6
With $>$ 15% duplication from a single source (significant) (No.)	2	1
With $>$ 15% duplication from a single source (significant) (% of total ORAs)	0.7	0.3

Table 2: Results of the Mann-Whitney	y U test to compare the difference between total
significant (>15%) single-source dupl	ications of ORA's in AJNR and Radiology in 2011

Group	No.	Rank Average	Sum of Ranks	U	Ζ	Р
AJNR 2011	11	12.59	138.5	59.5	0.61	.545
Radiology 2011	12	11.46	137.5			

ating all "Original Research" articles in 2 randomly selected issues of *AJNR* in 2012 (March and September) and comparing the results with those obtained for an entire year (2011).

Part 3

An analysis was performed to assess the total current and potential future annual cost of using iThenticate to evaluate all SORMSRs for self-plagiarism, factoring in the fixed costs (iThenticate annual subscription fee) and variable costs (iThenticate manuscript analysis fee and time of the *AJNR* staff) of screening.

RESULTS

Part 1

In 2011, *AJNR* was published 11 times and contained a total of 302 original research articles ($\bar{x} = 27.5$ per issue, $\sigma = 2.73$), while *Radiology* was published monthly with a total of 343 original research articles ($\bar{x} = 28.6$ per issue, $\sigma = 1.31$) (Table 1).

In *AJNR*, iThenticate found 54 articles (17.9%) that showed >15% duplication rates, while 12 (3.5%) such articles were found in *Radiology*. In *AJNR*, 12 articles (4%) were found to have duplication rates above 15% from a single source, while in *Radiology*, only 2 (0.6%) such articles were found. After individual evaluations of these articles, only 2 in *AJNR* and 1 in *Radiology* were judged by consensus agreement to contain significant "single-source" duplications consistent with self-plagiarism. In *AJNR*, this finding led the editors to contact the authors and request explanations, concluding that most duplication arose from similarities in the "Materials and Methods."

The total number of identified "single-source" duplications in original research articles in both journals was compared (Table 2).

Table 2 reveals the results of Mann-Whitney *U* test for the number of significant single-source duplications in original research articles published in *AJNR* and *Radiology* in 2011, which did not show any statistical difference (Z = 0.606; P = .5447). The rank average of the number of single-source duplications in original research articles published in *AJNR* was 12.59, and in *Radiology*, it was 11.45. The similar rank averages demonstrate that the number of significant single-source duplications

Part 2

Two *AJNR* issues from 2012, selected at random (March and September), contained 25 and 27 original research articles each. In these 2 issues, 5 and 4 articles surpassed our 15% duplication threshold, respectively, compared with an average of 4.9 (σ = 1.92) articles in all of 2011. The fact that this value was within the SD led us to believe that the rate of articles surpassing the 15% minimum threshold is likely constant throughout the year. The number of articles in which the 15% duplication rate originated from a single source was 3 in March and 1 in September (\bar{x} = 1.09 per

issue in 2011, $\sigma = 1.04$). Individual evaluation of these 4 articles demonstrated no significant self-plagiarism that accounted for the duplication rate above the 15% threshold.

Part 3

We evaluated the cost of analyzing an entire year of submitted manuscripts, factoring in the financial cost of using iThenticate and wage hours spent by *AJNR* staff to perform this task (Table 3).

In 2013, an average of 3 articles per issue (36 articles per year) were prospectively flagged for evaluation by using iThenticate due to previous duplication problems with authors at an annual cost of US \$840.72. If *AJNR* were to evaluate all submitted articles by using iThenticate, the annual cost would rise by US \$5963.76 to a total of US \$6804.48.

DISCUSSION

In comparing AJNR and Radiology, our results indicate that both journals published about the same number of articles under the category of "Original Research" per issue but that the rate of nonsignificant duplication was higher in AJNR. When the origin of the nonsignificant duplications was evaluated in AJNR, it was found to be mostly secondary to similarities in research methodologies used in previously published research by the same authors, and the discrepancy between the 2 journals may reflect differences in journal standards regarding screening of SORMSRs for the presence of self-duplication and re-publication of nonsignificant duplication of research methodologies. When these articles were individually studied for significant duplications, only 2 in AJNR and 1 in Radiology were considered exceeding the 15% threshold (0.66% of articles in AJNR and 0.29% of articles in Radiology). As a result of identified duplications, AJNR editors contacted the authors and requested an explanation (M. Castillo, MD, AJNR e-mail correspondence with manuscript authors regarding duplication, 2012). After careful deliberation by an internal committee, it was concluded that most similarities arose mostly from the "Materials and Methods" and matters were not pursued further.

In cases of proven self-plagiarism, the Committee on Publication Ethics guidelines recommends taking further action, which

Cost Estimate Variable	Cost		
Annual iThenticate subscription fee	\$550.00		
Annual iThenticate administrative fee	\$110.00		
Total annual fixed cost to use iThenticate (550.00 subscription fee $+$ \$110.00 administrative fee)			
iThenticate cost per analyzed manuscript	\$0.75		
Hourly rate of AJNR administrative staff	\$25.64		
Average time to submit manuscript to iThenticate (min)	10		
Average staff cost to submit manuscript to iThenticate	\$4.27		
Total variable cost to submit a single manuscript to iThenticate (\$0.75 cost per article + \$4.27 <i>AJNR</i> staff time cost per manuscript)	\$5.02		
Number of AJNR issues per year	12		
Current No. of SORMSRs analyzed by iThenticate per issue	3		
Total current annual cost to use iThenticate (\$660.00 annual fixed cost + [\$5.02 time cost/manuscript \times 3 manuscripts/ issue \times 12 issues/year])	\$840.72		
Total current cost per issue to use iThenticate (\$840.72 per year/12 issues)	\$70.06		
Average No. of SORMSRs per issue	102		
Total potential annual cost to use iThenticate to analyze all SORMSRs (\$660.00 annual fixed cost + [\$5.02 time cost/ manuscript × 102 manuscripts/issue × 12 issues/year])	\$6804.48		
Total potential cost per issue to use iThenticate to analyze all SORMSRs (\$6804.48 per year/12 issues)	\$567.04		
Additional potential annual cost to use iThenticate for all SORMSRs (\$6804.48 annual cost of iThenticate for all SORMSRs, \$840.72 annual cost of iThenticate for 3 SORMSRs/issue)	\$5963.76		

may lead to informing the author's institution (chairperson, dean of the medical school, institutional review board, and so forth), flagging the authors, scrutinizing their future submissions, and/or potentially prohibiting them from publishing in that journal for a time or indefinitely, and even retracting the article in question from PubMed.^{28,29} The economic costs of dealing with research misconduct are nonnegligible; while iThenticate has reported an estimate of between \$10,000 and \$50,000 in capital losses due to incidents of research misconduct in nearly 200 organizational clients surveyed, the direct costs of an investigation of a case of research misconduct have been estimated to approach \$525,000.^{30,31} The additional costs of published research misconduct are characterized by broader detrimental effects on the credibility and reputation of a scientific journal and the specialty that it represents.³² While these effects are difficult to quantify in monetary terms, their potentially long-standing and irreversible effects may be devastating.

Ideally, this sequence of events could be avoided by prospectively checking for duplications when articles are initially received. This may, however, delay the review process and be financially impossible for smaller journals and, in the case of *AJNR*, would result in checking about 75%–80% of SORMSRs that are ultimately rejected. Because of this, the second question we sought to answer was whether randomly studying a sample of published articles yields similar results to studying all articles published in 1 year.

In 2012, *AJNR* published an average of 27.45 original articles in each of its 11 issues. The average number of articles per issue surpassing the 15% duplication rate was 4.91, which nearly matched the number of articles in the same category in the 2 months randomly chosen (March: 5 and September: 4). Overall, the average monthly number of articles containing a >15% overlap from a single source (1.09) also nearly matched that found in the 2 individual months (March: 3 and September: 1). Therefore, in our study, sampling only 2 randomly selected months yields similar results to analyzing an entire year.

While the financial and time costs of prospectively evaluating

all *AJNR* submissions (approximately US \$6800 per annum) may not seem feasible for smaller journals or justified by the low rate of duplication identified in this study, these are relatively low compared with the potential economic costs of dealing with incidents of research misconduct, broader detrimental effects on journal credibility and reputation, and the professional and public perception and trust of the represented subspecialty, which may be long-standing and irreversible.

CONCLUSIONS

Using the suggested empiric duplication threshold of 15%, which is used by AJNR, the number of original articles found to have significant duplicated content was low for AJNR and Radiology. While AJNR had a greater number of articles showing >15% overlap compared with Radiology, most duplications were not considered significant. When these articles were individually studied, self-plagiarism was found to account for all of the duplicated content, and most of it arose from the "Materials and Methods" and was ultimately considered unimportant. Last, analyzing 2 randomly chosen issues of AJNR yielded no significant differences in articles with duplications of >15% compared with those published during an entire year. However, in light of the potentially enormous economic cost of dealing with incidents of research misconduct and adverse effects on the credibility of the journal, its reputation, and perception and trust of its represented subspecialty, more extensive screening of SORMSRs may provide a cost-effective safeguard.

REFERENCES

- 1. Roig M. Plagiarism in the sciences: conference highlights. *Science Editor* 2006;29:48–49
- Mervis J. NSF audit of successful proposals finds numerous cases of alleged plagiarism. Science Insider. March 8, 2013. http://news. sciencemag.org/2013/03/nsf-audit-successful-proposals-finds-numerouscases-alleged-plagiarism. Accessed November 23, 2013
- Department of Health and Human Services. Public Health Service Policies on Research Misconduct: Final Rule. CFR Parts 50 and 93.

Federal Register. Vol 70. No. 94. May 17, 2005:28386. Washington: Department of Health and Human Services

- 4. Bird SJ. Self-plagiarism and dual and redundant publications: what is the problem? Commentary on 'seven ways to plagiarize: handling real allegations of research misconduct.' *Science and Engineering Ethics* 2002;8:543–44
- 5. Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci U S A 2005;102:16569–72
- Marnet A. H-index: what it is and how to find yours. 2010. http:// www.benchfly.com/blog/h-index-what-it-is-and-how-to-find-yours/. Accessed November 23, 2013
- Berquist TH. Scientific integrity: has CrossCheck made a difference? AJR Am J Roentgenol 2013;201:703
- Roig M. Recycling portions of text from the same author's previously published papers in psychology: an exploratory study. In: Proceedings of the Conference of the Second Office of Research Integrity. Bethesda, Maryland. November, 16–18, 2002
- Roig M. Avoiding plagiarism, self-plagiarism, and other questionable writing practices: a guide to ethical writing. August 2006. http:// ori.hhs.gov/sites/default/files/plagiarism.pdf. Accessed November 24, 2013
- Scanlon PM. Song From Myself: An Anatomy of Self-Plagiarism. Ann Arbor: Michigan University Library; 2007:57–66
- iParadigms. http://www.iparadigms.com/. Accessed November 10, 2013
- Crosscheck. http://www.crossref.org/crosscheck/index.html. Accessed November 10, 2013
- Turnitin. Reduce plagiarism. http://turnitin.com/en_us/features/ originalitycheck. Accessed November 10, 2013
- iThenticate. http://www.ithenticate.com/content. Accessed November 13, 2013
- 15. iThenticate. Prevent plagiarism in published works. http://www. ithenticate.com/. Accessed November 3, 2013
- Turnitin. Instructor training about OriginalityCheck. http://turnitin. com/en_us/training/instructor-training/about-originalitycheck. Accessed November 10, 2013
- iThenticate. Text-only reports. http://www.ithenticate.com/training/ text-only. Accessed November 10, 2013
- Jocoy CL, DiBiase D. Plagiarism by adult learners online: a case study in detection and remediation. International Review of Research in Open and Distance Learning 2006;7:1–15
- The University of the West Indies. Guidelines for staff and students on plagiarism. October 2010. http://sta.uwi.edu/resources/documents/ postgrad/guidelines_staff_students_plagarism.pdf. Accessed November 3, 2013
- 20. University of Stellenbosch Business School. Contract conditions for

ad hoc Faculty teaching on USB-ED programmes, December 2013. Policy Annexure B: Interpreting Turnitin Results. 2013. http:// www.usb-ed.com/content/Downloadable%20documents/Faculty %20Contract%20Conditions%20USB%20ED%202014.pdf. Accessed November 17, 2013

- 21. Syracuse University answers. Turnitin and Blackboard. Interpreting Turnitin originality reports. August 28, 2013. https://answers. syr.edu/display/blackboard01/Interpreting+Turnitin+Originality+ Report. Accessed November 17, 2013
- 22. Tennessee University E-learning @ Tees Staff. Interpreting Turnitin originality reports. November 2013. https://eat.scm.tees.ac.uk/ bb8content/resources/recipes/interpretTurnitin.pdf. Accessed November 17, 2013
- Committee on Publication Ethics. Serial plagiarism by an experienced author. 2012. http://publicationethics.org/case/serial-plagiarismexperienced-author. Accessed November 10, 2013
- Meddings K. CrossCheck plagiarism screening: what's the magic number. Editorial Office News. August 2011. http://www.ismte.org/ File/Aug11CrossCheck.pdf. Accessed November 10, 2013
- 25. International Committee of Medical Journal Editors. Recommendations for the conduct, reporting, editing and publication of scholarly work in medical journals. December 2013. http://www.icmje.org/ icmje-recommendations.pdf. Accessed November 24, 2013
- Wager E. Committee on Publication Ethics. What to do if you suspect redundant (duplicate) publication. 2008. http://publicationethics. org/files/u2/01A_Redundant_Submitted.pdf. Accessed August 24, 2014
- Editors at BioMed Central. Committee on Publication Ethics. Text recycling guidelines. http://publicationethics.org/text-recyclingguidelines. Accessed August 24, 2014
- Bankier AA, Levine D, Sheiman RG, et al. Redundant publications in Radiology: shades of gray in a seemingly black-and-white issue. *Radiology* 2008;247:605–07
- Committee on Publication Ethics. What to do if you suspect plagiarism: (a) suspected plagiarism in a submitted manuscript. http://publicationethics.org/files/u7140/plagiarism%20A.pdf. Accessed November 10, 2013
- iThenticate. White paper. True costs of research misconduct. http:// www.ithenticate.com/resources/papers/research-misconduct. Accessed August 17, 2012
- 31. Michalek AM, Hutson AD, Wicher CP, et al. The costs and underappreciated consequences of research misconduct: a case study. *PLoS Med* 2010;7:e1000318
- Hoover GA. A game-theoretic model of plagiarism. Atlantic Economic Journal 2006;34:449–54